Virtual Reality for Mobile Knowledge Work

Von der Universität Bayreuth
zur Erlangung des Grades eines
Doktors der Naturwissenschaften (Dr. rer. nat.)
genehmigte Abhandlung

Kooperative Promotion mit der Hochschule für angewandte Wissenschaften Coburg

von

Verena Biener

aus Sigmaringen

1. Gutachter: Prof. Dr. Jörg Müller
2. Gutachter: Prof. Dr. Douglas Bowman

Tag der Einreichung: 08.03.2024

Tag des Kolloquiums: 23.07.2024

Acknowledgements

This work would not have been possible without the support of many people which I want to thank here.

First, I would like to thank my supervisor Jens Grubert for giving me the opportunity to do my PhD at Coburg University and for introducing me to many interesting people and research areas. He taught me how to do research independently and always provided valuable feedback and gave me great advice. Then, I would also like to thank Jörg Müller from Bayreuth University for agreeing to be my second supervisor, for providing valuable and constructive feedback and for introducing new perspectives on my work.

I also want to thank all students and employees of Coburg University and all the other people who participated in our studies. Without them, this research would not have been possible. Here, a special thanks goes to the participants of our long-term study who accepted the challenge of working in VR for a whole week. I especially want to thank Renate Lucke who not only participates in most studies conducted in our lab but also organizes great seminars and workshops for PhD students and always showed genuine interest in my work and provided valuable feedback both on a professional and a personal level.

I would also like to thank all my co-authors who supported me in all stages of my PhD. Eyal Ofek and Michel Pahud, who always had really good ideas on how to advance our prototypes. My collaborators from Cambridge University, John Dudley, Jinghui Hu, and Per Ola Kristensson for their great support when working on multiple publications. I want to particularly thank Per Ola for all the valuable and professional feedback on our work and also for his encouraging and supporting words. Then, I would like to thank my colleagues at Coburg University, Daniel Schneider, Travis Gesslein and Alexander Otte, who welcomed me in Coburg and taught me everything I need to know. I especially want to thank Travis for being an awesome programmer and making our prototypes look really nice. I would also like to thank my

"new" colleagues Snehanjali Kalamkar, Negar Nouri, Forouzan Farzinnejad, Navid Khezrian, Bastian Kuth, and Markus Hülß for all they have done to support me in pursuing my PhD and for the great time we had on various non-work activities. I especially thank Snehanjali and Negar for making work fun every day, for supporting my work, for the great time we had at the conferences in Singapore and Sidney, for providing moral support when needed and especially for not only being colleagues but awesome friends. I also want to thank my Slovenian colleagues Maheshya Weerasinghe, Cuauhtli Campos, Matjaž Kljun, Klen Čopič Pucihar, and all the others that I have met during my time in Koper, for being part of an unforgettable experience. Special thanks to Maheshya who also helped me conduct our long-term study with participants in Slovenia.

I also want to thank Max who always supported me, even though it meant not spending a lot of time with each other. And I would like to thank all of my friends who engaged me in a life outside of the academic world. Finally, I want to thank my parents for providing me with the opportunity to pursue academic goals, and my whole family for supporting me on my way.

Zusammenfassung

In den letzten Jahren sind mobile Arbeitskonzepte immer mehr auf dem Vormarsch. Diese sind im Allgemeinen sehr gut für Wissensarbeiter geeignet, die typischerweise nur einen Computer und eine Internetverbindung benötigen, um ihre Arbeit zu erledigen. Verschiedene Orte können jedoch gewisse Nachteile mit sich bringen, wie zum Beispiel, dass auf kleineren, portablen Geräten gearbeitet werden muss. Im Rahmen dieser Arbeit habe ich in einer Reihe von Studien untersucht wie die Virtuelle Realität (VR) für die Wissensarbeit genutzt werden kann, um solche Probleme zu lösen. Diese Studien haben gezeigt, dass Tätigkeiten, bei denen mehrere Ebenen mit Informationen neu angeordnet werden sollen, von einer dreidimensionalen Darstellung in VR profitieren und, dass der große virtuelle Anzeigeraum in VR die Suche nach einem präattentive Objekt beschleunigen kann. Des Weiteren stellt diese Arbeit multimodale Interaktionstechniken im Kontext der Wissensarbeit vor, die Touchscreens mit Eye-Tracking und einem räumlich getrackten Stift kombinieren and daher auch an beengten Orten genutzt werden können. Diese Interaktionstechniken wurden von den Probanden als brauchbar, nützlich und angenehm empfunden. Zusätzlich hat eine Studie gezeigt, dass die Kombination aus Blicksteuerung und Toucheingabe bei der Interaktion mit mehreren virtuellen Bildschirmen besser abschneidet als eine reine Toucheingabe. Eine Untersuchung zur Nutzung von VR zum Arbeiten über einen längeren Zeitraum, zeigte, dass dies zu schlechteren Bewertungen bei den meisten Messungen, wie zum Beispiel zur Simulatorkrankheit, zur Arbeitsbelastung oder zur visuellen Ermüdung, im Vergleich zu einem ähnlichen physischen Arbeitsplatz führt. Einige Beobachtungen weisen jedoch auch auf Akkommodationseffekte hin. Zusätzlich zeigte eine weitere Studie, dass sich Nutzer während dem Arbeiten an einem öffentlichen Ort mit VR weniger sicher fühlen als mit einem Augmented Reality Gerät oder einem Laptop, da sie ihre Umgebung weniger bewusst wahrnehmen. Insgesamt zeigen diese Ergebnisse, dass es praktikabel ist VR für die mobile Wissensarbeit zu nutzen, es aber immer noch Nachteile gibt, welche die Vorteile über-

wiegen könnten, besonders bei der Langzeitnutzung oder bei der Nutzung in der Öffentlichkeit. Während einige Nachteile nur mit verbesserter Hardware überwunden werden können, könnten andere durch sorgfältig entwickelte Anwendungen, Geräte oder Richtlinien zum Nutzen von VR verringert werden, mit dem Ziel Wissensarbeiter bestmöglich zu unterstützen.

Abstract

In recent years, mobile work concepts have been on the rise which are generally well-suited for knowledge workers, who typically only need some kind of computer and an internet connection. However, different locations could have certain drawbacks such as the need to work on smaller portable devices. In the scope of this thesis, I explored in a series of studies how virtual reality (VR) could be used as a tool for knowledge work to overcome such problems. The studies showed that tasks involving layers of information that need to be reordered can benefit from the three-dimensional visualization in VR and that the large virtual display space in VR can reduce the search time for preattentive targets. Moreover, this thesis presents multimodal interaction techniques in the context of knowledge worker tasks that combine touchscreens with eye-tracking and a spatially tracked pen and, therefore, can be used even in confined spaces. These techniques were found usable, useful and enjoyable by participants. In addition, a performance study showed that combining eye-gaze and touch outperforms a touch-only interface when interacting with multiple virtual displays. However, studying the use of VR for work for extended periods of time showed that it leads to worse ratings for most measures, such as simulator sickness, task load, or visual fatigue, compared to a similar physical setup. Yet, some observations also suggested accommodation effects. In addition, another study indicated that when working in VR, users feel less safe in a public environment than with an augmented reality device or a laptop because they are less aware of their surroundings. Overall, these results show that it is feasible to use VR for mobile knowledge work, but there are still drawbacks, especially for long-term use or use in public scenarios, that could outweigh the benefits. While some drawbacks can only be overcome with improved hardware, others could be reduced by carefully designing applications, devices, and guidelines for using VR, with the goal of supporting knowledge workers in the best possible way.

Contents

Chapter 1

Introduction

In recent years, more flexible and mobile work concepts are gaining popularity, fueled also by the COVID-19 pandemic which required more people, many of them knowledge workers, to work from home instead of offices [4]. The term "knowledge worker" was first used by Peter Drucker in 1959 [29] and is used throughout this thesis to describe workers whose main resource is their knowledge and who typically do not need access to special equipment other than computers and an internet connection. Even though they typically work in offices, they can work from different places as long as they have access to a computer, which can also be portable. This flexibility in choosing the workplace can have several benefits, such as taking care of family members, saving time on commutes, or using time spent in public transportation productively [9].

However, there can also be certain limitations when working outside the office. The available hardware can be limited, such that the workers need to use laptops, tablets, or even phones instead of large screens or a multi-monitor setup. Furthermore, input modalities might be limited to smaller keyboards, touch keyboards on tablets, or a touchpad instead of an external mouse. Such limitations could potentially impact productivity. Also, in spaces not designed for work, there could be distractions from the environment such as noise, or from other people, for example, other passengers on public transportation [76].

With new virtual reality (VR) devices being launched in recent years, many of them being affordable to the general population, a technology becomes available that could potentially change knowledge work and help to overcome certain limitations of mobile work. The majority of work presented in this thesis was done using VR devices. However, many findings are potentially transferable to other extended reality devices. Extended reality (XR) covers all immersive experiences ranging from VR to augmented reality (AR), including anything in between these two on the reality-virtuality continuum [74]. In AR, the real, physical environment is augmented with virtual content, while virtual reality VR provides a purely virtual experience even though true virtual reality that stimulates all senses with virtual content is currently not achievable, as mentioned by Skarbez et al. [101]. Therefore, with VR I refer to systems that mainly show virtual content even though parts of the physical environment could be included by recording them with cameras attached to the device. This means, with VR I mainly refer to the device, which can totally block out all visuals from the physical world.

XR has already been discussed as a tool for knowledge work for quite some time [89, 92]. This means, XR and specifically VR, could transform any environment into an office in which the user could potentially be more productive. For example, Ruvimova et al. [94] have shown that VR can be used to suppress visual and audio distractions to increase performance. However, VR can not just shield the user from distractions, but it could also further optimize the work environment, even beyond the limitations of a physical environment. These optimizations could reduce stress [5, 108], increase productivity [27], or extend or replace the physical space with virtual elements that could support the work tasks. For example, VR, or XR in general, can extend a small mobile screen with one or multiple large virtual screens [82] or totally replace physical screens as visualized in Figure 1.1 a). Also, the ergonomics of a workplace could be improved, as there are no limitations on how to position screens or other objects [64]. Such virtual work environments could provide a consistent work setting regardless of the physical environment, but also be adapted easily to varying requirements.

VR also opens up new possibilities on how to display and interact with information. Instead of the traditional two-dimensional (2D) display, it provides a large three-dimensional (3D) space that can display information in 3D as visualized in Figure 1.1 b). This has, for example, been used by Einsfeld et al. [32] and Dengel et al. [25], who have visualized documents and semantic relations between documents on a stereoscopic monitor. This opens up new possibilities on how to display, interact and work with information when the whole three-dimensional space around the user is available. For example, it is possible to actually view 3D models in a three-dimensional environment and interact with them, or arrange 2D information on planes at arbitrary positions in the 3D space.

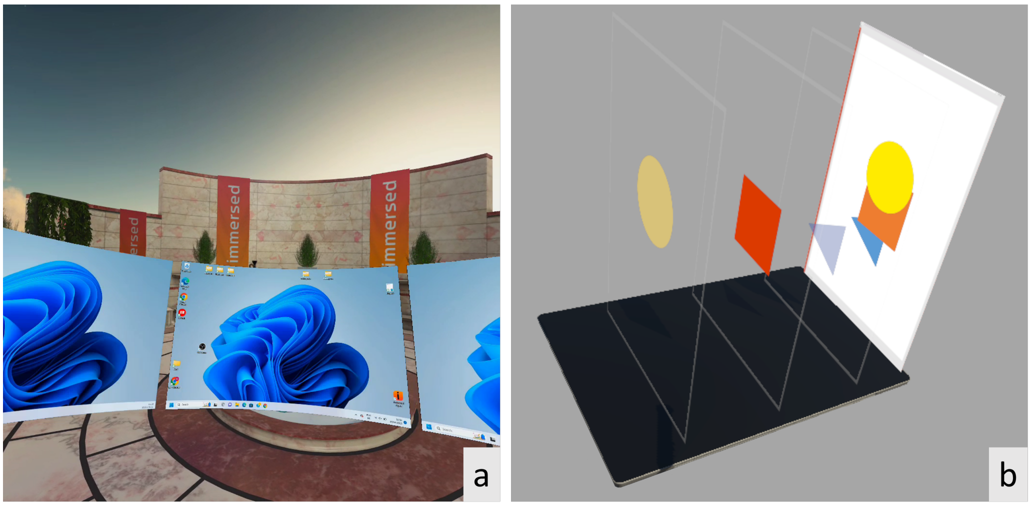

Figure 1.1: (a) The large display space in VR can, for example, show multiple windows. This image depicts the view in the VR app "Immersed" [53]. (b) The 3D display in VR can be used to display information in 3D, such as the different layers containing various shapes as displayed here (see also Chapter 5 [14]).

In addition, VR has the potential to connect workers from remote places and immerse them in a shared virtual environment, which is already possible using applications such as Mozilla Hubs [52] or Meta Horizon Workrooms [51]. This means colleagues can meet in a virtual space while working from remote locations. In such virtual spaces, users could potentially interact more naturally than through traditional technologies such as phones or video-conferencing systems. Employing VR instead of traveling for meetings could also be more environmentally friendly [83] and could potentially reduce travel costs.

When talking about VR using a head-worn display, as I do in this work, it also offers the benefit of increased privacy. As the display is generally only visible to the user, there are no problems in viewing sensitive data in the presence of other people [41]. In addition, it provides further possibilities to enhance security as shown by Schneider et al. [96], who changed the layout of a physical keyboard through VR so that bystanders can not monitor typed characters to gain the true password. Also, a row of applications already exist that show how VR can support certain tasks in areas such as data analytics, programming or creative tasks [121, 33, 27, 77]. In addition, VR has also been used in industry for several years [12].

Finally, VR also supports further interaction possibilities beyond what is currently commonly used by knowledge workers. This includes techniques such as eye-tracking [49, 86], hand-tracking and gestures [2], or spatially tracked devices [37]. These techniques can of course also be combined with traditional ones, including touchscreens, keyboards, and mice. For example, Surale et al. [105] used a tablet to interact with objects in VR and Schneider et al. [96] showed how VR can be used to enhance the capabilities of a physical keyboard. Similarly, AR HWDs can be combined with a smartwatch, as shown by Grubert et al. [40], with a smartphone to enlarge it's display space, as presented by Normand et al. [78], or to enhance each other bidirectionally, as shown by Zhu et al. [120]. Figure 1.2 visualizes how a pen (a), a touchscreen (a), a keyboard (b) and hand gestures (c) can be included into VR.

Yet, there are also challenges to overcome before VR can become a widely-spread tool for knowledge work. First, there are open questions such as how to properly transfer and display current 2D information in a 3D space. This will also require new interaction concepts to efficiently interact with the large three-dimensional display space. Many current interaction techniques for VR are based on in-air interactions with gestures or controllers, which might not be suitable for knowledge workers, as they can cause fatigue [47], lack social acceptability, or there might not be enough space to execute them in certain settings [65]. This also touches on a second aspect about how users might feel when using VR for work in public spaces such as public transportation, how other passengers would react, and how the VR-user can be aware of safety-relevant events happening in the physical world [65]. Also, even though VR devices, and also other XR devices, have evolved significantly in the recent years, there are of course still issues with the hardware, such as their weight, limited resolution, narrow field of view, and fixed focal distance [41]. The effects of using such devices, especially for longer periods of time, which would be necessary to do serious work, are not well explored.

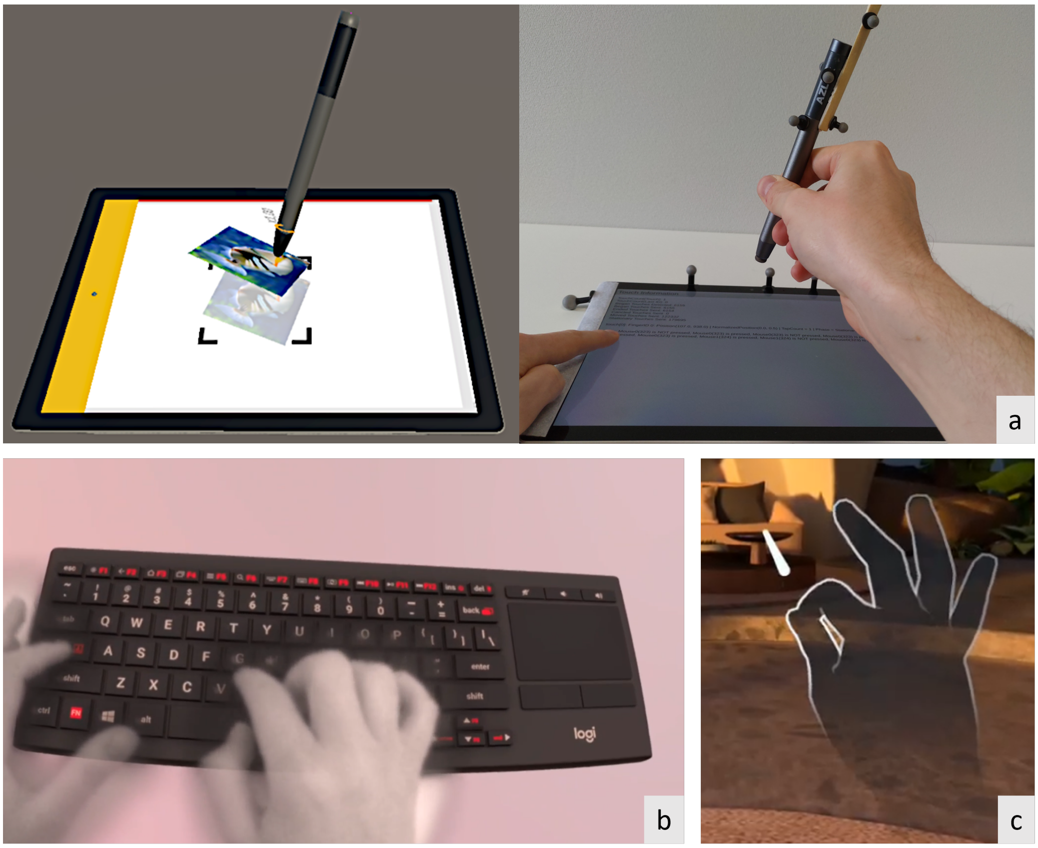

Figure 1.2: (a) The pen on the right is spatially tracked and displayed in VR, as seen left, and used to manipulate virtual objects (see also Chapter 5 [14]). (b) A keyboard and hands can be tracked by the XR device (in this case a Quest 2 [71]) and displayed in the virtual environment. (c) A hand is tracked by the device (Quest 2 [71]) and displayed in VR and can be used to perform gestures.

1.1 Research Questions

Reviewing the existing literature has provided me with an understanding of how knowledge work might benefit from VR, or more generally XR, and which aspects of this research area are still not well explored. A broad overview of this literature can be found in Chapter 3 [17]. Considering the potential benefits that VR could have for mobile knowledge work and the shortage of literature that is investigating them, this thesis aims at exploring them more closely. This is an important step to gain insights that can inform the design of VR applications that are capable of effectively supporting mobile knowledge work. Also, reviewing the literature showed that ecologically valid studies on the use of current VR devices for knowledge work are very scarce, especially regarding long-term usage. It is crucial to also explore this and uncover possible issues to develop guidelines that can support the design of new systems and provide recommendations for a responsible use of VR for knowledge work.

Therefore, I defined several research questions that are described in detail in the following subsections and can be grouped into two categories. First, I examine how mobile knowledge work can be facilitated using VR, and second, I evaluate how the use of VR devices in realistic use-cases affects knowledge workers.

1.1.1 Facilitating Mobile Knowledge Work Using VR

As described above, prior work suggests several advantages of using VR or XR for mobile knowledge work, such as a larger display space [64, 82], the possibility to visualize information in 3D [32, 25], new interaction modalities, including eye-tracking [49] and spatial interactions [2, 37], and the opportunity to personalize and optimize the virtual work environment beyond what is physically possible [94]. However, for the purpose of facilitating knowledge work through the use of VR, the potential benefits of the larger screen space and the three-dimensional display should be examined in the context of knowledge worker tasks. This is important to gain a good understanding of how and to what extend they can support knowledge work and to provide examples of how they can be effectively utilized. Similarly, as will be described in Chapter 3 [17], it is essential to find good interaction techniques to efficiently interact with the large 3D space, especially in a mobile context in which only a restricted input space is available.

What are the advantages of a large 3D display for knowledge work in VR? (Chapter 9.1.1)

(a) What are the benefits of viewing layers of information in 3D compared to showing them in 2D, one at a time? (Chapter 4 [18])
(b) What are the benefits of visualizing layered information in 3D compared to state-of-the-art 2D visualizations? (Chapter 5 [14])
(c) Can the extended display space in VR help users find targets faster? (Chapter 5 [14])

Are multimodal techniques that combine VR and established devices feasible for interacting with large 3D spaces? (Chapter 9.1.2)

(a) Does touch combined with gaze have benefits over touch-only when interacting across multiple screens? (Chapter 4 [18])
(b) Are applications that use the large 3D display of VR and multimodal interaction feasible? (Chapter 4 [18])
(c) Are multimodal techniques feasible that combine eye-tracking, a spatially tracked pen, touch, and information displayed in a large 3D space? (Chapter 5 [14])

Therefore, in a first step, this thesis looks at *the advantages of a large 3D display space in VR* in the context of mobile knowledge work. Such displays have two major properties. First, they provide a large display space that covers every direction in which users could look, as if they are looking at the surface of a sphere placed around their head. Second, they enable the presentation of three-dimensional information, which increases the display space from the surface of the previously mentioned sphere to an infinite three-dimensional room around the users. This provides the opportunity to display information at different depth levels.

Even though VR provides us with the possibility to display three-dimensional data, the prototypes presented in this thesis mainly focus on displaying two-dimensional data in a 3D space. I chose to do so to support familiar concepts, as current knowledge work is done in 2D and I want to maintain compatibility with existing applications and devices. Still, layers of 2D information can be presented to the users at different depth levels in 3D, which allows them to see multiple layers at once, although they might partially occlude each other. Therefore, one of the research questions addressed in Chapter 4 [18] is about evaluating the *benefits of viewing layers of information in 3D compared to a 2D visualization*, which means only showing one information layer at a time.

As the objective is to facilitate knowledge work through VR beyond what is possible with currently use mobile devices, it is also important to compare VR with current baseline systems. Therefore, one of the studies presented in Chapter 5 [14] *compares state-of-the-art 2D visualization techniques with a 3D visualization of layered information* in the context of authoring presentations. Authoring presentations was chosen as an example application that is commonly used by knowledge workers [90].

Independent of the capability to display information in 3D, VR also provides a very large display space which can be used to display much more information than on a physical mobile screen. This can potentially provide users with a better overview of their data and allow them to identify relevant information faster. To quantify this potential benefit, Chapter 5 [14] presents a study to explore the advantages of an increased display space in VR and specifically *if it can help users find targets faster.*

As will also be discussed in Chapter 3 [17], it is not trivial how to efficiently interact with the large 3D space in VR, given that the input space is generally much smaller. In addition, common interaction techniques for VR, such as gestures and controllers, might not be well suited for knowledge work because they can lead to fatigue [47], certain environments could lack the necessary space, and the gesture could lack social acceptability [65]. Therefore, the prototypes presented in this thesis focus on established devices that can be used in confined spaces, require little physical effort, and are socially acceptable. As these requirements are fulfilled by tablets, they were used in the studies described in Chapter 4 [18] and 5 [14]. However, the relatively small size of the tablet screen intensifies the problem of how to efficiently map the small input space to the large output space provided by VR. Also, it is not trivial how to support three-dimensional interactions on a 2D tablet screen. That is why I decided to design multimodal techniques that combine touch input with other techniques offered by VR that do not require much space or induce high fatigue. This includes techniques such as eye-tracking and spatially tracked pens that can be used to interact three-dimensionally, as their movement can not only be tracked on the screen but also in the air above it. In addition, multimodal techniques have reported to be beneficial, as they provide a higher flexibility to choose a modality depending on users needs and situations [111]. Therefore, I am exploring *if multimodal techniques that combine VR with established devices, that can be used in confined spaces and are already socially accepted, are feasible for interacting with large 3D spaces.*

I have already mentioned that interacting with this large screen-space while only using a small input area can be challenging. While a touchscreen can provide very accurate input, there is no trivial solution for how it can be used to efficiently switch between areas of information on a large screen. Eye-gaze-tracking on the other hand, which is implemented in many current VR devices, complements the touchscreen. It is less accurate but very efficient to switch between areas of information on large screen spaces because users simply need to look at the area of interest. The combination of touch and gaze has already been explored outside of VR or XR [85, 84] and shown to provide benefits, such as less physical effort. Therefore, Chapter 4 [18] presents a study that examines *if combining touch and gaze input has benefits over touch-only when interacting across multiple screens.*

To provide more concrete examples of how such interaction concepts could benefit knowledge work, Chapter 4 [18] also explores the *feasibility of several knowledge worker applications that use the large 3D display and multimodal interaction.* To provide a more sophisticated example of how knowledge work in VR could look like, Chapter 5 [14] presents a prototype for presentation authoring, representing a common knowledge worker task [90], and evaluates the *feasibility of multimodal techniques that combine eye-tracking, a spatially tracked pen, touch, and information displayed in a large 3D space.*

1.1.2 Effects of Current VR Devices on Knowledge Workers in Realistic Use-Cases

As described in the previous subsection, I have explored how knowledge work can be supported through VR by developing prototypes using current VR devices and additional technologies such as an Optitrack tracking system [81] for robust tracking of hands and devices such as tablets or pens. However, these prototypes are very expensive and are only feasible in a lab-based environment due to the external tracking devices. Current off-the-shelf devices, which often are portable and could therefore be used anywhere, might not provide all of these functionalities at the necessary level of precision. For example, Schneider et al. [95] found that finger-tracking accuracy greatly varies between devices. Therefore, the question remains if working in VR with currently available devices is feasible and how using such devices would affect knowledge workers in more realistic scenarios. This is a very important question, as it indicates the feasibility of current devices as a serious tool for knowledge workers and it can uncover issues that should be considered when designing the next generation of devices.

How do knowledge workers respond to prolonged use of VR?

(a) What are the psychological and physiological effects of working in VR for one week compared to a standard physical setup? (Chapter 6 [16], Chapter 9.2.1)
(b) How does participants' behavior differ between VR and a physical environment during long-term use of VR for work? (Chapter 7 [13], Chapter 9.2.1)
(c) How do behavioral patterns evolve over time when working in VR for extended periods of time? (Chapter 7 [13], Chapter 9.2.1)

What are the implications on users and bystanders when using XR for work in public?

- How do bystanders react when seeing an XR user in public compared to a laptop user? (Chapter 8 [15], Chapter 9.2.2)
- How does using XR or a laptop affect users when working in public? (Chapter 8 [15], Chapter 9.2.2)

Prior work in the area of VR for knowledge work has mostly examined only relatively short usage times, such as was also done in Chapter 4 [18] and 5 [14]. In addition, current literature on the prolonged use of VR is very scarce as also indicated in Chapter 3 [17]. So far, the longest systematic study by Guo et al. [42, 43] lasted 8 hours and otherwise there have only been anecdotal reports [20, 56] or a 24 hour self-experiment [104]. Therefore, this thesis reports on the, so far, longest systematic study in VR, lasting five days, which examined *how knowledge workers respond to prolonged use of a VR device.* In a first step, Chapter 6 [16] looks at the *psychological and physiological effects of working in VR for one workweek compared to a standard physical setup*, by closely monitoring participants, mainly through repeated questionnaires. Therefore, this article mainly reports on participants' subjective feedback. In a next step, Chapter 7 [13] reports on a video analysis to evaluate *how participants' behavior differed between VR and the physical environment* and *how behavioral patterns evolved over time when working in VR for extended periods of time.*

Even though studying the user behavior during a full week provides valuable insights, they are still gained in a lab-based study. As mentioned before, VR could be especially useful for mobile knowledge workers which often includes scenarios in public places. Therefore, it is important to explore the effects of using VR in public. As AR devices offer similar benefits as VR while providing the possibility to see the real surroundings and all people and objects in it, they might lead to significantly different experiences in a public setting, which can also involve social interactions, than VR devices. Therefore, the last study considered both AR and VR to identify *what the implications on users and bystanders are when using XR for work in public.* Several studies have explored bystanders' reactions by showing participants videos of XR users [87, 98], while others actually evaluated XR prototypes in public such as on the street [62] or in a cafe [110]. However, this has not yet been explored in the context of knowledge work and also there has not been a comparison between different XR devices and an already commonly used device, such as a laptop, in public places. This, however, is important to evaluate if XR is a feasible choice as a mobile tool for knowledge work. Therefore, Chapter 8 [15] investigates *how bystanders react when seeing an XR user in public compared to a laptop user* and also, *how XR or a laptop affects the user when working in public.*

1.2 Contributions

This thesis explores the field of VR for knowledge work in two directions that are also reflected by the structure of the research questions. After providing an overview of current literature in the area of VR for knowledge work and exploring how VR and multimodal interaction techniques can benefit knowledge work, this thesis investigates the effects on users when working with current VR devices in realistic use-cases.

First, this thesis contributes towards a better understanding of how mobile knowledge work can be facilitated through VR, by implementing several prototypes (Chapter 4 [18], Chapter 5 [14]). Chapters 4 [18] and 5 [14] show that the three-dimensional representation in VR, for example, when reordering layered information, can be advantageous and outperforms current state-of-the-art two-dimensional methods with regard to task completion time and usability. In addition, Chapter 5 [14] shows that the large virtual display space provided by VR can support search tasks by making them faster if the target is visually distinct from other information. In addition, both Chapters 4 [18] and 5 [14] show that users can benefit from multimodal interaction techniques that combine standard input modalities, such as touch, with new modalities provided by VR, such as eye-tracking or spatially tracked pens. Specifically, Chapter 4 [18] of this thesis contributes six diverse applications, such as a map navigation or window manager, and a novel input technique to efficiently interact across multiple screens by combining touch and eye-gaze. This technique proves to be faster, has a higher usability and leads to a lower task load and simulator sickness than a touch-only technique. In addition, Chapter 5 [14] contributes four techniques in the context of presentation authoring, including techniques for object manipulation and for creating animations, that illustrate how such techniques could be employed in future applications and show that participants find them useful and enjoyable.

Second, this thesis reports on the effects on users when working in VR, both for extended usage times (Chapter 6 [16], Chapter 7 [13]) as well as in public environments (Chapter 8 [15]). Most studies are only conducted for short periods of time and, so far, reports on long-term effects have been scarce in VR research. Therefore, the study described in Chapter 6 [16] provides new unique insights into how VR affects the users and how users react and cope with using the device for prolonged times. These findings show that for subjective measures, such as usability, frustration, or wellbeing, VR performs worse than a comparable physical setup. Observations from analyzing the videos obtained in this long-term study show that the behavior of participants changes over time such that they adjust the headset less at the end of the week and also take fewer but longer breaks. In addition, the videos showed that even though participants mostly kept the HWD on while interacting with the physical world by performing actions such as drinking, talking, or using the phone, they performed some of these actions less frequently than they did in the physical setup. This indicates that wearing VR HWDs makes participants deviate from their usual behavior. These combined findings can be used as a baseline for subsequent research in this area and I hope that it also encourages other researchers to conduct studies over longer periods of time. These insights can also inform the design of more ergonomic and comfortable HWDs and software that actively supports the users wellbeing. Additionally, a majority of VR and XR studies is conducted in closed lab environments. However, it is important to also assess systems in the environments in which they are envisioned to support users, which in this case includes public spaces, even though XR devices are still not very common in public. Therefore, Chapter 8 [15] of this thesis contains the results of a study that was conducted in a public cafeteria, which confirms that using XR devices in public spaces still makes users stand out. In addition, the results indicate that participants feel safer when working with an AR instead of a VR device because they are more aware of their physical surroundings. These findings can motivate the design of better systems that help users feel safe while using XR in public.

Overall, this thesis proposes concrete examples of how VR can be used to support mobile knowledge work and presents several prototypes that show how applications for knowledge work in VR can look like. In addition, this thesis provides evaluations of currently available HWDs in realistic use-cases, which show that at this stage long-term use could be problematic and there are still challenges when using XR in public. Viewed at together, these results provide an overview of benefits and challenges of using VR for mobile knowledge work which can contribute to the design of new systems and guidelines on how to efficiently use VR as a tool for knowledge work.

1.3 Thesis Outline

The thesis is structured as follows. Chapter 2 describes the methodology used for this research, which includes the design and implementation of the prototypes as well as the evaluation procedures and data analysis.

Chapter 3 [17] contains an article describing a thorough literature review of the area of VR, and also XR, for knowledge work.

The following Chapters 4 to 8 include the research articles that address different aspects of the previously mentioned research questions. Chapter 4 [18] evaluates the benefits of multimodal interaction and 3D visualization of layered information. Similarly, Chapter 5 [14] also addresses these topics, but in a more practical aspect, by presenting techniques for authoring presentations in VR.

Then, Chapter 6 [16] describes a study in which participants were using VR for work for a whole workweek (40 hours) and the effects of this extended use compared to a physical work environment. Chapter 7 [13] builds onto that by analyzing videos, recorded in the previous study, to gain a more nuanced understanding of how users behave during such a prolonged use of VR. Finally, Chapter 8 [15] reports on a study that was conducted in a public space by comparing different XR devices and a standard laptop and their effects on users and bystanders.

Then, the results of the previous chapters are discussed jointly in Chapter 9 and a final conclusion is presented in Chapter 10.

Chapter 2

Methodology

This chapter gives an overview of the methodology used in this thesis and the six articles included in Chapters 3 to 8. As the nature of these articles varies, different methodology was used which is specified in the following.

2.1 Literature Review

I conducted thorough, yet not strictly formal and complete literature reviews for all articles included in this thesis. Relevant literature was identified by using different combinations of search terms in multiple digital libraries, such as IEEE Explore, Google Scholar, and ACM Digital Library. This search was not limited to certain conferences, journals, or years of publication. In addition, I also included relevant papers that were referenced in other papers. These were then summarized to present the current research status in a certain area and to identify how my work can tie in with prior research.

A particularly extensive literature review was conducted for the article presented in Chapter 3 [17]. For this article papers were found to be relevant, if they met the criteria of including some kind of XR technology and are connected to the field of knowledge work. All identified papers were then clustered with respect to their main topics to summarize the current state of research and possible challenges.

2.2 Design and Prototyping

The prototypes that are described in the articles presented in Chapters 4 [18] and 5 [14] were designed to address some of the previously mentioned research questions. The design and implementation of these prototypes was an iterative process with formative evaluations [50] that involved repeated informal, internal testing and improvement to optimize certain parameters. Such parameters were, for example, the control-to-display gain for switching screens in Chapter 4 [18] or the amount of rotation for the occlusion handling technique in Chapter 5 [14]. I generally developed and implemented novel techniques. In addition, I usually also re-implemented existing or state-of-the-art techniques, for example, the standard PowerPoint reordering technique in Chapter 5 [14]. In a next step, these techniques could then be compared against each other.

The process for the articles of Chapters 6 [16] and 8 [15] differed from the previously described approach, as I used off-the-shelf hardware and software, and therefore, no implementation was needed. However, preparing the user study still involved an iterative process of internal testing and adjusting different options in hard- and software to set up the experiment. This included, for example, deciding on the type of HMD or selecting software that provides all features necessary for the study.

2.3 Evaluation

Except for Chapter 3 [17], all articles report on a summative evaluation of the proposed prototype or system. The general process for that is explained in the following.

2.3.1 Study Design

All prototypes in Chapters 4 [18] and 5 [14] and the setups described in Chapters 6 [16] and 8 [15] were evaluated in user studies. I always decided to do a within-subjects design. This helped me to control and lower the risk of individual differences between participants, which could lead to errors. In addition, it also enabled participants to compare or even rank their experiences in the different conditions. To avoid order-effects, which can be a problem in within-subjects designs, I counterbalanced the order of the conditions as far as possible, commonly using Latin squares.

Most studies were conducted in a lab in Germany, but due to Covid restrictions, parts of the study in Chapter 4 [18] were conducted in a home in the US. Due to difficulties in finding a large enough number of participants for the very long study presented in Chapter 6 [16], six participants were recruited in other labs in the UK (2) and Slovenia (4). Only the last study, described in Chapter 8 [15], was conducted outside of the lab, namely in a university cafeteria, as the objective was to study the implications of using XR in public. For Chapter 7 [13], I analyzed video data previously collected in the study of Chapter 6 [16], and therefore, no additional study was conducted.

2.3.2 Procedure

All studies generally started with the participant filling out a consent form and a demographic questionnaire, including questions about age and gender as well as their experience with study-relevant devices such as VR-HWDs or touchscreens. Participants were also reminded that they can quit the study at any time or take breaks whenever they need to.

For the studies that include performance evaluations (puzzle task and content transfer task of Chapter 4 [18] and search task and reordering task of Chapter 5 [14]), the following sequence was repeated for all conditions. First, the technique was explained and participants had some time to test it and get familiar before they started the actual test in which their performance was recorded. Then, they answered some questionnaires before continuing with the next condition. The study described in Chapter 8 [15] followed the same principle, but due to the length of one condition, they were distributed to three different days. The process was also slightly different for the study described in Chapter 6 [16], as one condition lasted a whole week. Therefore, questionnaires were answered in regular intervals throughout the week.

The articles of Chapters 4 [18] and 5 [14] additionally report on usability studies in which participants were walked through the functionalities of the prototypes and could explore them without completing certain tasks or recording any performance measures, similar to other prior work [23]. Afterwards, participants were then answering some questionnaires about their experiences with the prototype.

After finishing all conditions, participants were usually asked to rate the conditions and a short semi-structured interview [1] was conducted to gain more insights into how participants experienced the techniques. Participants were usually compensated with a gift card or, if they were employees of the university, could participate during paid worktime.

For the article presented in Chapter 7 [13], I did not run a user study but analyzed the videos previously recorded in the study of Chapter 6 [16]. The participant's behavior in the videos was annotated through a process of open and axial coding [117]. Finally, this resulted in quantitative data that could be analyzed as described in the following.

2.4 Data Analysis

Through the user studies, I generally collected both quantitative and qualitative data. The process of analyzing this data is described in the following.

2.4.1 Quantitative

If applicable, I collected different objective measures, such as task completion time, errors, accuracy, or heart rate. In addition, a range of subjective measures was obtained through questionnaires. Commonly used questionnaires were the system usability questionnaire (SUS) [19], NASA Task Load Index (NASA TLX) [44], and simulator sickness questionnaire (SSQ) [55]. Yet, also other questionnaires were used as needed in the specific study, such as questionnaires to measure flow [93], presence [97], anxiety [122], visual fatigue [46], and positive and negative affect [115]. To gain insights into certain aspects that are not covered by any established questionnaires, I also added individual questions myself, often inspired by related prior work.

As all studies involved either more than two conditions or multiple independent variables, I generally used a repeated measures analysis of variance (RM-ANOVA) to statistically analyze the data. Non-normal data was generally log-transformed before running RM-ANOVA and for multiple comparisons in post hoc tests, I used t-tests with Bonferroni adjustments. For subjective data from questionnaires and for data that could not be normalized through log-transformation, I used aligned-rank-transform (ART) [118] before appying RM-ANOVA. For this kind of data, I then used the Wilcoxon signed-rank test in post hoc comparisons, also with Bonferroni adjustments. More detailed descriptions of the statistical methods that were used can be found in the respective chapters.

2.4.2 Qualitative

At the end of each study, I typically conducted a short semi-structured interview [1]. I asked participants some predefined questions, usually about their preferences, what they liked or disliked, or if they had any suggestions on how to improve the systems. However, I would also react to the participants' answers to get more detailed statements or ask about certain observations made during the study.

Following an inductive approach [107], answers were then coded and clustered into groups of similar statements to present an overview of the most common responses for each question. Certain answers that were very representative of a group of statements or that I found particularly interesting or surprising were also directly cited in the papers. These answers were then generally also used to explain the results of the quantitative measures.

Chapter 3

Extended Reality for Knowledge Work in Everyday Environments

Summary

This first article presents an overview of current research on the topic of extended reality for knowledge work. It contains a literature review, including papers that are related to the field of knowledge work and extended reality. More specifically, these papers contain the keywords "virtual reality" and "augmented reality" in combination with work-related keywords, such as "knowledge work", "text entry", "collaboration", "office environment" and "long-term use". After collecting relevant papers, they were grouped into several categories representing important research areas in this field. This included interaction techniques that allow users to interact with 2D content in three-dimensional space, but also text-entry techniques, as text entry is an important part of knowledge work. The article also describes how extended reality can support collaboration between both collocated and remote workers and also how virtual environments can help to manage stress and increase productivity. In addition, current research on the long-term use of XR devices is reviewed and also specific applications that can support knowledge work. However, there are still many open questions in all these areas which need to be addressed in future work. Some key aspects are examined closely in the following chapters of this thesis, such as the feasibility of combining traditional devices with virtual reality, how to support specific knowledge worker tasks, using the authoring of presentations as an example, and finally, quantifying the effects of using virtual reality for extended periods or in public spaces.

Contribution Statement

The approach for the literature review presented in this article was discussed and agreed upon by all authors. The literature review, identifying relevant papers in the specified area, was mainly conducted by Verena Biener with some additions by all other authors. The classification of the papers and the structure of the article was proposed by Verena Biener and discussed and finalized by all authors. Figures were created by Verena Biener with Jens Grubert and Per Ola Kristensson providing some image-material. Verena Biener drafted the initial version of the article and rewriting was done by all authors.

Article [17]

Verena Biener, Eyal Ofek, Michel Pahud, Per Ola Kristensson, Jens Grubert. Extended Reality for Knowledge Work in Everyday Environments. In: Simeone, A., Weyers, B., Bialkova, S., Lindeman, R.W. (eds) Everyday Virtual and Augmented Reality. Human–Computer Interaction Series. Springer, Cham. 2023. `https://doi.org/10.1007/978-3-031-05804-2\_2`

Extended Reality for Knowledge Work in Everyday Environments

Verena Biener, Eyal Ofek, Michel Pahud, Per Ola Kristensson, and Jens Grubert

Abstract Virtual and augmented reality (*VR* and *AR*) have the potential to change information work. The ability to modify the workers senses can transform everyday environments into a productive office, using portable head-mounted displays (*HMDs*) combined with conventional interaction devices, such as keyboards and tablets. While a stream of better, cheaper, and lighter HMDs has been introduced for consumers in recent years, there are still many challenges to be addressed to allow this vision to become reality. This chapter gives an overview of the state of the art in the field of extended reality for knowledge work in everyday environments, identifies challenges and proposes steps to address the open challenges.

1 Introduction

Extended reality (XR) covers a spectrum of diverse technologies ranging from augmented physical (*AR*) environments to fully virtual (*VR*) environments. While XR technologies have been studied for decades in laboratory settings, the shift toward affordable consumer-oriented products allows exploring the positive and negative qualities of such technologies in everyday environments. Among diverse application

---

V. Biener (✉) · J. Grubert
Coburg University of Applied Sciences and Arts, Coburg, Germany
e-mail: verena.biener@hs-coburg.de

J. Grubert
e-mail: jens.grubert@hs-coburg.de

E. Ofek · M. Pahud
Microsoft Research Redmond, Redmond, USA
e-mail: eyalofek@microsoft.com

M. Pahud
e-mail: mpahud@microsoft.com

P. O. Kristensson
University of Cambridge, Cambridge, UK
e-mail: pok21@cam.ac.uk

A. Simeone et al. (eds.), *Everyday Virtual and Augmented Reality*, Human–Computer Interaction Series, https://doi.org/10.1007/978-3-031-05804-2_2

domains, spanning entertainment, medical and industrial use (Slater and Sanchez-Vives 2016; Billinghurst et al. 2015), supporting knowledge work has attracted increasing interest in recent years.

The notation '*Knowledge Work*' follows the definition initially coined by Drucker (1966) where information workers (or *IW*s) apply theoretical and analytical knowledge to develop products and services. Much of the work might be detached from physical documents, artifacts, or specific work locations and is mediated through digital devices such as laptops, tablets, or mobile phones, connected through the Internet. This unique nature of knowledge work enables better mobility—the ability to work far from the physical office—and raises new possibilities to both overcome difficulties and enable services that are natural for physical offices, such as streamlined communication and collaboration and environments designed for creativity, privacy, and other factors.

In this chapter, we intend to explain our vision that moving the worker further into a digital immersive environment opens up options not limited by the physicality of the worker's devices and environments. We hypothesize that a virtual work environment enables workers to do more than they could before, with less effort, and may also eventually level the playing field between workers regardless of their physical locations, physical limitations, or the quality of their physical work environment. Doing knowledge work in XR is still an emerging field and current devices entail quite a few drawbacks compared to current physical work environments. Therefore, we summarize both existing research on supporting knowledge work through XR technologies and challenges that should be addressed to move the field forward.

The focus of this work is to explore research in the area of extended reality that takes place in a knowledge work context.

Our search includes diverse papers from journals and conferences and is not limited to certain venues or years. Most papers were found through ACM Digital Library, IEEE Explore, and Google Scholar. We did not conduct a strictly formal and complete literature review, as the topic is an emerging field and terminology is not yet consistent enough to find all relevant papers with a predefined set of keywords. The following search terms were used in different combinations: virtual reality, augmented reality, knowledge work, text entry, collaboration, office environment, and long-term use. These were used in the advanced search engines of aforementioned databases. Additionally, we considered papers that referenced, or were referenced, by relevant papers, that would not have been included through the keyword search alone. We primarily examined the title and abstract to decide, based on our own expertise in the field, whether a paper was relevant. We used the following criteria: (1) the paper included an XR technology which means systems or devices that enhance or replace the physical surrounding, and (2) the paper had a connection to the field of knowledge work.

This chapter is divided into several sections, each investigating a different aspects of knowledge work in XR. As this is an overview of literature in an emerging field the literature does not provide an agreed upon set of sub-categories. We identified our subsections during the literature review process, by clustering the papers according to their main contents. Section 2, *Interaction Techniques*, reviews research on interaction techniques that can facilitate knowledge work in XR, including techniques for text entry (Sect. 2.3)—a crucial task in knowledge work. Section 3, *Collaboration*, explores the collaborative aspects of information work, which are crucial to allow the worker to maintain group working from remote and new environments. Section 4, *Environment*, considers the influence of the environment on the knowledge worker and how XR can help optimize it, as well as the social implications arising from the use of XR. Section 6, *Application*, presents practical applications of XR in the area of knowledge work. Finally, we envision the worker to use the XR environment as an alternative to the limiting physical environment, resulting in the use of XR for extended time periods. Therefore, Sect. 5, *Long-term Immersion*, provides insights into the current research status in this area. Toward the end of this chapter, we provide a summary, discuss future challenges, and synthesize our main conclusions.

2 Interaction Techniques

Envisioning work in XR spaces, we look at ways that information workers are doing their task today, and how they can be done in XR space. The transition to XR space, where the user's senses are being modified by devices such as head-mounted displays (HMDs) and different interfaces for the hands are challenging and require new interaction concepts (Bowman et al. 2004). Some tasks such as typing that are crucial in today's work are dependent on good combination of senses (such as vision, proprioception, and haptics).

Another aspect of interest is embedding the, mostly 2D, information work in three-dimensional space. The use of a 3D immersive space enables several capabilities that were not available to workers using standard 2D monitors and input devices, such as a very wide display space around the user, a depth display that is not limited to one plane as most monitors, and storing data in the same space as the user's body, enabling new and direct methods to interact with data using natural gestures (Fig. 1). Different works look at how to efficiently map user's 3D motions in space to the 2D space of the task (Andujar and Argelaguet 2007; Brasier et al. 2020) or documents (e.g., spreadsheets or letters), while other looked at how to use the additional dimension of display given by XR to expose more meta-data about the task that can help the worker (Biener et al. 2020; Gesslein et al. 2020).

2.1 Working with 2D Content in 3D

Toady's most common knowledge worker tasks, such as document editing or spreadsheet applications, are done in 2D on 2D displays. Therefore, in order to leverage on the existing data, tools, and user's familiarity, many approaches that combine knowledge work with XR three-dimensional environments keep the 2D nature of such tasks and develop techniques for interacting with 2D content in a 3D environment.

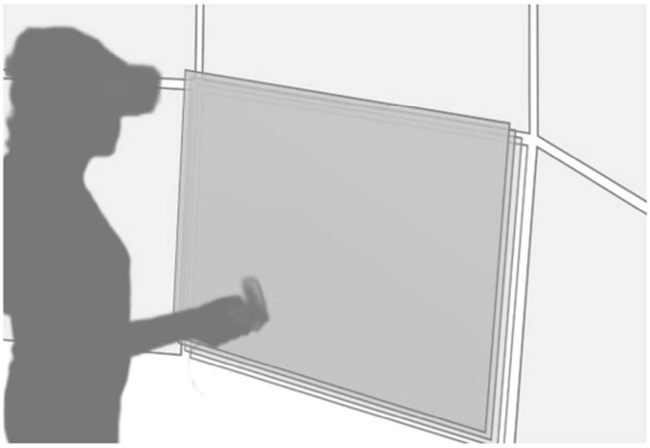

Fig. 1 Using a head-mounted display adds several capabilities over conventional 2D displays. First, the effective display field of view engulfs the user; second, the HMD's stereo display enables the position of data at different distances from the user, and is not limited to a planar display. Finally, positioning the user and the data in the same space enables new and direct ways for the user to interact with the data using natural gestures

Immersive 3D environments offer a large space around the user that can be used to display more information. However, directly interacting with the content, for example, through gestures, can be fatiguing, especially for large displays. This problem can be tackled by indirect interaction. Simple interactions on 2D data in a 3D virtual environment could be done by simply using existing devices like mice or touchpads. This is for example supported by Oculus Infinite Office (Oculus 2021) where a bluetooth keyboard including a touchpad can be connected to the Oculus Quest 2, is visualized in VR and can be used to interact with different applications. However, such devices might not scale well from use on standard displays to the large interaction space XR. Also, their interaction space is limited to the 2D plane on which they operate, limiting the possibilities provided by XR, like depth visualization. Huang et al. (2017) compared mouse input with gesture input for manipulating graphs in VR and results from their user study indicate that gestures can be more efficient for complicated graphs. To interact with data visualized in 3D, one could also employ 3D mice such as presented by Perelman et al. (2015) , which allows interaction on a two 2D plane as well as in 3D space.

For example, Andujar and Argelaguet (2007) proposed to interact with 2D windows inside virtual reality by decoupling motor space from visual space. To this end, users interacted with a controller pointing on a virtual pad, which mapped movements to the respective 2D window (see Fig. 2). They compared the technique with direct manipulation raycasting and the performance results indicated that there is a small overhead. However, authors argue that it is a good trade-off for more flexibility and comfort. Later, this idea was also studied within AR (Brasier et al. 2020) where results suggested that indirect input can perform equally to direct hand raycast and produces less fatigue. Also, Hoppe et al. (2020), proposed a set of tools to further

Fig. 2 Sketch of the virtual pad as proposed by Andujar and Argelaguet (2007). On the left, the user directly interacts with the 2D content through raycasting. On the right, the user casts the ray onto a virtual pad which redirects the movement to the 2D content and allows a more comfortable position. Image courtesy by Ferran Argelaguet

interact with 2D Window content, such as a workbench tool for copying and pasting 2D content within 3D or a macro tool, inside 3D virtual environments.

Normand and McGuffin (2018) used augmented reality to extend the display space of a smartphone and additionally presented new mid-air interaction techniques. The results of their study indicate that the extended display space was superior to using the smartphone alone. Additionally, input on the phone performed better than the proposed mid-air interaction techniques. Le et al. (2021) presented VXSlate, combining head tracking and a tablet to perform fine-tuned manipulations on a large virtual display. A virtual representation of the users hand and the tablet is presented on the large display. Kern et al. (2021) introduced a framework for 2D interaction in 3D including digital pen and paper on physically aligned surfaces, and a study ($n = 10$) showed that the technique resulted in low task load and high usability.

Biener et al. (2020) investigated the joint interaction space of tablets and immersive VR HMDs for supporting the interaction of knowledge workers with and across multiple 2D windows embedded in a 3D space (see Fig. 3). Specifically, they designed techniques, including touch and eye gaze, and therefore combined indirect and direct techniques. The goal was to be unobtrusive and compatible with small physical spaces within everyday environments like trains and planes. Using physical devices like tablets or pens can also minimize motions and provide hand support for long hours of work and therefore reduce fatigue. For example, Gesslein et al. (2020) proposed pen and gaze-based interaction techniques for supporting interaction with spreadsheets in VR. Aside from the large display space, they also made use of the 3D view to show additional information above the 2D screen. Einsfeld et al. (2006) also presented a semantic information visualization of documents in a 3D interface, which visualizes documents, meta-data, and semantic relations between documents. Dengel et al. (2006) extended this work through interaction techniques for searching and navigating large document sets with a stereoscopic monitor and a data glove.

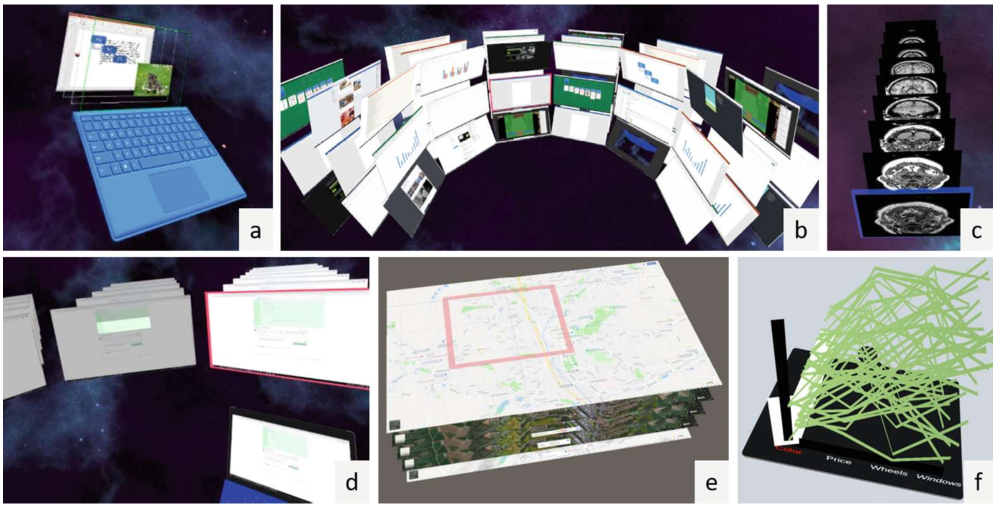

Fig. 3 Applications proposed by Biener et al. (2020) using the joint interaction space of tablets and immersive VR HMDs for tasks like presentation editor (**a**), window manager (**b**), medical imaging (**c**), code version control (**d**), map navigation (**e**), and information visualization (**f**)

Evaluating these techniques indicated that they are fun and enable an efficient interaction. Further, Deller et al. (2008) introduced interaction techniques for managing 2D documents within a virtual work desk.

As has been seen, many proposed techniques include a touch surface (Normand and McGuffin 2018; Biener et al. 2020; Gesslein et al. 2020; Le et al. 2021), which enables fine grained sensing of writing and touch, while supporting the users fingers or stylus, and has found to perform better than mid-air techniques (Normand and McGuffin 2018; Gesslein et al. 2020; Romat et al. 2021; Kern et al. 2021).

The large display space of XR enables display of more data than typical physical displays, yet the need to manipulate data over such large space efficiently without generating fatigue of the user is a challenge. Indirect input has been shown to perform equally to direct while inducing less fatigue (Brasier et al. 2020). The user may execute smaller gestures, and may support their hands to enable long hours of work, and maps their action to the large display space. This is not a new concept for information workers, used for indirect mapping of mouse inputs.

2.2 *System Control*

While many techniques have been proposed for task such as object manipulation (for a recent survey we refer to Mendes et al. (2019)), research on system control has not been as much in the focus of attention. Within the context of seated virtual reality, Zielasko et al. (2019a), studied the effects of passive haptic feedback on touch-based menus. They found a mid-air menu with passive haptic feedback to outperform desk-

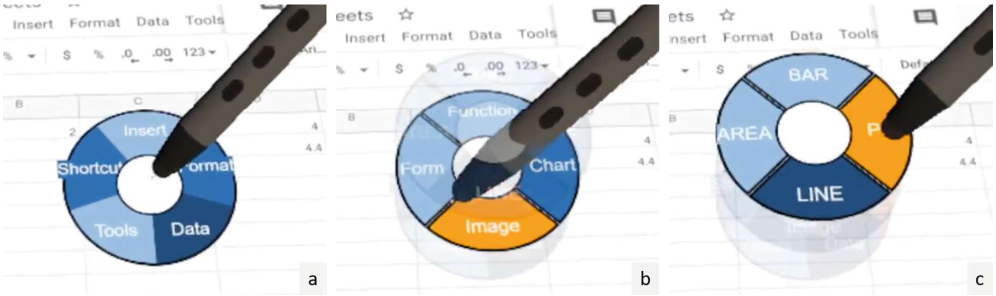

Fig. 4 Stacked radial menu in the spreadsheet application presented in Gesslein et al. (2020)

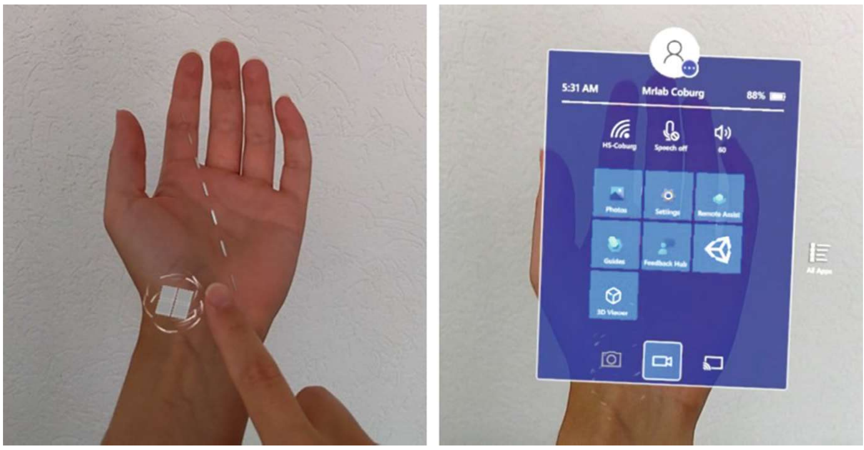

Fig. 5 Touching a wrist with a finger is used to summon a menu in HoloLens2 HMD

aligned alternatives (with and without haptic feedback). However, they also noted hardware requirements for supporting passive haptic feedback in VR, which might be challenging to achieve in everyday environments. Bowman and Wingrave (2001) compared a menu system using pinch gloves (TULIP) with floating menus and pen and tablet menus. They found that users had a preference for TULIP although pen and tablet was significantly faster. This was due to the TULIP interface providing less affordances (Fig. 4).

In current HMDs system, control is usually realized via gestures, controllers and pointing. For example, the Oculus Quest opens a rectangular menu-window upon pressing the menu-buttons on the controller; then, the user can navigate through the menu using raycasting and the controllers trigger. Alternatively, a hand gesture mode is available where the user can open the menu by performing a pinch gesture and navigate using raycasting originating from the hand and pinch gestures to select. Tapping with a finger on the wrist of the other hand will open the menu in a HoloLens 2. Navigation is then done by tapping directly on the buttons on the interface, similar to a touchscreen (see Fig. 5).

The work by Zielasko et al. (2019a) shows that mid-air menus are preferred over menus aligned with the desk. This is also reflected in implementations of current HMDs that present the menu vertically in front of the user. It is also beneficial to have haptic feedback which is also given in pen and tablet techniques that have shown to be usable for system control.

Many of the works mentioned above focus on special gestures for summoning a system control window when interacting with 3D space around the user. The gestures for summoning are quite large, sometimes using two hands, and they are mostly context-free. Information workers, working mostly sitting next to a desk and using small supported hand motions, may not be able to use some of these gestures. Furthermore, it is of interest to use very small gestures that can help the user to summon menus within the context of the data. Gesslein et al. (2020) uses the depth display of HMDs to separate between the planar display of a tablet as the location of the original data and the space above the plane as the place for meta-data and menus. The researchers render multiple layers of pie menus in context of 2D spreadsheets (follow the radial menus design of Gebhardt et al. (2013)) when displayed in VR. The motion of the user's pen above the 2D tablet screen is used to access stacked radial menus (see Fig. 4).

2.3 Text Entry

Text entry and editing is a major task of information work, which is a spatial challenge for XR users. The use of near-eye displays can block the view on physical peripherals such as physical keyboards or touch keyboards and may interfere with the view of the users own palms and the hand-eye coordination, required for efficient typing.

2.3.1 Representing Keyboards and Hands

The best and most popular way to enter text is the full-scale physical keyboard that has hardly changed in the last century and a half since its introduction. The keyboard supports the user's fingers and gives haptic feedback, as the user presses each key. To leverage on the massive install base of physical keyboards for XR users, there is a need to track its position and orientation in space and represent it inside the HMD display, as well as the user's hands.

Jiang et al. (2018) present a technique called HiKeyb. The keyboard is being recognized in the depth video and is represented in the virtual space by a corresponding virtual model, while the user's hands are segmented from the depth video and included in the virtual environment as a planar billboard. Users using this technique were measured reaching typing speed of 23.1 words per minute. McGill et al. (2015) used a color video input and blend the real-time video of the user's hands over the physical keyboard, as a window within the HMD display, showing that users were able to type using this input although their typing speed was slower compared to their natural typing speed in the real world. Part of this difference might be due to the novelty of the XR environment, and part maybe due to inherent latency or inaccuracies of the system.

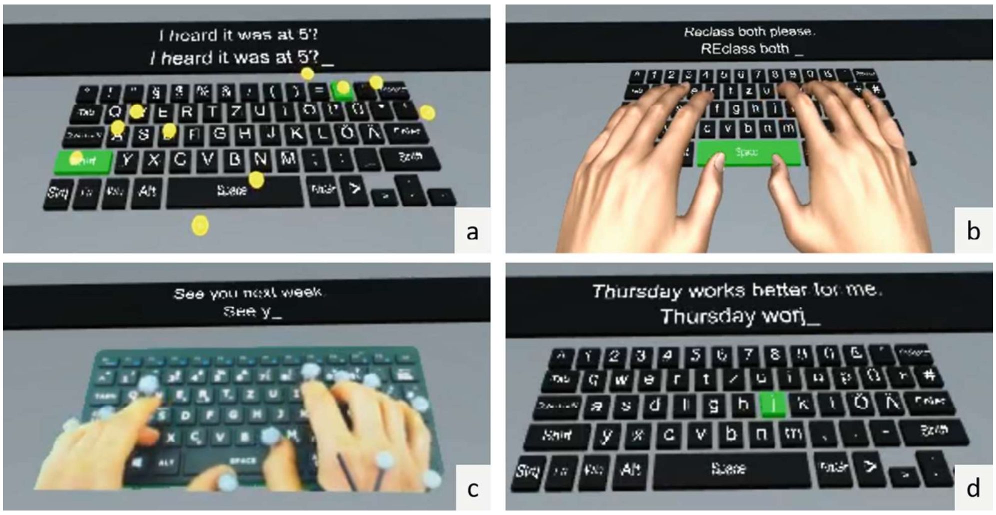

Fig. 6 Different representations of the user's hands for typing in a virtual environment (Grubert et al. 2018a), including representing fingertips as spheres (**a**), using an inverse kinematic model (**b**), video stream (**c**), and no hand representation (**d**)

Capturing a video of the real-world hands may not be a solution that fits every application. Its styling may break the immersion of a VR experience, and it is limited by the ability to capture the real hands (real-world illumination, real-world visibility). Researchers have explored other ways of sensing the user's hands actions, from 3D scanning to none at all.

Walker et al. (2017) rendered a virtual keyboard in the HMD's virtual environment and did not render the user hands at all. Upon a finger pressing a key on the physical keyboard, the corresponding virtual key lights up. They combined this approach with an auto-correction algorithm and showed that users could reach a speed of 40 words per minute. A similar approach by Otte et al. (2019) used touch-sensitive keyboards enabling highlighting touched keys in virtual displays prior to pressing them. They compared fingertip visualization with a touch-sensitive keyboard and found that they are similarly efficient.

Researchers looked at reconstructing the user's hands geometry and render them in the HMD virtual space as 3D models that follow the user's hand motions and position. Knierim et al. (2018) present a system that tracks the physical keyboard and the users hands. They compared different renderings of the hand's models from a realistic rendering, abstract and fingertip rendering both as full opaque objects as well as semi-transparent objects. The performance of experienced typists, which less rely on view of their hands, was not significantly influenced by the visualizations, while inexperienced typists needed some kind of visualization and transparency had no significant influence. For all typists realistic rendering of their hands resulted in higher presence and lower workload.

Grubert et al. (2018a) compared four different hand representation techniques in the context of text entry: no representation and video inlay (following the work of McGill et al.), and using tracking of fingertips in 3D space, they rendered two types of 3D models: A minimalist model of the fingertips only, leaving most of the user palm transparent, and a full 3D animated model of the palms. Since the researchers tracked only the fingertips positions, an inverse kinematics technique was used to animate the finger's joints (see Fig. 6). Their study showed no significant differences in typing speed between different renderings. However, using video inlay and fingertip visualization resulted in significantly lower error rates compared to the other two techniques. Surprisingly, using fully animated models increased the error rate of typing, almost as much as using no visualization of the hands at all, probably due to the accumulated effects of small latency and differences between the recovered model and the real hands. Interesting, the actual speed of typing was not affected by the representation; only, the error rate, however, the representation can influence subjective measures like presence and workload.

Other works looked at the effect of the keyboard peripheral and the way they are rendered on the text entry quality. The use of XR also opens new possibilities that were not possible in the physical world. For example, since the user does not see their own hands and the keyboard, those may be rendered to the user in new locations, for example, closer to the document or the gaze direction of the user. Dube and Arif (2020) explored the effect of rendering virtual keyboards' keys shapes and choose 3-dimensional square keys. Grubert et al. (2018b) visualized the keys of a physical keyboard in VR and the fingertips of the user. The researchers also studied the ability to relocate the keyboard and user hands from their physical location to a position in front of the user's gaze. While physical keyboards performance, where the user's fingertips stays lying on the physical keyboard, where not affected by the relocation, soft keyboards, requiring the user's fingertips to raise above the touch surface showed some reduction in speed yet kept a reasonable performance.

Another way to sense and render the user hands is to use point clouds, generated by depth cameras and LiDARs. While displaying a 3D representation of the hands which might improve hand-eye-coordination, they do not require computational heavy and error-prone processes of recovery like for a full 3D-articulated model of the hands. Pham and Stuerzlinger (2019) compared visualizations for typing on a physical keyboard in VR. They compared no VR with the following: no keyboard representation, hand represented as point cloud; keyboard represented as rectangular frame, hand represented as point cloud; virtual model of keyboard, hand represented as point cloud; keyboard and hands shown as video; keyboard and hands represented as point cloud. Authors concluded that the video-see-through is the best option because it is easy to implement and achieves a good entry speed. The point-cloud solution was also found to be competitive; however, it is more complex to implement.

Even though using a physical keyboard for text entry in XR allows for efficient text entry, it demands that there is a physical keyboard available. This might not be suitable for some scenarios like mobile applications or in limited spaces.

2.3.2 Mobile Text Entry

While most large text entry tasks are still best to be done near a working desk, using a full-size keyboard, the use of wearable HMDs enables the users to enter text also on the go without the constraints of a physical environment and a physical keyboard. Researchers looked at using phones as text entry devices for XR users that are common and mobile. One challenge of current phone keyboards is being based on touch, so they require the user's visual sense to guide the fingertips before they touch keys on the phone's screen. To overcome this difficulty, Kim and Kim (2017) use a phone with hovering sensing (sensing finger tips at some distance prior to touching the phone's screen) to visualize both the phone's keyboard and the nearby user's fingertip. Son et al. (2019) used two touch pads with hover function for typing in VR with two thumbs, resulting in a typing speed of 30 WPM. Knierim et al. (2020) focused on a portable solution and compared the on-screen smartphone keyboard with a desktop-keyboard connected to a smartphone and with a VR-HMD that shows the physical keyboard via video-pass-through. Results indicate a higher input speed in the HMD condition compared to smartphone only, but lower speed compared to a smartphone combined with a physical keyboard. This shows that HMDs with physical keyboards perform better than virtual touchscreen keyboards but worse than physical keyboards without HMDs.

2.3.3 Gaze-Based Text Entry

Using XR HMDs opens the possibilities to use new modalities to aid text entry. In recent years, several commercial AR and VR HMDs have introduced integrated eye-tracking functionality. Ahn and Lee (2019) and Kumar et al. (2020) combined gaze and touch to input text and Rajanna and Hansen (2018) combined gaze and a button click. The additional touch modality is used to select a key and speeding up eye tracking for text entry, usually using dwell time over a key to confirm inputs. Lu et al. (2020) explored a hands-free text input technique and compared sensing blinking and neck movements as alternatives to dwell time. Results showed that blinking performed best. Ma et al. (2018) added a brain–computer interface as a selection mechanism.

To date, these approaches allow much slower entry speeds compared to physical keyboards and are currently not the first choice for extensive text entry tasks.

2.3.4 On-Surface and Mid-air Text Entry

As XR modifies the user senses, it is possible to render virtual keyboards and turn any physical surface in the environment into a keyboard, enjoying the support and haptic feedback of the surface. Richardson et al. (2020) present a technique which combines hand tracking and a language model to decode text from the hand motions using a temporal convolutional network. Participants of a study reached a speed of 73 words per minutes (WPM), comparable to using physical keyboards. To achieve such speed, the current system was trained for each user for about an hour. Fashimpaur et al. (2020) also used a language model to disambiguate text entered by pinching with the finger that would normally be used to type a character. This approach, however, achieved a much lower performance (12 WPM).

Typing on a virtual keyboard in mid-air lacks haptic feedback. Gupta et al. (2020) explored different tactile feedback techniques for a mid-air keyboard in VR. In their study, they compared audio–visual feedback to vibrotactile feedback on the fingers, as well as spatialized and non-spatialized feedback on the wrist. Performance of the four techniques was comparable, but participants preferred tactile feedback. Results also indicated a significantly lower mental demand, frustration, and effort for the tactile feedback on fingers. Participants also preferred the spatial feedback on the wrist over the non-spatial. Dudley et al. (2019) compared typing in mid-air to typing on a physical surface (Fig. 7), both using only the index finger or all ten fingers and found that users are significantly faster when typing on a surface compared to mid-air. They also reported that participants could not effectively use all ten fingers in the mid-air condition, resulting in a lower speed than the index finger condition.

Other text entry techniques in XR use purely virtual keyboards. This offers higher mobility because no extra hardware needs to be carried around. Xu et al. (2019b) and Speicher et al. (2018) evaluated different pointing methods for typing on virtual keyboards, including controllers, head and hand pointing. Both concluded that controllers are usually the best choice. Research was also done in evaluating different keyboard layouts for virtual keyboards. For example, using a circular keyboard that can be controlled via head-motion (Xu et al. 2019a) or a pizza-layout using dual thumbsticks (Yu et al. 2018). Instead of keyboards, different devices for text entry were explored, like a cube with keys (Brun et al. 2019), a ring worn on the index

Fig. 7 Keyboard and hand visualization as presented by Dudley et al. (2019). In their study, the keyboard was once positioned in mid-air and once on the table. In both cases, the fingers were visualized as purple spheres

finger (Gupta et al. 2019), a circular touchpad (Jiang and Weng 2020) or a glove that uses chords to represent letters. Also González et al. (2009) compared six text input techniques for VR and found that the mobile phone resulted in the highest typing speed, followed by pen-based QWERTY and pinch keyboard.

However, the performance of such techniques is much lower than for physical keyboards which makes it less suited for longer text entries. Yet, they can still be useful in a mobile scenario, when, for example, writing short messages.

2.3.5 Speech Input

Speech input is another option for text entry in XR. It is already used with many everyday devices today, for example, on phones. Using speech input in XR has also been explored in previous work and Bowman et al. (2002) even found it outperforms a pinch keyboard, a one-hand chord keyboard and a pen on tablet technique with regards to speed. Yet, none of the examined techniques resulted in high levels of performance or usability. Hoste et al. (2012) present a multimodal approach for text entry combining speech input with body gestures for error corrections. In a user study, participants have rated it the most efficient technique compared to controller input, pure speech input, and pure gesture input. However, with 11.5 WPM, it was not the fastest of the compared techniques. Even though speech input can be faster than some other techniques, there are several drawbacks of using it for knowledge work in everyday environments. Common use cases for speech input like voice assistants at home or in a car, which use speech input mainly for commands, are very different from use cases for knowledge work where it is very common to write longer texts. Use of speech input can issue privacy concerns whenever the users are not located in a private space on their own. For example, when entering a password or handling sensitive data, speech input can not be used without revealing this data to any bystanders. Regardless of the content, speech input can also be disturbing to other people, for example, colleagues in an office or other passengers when using the system in public transportation. Another more fundamental problem is the distinction between commands and transcription which is problematic in all speech input systems, not just in VR. The system needs to figure out which words belong to a command like "open new document" and which words should be transcribed in said document. A similar problem also arises in collaborative settings like in a meeting, where it can be difficult for the speech input system to discern between words that should be transcribed and words that are just part of the discussion with colleagues. Such situations could be disambiguated by a mode switch, for example, only transcribe words while pressing a certain button. Another concern is that speech might interfere with productivity.

2.3.6 Using Keyboards Beyond Text Entry

Additionally, combining a keyboard with VR can open up new possibilities not available in real world. Schneider et al. (2019) explored different input and output

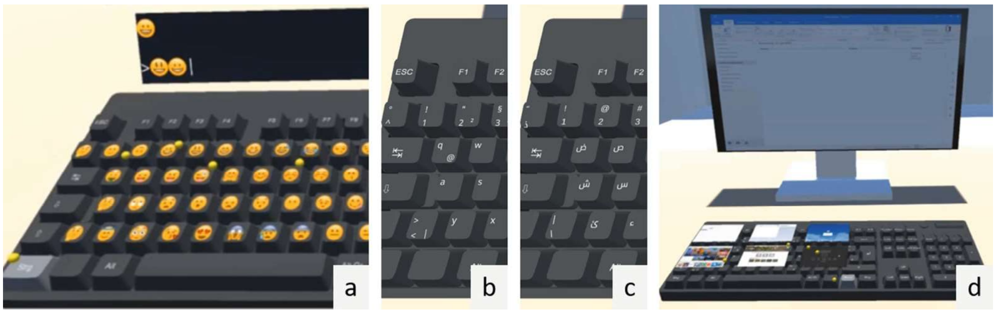

Fig. 8 Reconfiguring the layout of a physical keyboard in VR. Examples are enabling emoji input (**a**), input characters from different languages (**b**, **c**), and window manager (**d**)

mappings enabling different keyboard layouts and functionalities like emoji entry, text processing macros, secure password entry by randomizing the keys, window manager or entering characters from different languages (see Fig. 8).

3 Collaboration

XR as immersive technology has a great potential to connect remote coworkers as if they are immersed in the same space. XR's need to track the user's head and hands can be used to animate their avatars, representing their body language.

On the other hand, XR generates a unique challenge for communication between coworkers by physically occluding the users faces, their gaze, and facial expression by the head-mounted display. While we are optimistic that in long term, HMDs are going to look like regular prescription glasses and not occlude important parts of the users faces; collaboration software has to bridge this current gap, while using the unique abilities of the immersive displays. Next, we discuss approaches for collaboration in VR and AR.

3.1 Collaboration in Virtual Reality

Virtual reality blocks the full view of users and replaces them with a virtual world, generating a special challenge to combine coworkers in the same physical space. Without exact representation of the other users and any geometry in the environment, users may accidentally hit them.

Programming in a pair, working together at the same time on the same code, has shown several advantages, such as increased knowledge transfer, generation of higher quality code, increased code comprehension, and team bonding. Lately, Dominic et al. (2020) compared the use of a state-of-the-art remote video sharing to a VR system that represented the avatars of programmers, as well as their keyboard and mice. They found the VR system enables participants to find twice as much bugs and reduce the time needed to solve them. Programmers may have been more focused thanks to reduction of external distractions, and they might also have been affected by the immersive feeling of sharing the same space. The authors suggest to further explore the physical and mental demands of collaboration in VR and the effects on productivity and frustration. Sharma et al. (2011) developed the *VirtualOffice* system to connect virtually between remote workers along the full work day, including awareness of side conversations, social happening, and additional opportunities to generate informal communication. The system supported different alert mechanism from text to 3D displays of the virtual office. The authors used Greenhalgh and Benford's concept of an "aura" attached to each operation to publicized actions and attract coworkers to get involved. The system was designed to use real offices as a base, so it was limited to connected remote workers that work in remote physical offices. Nguyen et al. (2017a) present CollaVR which enables collaborative reviewing of VR videos using VR HMDs. In addition to watching the video, it allows the users to take notes and exchange feedback. In a preliminary study, experts were positive about using VR for such tasks. In the study, only up to three people worked together; therefore, the authors proposed to do further research on awareness visualizations that are scalable. Additionally, it would be helpful to explore asymmetric hardware setups to include collaborators without access to specific hardware.

While most of the works we discuss in this chapter are focused on enabling collaborations in the context of specific tasks, the office environment enables more than a designed environment for performing work tasks. The joining of coworker together enables generation of informal interactions through chance conversations that are not well supported by current pre-planned teleconferencing. Chow et al. (2019) investigated asynchronous interaction in a virtual office environment through virtual recoding and replay of one's action. They also identified and addressed challenges in such a setting such as intrusion of personal space or awareness of actions.

In summary, it has been shown that VR has the ability to increase productivity in collaboration tasks, which is potentially caused by reduced distractions and a feeling of sharing the same space. A benefit of using VR in knowledge work is to connect workers at remote locations and enable natural, informal interactions that are not possible via teleconferencing systems.

3.2 Collaboration in Augmented Reality

Butscher et al. (2018) present ART, a tool for collaboratively analyzing multidimensional data in augmented reality. The visualization is placed on a touch-sensitive tabletop to facilitate existing interaction techniques. They evaluated their design using expert walkthroughs which revealed that AR can support collaboration, a fluid analysis process and immersion in the data. Dong and Kamat (2011) present another tool for collaboration using augmented reality and tabletops. They developed ARVita where collaborators can observe and interact with simulations in the context of construction planning. AR HMDs were also used alongside a tabletop display by Ens et al. (2020) to support casual collaborative visual analytics. They evaluated their prototype, uplift, with expert users and concluded that systems like that have the potential to support collaborative discussions, presenting models or analyses to other high-level stakeholders.

Pejsa et al. (2016) used projection-based spatial AR to generate a life size view of collaborators, rendered from the point of view of each participant (called Room2Room). Avoiding the use of HMDs, enabled to present the participants in a natural view, life size, using hand gestures that related to the room and using face expressions as part of the communication. As the system did not support stereo rendering, the view of the collaborators was set to align with 3D objects in the room, such as sofas or walls. While the presence enabled by Room2Room has not reached that of a physical face to face meeting, the participants reported much more presence than using video conferencing and were able to finish physical arrangement task faster, aided by expressive direction pointing and hand gestures.

Park et al. (2000) examined a CAVE-based approach for collaboratively visualizing scientific data. They found that users mostly work alone and used localized views to test things without disturbing the overall view and then global views to discuss results with the others. An open question was how the number of participants effects collaborative visualizations. Jing et al. (2019) presented a handheld AR application to support synchronous collaborative immersive analytics tasks such as querying information about content depicted on a physical whiteboard.

Cordeil et al. (2016) compared HMD and CAVE for collaboratively analyzing network connectivity. They found that HMDs were faster than CAVE in tasks like searching for the shortest path or counting triangles. But, no differences in accuracy or communication were found. Additionally they stress the fact that HMDs are less expensive and more readily available.

These works show that AR increases immersion and presence and can support collaborative analyses and presentations. It has also been shown that users value private spaces before sharing their work with others to not clutter the shared space.

3.3 Hybrid Collaboration

We have seen examples of collaboration in VR and AR. However, specific devices are not always available to all collaborators. In the following, we present approaches that combine different technologies.

If collaborators do not have access to an XR device or are in situations where using them might be inappropriate, it is useful to enable collaboration between XR and standard applications on desktops or phones. Reski et al. (2020) presented a system for synchronous data exploration using a hybrid interface. It consists of an immersive virtual reality application and a non-immersive Web application. They validated the approach in a user study representing a real-world scenario. They concluded that the system fostered experiences of shared discovery and therefore has potential for collaborative data exploration tasks. Future work could have a closer look at the collaborative and communicative behavior of coworkers inside and outside of VR.

Norman et al. (2019) presented a study involving collaboration between two local mixed reality users and one remote desktop user. They found that the remote user was more engaged while in the role of a coordinator. Therefore, it is suggested that remote users have specific roles to increase their participation.

Fender and Holz (2022) presented a system that allows synchronous collocated and remote collaboration, as well as asynchronous collaboration. They demonstrate scenarios in which a VR user interacts with another user represented in VR as a live 3D reconstruction either collocated or remotely. Additionally, they also present a asynchronous scenario, where the VR user blocks out everything from the real environment, but all events are recorded and can be replayed later, showing, for example, a 3D reconstruction of a coworker dropping something off. The asynchronous method, however, raises the question such as how the VR user can respond appropriately, if the coworker is no longer present when the event is replayed.

Tang et al. (2010) explored the communication channels during a three-way collaborative meeting. The channels include the person space, the reference space using hand shadows, and the shared task space. Although in this work the setup was not studied in hybrid setting, it can inform how to create hybrid experiences with XR where participants in XR could see virtual hands as reference space and participants without XR could see hand shadows displayed over the task space.

In some cases, different technologies were also combined to construct new environments. Cavallo et al. (2019) presented a co-located collaborative hybrid environment for data exploration. This means they combine high-resolution displays, table projections, augmented and virtual reality HMDs, and mobile devices like laptops. Their evaluation results indicate that integrating AR can increase the speed of the analysis process. However, they also state that there are still limiting factors like resolution and field of view of current HMDs.

It can be seen that collaboration is also viable between collaborators using different technologies. In such cases, it has found to be useful to assign roles to increase participation. Combining different technologies like HMDs, mobile devices and displays provide possibilities to include collocated users with different devices. Yet, the behavior of collaborators in different environments needs further research.

4 Environments

XR has the potential to change the work environment of users, beyond the limitations of their physical environment, and to be designed to reduce stress and increase productivity which are important factors in everyday work as stress has been found to affect physical (Heikkilä et al. 2020; Eijckelhof et al. 2013) and mental health (Stansfeld and Candy 2007). We will discuss the importance of including parts of the physical world into the virtual environment (*VE*) and how this can be achieved.

This also leads to social implications, which can be caused by not being aware of people outside the virtual environment or because bystanders are not aware of what the XR user is doing.

4.1 Managing Stress and Productivity

Ruvimova et al. (2020) evaluated the use of VR to reduce distractions induced by open office environments. They compare four different environments for a task of visual programming: a closed office without VR, an opened office without VR, a VR beach environment while the participant was located in a real open office (see Fig. 9), a VR open office environment while the participant was located in a real open office. Results indicated that both the closed physical office and the VR beach virtual environment were equally successful in reducing distraction and inducing flow and were preferred over the two open office environments. This suggests that VR can be used to stay focused in open offices and the VR environment can be customized to every user's needs. This study focused on single person tasks; the effect of using VR on social interactions between colleagues who share the space is still an open question. Anderson et al. (2017) and Thoondee and Oikonomou (2017) showed that their VR application, showing a nature environment, can reduce stress at work and improve mood. Pretsch et al. (2020) showed that using VR to experience natural landscapes can significantly reduce stress as perceived by participants and has a significantly higher effect than video streaming similar images. Valtchanov et al. (2010) immersed users after stress-induction task in an explorable VR nature settings and showed that interactive VR nature reduces stress and has positive effect beyond passive VR nature displays.

Li et al. (2021) investigated the influence of physical restraints in a virtual working environment in a study where participants did a productivity task in a car while being exposed to different environment. Their findings suggest that users perform better in familiar working environments like an office but prefer secluded unlimited nature environments. They also showed how virtual borders can guide the user to touch the cars interior less.

Lee et al. (2019) used augmented reality as visual separators in open office environments to address visual distractions. The results of their study suggest that this technique reduces visual distractions and improves the experience of shared workspace by enabling users to personalize their environment.

Pavanatto et al. (2021) compared physical monitors with virtual monitors displayed in augmented reality and concluded that it is feasible to use virtual monitors for work, yet technically they are still inferior to physical ones. They suggest mixing physical and virtual displays to utilize the enlarged space of the virtual reality and the familiarity of the physical monitors.

All this research shows that XR has the potential to improve the work space of a knowledge worker, by increasing productivity and reducing distractions and stress. XR allows the user to create an optimal working environment which can be easily

Fig. 9 VR replaces the working environment of a worker in an open place to a beach scene, helps the worker stays focused

adapted to different situations and requirements. VR may currently be ahead of AR in this aspect, as it replaces the entire environment of the user, while AR displays are compared to the real-world quality, yet both can be used to optimize the workplace, by adding additional displays or visual separators. Multiple studies have shown how stress can be reduced by immersing oneself in a computer-generated nature in VR. It has been shown to outperform 2D displays, and interactions with the display increase immersion and relaxation performance. However, it is not totally clear how it compares to conventional relaxing methods.

4.2 Including Reality

Virtual reality replaces the visual and audio sensing of the users with a new virtual environment. As seen above, this new environment may help users to better focus or relax. However, there are objects in the physical environment of the worker that might be of importance for her, like tea cups, desks, or coworkers. This subsection looks at research that examines how much of the physical environment users should be aware of while in a virtual environment.

McGill et al. (2015) stated that VR users have problems when interacting with real-world elements and being aware of the real-world environment. They addressed these issues in three studies. As already mentioned in Sect. 2.3, they showed that typing performance can be significantly improved when enabling a view of reality instead of showing no hands and no keyboard. In another study, they investigated how much reality can be shown and still enable the user to feel present in the virtual environment. They concluded that it was optimal to present the user with reality when the user is currently interacting with real-world objects. Then, they studied how this approach could be applied to people instead of objects in social environments to make the user aware of the presence of others. They concluded that it is important to include some aspects of the physical environment into the virtual. Otherwise, the usability of HMDs in everyday life would be reduced. Therefore, they propose to blend in relevant parts of reality to preserve immersion while allowing the user to accomplish important actions in the physical environment, like using certain objects, drinking, or being aware of other people.

OneReality (Roo and Hachet 2017) looks at blending objects over the continuum of virtuality and describe it as a design space. RealityCheck (Hartmann et al. 2019) takes McGill's approach further by doing real-time 3D reconstruction of the real world that is combined with existing VR applications for the purposes of safety, communication, and interaction. The system allows users to freely move, manipulate, observe, and communicate with people and objects situated in their physical space without losing the sense of immersion or presence inside their VR applications.

O'Hagan and Williamson (2020) presented "Reality-aware VR headsets" that can identify and interpret elements in the physical environment and respond to them. They evaluated four different notification methods to inform users about the presence of bystanders. Two methods only reported the existence of a bystander and the other two also indicated the position. They detected that some participants were uncomfortable when informed about a bystander with no position information. In a second study, they explored dynamic audio adjustment to react to the real world. They either decreased volume to direct attention to a sound in the physical environment or increased volume to block-out noise. Results showed that decreasing the volume was effective, but not increasing.

Simeone (2016) used a depth camera to detect bystander's positions and show the information to the user in a "VR motion tracker" which indicated bystanders as a dots in a triangular area that represented the Kinect's field of view. Participants of a preliminary study considered it useful and not distracting.

In addition to showing a live 3D reconstruction of a bystander in VR, Fender and Holz (2022) presented a different approach, namely asynchronous reality, which allows the user to enter a "Focus Mode" that blocks all distractions from the environment. However, events like a coworker dropping something off are recorded and can be played back later, so the user does not miss important events.

Zielasko et al. (2019b) compared how substituting a physical desk during a seated task in the virtual environment influences cyber-sickness, performance, and presence. They did not find a difference between showing and not showing a desk.

However, they argue that showing a desk allows seamless integration of other elements like a keyboard, for which we have seen in Sect. 2.3 that it is useful to visualize them.

Knowledge worker use phones regularly. Phones are an important connection to distant coworkers and may be even more important in a virtual reality, where users are completely isolated from their surroundings. However, phones, like other physical objects, are usually not included in the virtual environment. Several studies considered enabling using mobile phones in VR. Desai et al. (2017) presented a system (SDSC) that can detect a smartphone in a video stream, embed it into the virtual environment and fit screenshots sent from the smartphone onto it, allowing VR users to interact with their smartphones while being immersed in the virtual environment. Alaee et al. (2018) use depth-based video segmentation to show a smartphone and the user's hands as video-pass-through blended in to the virtual environment (NRAV). They compared the technique with SDSC and concluded that using NRAV participants could perform several tasks with the same efficiency as when not wearing a HMD and were faster with NRAV than with SDSC. Some users preferred to remove HMDs, but the authors argue that acceptance of the technology will further increase with technological improvements. Bai et al. (2021) presented a technique that brings a virtual representation of the phone and the users hand into VR. Evaluating this technique showed that it successfully connected the real phone with the virtual world, but the experience differed from using a phone in the physical world, with decreased usability caused by hardware and software limitations.

Previous research has shown that it can be very helpful to include parts of the reality in the virtual environment. This is especially true for knowledge workers being immersed for extended periods of time where they need to interact with physical objects like phones or a water bottle or with other people in their surroundings. There are different possibilities to achieve this, similar to including physical keyboards, from virtual replicas to blending in a video stream. For including smartphones, however, the current hardware of HMDs is a limiting factor, as the resolution is not high enough to properly display the content. Regarding collocated people, it has been shown to be important to not only indicate their presence but also their position to make the XR user feel comfortable.

4.3 Social Implications

When working in everyday environments, there are, in many cases, other people around. Therefore, it is important to also have a look at social implications. This includes how the knowledge worker feels when using XR devices, but also how people around the XR user feel and behave because it could potentially keep people from using XR in public, if they know bystanders are uncomfortable. Just looking at the XR user does not provide a realistic evaluation of everyday scenarios.

Bajorunaite et al. (2021) conducted two surveys to explore passenger needs in public transportation that might prevent them from using VR devices. One survey was aimed at an airplane scenario ($n = 60$) and one at ground public transportation, like buses or trains ($n = 108$). For both scenarios, participants expressed concerns about accidental interactions with other passengers and loss of awareness of their surrounding. They suggest to provide cues from the reality and find ways to achieve an engaging experience with less movement. This is in line with the findings presented in Sect. 4. The results from the surveys also indicate that participants are very conscious of their self-image and how they are perceived by other passengers while using a VR device. They express concerns about being judged because they block out reality.

George et al. (2019) had a closer look at bystanders and explored their ability to identify when HMD users switch tasks by observing their gestures. This could help them find good moments for interruptions. In their study, there was a set of tasks that the HMD user performed (authentication, reading, manipulation, typing, watching video). The bystanders, who were aware of the task set, in the study could identify the task type in 77% of the time and recognize task switches in 83% of the time. The authors suggest future work to find out if it has a positive effect on social acceptability when bystanders are able to retrieve meaning from the interactions of the HMD user.

O'Hagan et al. (2020) studied how comfortable bystanders are in interrupting VR users. Their results indicate that the level of comfort and the acceptability of the interruption strategy is more influenced by the relationship to the VR user than the setting.

Hsieh et al. (2016) presented techniques for socially acceptable text entry, scrolling, and point-and-select realized through hand orientation and finger movement detected by a sensor-equipped haptic glove which can independently track mid-air hand gestures. These interaction techniques were considered unobtrusive and socially acceptable.

Research shows that XR users are self-concious of the social aspect of their work, and the fear of being judged can limit their use of XR, which may be a subject for future research. It has also been indicated that it could be helpful if bystanders can retrieve some meaning from the behavior of XR users. This would allow them to better understand them, just like it helps the XR user to be aware of some aspects of the reality to avoid conflicts that disturb bystanders.

We believe that HMDs could be socially acceptable when they become almost not noticeable like regular glasses. Then, the issue might be that when the HMDs looks like glasses and become unnoticeable, if the user is sitting at a cafe and doing strange gestures in the air it could look awkward and not acceptable. This is were having subtle indirect gestures from another device such as a tablet could be more socially acceptable (and also require less energy and be more appropriate on crowded space like inside an airplane). However, some tasks may require large gestures that can not be hidden. Then, this is an issue of the population getting used to it, similarly to how we accept people talking to mid-air when using an ear piece. If more people use VR HMDs in public, it will probably become more accepted, like using a laptop.

In addition, the progress in display resolution could contribute to create HMDs experiences where the main display at the front of the user can contain most of the information in it thus require less head turning on the side or minimize the head turning angle.

Another social implication is the fact that HMDs have cameras that could violate the privacy of others. If we can built HMDs that look like glasses (HoloLens is a step in that direction), we should learn from the privacy lessons from Google Glass.

5 Extended Exposure

In recent years, we have seen a surge in research of knowledge work in XR (AR and VR), showing the possibilities provided by the immersive space. Different works are aimed at enabling users to work in XR for longer periods of time. For example, one possibility of VR, which hides the user's own body from her sight, includes the ability to use smaller physical motions, while the self-avatar completes a full motion (CoolMoves 2021), reducing fatigue while enabling large interaction spaces.

However, research on the effect of long-term immersion over full work days, multiple days a week, is very limited.

Earlier works on eye strain caused by physical displays (Stewart 1979; Jaschinski et al. 1998) suggest individuals are affected differently. Stewart (1979) argued that eye strain results from different factors (i.e., visual, postural, environmental, personal) and that these problems can be solved by considering ergonomics when designing visual displays, suggesting the design of HMDs can be improved to reduce potential problems.

Later research on long periods of using XR has been focusing on the context of manufacturing; therefore, we address these works here, even though they do not actually cover knowledge work. Grubert et al. (2010) conducted a study with 19 participants doing an order picking task for 4 h with and without AR support. Results showed that using AR does not increase the overall objective and subjective strain. However, some participants perceived higher eye discomfort in AR. This was more likely for users with visual deficiencies. They also reported a higher work efficiency in AR compared to non-AR.

Wille et al. (2014) conducted several studies comparing the work with HMDs over several hours with other technologies like tablets or monitors. Objective measures indicated no physiological effects on the visual system and only limited influence of the HMDs weight on the neck muscles which contrasts with the subjective ratings. The authors speculate that the unfamiliar technology influences the subjective ratings.

Funk et al. (2017) studied an industrial assembly task with instructions projected directly on workpiece, and using a depth camera (Kinect v2) to verify assembly correctness. All participants used the system for at least 3 full working days. The results showed that the instructions were helpful for untrained workers; however, it also slowed down the performance and increased cognitive load, especially for expert workers.

Steinicke and Bruder (2014) conducted a 24-h self-experiment in virtual reality (using Oculus Rift DK1 HMD) with one participant, who worked, ate, slept and entertained himself with music or movies. They reported higher simulator sickness after periods involving many movements and lower values when resting. The participant reported limitations of the HMD due to latency when moving and a limited resolution when working. It was also reported that the participant was sometimes confused about whether he was in VR or a real environment and that the perceived accommodation seemed to vary after a few hours.

Nordahl et al. (2019) had two participants using VR HMDs for 12 h. They used Oculus Rift CV1 HMD for 6 h and the HTC Vive HMD for the rest. They only took the HMDs off for switching them after 6 h. While in VR, the participants used different applications. The reported results indicate that simulator sickness symptoms were mild with a peak after 7 h that is difficult to explain because the experiment was not fully controlled.

The most extensive research on long-term immersion has been conducted by Guo et al. (2019a, b, 2020) and Shen et al. (2019). They applied Maslows Hierarchy of Needs to guide the design of a VR office. First, they conducted a short-term study ($n = 16$) (Guo et al. 2019b) using a text input and image processing task which was done in the virtual and a physical environment. They concluded that the designed VR office was comfortable and efficient and therefore, in a next step, used it for a long-term study.

In the long-term experiment (Guo et al. 2019a), 27 participants were in the virtual and physical office for 8 h each, doing knowledge worker tasks like document correction, keyword searching, text input, and image classification. They compared the results with the short-term study and concluded that physiological needs like drinking, belongingness needs like communication, temporal presence, and self-presence are important for long-term immersion but can be ignored for short term. Safety needs and emotional needs on the other hand must be met in both conditions.

This 8-h experiment was also used to investigate mental fatigue differences between the virtual and physical work spaces (Shen et al. 2019). Participants performed a psycho-motor vigilance task (PVT) 6 times during the experiment, and results showed that there were significantly less PVT lapses in the physical environment, and the reaction times were slower in VR, indicating a higher mental fatigue in VR. The authors propose two explanations. Either the additional visual information processing in VR occupies more attention resources, or that VR can increase attention of participants more effectively which lets them allocate higher attention resources.

The same experimental setup was also used to explore the difference in visual discomfort between working in VR and physical environment (Guo et al. 2020). The results showed that subjective visual fatigue, pupil size, and accommodation response changes with time in both conditions. They also detected a gender difference suggesting that female participants suffer more from visual fatigue in VR than male participants. However, this could be caused by male participants in the sample having more experience with VR than the female participants. There was also no significant difference in nausea which can be explained by the static content. The authors recorded no significant eye strain difference between VR and the physical environment, yet focus difficulty was significantly higher in VR which might be caused by the accommodation vergence conflict.

These studies showed only mild simulator sickness symptoms which can be explained by displaying relatively static content in knowledge worker tasks. A large part of physical discomfort is due to the weight and form factor of HMDs, which will hopefully become much more comfortable in the near future.

Shen et al. (2019) reported higher mental fatigue in VR compared to a physical environment. Additionally, in VR, the reaction time significantly increased over time but not in the physical environment. The authors propose two explanations. Either the additional visual information processing in VR occupies more attention resources, or VR can increase attention of participants more effectively which lets them allocate higher attention resources.

From those studies, it can be concluded that simulator sickness symptoms are rather mild (Nordahl et al. 2019; Guo et al. 2020) which can be explained by the relatively static content in knowledge worker tasks. A large part of physical discomfort is clearly due to the weight and form factor of HMDs. As HMD hardware is being improved, mainly on resolution, frame rate, weight and even dynamic focal distance, we expect the VR experience become more comparable to physical environment one, and more such long-term studies will be required. Existing research points to important directions when designing applications for long-term use like latency (Steinicke and Bruder 2014), physiological needs, emotional needs, safety, presence, and belongingness (Guo et al. 2019a). Experience can be heavily influenced by personal traits like impaired vision or XR past experience. Therefore, longer studies are needed, so the participants get accustomed to the new technology which will make it much more comparable to a physical environment.

6 Applications

This section will have a closer look at examples of applications that support knowledge work in XR.

There are several application examples involving data analysis. Zielasko et al. (2017) discussed the potentials and benefits of using VR HMDs in a desktop configuration, i.e., when being seated in front of a table. They described scenarios and use cases in the domain of data analysis. A similar proposal was made by Wagner Filho et al. (2018) and Wagner Filho et al. (2019). On both instances, users interacted with data using mid-air interaction using their hands.

Another application scenario that benefits from XR is programming. Elliott et al. (2015) presented concepts for supporting software engineering tasks such as live coding and code reviews using VR HMDs in conjunction with mouse and keyboard interaction. Dominic et al. (2020) compared remote pair programming and code comprehension in VR against a standard shared display. In their study ($n = 40$), the average time to solve a bug was lower in VR, and they solved nearly twice as many. Authors explain this difference through reduced distractions in VR and a sense of being collocated with the collaborator.

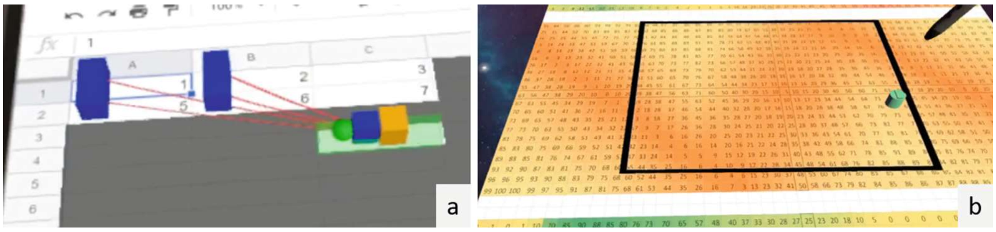

Fig. 10 The left image shows relations between cells. A function is applied to the cells at the top which results in the value shown in the bottom cell. The right image shows how the spreadsheet is visualized beyond the boundaries of the tablet used for interaction which is outlined in black

XR has also been explored for creative tasks, for example Nguyen et al. (2017b), proposed a video editing system for VR.

As the resolution, size and weight of HMDs evolves to more consumer-oriented glasses-like form factor; HMDs becomes an attractive alternative to physical monitors (Pavanatto et al. 2021), being more mobile, large, and private. In recent years, we see research that looks at ways that common 2D office tasks may be enhanced when done via near-eye displays.

Gesslein et al. (2020) enhanced a spreadsheet application in VR which enabled new functionalities not possible in 2D, like visualizing relations between cells above the screen or extending large sheets beyond what is possible on physical screens (see Fig. 10). O'Brien et al. (2019) presented a virtual environment for browsing wikipedia content. Biener et al. (2020) propose many different applications that leverage the possibilities of VR in the context of mobile knowledge workers. For example, they present an array of layered screen arranged around the user that expands the users small portable display and an application that visualizes 3D data on top of a touchscreen device (Fig. 3). Dengel et al. (2006) and Deller et al. (2008) presented applications for working with documents on a stereoscopic display.

There are also commercially available solutions that support knowledge work in virtual reality, such as Oculus Infinite Office (Oculus 2021) where the user can open virtual browser windows and interact with them using a controller, gestures or a physical keyboard which is included in the virtual environment as a 3D model and a video-pass-through visualization of the hands while typing. An example for AR would be spatial.io Spatial.io (2021) which allows collaborative meetings by representing remote coworkers as avatars.

7 Challenges

Informed by the presented review of literature, we see several challenges that need to be addressed in the future to make every day virtual and augmented reality a compelling experience that every knowledge worker wants to use and possibly prefer over a real office. This includes evolving HMDs, by also embracing other devices and peripherals around the user and improving the overall user experience in XR.

7.1 Challenges in the Next Generation of HMDs

The technology of consumer available XR HMDS has progressed by leap and bounds in recent years. It is possible nowadays to see HMDs with a resolution of 5K, and refresh rates of up 144 Hz. However, for being good monitor replacements for information workers, there is a need to display text at quality that is comparable to 2D monitors. Current lenses that are used in commercial HMDs are simple Fresnel lenses that reduce the display quality further away from the center of the display, many times to a level where text is unreadable. Multilayered lenses and other optic solutions that are able to generate sharper displays might still be too expensive for current HMDs that are used mostly for entertainment. Also, the combination of eye tracking and changeable focal distance of the display or light field display can reduce fatigue originated from inaccurate convergence/focus rendering.

Another aspect that limits the everyday use of XR HMDs, especially for longer periods of time, is the form factor which can impact comfort and social acceptability. It has been seen in Sect. 5 that a great part of users discomfort during the use of HMDs is the weight. As manufacturers are working on lighter and thinner HMDs, we hope they will be comfortable for extended use. Glasses-like form factors may also help with social acceptability as it is less disconnecting workers from their environment and make them easier to communicate and collaborate.

As has been mentioned in Sect. 4, when using XR for longer periods, there is more importance for sensing around the headset, from the structure of the environment around the user, enabling the user to move around and use different physical resources, to interacting with coworkers, and updating of the XR space and applications according to the changing conditions around the worker. Detection of peripherals such as physical keyboards (e.g., supported by Oculus Quest 2) enables representation of them in the HMD display, where XR applications can use them and even modify them for their needs (Schneider et al. 2019). Better environment sensing via cameras and depth sensors enables users to move naturally in virtual offices while being in an uncontrolled dynamic environment and avoiding obstacles (Cheng et al. 2019; Yang et al. 2019). And precise and responsive hand tracking also supports a range of interaction possibilities like pointing, gesturing, or text entry.

7.2 Challenges in Future XR Experience for Knowledge Worker

There are also several design challenges that need to be addressed to make XR experience appealing to knowledge workers.

One of the advantages of using XR for work is the independence of the physical conditions in the user's environment. When workers travel or as they work from home, they might find themselves in physical environments that are less than optimal.

In many cases, users need to limit themselves to small spaces, such as an airplane seat, while they are expected to work at their best capability. Working applications should be aware of the user's limitations and display a virtual working environment that is fully controlled by the users given their limited input space. While current XR applications require users to prepare an empty physical environment to enable uninterrupted and safe working spaces, future applications will have to be flexible and adaptable to a variety and dynamically changing environments (Gal et al. 2014; Yang et al. 2019).

Given the possibly limited input space and the need to reduce fatigue of the worker, new mappings will enable users to execute small motions, mostly supported, according to their or the environments limitations, and have them mapped to large effects in the virtual space (Ahuja et al. 2021; Wentzel et al. 2020). The environment around the user may exhibit the change of tasks and enable fast switching of context. It should help relax the users when it is needed and keep them aware of the presence of other people in the work vicinity both physical and remote. The representation of the users should bridge the difference in presence between local and remote users, for example, by representing physical affordances such as white boards and meeting rooms in the virtual domain too. Workers without XR displays need also to be aware of the presence of remote users, with range of methods, from large displays and representation robots to projections (Pejsa et al. 2016).

8 Summary and Conclusion

In this chapter, we have presented an overview of interaction techniques that can support knowledge workers in XR.

Techniques for interacting with 2D content in 3D included controllers, gloves, hand tracking, tablets, eye gaze, pens, and more. Studies have shown that it can be beneficial to use indirect input which reduces fatigue while preserving performance. Researchers are already working on socially acceptable techniques, and future research should also keep that in mind because this is an important aspect in everyday environments. It has also been found that touch surfaces perform better than mid-air which is in line with the findings from exploring text entry techniques. In that context, it has also been shown that users prefer tactile feedback, for example, provided by a physical surface, over mid-air typing. For text entry, it is very popular to include physical keyboards in the virtual environment and different visualizations for keyboards and hands have been explored like video-pass-through, point clouds, virtual models, and abstract representations. It has been shown that the type of visualization influenced workload and presence, but not performance, as long as there is some kind of visualization. Typing on any physical surface has also been shown to be effective when combined with language models and convolutional networks. However, currently, such techniques need to be calibrated for each user which decreases usability. Additionally, many other techniques and devices have been explored including eye-gaze-typing and phones, usually resulting in much lower entry speeds. While this could be negligible for some use cases, it is an important factor for knowledge workers. Future research on interaction techniques for knowledge work in XR could focus on how the system can adapt to changes in the physical environment, by providing appropriate input techniques for different scenarios.

We also presented approaches on how collaboration can be supported in XR for tasks like pair programming, video reviewing, or exploring information. It has been shown that such systems can perform better than standard video tools and increase presence. A problem that arises is how the presented systems can be scaled to allow collaboration of more people, as usually the studies only involved two or three users. We also presented solutions for including users without XR devices, for example, via a Web application. However, it is important to design such applications in a way that engage participation from the non-VR user.

Research on virtual environments has shown that VR can be used to reduce distractions in office environments because the virtual environment can be customized to the user's needs. VR has also been found to be able to reduce stress, for example, by showing natural scenes and that users prefer such scenes even though they perform better in familiar environments like offices. Future work could investigate which environments are suited best for which situations. Research also suggests to include some parts of the physical environment in the virtual environment. This can be done by blending in relevant parts, like tables or objects on the table. They can be included as video stream or as virtual replicas while preserving immersion. Similar to keyboards this can also be done for phones. However, the limited resolution of today's HMDs makes it hard to use such small devices. Besides objects, it is also helpful to include information about bystanders, like presence or position. This can help the VR user to avoid undesired bahavior like accidental interactions. However, further research is needed on how VR user and bystanders feel and how they are affected.

We have shown that some research has been conducted on the effects of extended usage of XR devices such as VR and AR HMDs. Unsurprisingly, a main issue is physical discomfort caused by weight and form factor of HMDs which will hopefully improve as technology advances. Studies also suggest that the experience can be influenced by factors like latency, safety, presence, impaired vision, or experience. To date, the longest study was 24 h; therefore, more research is needed on longer even periods of time.

Finally, we presented some applications that show the value of using XR for knowledge work. Several works show that XR can be beneficial in the area of data analysis, including spreadsheet applications. XR has also been shown to support collaborative programming, and it can be utilized to increase limited screen space in mobile scenarios.

References

Ahn S, Lee G (2019) Gaze-assisted typing for smart glasses. In: Proceedings of the 32nd annual ACM symposium on user interface software and technology, pp 857–869

Ahuja K, Ofek E, Gonzalez-Franco M, Holz C, Wilson AD (2021) Coolmoves: user motion accentuation in virtual reality. Proc ACM Interact Mob Wearable Ubiquitous Technol 5(2). https://doi.org/10.1145/3463499

Alaee G, Deasi AP, Pena-Castillo L, Brown E, Meruvia-Pastor O (2018) A user study on augmented virtuality using depth sensing cameras for near-range awareness in immersive VR. In: IEEE VR's 4th workshop on everyday virtual reality (WEVR 2018), vol 10

Anderson AP, Mayer MD, Fellows AM, Cowan DR, Hegel MT, Buckey JC (2017) Relaxation with immersive natural scenes presented using virtual reality. Aerosp Med Hum Perform 88(6):520–526

Andujar C, Argelaguet F (2007) Virtual pads: decoupling motor space and visual space for flexible manipulation of 2d windows within VEs. In: 2007 IEEE symposium on 3D user interfaces. IEEE

Bai H, Zhang L, Yang J, Billinghurst M (2021) Bringing full-featured mobile phone interaction into virtual reality. Comput Graph 97:42–53

Bajorunaite L, Brewster S, Williamson JR (2021) Virtual reality in transit: how acceptable is VR use on public transport? In: 2021 IEEE conference on virtual reality and 3D user interfaces abstracts and workshops (VRW). IEEE, pp 432–433

Biener V, Schneider D, Gesslein T, Otte A, Kuth B, Kristensson PO, Ofek E, Pahud M, Grubert J (2020) Breaking the screen: interaction across touchscreen boundaries in virtual reality for mobile knowledge workers. arXiv preprint arXiv:2008.04559

Billinghurst M, Clark A, Lee GA (2015) A survey of augmented reality. Found Trends Hum Comput Interact 8:73–272

Bowman DA, Wingrave CA (2001) Design and evaluation of menu systems for immersive virtual environments. In: Proceedings IEEE virtual reality 2001. IEEE, pp 149–156

Bowman DA, Rhoton CJ, Pinho MS (2002) Text input techniques for immersive virtual environments: an empirical comparison. In: Proceedings of the human factors and ergonomics society annual meeting, vol 46. SAGE Publications Sage CA, Los Angeles, CA, pp 2154–2158

Bowman D, Kruijff E, LaViola JJ Jr, Poupyrev IP (2004) 3D user interfaces: theory and practice. Addison-Wesley, CourseSmart eTextbook

Brasier E, Chapuis O, Ferey N, Vezien J, Appert C (2020) Arpads: mid-air indirect input for augmented reality. In: 2020 IEEE international symposium on mixed and augmented reality (ISMAR). IEEE, pp 332–343

Brun D, Gouin-Vallerand C, George S (2019) Keycube is a kind of keyboard (k3). In: Extended abstracts of the 2019 CHI conference on human factors in computing systems, pp 1–4

Butscher S, Hubenschmid S, Müller J, Fuchs J, Reiterer H (2018) Clusters, trends, and outliers: how immersive technologies can facilitate the collaborative analysis of multidimensional data. In: Proceedings of the 2018 CHI conference on human factors in computing systems, pp 1–12

Cavallo M, Dolakia M, Havlena M, Ocheltree K, Podlaseck M (2019) Immersive insights: a hybrid analytics system forcollaborative exploratory data analysis. In: 25th ACM symposium on virtual reality software and technology, pp 1–12

Cheng L-P, Ofek E, Holz C, Wilson AD (2019) Vroamer: generating on-the-fly VR experiences while walking inside large, unknown real-world building environments. In: 2019 IEEE conference on virtual reality and 3D user interfaces (VR), pp 359–366. https://doi.org/10.1109/VR.2019.8798074

Chow K, Coyiuto C, Nguyen C, Yoon D (2019) Challenges and design considerations for multimodal asynchronous collaboration in VR. Proc ACM Hum Comput Interact 3(CSCW):1–24

Cordeil M, Dwyer T, Klein K, Laha B, Marriott K, Thomas BH (2016) Immersive collaborative analysis of network connectivity: cave-style or head-mounted display? IEEE Trans Vis Comput Graph 23(1):441–450

Deller M, Agne S, Ebert A, Dengel A, Hagen H, Klein B, Bender M, Bernardin T, Hamann B (2008) Managing a document-based information space. In: Proceedings of the 13th international conference on intelligent user interfaces, pp 119–128

Dengel A, Agne S, Klein B, Ebert A, Deller M (2006) Human-centered interaction with documents. In: Proceedings of the 1st ACM international workshop on human-centered multimedia, pp 35–44

Desai AP, Pena-Castillo L, Meruvia-Pastor O (2017) A window to your smartphone: exploring interaction and communication in immersive VR with augmented virtuality. In: 2017 14th Conference on computer and robot vision (CRV). IEEE, pp 217–224

Dominic J, Tubre B, Ritter C, Houser J, Smith C, Rodeghero P (2020) Remote pair programming in virtual reality. In: 2020 IEEE international conference on software maintenance and evolution (ICSME). IEEE, pp 406–417

Dong S, Kamat VR (2011) Collaborative visualization of simulated processes using tabletop fiducial augmented reality. In: Proceedings of the 2011 winter simulation conference (WSC). IEEE, pp 828–837

Drucker P (1966) The effective executive harper & row. New York, NY

Dube TJ, Arif AS (2020) Impact of key shape and dimension on text entry in virtual reality. In: Extended abstracts of the 2020 CHI conference on human factors in computing systems, pp 1–10

Dudley J, Benko H, Wigdor D, Kristensson PO (2019) Performance envelopes of virtual keyboard text input strategies in virtual reality. In: 2019 IEEE international symposium on mixed and augmented reality (ISMAR). IEEE, pp 289–300

Eijckelhof B, Huysmans M, Bruno Garza J, Blatter B, Van Dieën J, Dennerlein J, Van Der Beek A (2013) The effects of workplace stressors on muscle activity in the neck-shoulder and forearm muscles during computer work: a systematic review and meta-analysis. Eur J Appl Physiol 113(12):2897–2912

Einsfeld K, Agne S, Deller M, Ebert A, Klein B, Reuschling C (2006) Dynamic visualization and navigation of semantic virtual environments. In: Tenth international conference on information visualisation (IV'06). IEEE, pp 569–574

Elliott A, Peiris B, Parnin C (2015) Virtual reality in software engineering: affordances, applications, and challenges. In: 2015 IEEE/ACM 37th IEEE international conference on software engineering, vol 2. IEEE, pp 547–550

Ens B, Goodwin S, Prouzeau A, Anderson F, Wang FY, Gratzl S, Lucarelli Z, Moyle B, Smiley J, Dwyer T (2020) Uplift: a tangible and immersive tabletop system for casual collaborative visual analytics. IEEE Trans Vis Comput Graph

Fashimpaur J, Kin K, Longest M (2020) Pinchtype: text entry for virtual and augmented reality using comfortable thumb to fingertip pinches. In: Extended abstracts of the 2020 CHI conference on human factors in computing systems, pp 1–7

Fender AR, Holz C (2022) Causality-preserving asynchronous reality. In: CHI conference on human factors in computing systems, pp 1–15

Funk M, Bächler A, Bächler L, Kosch T, Heidenreich T, Schmidt A (2017) Working with augmented reality? A long-term analysis of in-situ instructions at the assembly workplace. In: Proceedings of the 10th international conference on pervasive technologies related to assistive environments, pp 222–229

Gal R, Shapira L, Ofek E, Kohli P (2014) Flare: fast layout for augmented reality applications, pp 207–212. https://doi.org/10.1109/ISMAR.2014.6948429

Gebhardt S, Pick S, Leithold F, Hentschel B, Kuhlen T (2013) Extended pie menus for immersive virtual environments. IEEE Trans Vis Comput Graph 19(4):644–651

George C, Janssen P, Heuss D, Alt F (2019) Should I interrupt or not? Understanding interruptions in head-mounted display settings. In: Proceedings of the 2019 on designing interactive systems conference, pp 497–510

Gesslein T, Biener V, Gagel P, Schneider D, Kristensson PO, Ofek E, Pahud M, Grubert J (2020) Pen-based interaction with spreadsheets in mobile virtual reality. In: 2020 IEEE international symposium on mixed and augmented reality (ISMAR). IEEE, pp 361–373

González G, Molina JP, García AS, Martínez D, González P (2009) Evaluation of text input techniques in immersive virtual environments. In: New trends on human-computer interaction. Springer, Berlin, pp 109–118

Grubert J, Hamacher D, Mecke R, Böckelmann I, Schega L, Huckauf A, Urbina M, Schenk M, Doil F, Tümler J (2010) Extended investigations of user-related issues in mobile industrial AR. In: 2010 IEEE international symposium on mixed and augmented reality. IEEE, pp 229–230

Grubert J, Witzani L, Ofek E, Pahud M, Kranz M, Kristensson PO (2018a) Effects of hand representations for typing in virtual reality. In: 2018 IEEE conference on virtual reality and 3D user interfaces (VR). IEEE, pp 151–158

Grubert J, Witzani L, Ofek E, Pahud M, Kranz M, Kristensson PO (2018b) Text entry in immersive head-mounted display-based virtual reality using standard keyboards. In: 2018 IEEE conference on virtual reality and 3D user interfaces (VR). IEEE, pp 159–166

Guo J, Weng D, Zhang Z, Jiang H, Liu Y, Wang Y, Duh HBL (2019a) Mixed reality office system based on Maslow's hierarchy of needs: towards the long-term immersion in virtual environments. In: 2019 IEEE international symposium on mixed and augmented reality (ISMAR). IEEE, pp 224–235

Guo J, Weng D, Zhang Z, Liu Y, Wang Y (2019b) Evaluation of maslows hierarchy of needs on long-term use of HMDs–a case study of office environment. In: 2019 IEEE conference on virtual reality and 3D user interfaces (VR). IEEE, pp 948–949

Guo J, Weng D, Fang H, Zhang Z, Ping J, Liu Y, Wang Y (2020) Exploring the differences of visual discomfort caused by long-term immersion between virtual environments and physical environments. In: 2020 IEEE conference on virtual reality and 3D user interfaces (VR). IEEE, pp 443–452

Gupta A, Ji C, Yeo H-S, Quigley A, Vogel D (2019) Rotoswype: word-gesture typing using a ring. In: Proceedings of the 2019 CHI conference on human factors in computing systems, pp 1–12

Gupta A, Samad M, Kin K, Kristensson PO, Benko H (2020) Investigating remote tactile feedback for mid-air text-entry in virtual reality. In: 2020 IEEE international symposium on mixed and augmented reality (ISMAR). IEEE, pp 350–360

Hartmann J, Holz C, Ofek E, Wilson AD (2019) RealityCheck: blending virtual environments with situated physical reality. Association for Computing Machinery, New York, NY, USA, pp 1–12

Heikkilä K, Pentti J, Madsen IE, Lallukka T, Virtanen M, Alfredsson L, Bjorner J, Borritz M, Brunner E, Burr H et al (2020) Job strain as a risk factor for peripheral artery disease: a multi-cohort study. J Am Heart Assoc 9(9):e013538

Hoppe AH, van de Camp F, Stiefelhagen R (2020) Enabling interaction with arbitrary 2d applications in virtual environments. In: International conference on human-computer interaction. Springer, Berlin, pp 30–36

Hoste L, Dumas B, Signer B (2012) Speeg: a multimodal speech-and gesture-based text input solution. In: Proceedings of the international working conference on advanced visual interfaces, pp 156–163

Hsieh Y-T, Jylhä A, Orso V, Gamberini L, Jacucci G (2016) Designing a willing-to-use-in-public hand gestural interaction technique for smart glasses. In: Proceedings of the 2016 CHI conference on human factors in computing systems, pp 4203–4215

Huang Y-J, Fujiwara T, Lin Y-X, Lin W-C, Ma K-L (2017) A gesture system for graph visualization in virtual reality environments. In: 2017 IEEE Pacific visualization symposium (PacificVis). IEEE, pp 41–45

Jaschinski W, Heuer H, Kylian H (1998) Preferred position of visual displays relative to the eyes: a field study of visual strain and individual differences. Ergonomics 41(7):1034–1049

Jiang H, Weng D (2020) Hipad: text entry for head-mounted displays using circular touchpad. In: 2020 IEEE conference on virtual reality and 3D user interfaces (VR). IEEE, pp 692–703

Jiang H, Weng D, Zhang Z, Bao Y, Jia Y, Nie M (2018) Hikeyb: high-efficiency mixed reality system for text entry. In: 2018 IEEE international symposium on mixed and augmented reality adjunct (ISMAR-adjunct). IEEE, pp 132–137

Jing A, Xiang C, Kim S, Billinghurst M, Quigley A (2019) Snapchart: an augmented reality analytics toolkit to enhance interactivity in a collaborative environment. In: The 17th international conference on virtual-reality continuum and its applications in industry, pp 1–2

Kern F, Kullmann P, Ganal E, Korwisi K, Stingl R, Niebling F, Latoschik ME (2021) Off-the-shelf stylus: using XR devices for handwriting and sketching on physically aligned virtual surfaces. Front Virtual Reality 2:69

Kim YR, Kim GJ (2017) Hovr-type: smartphone as a typing interface in VR using hovering. In: 2017 IEEE international conference on consumer electronics (ICCE). IEEE, pp 200–203

Knierim P, Schwind V, Feit AM, Nieuwenhuizen F, Henze N (2018) Physical keyboards in virtual reality: analysis of typing performance and effects of avatar hands. In: Proceedings of the 2018 CHI conference on human factors in computing systems, pp 1–9

Knierim P, Kosch T, Groschopp J, Schmidt A (2020) Opportunities and challenges of text input in portable virtual reality. In: Extended abstracts of the 2020 CHI conference on human factors in computing systems, pp 1–8

Kumar C, Hedeshy R, MacKenzie IS, Staab S (2020) Tagswipe: touch assisted gaze swipe for text entry. In: Proceedings of the 2020 CHI conference on human factors in computing systems, pp 1–12

Le K-D, Tran TQ, Chlasta K, Krejtz K, Fjeld M, Kunz A (2021) Vxslate: exploring combination of head movements and mobile touch for large virtual display interaction. In: Designing interactive systems conference, pp 283–297

Lee H, Je S, Kim R, Verma H, Alavi H, Bianchi A (2019) Partitioning open-plan workspaces via augmented reality. Pers Ubiquit Comput 1–16

Li J, George C, Ngao A, Holländer K, Mayer S, Butz A (2021) Rear-seat productivity in virtual reality: investigating VR interaction in the confined space of a car. Multimodal Technol Inter 5(4):15

Lu X, Yu D, Liang H-N, Xu W, Chen Y, Li X, Hasan K (2020) Exploration of hands-free text entry techniques for virtual reality. In: 2020 IEEE international symposium on mixed and augmented reality (ISMAR). IEEE, pp 344–349

Ma X, Yao Z, Wang Y, Pei W, Chen H (2018) Combining brain-computer interface and eye tracking for high-speed text entry in virtual reality. In: 23rd International conference on intelligent user interfaces, pp 263–267

McGill M, Boland D, Murray-Smith R, Brewster S (2015) A dose of reality: overcoming usability challenges in VR head-mounted displays. In: Proceedings of the 33rd annual ACM conference on human factors in computing systems, pp 2143–2152

Mendes D, Caputo FM, Giachetti A, Ferreira A, Jorge J (2019) A survey on 3d virtual object manipulation: from the desktop to immersive virtual environments. Comput Graph Forum 38(1):21–45

Nguyen C, DiVerdi S, Hertzmann A, Liu F (2017a) Collavr: collaborative in-headset review for VR video. In: Proceedings of the 30th annual ACM symposium on user interface software and technology, pp 267–277

Nguyen C, DiVerdi S, Hertzmann A, Liu F (2017b) Vremiere: in-headset virtual reality video editing. In: Proceedings of the 2017 CHI conference on human factors in computing systems, pp 5428–5438

Nordahl R, Nilsson NC, Adjorlu A, Magalhaes E, Willemsen S, Andersson NS, Wang J, Serafin S (2019) 12 hours in virtual reality: two cases of long-term exposure to consumer-grade virtual reality. In: 2019 IEEE conference on virtual reality and 3D user interfaces (VR)

Norman M, Lee GA, Smith RT, Billingurst M (2019) The impact of remote user's role in a mixed reality mixed presence system. In: The 17th international conference on virtual-reality continuum and its applications in industry, pp 1–9

Normand E, McGuffin MJ (2018) Enlarging a smartphone with AR to create a handheld vesad (virtually extended screen-aligned display). In: 2018 IEEE international symposium on mixed and augmented reality (ISMAR). IEEE, pp 123–133

O'Brien E, Jacquouton B, Moineau A, Campbell AG (2019) Wikipedia in virtual reality and how text-based media can be explore in virtual reality. In: Proceedings of the 2019 international conference on artificial intelligence and advanced manufacturing, pp 1–9

Oculus. http://www.oculus.org. Accessed: 28 July 2021

O'Hagan J, Williamson JR (2020) Reality aware VR headsets. In: Proceedings of the 9th ACM international symposium on pervasive displays, pp 9–17

O'Hagan J, Williamson JR, Khamis M (2020) Bystander interruption of VR users. In: Proceedings of the 9th ACM international symposium on pervasive displays, pp 19–27

Otte A, Schneider D, Menzner T, Gesslein T, Gagel P, Grubert J (2019) Evaluating text entry in virtual reality using a touch-sensitive physical keyboard. In: 2019 IEEE international symposium on mixed and augmented reality adjunct (ISMAR-adjunct), pp 387–392

Park KS, Kapoor A, Leigh J (2000) Lessons learned from employing multiple perspectives in a collaborative virtual environment for visualizing scientific data. In: Proceedings of the third international conference on collaborative virtual environments, pp 73–82

Pavanatto L, North C, Bowman DA, Badea C, Stoakley R (2021) Do we still need physical monitors? An evaluation of the usability of AR virtual monitors for productivity work. In: 2021 IEEE virtual reality and 3D user interfaces (VR). IEEE, pp 759–767

Pejsa T, Kantor J, Benko H, Ofek E, Wilson A (2016) Room2room: enabling life-size telepresence in a projected augmented reality environment. In: Proceedings of the 19th ACM conference on computer-supported cooperative work and social somputing (CSCW'16)

Perelman G, Serrano M, Raynal M, Picard C, Derras M, Dubois E (2015) The roly-poly mouse: designing a rolling input device unifying 2d and 3d interaction. In: Proceedings of the 33rd annual ACM conference on human factors in computing systems, pp 327–336

Pham D-M, Stuerzlinger W (2019) Hawkey: efficient and versatile text entry for virtual reality. In: 25th ACM symposium on virtual reality software and technology, pp 1–11

Pretsch J, Pretsch E, Saretzki J, Kraus H, Grossmann G (2020) Improving employee well-being by means of virtual reality-realex: an empirical case study. Eur J Econ Bus Stud 6(1):95–105

Rajanna V, Hansen JP (2018) Gaze typing in virtual reality: impact of keyboard design, selection method, and motion. In: Proceedings of the 2018 ACM symposium on eye tracking research & applications, pp 1–10

Reski N, Alissandrakis A, Tyrkkö J, Kerren A (2020) "oh, that's where you are!"—towards a hybrid asymmetric collaborative immersive analytics system. In: Proceedings of the 11th Nordic conference on human-computer interaction: shaping experiences, shaping society, pp 1–12

Richardson M, Durasoff M, Wang R (2020) Decoding surface touch typing from hand-tracking. In: Proceedings of the 33rd annual ACM symposium on user interface software and technology, pp 686–696

Romat H, Fender A, Meier M, Holz C (2021) Flashpen: a high-fidelity and high-precision multi-surface pen for virtual reality. In: 2021 IEEE virtual reality and 3D user interfaces (VR). IEEE, pp 306–315

Roo JS, Hachet M (2017) One reality: augmenting how the physical world is experienced by combining multiple mixed reality modalities. In: Proceedings pf UIST 2017, UIST'17. Association for Computing Machinery, New York, NY, USA, pp 787–795. https://doi.org/10.1145/3126594.3126638

Ruvimova A, Kim J, Fritz T, Hancock M, Shepherd DC (2020) "transport me away": fostering flow in open offices through virtual reality. In: Proceedings of the 2020 CHI conference on human factors in computing systems, pp 1–14

Schneider D, Otte A, Gesslein T, Gagel P, Kuth B, Damlakhi MS, Dietz O, Ofek E, Pahud M, Kristensson PO et al (2019) Reconviguration: reconfiguring physical keyboards in virtual reality. IEEE Trans Vis Comput Graph 25(11):3190–3201

Sharma G, Shroff G, Dewan P (2011) Workplace collaboration in a 3d virtual office. In: 2011 IEEE international symposium on VR innovation. IEEE, pp 3–10

Shen R, Weng D, Chen S, Guo J, Fang H (2019) Mental fatigue of long-term office tasks in virtual environment. In: 2019 IEEE international symposium on mixed and augmented reality adjunct (ISMAR-adjunct). IEEE, pp 124–127

Simeone AL (2016) The VR motion tracker: visualising movement of non-participants in desktop virtual reality experiences. In: 2016 IEEE 2nd workshop on everyday virtual reality (WEVR). IEEE, pp 1–4

Slater M, Sanchez-Vives MV (2016) Enhancing our lives with immersive virtual reality. Front Robot AI 3:74

Son J, Ahn S, Kim S, Lee G (2019) Improving two-thumb touchpad typing in virtual reality. In: Extended abstracts of the 2019 CHI conference on human factors in computing systems, pp 1–6

Spatial.io. https://spatial.io. Accessed: 29 July 2021

Speicher M, Feit AM, Ziegler P, Krüger A (2018) Selection-based text entry in virtual reality. In: Proceedings of the 2018 CHI conference on human factors in computing systems, pp 1–13

Stansfeld S, Candy B (2007) Psychosocial work environment and mental health-a meta-analytic review. Scand J Work Environ Health 32:443–462. https://doi.org/10.5271/sjweh.1050

Steinicke F, Bruder G (2014) A self-experimentation report about long-term use of fully-immersive technology. In: Proceedings of the 2nd ACM symposium on spatial user interaction, pp 66–69

Stewart T (1979) Eyestrain and visual display units: a review. Displays 1(1):17–24

Tang A, Pahud M, Inkpen K, Benko H, Tang JC, Buxton B (2010) Three's company: understanding communication channels in three-way distributed collaboration. In: Proceedings of the 2010 ACM conference on computer supported cooperative work, CSCW'10. Association for Computing Machinery, New York, NY, USA, pp 271–280. https://doi.org/10.1145//1718918.1718969

Thoondee KD, Oikonomou A (2017) Using virtual reality to reduce stress at work. In: 2017 Computing conference. IEEE, pp 492–499

Valtchanov D, Barton KR, Ellard C (2010) Restorative effects of virtual nature settings. Cyberpsychol Behav Soc Networking 13(5):503–512

Wagner Filho JA, Freitas CMDS, Nedel L (2018) Virtualdesk: a comfortable and efficient immersive information visualization approach. Comput Graph Forum 37(3):415–426

Wagner Filho JA, Freitas CM, Nedel L (2019) Comfortable immersive analytics with the virtualdesk metaphor. IEEE Comput Graph Appl 39(3):41–53

Walker J, Li B, Vertanen K, Kuhl S (2017) Efficient typing on a visually occluded physical keyboard. In: Proceedings of the 2017 CHI conference on human factors in computing systems, pp 5457–5461

Wentzel J, d'Eon G, Vogel D (2020) Improving virtual reality ergonomics through reach-bounded non-linear input amplification. Association for Computing Machinery, New York, NY, USA, pp 1–12

Wille M, Adolph L, Grauel B, Wischniewski S, Theis S, Alexander T (2014) Prolonged work with head mounted displays. In: Proceedings of the 2014 ACM international symposium on wearable computers: adjunct program, pp 221–224

Xu W, Liang H-N, Zhao Y, Zhang T, Yu D, Monteiro D (2019) Ringtext: dwell-free and hands-free text entry for mobile head-mounted displays using head motions. IEEE Trans Vis Comput Graph 25(5):1991–2001

Xu W, Liang H-N, He A, Wang Z (2019b) Pointing and selection methods for text entry in augmented reality head mounted displays. In: 2019 IEEE international symposium on mixed and augmented reality (ISMAR). IEEE, pp 279–288

Yang JJ, Holz C, Ofek E, Wilson AD (2019) Dreamwalker: substituting real-world walking experiences with a virtual reality. In: Proceedings of the 32nd annual ACM symposium on user interface software and technology, UIST'19. Association for Computing Machinery, New York, NY, USA, pp 1093–1107. https://doi.org/10.1145/3332165.3347875

Yu D, Fan K, Zhang H, Monteiro D, Xu W, Liang H-N (2018) Pizzatext: text entry for virtual reality systems using dual thumbsticks. IEEE Trans Vis Comput Graph 24(11):2927–2935

Zielasko D, Weyers B, Bellgardt M, Pick S, Meibner A, Vierjahn T, Kuhlen TW (2017) Remain seated: towards fully-immersive desktop VR. In: 2017 IEEE 3rd workshop on everyday virtual reality (WEVR). IEEE, pp 1–6

Zielasko D, Krüger M, Weyers B, Kuhlen TW (2019a) Passive haptic menus for desk-based and HMD-projected virtual reality. In: 2019 IEEE 5th workshop on everyday virtual reality (WEVR). IEEE, pp 1–6

Zielasko D, Weyers B, Kuhlen TW (2019b) A non-stationary office desk substitution for desk-based and HMD-projected virtual reality. In: 2019 IEEE conference on virtual reality and 3D user interfaces (VR). IEEE, pp 1884–1889

Chapter 4

Breaking the Screen: Interaction Across Touchscreen Boundaries in Virtual Reality for Mobile Knowledge Workers

Summary

As mentioned earlier, Virtual Reality can provide a large, three-dimensional display space that knowledge workers could benefit from. This can be specifically helpful in mobile contexts, where it is not possible for users to have their ideal work-setup with them. Therefore, this article describes a system in which the user can use a tablet, which is a standard portable device, and use it to interact with information displayed through a VR HWD. Using the small surface of the tablet to interact with the large display space in VR opened up the question of how to efficiently map the input to the virtual screen. Therefore, a multimodal technique combining eye-gaze-tracking with touch input was compared to a touch-only technique in a user study. The results showed that the multimodal technique outperforms the touch-only technique. A second study also confirmed that users can benefit from the three-dimensional display, as compared to a two-dimensional, when handling information on multiple layers. Finally, the article also presents a set of example applications that combine the touch-input with VR for specific knowledge worker tasks such as a window manager, code version control or a parallel coordinates plot. In another user study, they were presented to participants who generally found them easy to use, useful, and enjoyable.

Overall, this article showed that the combination of VR and a tablet is feasible and through the small form factor could be useful for mobile knowledge workers, especially in constrained spaces such as public transportation.

Contribution Statement

Per Ola Kristensson, Eyal Ofek, Michel Pahud and Jens Grubert worked out the design space. Verena Biener was mainly responsible for the design and implementation of the content transfer task and Daniel Schneider for the puzzle task respectively. Both tasks were repeatedly discussed with each other and Per Ola Kristensson, Eyal Ofek, Michel Pahud and Jens Grubert. The user studies for both tasks were conducted by Verena Biener, Daniel Schneider and Eyal Ofek. The statistical analysis for both tasks as well as the corresponding figures was mainly contributed by Verena Biener. The six applications described in the article were designed by Per Ola Kristensson, Eyal Ofek, Michel Pahud and Jens Grubert and implemented and evaluated in a user study by by Daniel Schneider, Travis Gesslein, Alexander Otte and Bastian Kuth. The first draft of the article was written by Verena Biener. and Jens Grubert and it was then rewritten by all authors.

Article [18]

© 2020 IEEE. Reprinted, with permission, from Verena Biener, Daniel Schneider, Travis Gesslein, Alexander Otte, Bastian Kuth, Per Ola Kristensson, Eyal Ofek, Michel Pahud & Jens Grubert. Breaking the Screen: Interaction Across Touchscreen Boundaries in Virtual Reality for Mobile Knowledge Workers. In IEEE Transactions of Visualization and Computer Graphics (TVCG), vol. 26, no. 12, Dec. 2020, pp. 3490-3502. `https://doi.org/10.1109/TVCG.2020.3023567`.

In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of Coburg University's and University of Bayreuth's products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to `http://www.ieee.org/publications_standards/publications/rights/rights_link.html` to learn how to obtain a License from RightsLink. If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.

Breaking the Screen: Interaction Across Touchscreen Boundaries in Virtual Reality for Mobile Knowledge Workers

Verena Biener[1] Daniel Schneider[1] Travis Gesslein[1] Alexander Otte[1] Bastian Kuth[1]
Per Ola Kristensson[3] Eyal Ofek[2] Michel Pahud[2] Jens Grubert[1]*

[1]Coburg University of Applied Sciences [2]Microsoft Research
[3]University of Cambridge

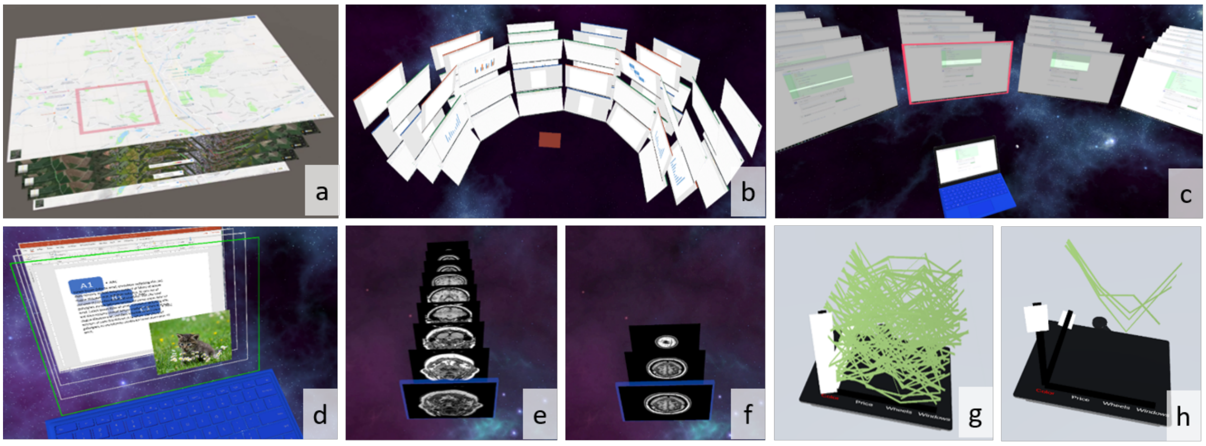

Fig. 1. Mobile knowledge worker applications using a touchscreen in VR implemented in this paper. a: map navigation, b: window manager, c: code version control, d: presentation editor, e-f: medical imaging, g-h: information visualization.

Abstract— Virtual Reality (VR) has the potential to transform knowledge work. One advantage of VR knowledge work is that it allows extending 2D displays into the third dimension, enabling new operations, such as selecting overlapping objects or displaying additional layers of information. On the other hand, mobile knowledge workers often work on established mobile devices, such as tablets, limiting interaction with those devices to a small input space. This challenge of a constrained input space is intensified in situations when VR knowledge work is situated in cramped environments, such as airplanes and touchdown spaces.
In this paper, we investigate the feasibility of interacting jointly between an immersive VR head-mounted display and a tablet within the context of knowledge work. Specifically, we 1) design, implement and study how to interact with information that reaches beyond a single physical touchscreen in VR; 2) design and evaluate a set of interaction concepts; and 3) build example applications and gather user feedback on those applications.

Index Terms—virtual reality, knowledge work, mobile office, window management, eye tracking, multimodal interaction

1 INTRODUCTION

Recent progress in virtual reality (VR) technology makes it possible to provide knowledge workers with a portable virtual office. This office can provide many potential advantages for mobile work, such as 1) providing a well-illuminated, private environment with wide display areas regardless of the physical surroundings; 2) enabling a virtual recreation of a consistent spatial workspace; and 3) relieving the user of physical limitations, using large and even distant displays while the user is remaining seated and the hands are resting on a table [39, 81].

The vision of a spatial user interface supporting knowledge work has been investigated for many years (e.g. [92, 94]). However, the recent emergence of consumer VR headsets now make it feasible to explore the design of deployable user interface solutions.

However, while VR is promising for mobile office work it also brings its own additional challenges. For example, mobile workers often work in front of touchscreens, such as tablets or laptops. Within this work, we propose to extend the limited output and input space of such configurations by utilizing VR head-mounted displays (HMDs). Specifically, recent HMDs with inside out camera-based sensing allow for pose tracking of the user's head and hands. Through access to HMDs cameras, spatial tracking of further objects, such as screens or keyboards become feasible.

Given these opportunities, this paper explores how to leverage the large display area and additional three-dimensional display volume around the user afforded by an HMD while using established capacitive touch interfaces provided with tablets and some laptops.

This paper addresses this scenario with the following contributions. First, we explore and evaluate how to spatially arrange and manipulate information within the joint interaction space of HMD-tablet interaction. Second, we design, implement and evaluate six applications (window manager, code version control, parallel coordinates exploration, map navigation, volumetric image viewer, and a virtual PowerPoint), see Figure 1.

---

*contact author: jens.grubert@hs-coburg.de

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

2 RELATED WORK

The work in this paper is underpinned by the following research areas: 1) mixed reality (MR) for knowledge work; 2) information windows in spatial environments; and 3) spatial interaction.

2.1 Mixed Reality for Knowledge Work

Recent prior work has begun exploring how to support knowledge work using MR [39, 44, 81, 96]. Early work investigated the use of projection systems to extend physical office environments for interaction with physical documents (e.g., [54, 88, 94, 122]). VR and AR HMDs have also been explored for interacting with physical documents (e.g., [36, 62]). Most prior work has focused on annotating documents displayed on a 2D surface while this work investigates the use of space surrounding a planar piece of information.

In addition to enhanced document interaction, prior work has also explored remote presence applications (e.g., [83, 92]). There is a number of publications investigating the use of VR in desktop-based environments for tasks such as text entry (e.g., [41, 53, 70]), system control [132, 133] and visual analytics [120]. Büschel et al. [15] surveyed a wide range of immersive interaction techniques for visual analytics. Previous research on productivity desktop-based VR has concentrated on the use of physical keyboards [99], controllers and hands [55, 133], and, recently, tablets [111]. The closest work to this paper is by Surale et al. [111], which focuses on spatial manipulation and creation of 3D objects. In contrast, we use the tablet for information management on virtual displays.

Complementary to this prior research, this paper aims to support mobile knowledge workers by extending commonly used tools such as tablets and notebooks through HMDs.

2.2 Information Windows in Spatial Environments

In 1993, Feiner et al. [31] introduced head-surrounding and world reference frames for positioning 3D windows in VR. In 1999, Mark Billinghurst and Thad Starner [8] introduced the spatial display metaphor, in which information windows are arranged on a virtual cylinder around the user.

Since then, virtual information displays have been explored in various reference systems, such as world-, object-, head-, body- or device-referenced [57]. Specifically, interacting with windows in body-centered reference systems [119] has attracted attention, for instance to allow fast access to virtual items [19, 61], mobile multi-tasking [27, 29] and visual analytics [28]. Gugenheimer et al. [43] introduced face touch, which allows interacting with display-fixed user interfaces (using direct touch) and world-fixed content (using raycasting). Yu et al. [129] investigated the use of body motions for switching interfaces in VR. Lee et al. [58] investigated positioning a window in 3D space using a continuous hand gesture. Petford et al. [84] compared the selection performance of mouse and raycast pointing in full coverage displays (not in VR). Recently, Jetter et al. [50] proposed to interactively design a space with various display form factors in VR. Wagner et al. proposed a desk metaphor for controlling visual analytics that reappropriates a physical desk in VR [120].

Prior work has also explored how to support user interaction across HMDs and mobile and wearable touch displays. Grubert et al. [37] presented body-aligned, device-aligned, and side-by-side modes for interacting between a touch display (smartphone or smartwatch) and an optical see-through HMD. Similar explorations have followed suit using video-see-through HMDs [80], an extended set of interaction techniques [131], using smartwatches [66, 123, 126], or with a focus on understanding smartphone-driven window management techniques for HMDs [95]. In a similar vein, prior work has studied the interaction across HMDs and stationary displays, such as tabletops [16, 93, 102].

Most prior work relate to the issue of field of view, that is, how to display and access more information by spreading it around the user in multiple windows. In this research, we are additionally interested in extending the information display in the depth direction. Such a display is suited for displaying different views of information (or layers) that are semantically connected by their 2D location. On the other hand, we use a very limited input space: While most referred prior work span all the angular range of the display as an input space, we only use the interaction space on or near the tablet in order to support interaction in constrained physical spaces [39, 71].

2.3 Spatial Interaction

A large number of techniques for selection, spatial manipulation, navigation, and system control have been proposed to support spatial user interfaces [57]. Regarding object selection, Argelaguet et al. [2] presented a survey on 3D selection techniques. For a recent survey on 3D virtual object manipulation, we refer to Mendes et al. [72]. Finally, recent surveys [14, 38] have extensively reviewed spatial interaction across mobile devices, mostly in non-VR settings.

In addition to unimodal techniques the combination of touch with mid-air has attracted attention from researchers. For example, outside of VR, Müller et al. [75] investigated the use of touch and mid-air interaction on public displays, Hilliges et al. [46] studied tabletop settings. Several researchers have proposed to use handheld touchscreens in spatial user interfaces for tasks such as sketching, ideation and modeling (e.g., [3, 23, 25, 33, 48, 91]), navigation of volumetric data [106], 3D data exploration [65] and system control [10]. Spatial manipulation has mostly been studied in single-user settings (e.g., [5, 51, 63, 68, 74, 111, 112]) but also in collaborative settings [35].

Evolving from the magic lens [7] and tangible interaction concepts [117], tangible magic lenses allow users to access and manipulate otherwise hidden data in interactive spatial environments. A wide variety of interaction concepts have been proposed within the scope of information visualization (e.g., recent surveys [114, 115]). Both rigid shapes (e.g., rectangular [107]) or circular [108] and flexible shapes (e.g., [109]) have been used, as well as various display media (e.g., projection on cardboard [17, 107]), transparent props [12, 98], handheld touchscreens [40, 59], or virtual lenses [64, 82]. In addition, the combination of eye-gaze with other modalities such as touch [85, 86], mid-air gestures [87, 97, 101] and head-movements [56, 103, 104] has been recently investigated for interaction in spatial user interfaces. For a recent survey on gaze-based interaction in AR and VR, see Hirzle et al. [47].

We draw on these rich sources of interaction ideas and adopt techniques in the context of VR interaction with touchscreens for mobile knowledge workers. Our work complements multimodal techniques combining touch and mid-air [46, 75], gaze-based techniques [85] and ideas for combining HMDs with touchscreens [37, 80, 131] through novel techniques for accessing virtual windows around or behind a physical touchscreen.

3 DESIGN SPACE

Currently most knowledge workers' applications are designed for 2D displays, which are the vast majority of displays in the world. VR and AR displays enable mobile workers to take their displays with them on the go. In addition, since HMDs are stereoscopic, they enable information presentation beyond the 2D display and the ability to manipulate this information by spatial manipulation. Researchers have already proposed various schemes for arranging (planar) information in a spatial context relative to the user. Specifically, prior work has proposed various reference systems, such as world-, object-, head-, body- or device-referenced [9], spatial window layouts in these reference frames, for example, scaffolding geometry such as spheres or cylinders [27, 28], and input modalities to access and manipulate information in these spaces (such as touch, gaze, mid-air interaction or multimodal interaction, c.f. [15, 47, 119]).

The guiding principle in this paper is that an information touchscreen display can break out of the screen in VR and transition from 2D to 3D—yet still be controlled by manipulating the original touchscreen. This is in part motivated by prior work reviewed in the previous section and in part motivated by the fact that a touchscreen allows the user to provide precise 2D input while displaying information in an HMD provides the ability to display information in 3D. This screen-centric manipulation using touch can be complemented with additional modalities, such as gaze tracking.

Two important aspects arising in this context are how to *spatially arrange* information relative to the touchscreen and how to *map the input from the physical touchscreen to the information in the virtual world*, given the joint input capabilities of 2D touch sensing (on the screen) and 3D spatial manipulation (around the screen).

3.1 Spatial Arrangement

Prior work has investigated options on how to align multiple 2D information windows around the user (e.g., [27,29]). One common approach is to arrange windows on a scaffolding geometry, such as spheres or cylinders [8](see Figure 2, b-d). In a spatially unconstrained environment, accessing this information is often realized using virtual hand or raycasting techniques. While using such extended display areas is possible, we should be aware that in constrained spaces, such as on an airplane or in touchdown spaces, it may not be usable for interaction as ample spatial movements can be unsuitable either due to social acceptability [1, 124] or simply due to lack of physical space to operate [39, 71].

An alternative display option is extending information along with the depth of the screen, extruding it into 3D (see Figure 2, a). This extrusion might be specifically suitable for additional information that is directly related to the document displayed on the physical screen. This information is often semantically different and should ideally be displayed separately from the main screen but with a spatial relation (specifically with corresponding x- and y-coordinates) to the document. For example, when reviewing a document, a common task consists of inserting comments. However, there is no natural space for the added comments within the document, and adding them in place can disrupt the layout of the document. Adding the comments in a separate layer, hovering in front of the main document, maintains the contextual relevance while enabling the user to focus on the layer of interest.

3.2 Input-Output Mapping

The touchscreen allows one layer of information at a time to align with the surface of the screen. This layer enjoys the easiest and most accurate input, as the screen supports the user's hand and enables the highest positional sensing, and, sometimes, pressure sensing. Next, we discuss, how to access layers of information that break out of the bounds of the physical screen.

3.2.1 Around the Screen Interaction

One challenge when interacting across screens is how to support both fine-grained selection on an active screen as well as efficient switching between screens. Naive implementations that use a fixed control-to-display gain (CD gain), such as stitching or pointer warping for bridging the space between displays [77, 121, 127], do not scale well to large displays due to the required trade-off between high precision pointing in a given region of interest on a display and fast switching between multiple regions of interest [78]. Hence, besides raycasting, prior work has proposed multiple strategies for controlling a large output space using a small input space on a touchscreen (e.g., [32, 69, 78, 79]).

Inspired by those techniques, we propose using a *bimanual selection technique* for allowing both precise input and efficient switching of the active window. This way, the user can move a cursor and select or deselect items inside the active window by simply using the tablet as a touchpad with a suitable CD gain for single screen interaction, see Figure 2, b. If the user touches the bezel of the touchscreen, which has a width of 2 cm, with their non-dominant hand, a coarse CD gain is activated to enable fluid switching between multiple windows, see Figure 2, c.

The CD gain for single screen interaction is a one-to-one mapping of the movement on the touchscreen. The CD gain for switching between multiple windows is set in such a way that the user has to move their finger 2 cm to switch to the next screen. We evaluated this CD gain along with values of 1 cm and 3 cm in an informal user study with five participants and found that 2 cm was the fastest and preferred by three participants while the other two CD gains were only preferred by one participant each.

As VR headsets also support head-pointing and partially eye-gaze tracking we implemented a second technique: *combined gaze and touch interaction*. Inspired by previous work on combining gaze and manual pointing [85, 130], we devised a technique in which the combined head and eye gaze provides the context for touch interaction, see Figure 2, d. With this technique regions of interest are discrete (the individual virtual screens). Hence, when a user gazes at a specific window, the cursor is transferred to that new screen, maintaining the same absolute position in the new screen as in the old screen (i.e., the same (x,y) coordinate).

In order to avoid unwanted cursor movements it is possible to use temporal thresholding (e.g., 200-300 ms [85, 130]), although we did not use this in our evaluation in order to allow for faster task completion times. To allow selection of items at the screen boundary, we use spatial thresholding with an empirically determined threshold of 5% of the window size in all directions. Specifically, to switch from one screen to another, the user first has to move their eyes beyond the second display boundary by that threshold. While we did not evaluate this threshold parameter in a formal user study, internal testing revealed that this threshold circumvents unwanted display switches and still allows for comfortable switching between display boundaries.

For both techniques, interacting with the now active virtual display can either 1) happen at the original location of the display, essentially turning the touchscreen into an indirect input device; or 2) occur by aligning the virtual display with the physical touchscreen. For example, dragging the finger on the bottom bezel of the touchscreen (or using a two-finger swipe) can rotate virtual windows around the x-axis of a scaffolding geometry (either display-by-display in discrete steps or continuously).

3.2.2 Depth Interaction

Accessing virtual displays along the depth dimension of the touchscreen can happen either in front, on, or behind the physical screen (see Figure 2, a). For virtual displays in front of the screen, direct mid-air selection might be used, but this may cause occlusion of content at layers further away and hinder interaction on the physical screen [13]. Accessing content behind the screen has to take into account the physical presence of the touchscreen, preventing direct access using virtual hand techniques. Users would either need to grab behind the screen using a virtual hand or use raycasting techniques that can handle object occlusion (e.g., [30]). Instead, we propose to use the touchscreen to select a desired virtual display. Depending on the number of virtual displays spread behind or in front of the physical screen, different mappings can be appropriate. We experimented with swiping along the bezel of the screen and two-finger swipe gestures, commonly available on multi-touch trackpads. Depending on the number of layers an absolute mapping for quick random access to layers or a relative mapping for navigating between adjacent layers might be appropriate. In an informal user study with three users, we found that swiping with two fingers outperformed bezel swipes and a relative mapping outperformed an absolute mapping for four and ten layers. Hence, we used this input mapping for the further development of the applications. We used the same CD gain as for switching between individual screens spread around the active screen, as described above.

While the input technique allows to quickly switch between layers it does not completely resolve potential occlusion issues. Occlusion can be mitigated (but not completely resolved) utilizing parallax if the user moves the head. We experimented with two visualization techniques to further mitigate potential occlusion. One option is to amplify the parallax effect when moving the head to the side of the screen, see Figure 7, f. Another option, inspired by explosion diagrams, is to temporarily scatter virtual screens around the physical screen, see Figure 7, e. While the explosion diagram technique allows the user to view all layers simultaneously and unoccluded, the number of concurrently visible layers is constrained by the available screen space. Also for both techniques, potential dependencies between layers are impacted (e.g., when there is a need for carefully arranging multiple graphical elements into a composite image). Another limitation of the explosion diagram technique is the reduction of the ability to associate

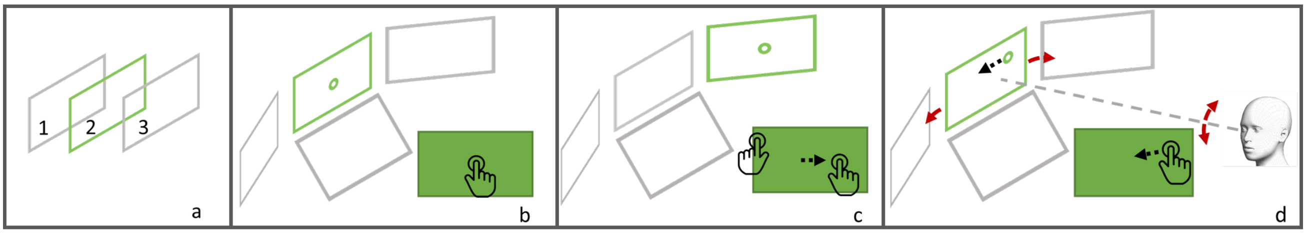

Fig. 2. a: Virtual displays (grey) arranged relative to the physical screen (green): behind (1), with (2), in front of (3). b and c: Changing between pointing on a single screen and switching between screens using a bimanual technique. b: If only a single finger is used on the touchscreen (filled green rectangle), it controls the cursor on the active screen (green boundary) with a suitable control-display gain. c: The screens are switched by pressing with a finger of the non-dominant hand on the bezel of the touchscreen and moving the finger of the dominant screen towards the second screen (in this case the right screen) using a second CD-ratio. d: Changing between pointing on a single screen and switching between screens using a combination of touch and gaze. The finger on the touch screen (filled green rectangle) controls the cursor on the active screen (green boundary). Gaze provides the region of interest. If the user gazes to the side, up or down, the corresponding screen gets activated (with appropriate temporal and spatial thresholds).

different objects between different layers based on their common 2D locations.

3.2.3 Single Screen Interaction

Even an individual layer that is aligned with the physical screen does not need to retain its physical bounds. The ability to display the view frustum according to the rotation of the user's head enables a virtual display that potentially spreads over a very large angular part of the user's field of view. It enables the simulation of screens that may be bigger than physical screens available to the user (see Figure 7, d). In this case, the mapping of the input space of the physical touchscreen can become indirect. For example, it is possible to use a different CD gain for absolute mapping of the physical finger position to the virtual finger position (following ideas from haptic retargeting [4, 20]), see Figure 7, or the touchscreen can be operated using a relative mapping. In some cases, it might be desirable to retain a direct mapping between the input and the output space. In this case, only the portion co-located with the physical touchscreen allows interactivity. Changing this active input area can either be realized by two-finger swipes or input on the bezels of the screen (analogous to implementations in desktop sharing applications, such as TeamViewer) or by redefining the active area by gaze (again requiring an appropriate clutch mechanism).

4 EVALUATION OF DESIGN PARAMETERS

While we designed the techniques following an iterative approach with multiple design iterations consisting of conceptualization, implementation, and initial user tests (eating your own dog food) [24, 118], we aimed to understand properties of the proposed interaction space in this paper in more detail. Within the proposed interactive space, users can both extend the display area around the current display on a two-dimensional proxy geometry as well as in-depth in front or behind the physical touch screen. For the evaluation, we investigated these two properties (around and in-depth) separately.

To this end, we first aimed at exploring, if using combined touch and gaze interaction has benefits over touch-only interaction when interacting across multiple virtual screens arranged on a proxy geometry around the touchscreen. Second, we wanted to quantify the benefits of viewing multiple stacked information layers in-depth behind a touchscreen compared to only showing a single layer at a time as is common in many applications today. While adding depth cues has been indicated to improve task performance in various settings [6, 89, 90, 105], we aimed at quantifying any performance improvement within the scope of the joint tablet-HMD interaction.

Hence, we investigated those two aspects in a user study. In the first part (subsequently called the CONTENT TRANSFER TASK), we compared the performance of the *bimanual selection technique* with the COMBINED GAZE AND TOUCH INTERACTION technique in a content transfer task for a small and large number of screens.

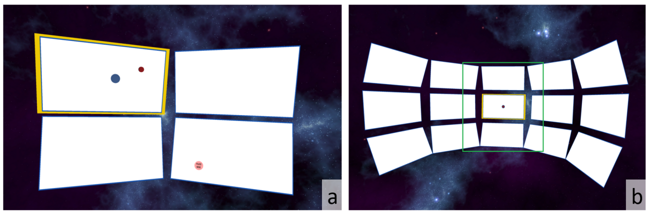

Fig. 3. Arrangements of the virtual screens for the CONTENT TRANSFER TASK. a: FOUR SCREENS: the blue dot can be acquired by the red cursor in the upper left screen and is to be placed on the goal in the lower right screen. b: FIFTEEN SCREENS: the green border indicates the field of view of the user.

In the second part (subsequently called the PUZZLE TASK), participants were asked to solve a puzzle task where each puzzle piece was displayed in an individual layer, mimicking tasks when composing presentations using multiple shapes or images using multiple layers in an image editing application. Note that we used the baseline technique described in Section *Depth Interaction* without amplified parallax or explosion diagram visualization due to the nature of the task having strong spatial dependencies between the x- and y position of the individual puzzle pieces.

For the CONTENT TRANSFER TASK, we followed a 2×2 within-subjects design. The first independent variable was INTERACTION TECHNIQUE for selecting the screens, which was done either GAZE-BASED or BIMANUAL as described above. The second independent variable was NUMBER OF SCREENS presented to the participant which was either FOUR SCREENS (Figure 3, a) or FIFTEEN SCREENS (Figure 3, b). We chose these configurations to better understand the performance of the techniques in presence of a few or many screens while still considering the field of view that can be comfortably covered when the user turns their head. This experimental design results in four different conditions, which are depicted in Figure 3. Dependent variables included tasks completion time (the duration from acquiring the content in one screen to placing it at the marked spot in another screen), accuracy (measured as Euclidean distance between the released item and the actual target location), usability as measured by the System Usability Scale (SUS) [11], workload as measured by NASA TLX (unweighted version) [45], simulator sickness (SSQ) [52] as well as user preferences. We hypothesized that the GAZE-BASED would outperform BIMANUAL in terms of task completion time, usability and workload, but not in terms of simulator sickness or accuracy (as the final movements were conducted with the finger of the dominant hand in both techniques).

For the PUZZLE TASK, we also followed a 2×2 within subjects design. The independent variables were VISUALIZATION with two levels: FLAT, a baseline condition in which all layers were displayed

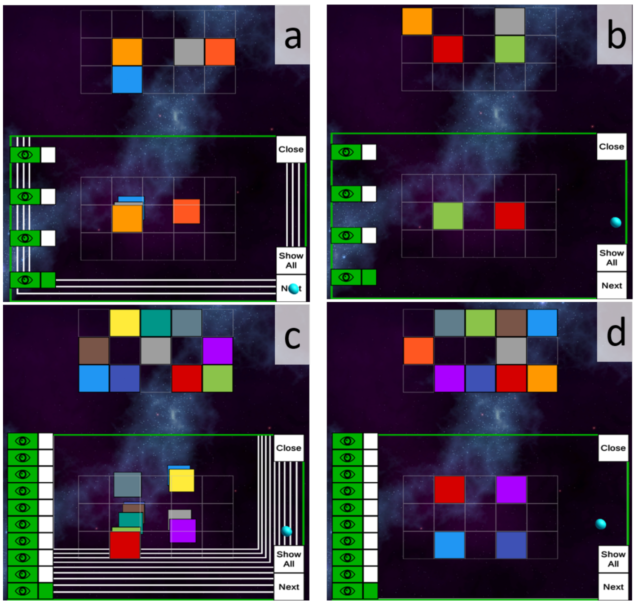

Fig. 4. Conditions for the PUZZLE TASK. a: DEPTH with FOUR LAYERS, b: FLAT with FOUR LAYERS, c: DEPTH with TEN LAYERS, d: FLAT with TEN LAYERS.

at the same depth, see Figure 4, b and d, and DEPTH, where each layer was displayed with increasing z-distance, see Figure 4, a and c. The second independent variable was NUMBER OF LAYERS with two levels: FOUR LAYERS with four layers displayed (Figure 4, a) and TEN LAYERS in which ten layers were displayed (Figure 4, c). This was done to investigate if an increasing number of layers has an effect on the performance of the visualization techniques.

As can be seen in Figure 4, the active layer is highlighted with a green frame, the user's fingertip is visualized with a turquoise sphere. The target arrangement is shown above the puzzle's layered display as a flat image. To the left of the layered display there is an overview widget, inspired by similar views in image editing applications. The eye symbol toggles the visibility of the respective layer (green: visible) and the rectangles to the right of the eye symbol allow direct selection of a layer (green: active layer, white: inactive layer). By selecting a layer (pressing on the green square associated with it), the system moved this layer to align with the screen depth and made all layers in front of it transparent, thereby maximizing the current layer visibility. Switching between two adjacent layers can also be achieved by swiping vertically with two fingers on the touchpad. The button "Show All" collapses all layers into a single layer and activates the visibility of all layers. The button "Next" switches to the next task. The button "Close" was only visible during training and allowed participants to end the training session when they felt comfortable with conducting the task.

The dependent variables were, as in the CONTENT TRANSFER TASK, task completion time, usability (SUS), workload (TLX) as well as simulator sickness (SSQ). In addition, the system logged the number of incorrectly placed puzzle items. While we hypothesized that DEPTH would outperform FLAT in terms of task completion time due to the added depth cues, we were specifically interested in quantifying this difference. We also hypothesized that DEPTH would lead to significantly higher usability rating and lower workload compared to FLAT with no difference in simulator sickness. We still included SSQ to check if severe symptoms would occur.

4.1 Participants

We recruited 14 participants (5 female, 9 male, mean age 30.07 years, sd = 10.59, mean height 176.57 cm, sd = 8.54). All of them indicated prior VR experience. Four participants used head mounted VR devices very frequently, two often, four sometimes, three rarely and one only once. One participant indicated she does not play video games, two rarely, four participants sometimes, three participants often and four participants very frequently. Seven participants wore contact lenses or glasses with corrected to normal vision. Eleven participants were right handed while three were left handed. All but one used their dominant hand to control the cursor and their non-dominant hand on the bezel. Thirteen participants used their index finger on the touchscreen and one used her middle finger.

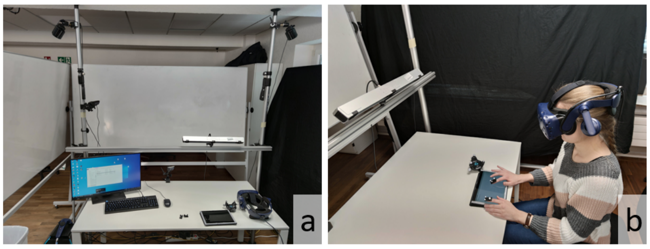

Fig. 5. The study setup. a: setup with motion tracked tablet, fiducials for the fingers and HTC Vive Pro Eye to the right. The lighthouse systems were placed on the ceiling around the table and the Optitrack system on a bar above the table. The monitor, keyboard and mouse to the left were operated by the experimenter to start individual conditions. b: Participant during the study.

4.2 Apparatus

The experiment took place in two locations: a laboratory environment in Europe and a home in the US. The study setup (except the PC) was replicated in both environments. A HTC Vive Pro Eye was used as VR HMD with built-in eyetracking. For touch input a Huawei Media Pad T5 was used (screen diagonal 10.1 inches, 16:10 aspect ratio). Velcro was attached to the left, right and bottom bezel of the tablet to support participants in identifying the touchscreen boundaries. The tablet was placed on a table in front of the participant in such a way that participants sitting in front of it on a chair could comfortably use it. The system was implemented in Unity 2019.3 and deployed on a PC with windows 10, an AMD Ryzen Threadripper 1950X 16-core processor with 3.4 Ghz, two Nvidia GeForce RTX 2080 graphics cards and 64 GB RAM in Europe and on a PC with Intel Core i9-9980HK 8-core processor, Nvidia GeForce RTX 2080 graphics card, 16 GB RAM in the US.

An OptiTrack V120 Trio tracking system was used for spatial tracking of index finger tips by using two rigid bodies with retro-reflective markers attached to them. Note that motion trajectories of the fingers were logged for comparative analysis in future experiments only, they were not analyzed within the scope of this paper. The Vive Lighthouse tracking system was used to track the VR HMD and the tablet. However, while the position of the tablet could have been changed by the participants during the experiment, no participant chose to do so. The study setup is shown in Figure 5.

For the CONTENT TRANSFER TASK, multiple screens were placed around the participant's head as shown in Figure 3, a and b. The virtual screen size was 24 inches and virtual screens were initially placed at a distance of 90 cm around the participant's head (if the participant moved their head forwards or backwards this distance changed). The horizontal angle spanned by the screens was at most 190 degrees and the vertical angle less than 30 degrees upwards and less than 35 degrees downwards. This allowed the participant to comfortably look at all screens. The cursor position was indicated as a disc with a diameter of 2.5 cm. For the present study, we chose to use no temporal threshold in order to obtain a comparable switching effect between both techniques.

The PUZZLE TASK used the same screen size and the same distance to the participant. The individual layers were placed 2.5 cm apart (this parameter was determined empirically).

4.3 Tasks

For the CONTENT TRANSFER TASK several windows were placed in a circle around the participant's head. Then they had to select a disk with a diameter of about 4 cm on one screen to a target area with the same diameter on another screen as shown in Figure 3 a. The window with the next target area was randomly selected while making sure that both small and longer distances were included during the condition. The disk and the target area appeared at one of eight randomly selected places around the center of a window. Both INTERACTION TECHNIQUES used the tablet as a touch pad to control the cursor within the active screen and to grab and release the dot via long-press. After successful placement of the disk onto the target area (with at least a partial overlap of the two), the task was repeated with a different arrangement of start and target areas.

For the PUZZLE TASK, participants were asked to arrange puzzle pieces (5×5 cm) as indicated by a template image, see Figure 4. Each puzzle piece was placed on a separate layer. The puzzle pieces snapped into a predefined 5×3 grid to facilitate accurate placement.

Participants were asked to advance to the next puzzle if they felt they had completed the task. The task was then repeated with a different template.

4.4 Procedure

After an introduction, participants were first asked to fill out a demographic questionnaire. Then the HTC Vive Pro Eye eye-calibration procedure was performed to ensure a working eye-tracking calibration for each participant. Thereafter the participants either started with the CONTENT TRANSFER TASK or the PUZZLE TASK (counterbalanced). For both tasks, the conditions were ordered using a balanced Latin square. This resulted in four permutations which could not be equally distributed among 14 participants. However, no significant order effects were detected. For each condition in the CONTENT TRANSFER TASK, participants performed a training phase where they completed ten content transfers. After that, they performed the actual task 32 times. The number of repetitions was chosen so that each distance and orientation in the FIFTEEN SCREENS condition appeared once (1–4 screens horizontally × 1–2 screens vertically × two directions each = 8 × 2 × 2 = 32). For each condition in the PUZZLE TASK, participants also performed a training phase where they completed puzzle tasks until they felt comfortable (most users only conducted one practice task). Thereafter, they performed the actual puzzle task ten times per condition. In both tasks, after each condition, participants completed the SSQ, SUS and NASA TLX questionnaires. At the end of each block they also answered a questionnaire about their preferences and participated in a semi-structured interview. Finally, the participants were thanked and compensated with voucher worth 10 Euro.

For the CONTENT TRANSFER TASK, we collected 32 repetitions × 4 conditions × 14 users = 1792 data points for the performance data and 4 conditions × 14 users = 56 data points for the subjective feedback. For the PUZZLE TASK, we collected 10 repetitions × 4 conditions × 14 users = 560 data points for the performance data and 4 conditions × 14 users = 56 data points for the subjective feedback.

4.5 Results

Statistical significance tests for log-transformed target acquisition time was carried out using general linear model repeated measures analysis of variance (RM-ANOVA) with Holm-Bonferroni adjustments for multiple comparisons at an initial significance level $\alpha = 0.05$. We indicate effect sizes whenever feasible (η_p^2).

For subjective feedback, or data that did not follow a normal distribution or could not be transformed to a normal distribution using the log-transform (such as errors), we employed the Aligned Rank Transform [125] before applying RM-ANOVAs.

The analysis of results did not reveal significant differences between the participant pool in the US and the one in Europe. Hence, we will solely report the joint results.

Due to logging errors in the CONTENT TRANSFER TASK, we only obtained data from 31 instead of 32 tasks. For one participant we only received about half of the data in one GAZE-BASED condition due to technical problems. Four values for the distance between goal and disc were excluded, because they were much higher than should be possible when successfully placing the disc. Note that slightly fewer values for some participants should not affect the overall results as we always used the mean performance for each participant in each condition.

The results in the following sections can be summarized as follows: For the CONTENT TRANSFER TASK, participants acquired targets significantly faster with the GAZE-BASED method compared to BIMANUAL (ca. 30%) and the GAZE-BASED resulted in significantly higher SUS ratings. No significant differences regarding accuracy, number of errors or simulator sickness ratings were detected between conditions. All but one participant preferred the GAZE-BASED technique.

For the PUZZLE TASK, participants performed the task significantly faster with DEPTH visualization compared to FLAT (approximately 15 %) but also made significantly more errors. The DEPTH visualization resulted in significantly lower mental demand and resulted in a significantly higher usability rating compared to FLAT.

4.5.1 Performance

For the CONTENT TRANSFER TASK there was a significant main effect of INTERACTION TECHNIQUE for task completion time, such that the GAZE-BASED method ($M = 3.70\ s, SD = 0.98$) resulted in a shorter task completion time than the BIMANUAL method ($M = 5.49\ s, SD = 2.33$). As expected, the main effect of the NUMBER OF SCREENS on the task completion time was also significant such that FOUR SCREENS ($M = 3.66\ s, SD = 1.20$) resulted in a shorter task completion time than FIFTEEN SCREENS ($M = 5.52\ s, SD = 2.20$). This was predictable because moving the target across multiple columns and rows takes longer than moving it only across two columns and rows. There was no significant interaction effect between INTERACTION TECHNIQUE and NUMBER OF SCREENS with respect to task completion time. There was no significant differences between the conditions for accuracy in placing the target on the goal. The performance results for the CONTENT TRANSFER TASK can be seen in Figure 6 and the results of the RM-ANOVA in Table 1.

For the PUZZLE TASK there was a significant main effect of VISUALIZATION for task completion time such that the DEPTH method ($M = 37.17\ s, SD = 22.79$) resulted in a shorter task completion time than the FLAT method ($M = 43.31\ s, SD = 29.52$). As expected, the main effect of NUMBER OF LAYERS for task completion time was significant such that the FOUR LAYERS ($M = 21.28\ s, SD = 8.87$) resulted in a shorter task completion time than TEN LAYERS ($M = 59.19\ s, SD = 24.23$). No significant interactions have been found between NUMBER OF LAYERS and VISUALIZATION. Analysis of the error data (using the total number of errors across all ten repetitions) indicated that both the NUMBER OF LAYERS and the VISUALIZATION method had a significant impact on the number of errors made. The conditions using the DEPTH method ($M = 0.29$, $SD = 0.6$) had significantly more errors than the conditions using the FLAT method ($M = 0.07$, 0.26). Furthermore, conditions with FOUR LAYERS ($M = 0.04$, $SD = 0.19$) resulted in significantly fewer errors than conditions with TEN LAYERS ($M = 0.32$, 0.61).

4.5.2 Simulator Sickness, Workload, Usability

For the CONTENT TRANSFER TASK we found that the Total Severity aspect of the Simulator Sickness Questionnaire was significantly influenced by the INTERACTION TECHNIQUE such that BIMANUAL ($M = 14.43$, $SD = 20.73$) resulted in a higher total severity than GAZE-BASED ($M = 13.35$, 20.57). Also the Oculo-motor aspect was lower for the GAZE-BASED than for the BIMANUAL. However, the values are very low and pairwise comparisons showed no significant differences. There was a significant difference for the overall TLX results, such that BIMANUAL resulted in a higher taskload than GAZE-BASED. Also the mental, physical, effort and frustration results of the TLX were significantly higher for the BIMANUAL conditions than for the GAZE-BASED conditions. The perceived performance of the participants, however, was significantly higher for the BIMANUAL method ($M = 27.14$, $SD = 19.31$) compared to the GAZE-BASED method ($M = 21.79$, $SD = 21.65$). This is in contrast to the findings

Content Transfer Task

	Task Completion Time					Accuracy					TS-SS					SUS					Overall Taskload				
	df_1	df_2	F	p	η_p^2	df_1	df_2	F	p	η_p^2	df_1	df_2	F	p	η_p^2	df_1	df_2	F	p	η_p^2	df_1	df_2	F	p	η_p^2
I	**1**	**13**	**63.66**	**$<$ .001**	**70.83**	1	13	.79	.39	.06	**1**	**13**	**11.11**	**.005**	**.46**	**1**	**13**	**38.42**	**$<$ .001**	**.75**	**1**	**13**	**29.81**	**$<$ .001**	**.7**
S	**1**	**13**	**99.41**	**$<$ .001**	**.88**	1	13	.53	.48	.04	1	13	.60	.45	.04	1	13	3.79	.07	.23	1	13	3.02	.11	.19
I × S	1	13	.86	.37	.06	1	13	1.2	.29	.08	1	13	2.0	.18	.13	1	13	.87	.37	.06	1	13	.14	.72	.01

Puzzle Task

	Task Completion Time					Errors					TS-SS					SUS					Overall Taskload				
	df_1	df_2	F	p	η_p^2	df_1	df_2	F	p	η_p^2	df_1	df_2	F	p	η_p^2	df_1	df_2	F	p	η_p^2	df_1	df_2	F	p	η_p^2
V	**1**	**13**	**5.32**	**.04**	**.29**	**1**	**13**	**8.97**	**.04**	**.28**	**1**	**13**	**12.45**	**.004**	**.49**	**1**	**13**	**10.39**	**.007**	**.44**	**1**	**13**	**12.34**	**.004**	**.49**
L	**1**	**13**	**1574.78**	**$<$ .001**	**.99**	**1**	**13**	**5.14**	**.04**	**.28**	1	13	4.04	.07	.24	1	13	.61	.45	.04	**1**	**13**	**15.4**	**.002**	**.54**
V × L	1	13	1.23	.29	.09	**1**	**13**	**4.7**	**.05**	**.03**	1	13	1.86	.2	.13	1	13	.08	.78	.0	1	13	.03	.86	.0

Table 1. RM-ANOVA results for both tasks. Gray rows show significant findings. I = INTERACTION TECHNIQUE, S = NUMBER OF SCREENS, V = VISUALIZATION, L = NUMBER OF LAYERS. TS-SS: Total Severity Dimension of the Simulator Sickness Questionnaire. SUS: System Usability Scale $df_1 = df_{effect}$ and $df_2 = df_{error}$.

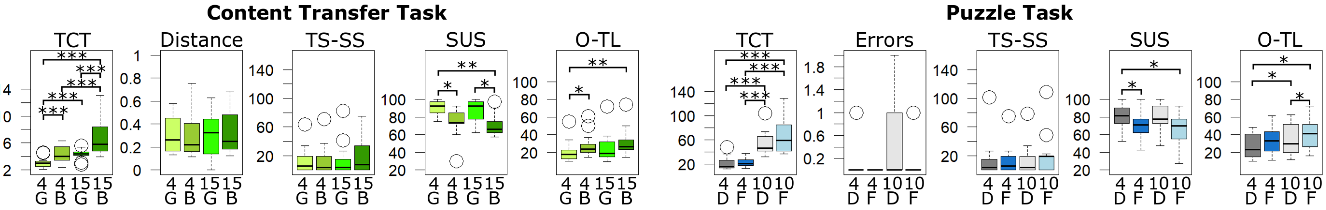

Fig. 6. Results of the task completion time (TCT), the Total Severity aspect of the Simulator Sickenss Questionnaire (TS-SS), the System Usability Scale (SUS) and the overall taskload (O-TL) for the CONTENT TRANSFER TASK and the PUZZLE TASK. Also for the CONTENT TRANSFER TASK the distance between disc and goal in cm upon task completion is shown and for the PUZZLE TASK the mean number of total errors. On the x-axis the digits indicates the NUMBER OF SCREENS (4 or 15) or the NUMBER OF LAYERS (4 or 10) as well as the INTERACTION TECHNIQUE G = GAZE-BASED, B = BIMANUAL and the VISUALIZATION F = FLAT, D = DEPTH. The number of stars indicates the level of significance between the conditions (*** < 0.001 ** < 0.01 * < 0.05).

of the task completion time and distance to goal shown above, where task completion time was significantly shorter for GAZE-BASED, and no significant difference was found in distance to the goal. Furthermore, INTERACTION TECHNIQUE had a significant impact on the SUS score in such a way that the usability of the GAZE-BASED method ($M = 88.75, SD = 10.26$) was higher than the usability of the BIMANUAL method ($M = 72.59, SD = 13.63$). The main results of the three questionnaires are shown in Figure 6. Plots and detailed results on the further dimensions of simulator sickness and task load can be found in the supplementary material.

For the PUZZLE TASK, the results of the simulator sickness questionnaire indicated that VISUALIZATION had a significant impact on the Total Severity such that it was significantly lower for DEPTH ($M = 14.29$, $SD = 23.85$) than for FLAT ($M = 16.7$, $SD = 24.76$). This was also true for the Oculo-motor and disorientation aspect. Also, the NUMBER OF LAYERS significantly influenced the disorientation aspect, such that the FOUR LAYER had a significantly lower value than TEN LAYER. However, the values are again very low and no significant differences were found in pairwise comparisons. The NASA TLX results indicated that the VISUALIZATION significantly influenced the overall task load such that it was significantly higher for the FLAT visualization ($M = 38.3$, $SD = 16.06$) than for the DEPTH visualization ($M = 30.92$, $SD = 15.92$). This was also the case for the mental and physical task load, as well as the frustration perceived by the participants. The perceived performance, however, was significantly higher for FLAT ($M = 29.64$, $SD = 21.73$) than for DEPTH ($M = 22.86$, $SD = 22.17$) which again, is in contrast to the results of the task completion time. In addition, we found that the number of layers had a significant impact on the overall task load such that it was lower for FOUR LAYERS ($M = 30.83$, $SD = 14.67$) than for TEN LAYERS ($M = 38.39$, $SD = 17.17$). The same was observed for frustration, physical task load and the temporal aspect. Also, we found that VISUALIZATION had a significant impact on the SUS results, in such a way that DEPTH ($M = 80.09$, $SD = 14.05$) resulted in a higher usability than FLAT ($M = 68.48$, $SD = 17.00$). The main results of the three questionnaires is shown in Figure 6 and Table 1. Plots and detailed results on the further dimensions of simulator sickness and task load can be found in the supplementary material.

4.5.3 Preference, Open Comments and Observations

For the CONTENT TRANSFER TASK, all but one participant preferred the GAZE-BASED method over the BIMANUAL, regardless of the number of screens. However, when asked about their preference for each NUMBER OF SCREENS separately this participant also liked GAZE-BASED more for FIFTEEN screens. When asked which interaction method the participants perceived as faster, all participants chose GAZE-BASED. When asked which method they could interact with more precisely, all participants chose GAZE-BASED for the conditions with FIFTEEN SCREENS. However, in the condition with FOUR SCREENS, only four participants thought that GAZE-BASED allowed for a more precise interaction and three participants felt that BIMANUAL was more precise. This is also supported by the objective measurements shown in Figure 6, even though no significant difference between conditions was indicated for distance to the target. We conjecture that these differences are due to the fact that in the GAZE-BASED method the selection of the target screen, and moving to the target within the screen, *can* happen concurrently. In contrast, in the BIMANUAL condition the selection of the target screen is *always* separated by an explicit mode switch (raising the finger of the non-dominant hand from the bezel) from the subsequent phase of moving towards the target within that screen. The lower precision in the BIMANUAL condition with FIFTEEN SCREENS compared to FOUR SCREENS might also be due to higher fatigue when moving the item across more screens. The participant who preferred BIMANUAL explained that it reminds him of the keyboard (alt+tab). Most other participants explained their preference for GAZE-BASED due to it being faster. One participant found GAZE-BASED more intuitive, and another participant found it less fatiguing. One participant suggested a hybrid solution where GAZE-BASED is only enabled when touching the bezel.

For the PUZZLE TASK, all but one participant preferred the DEPTH condition. However, one participant preferred FLAT. This participant later mentioned having trouble focusing the right eye and therefore found it more difficult to match the 3D state with the 2D target. Two participants thought the FLAT visualization was faster for FOUR LAYERS than the DEPTH visualization. All other participants felt that DEPTH was faster. Regarding the DEPTH visualization, participants also mentioned that “If you make a mistake it is easier to find the layer with the wrong element” (Participant P01), “I can better see which puzzle piece is where” (P02), “you directly see where the puzzle piece is” (P04) and that DEPTH “provides more information” (P06).

Participants used different strategies to solve the puzzle task for the FLAT conditions. Three participants mainly identified the element belonging to the current layer by trial and error. They tried to move the element and if it was not possible they tested the next element until they found the element of that layer. Four participants started using this strategy but then switched to a more efficient one. One participant used it sporadically in the condition with FOUR LAYERS.

Some participants moved to the last layer (furthest away) at the start of each puzzle and continued to the front. In the last layer, only one element (the element of that layer) was visible, therefore the participant did not have to guess the correct element. When moving one layer, a new element appeared on the screen, and the participant could identify the next element with ease. Nine participants mainly used this strategy.

There were also participants who started from the front layer but shortly switched to the next layer and back in order to detect the active element as this element will then disappear and reappear. Two participants switched to this strategy after first using another strategy. Also, this strategy was basically always used when participants wanted to correct a mistake.

The fourth strategy observed was to click on the eye-symbols in the menu at the left to deactivate and activate the current layer. This would also show the user which element would disappear and reappear and therefore was the element of the active layer. This strategy was only used by one participant in his first FLAT condition. Thereafter he switched strategy. Two people also tried this strategy during the training phase.

5 APPLICATIONS

Based on the insights presented in the previous sections, we implemented six applications that we envision are applicable in mobile knowledge worker scenarios.

Window Manager: We implemented a window manager that allows the arrangement of multiple windows around the user (see Figure 1, b). Interaction with those windows is supported by joint spatial and touch interaction. For example, the active window is selected by head gaze (indicated by a red border around the window in Figure 7 a and b). The input space of the physical touchscreen (depicted as a red rectangle in Figure 7, b) is then mapped to the output space of the virtual window. Contrary to the content transfer task in the evaluation of design parameters, for the window manager, the selected window stays selected even if the user’s gaze is switching to another window for as long as the user is touching the physical touchscreen. This design lets the user glance at other windows without losing the context of the task. Other clutching mechanisms are also feasible, such as locking the gaze-selected window on touch-down but then delaying the release on touch-up to let the user interact with the touchscreen.

To switch depth layers, the user can either swipe along the bezel or use a two-finger swipe, as described in the Section “Depth Interaction”. Alternatively, to retrieve a temporary preview of the otherwise hidden layers, the user can lean towards the virtual screen to peak through (Figure 7, c). For interaction with virtual windows that are larger than the physical touchscreen, retargeting can be used, as described in the Section “Single Screen Interaction” (Figure 7, d).

Code Version Control: We implemented an interface for code version control (Figure 1, c) that uses a spatial layout around the physical screen, see Figure 1, c. Using on-screen pinches the user can select different scopes of code changes (line, block, function, class, file), swipe through the commits using the selected scope, and swipe the desired commit down to the physical screen for further editing.

Parallel Coordinates: Inspired by recent research in immersive analytics [21, 22, 67, 113], we built a parallel coordinates plot that is grounded on the tangible physical screen, see Figure 1, g, and h. Subranges of variables can be selected using on-surface swipes for the extent of the variable range and on-surface drags for the center of the variable range. Individual variables are selected by touch. Switching between coordinate axes (z, protruding the display, for the variable range, y, along the height of the tablet, for the data items) can be achieved explicitly using a mode switch (e.g., pressing a soft button) or by mapping the swipe and drag along the x- and y-axis of the touchscreen.

Map: In the map application, users can navigate on a single layer without the map being cropped at the boundaries of the physical screen, see Figure 1, a . Available alternative map layers can be previewed by tilting the physical screen and be selected by swiping along the bezel of the screen.

Medical Imaging: In the medical imaging application, users can also swipe through different layers of information, see Figure 1, e–f. Additionally, the image slices can be previewed when the user moves their head to the side of the screen to look behind it.

PowerPoint: In the VR PowerPoint application users can arrange graphical items on a canvas, see Figure 1, d. Each item is associated with a separate layer. In the cases when items are occluded, they can be quickly accessed by swiping through the layers.

6 MOBILE IMPLEMENTATION

The system was implemented using Unity 2019.3. We used a HTC Vive Pro Eye as the HMD, which also enables hand tracking[1] as well as access to the camera streams[2]. While we used external OptiTrack outside-in tracking systems due to superior hand tracking accuracy compared to HMD-based inside out tracking [100] for our evaluations, we also implemented a system to track the user’s screen using the HMD cameras, see Figure 7, g-h. While it is possible to build a computer vision algorithm to track the specific laptop model used by users, we used a model-independent approach. Since the laptop or tablet display screen is hidden from the user, wearing an HMD, we can use the screen to display a standard tracking pattern, such as ARUCO markers [76]. Displaying the pattern enables a robust detection and orientation using multiple wide-field-of view cameras of the screen and information about the laptop dimension are sufficient to display a virtual model that can guide the user touch gestures (see accompanying video). Other solutions, such as using an external camera, such as a laptop camera, to track the HMD are feasible but require additional instrumentation of the HMD [73].

To facilitate the community in further research and development of joint interactions between touchscreens and VR headsets, we make our code available under https://gitlab.com/mixedrealitylab/breakingthescreen.

7 USER FEEDBACK ON APPLICATIONS

The purpose of this user study was to learn from initial user reactions on the applications we previously introduced. To this end we demonstrated individual experiences to the user instead of carrying out task-based evaluations (following the approaches used in prior work, e.g., [18, 99]). We recruited 17 participants (8 female, 9 male, mean age 31.9 years, sd = 9.2, mean height 173.9 cm, sd = 6.8) from a university campus with diverse study backgrounds. All participants were familiar with touch-sensitive screens.

7.1 Apparatus, Procedure, and Task

The same apparatus as in the evaluation of the design space parameters were used with the following change: instead of the Android tablet, a motion-tracked Surface Pro 4 tablet was used.

After an introduction, participants were asked to fill out a demographic questionnaire. Thereafter, we fixated OptiTrack trackers on the index finger of the dominant hand and the thumb of the other hand. Then individual applications were presented (counterbalanced). After trying out the applications the users rated them using a short three-item questionnaire capturing the user experience dimensions of ease of use, utility, and enjoyment (in line with similar approaches in prior work [18, 60]). After trying out the applications, participants were asked

---

[1] https://developer.vive.com/resources/knowledgebase/vive-hand-tracking-sdk/ Last accessed April 24th, 2020.

[2] https://developer.vive.com/resources/knowledgebase/intro-vive-srworks-sdk/ Last accessed April 24th, 2020.

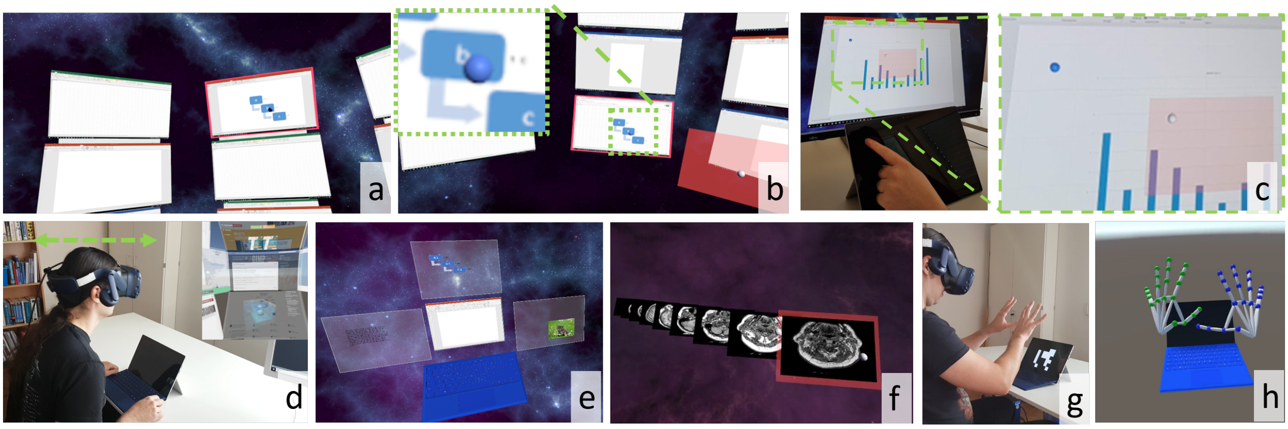

Fig. 7. a-d: Window Manager. a: An image is selected for movement to another window by head-gaze (selecting the window) + touch (selecting the image inside the window). b: The image is released on a different window at a different layer. The grey sphere on the red rectangle indicates the physical finger position. The blue sphere on the image depicts the projected finger position on the current window. c: The input on and above the physical touchscreen (red transparent rectangle, fingertip on touchscreen indicated by a white sphere) is retargeted to the virtual display (indicated by a blue sphere). d: A user previews windows in a hidden layer by leaning forward. e-f: Options to temporarily rearrange information layers. e: Virtual displays are temporarily arranged around the physical screen. f: If a user peeks behind the physical screen, virtual displays are extruded to that side. g-h mobile implementation: g: user wearing an HTC Vive Eye Pro looking on a tablet PC, which displays an Aruco marker. h: the user’s fingers are spatially tracked with the HTC Vive Hand tracking SDK

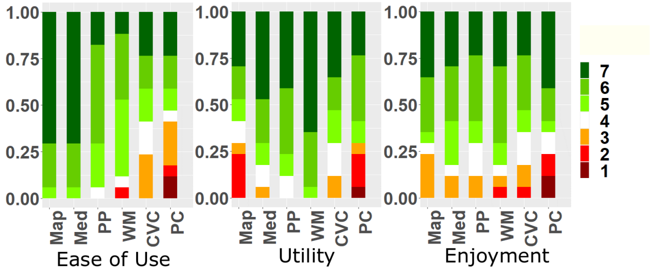

Fig. 8. Ease of Use, Utility and Enjoyment ratings for the evaluated applications on a range of 1 (strongly disagree) to 7 (strongly agree). Map: Map Interaction, Med: Medical, PP: PowerPoint, WM: WindowManager, CVC: CodeVersionControl, PC: ParallelCoordinates. The y-axis depcits the number of participants in percent (1.0 = 100% of participants).

to participate in a semi-structured interview. Finally, the participants were compensated with a 10 Euro voucher.

The order of the applications was balanced across participants, insofar as possible (a full permutation was not possible due to the number of applications).

7.2 Results

Figure 8 shows user ratings on seven-item Likert scales for questions on ease of use (“I found the application easy to use”), utility (“I found the application to be useful”) and enjoyment (“I had fun interacting with the application”). Note that we did not run null hypothesis significance tests on these ratings as they should serve as a descriptive indication of these user experience dimensions only. Participants were asked to comment on individual applications. We followed top-down qualitative coding and structuring procedures to identify the benefits and drawbacks of the individual applications [110]. For WindowManager the participants appreciated that the need for multiple physical monitors was mitigated (“Avoids the need of a monitor wall.” participant P13) but also the layered view of the virtual displays (“You can view dozens of documents.” P03) and it’s ease of use (“It was easy to understand even for non-experts.” P04). Regarding CodeVersionControl, one participant mentioned that she had an “overview like in a airport control tower.” P13 and another one saw applicability to another domain: “I would like that for functional subunits in engineering” P17. Similarly, for PowerPoint the participants saw a transfer to other domains with one mentioning that “these layers could also be used in drawing tools like Gimp or Photoshop” P14. Regarding ParallelCoordinates a participant mentioned that “it is a more entertaining way to present boring data” P06 and another participant, “The new visualization was nice” P13. Regarding Map a participant mentioned that “I could need that for driving.” P09 and regarding Medical participants mentioned that “a 3D Model from the layers would be awesome” P14 but also that “It would be cool to pick the layers directly” P04.

7.3 Discussion

Using immersive HMDs enables mobile knowledge workers to bring their working environment with them everywhere they go [39]. Relative to the compact input and output space of today’s mobile devices, the large field of view of HMDs allows to substantially extend the available display and increase privacy. Our design explorations and evaluations, explored different dimensions of this extended display space while focusing the input to the small area of a tablet screen. The first question we investigated was handling the large field of view of the virtual environment while using a very limited input space. To handle the large scale ratio between input and display, we tested two different techniques, One where we combined noisy eye gaze for large scale motions and touch for fine manipulation and compared it with a multi-touch technique. The results of the design parameter evaluation indicated that combination of gaze and touch input on a tablet as realized in the Gaze-based input method are beneficial for interaction in virtual multi-window environments. Specifically, the Gaze-based method outperformed the Bimanual method by approximately 30%. The second parameter we studied was the use of the HMD’s depth display for interaction above the touch screen. Using multiple layers has benefits of extending the amount of information that can be displayed and manipulated using a 2D screen. On the other-hand depth parallax and occlusions may hinder the user’s performance. Our second study was designed to test it, by asking users to do a task both in 2D and in 3D. As indicated by the results, the use of depth can be both efficient and usable when compared to a 2D touch interaction. However, we were surprised by different strategies employed by some users to compensate for the lack of the depth structure in Flat visualization, such as ordering their manipulations from back to front. Pairwise comparisons showed no significant differences between the conditions with regard to errors and the overall number of errors was low (5.7% for DEPTH and 1.4% for FLAT). During interviews with participants, the five participants generating errors stated that they did not use the option to show all layers at the end of the task using the dedicated button shown in Figure 4. By design, the layers were arranged behind the physical screen, allowing the user to navigate them front to back. As previously noted, in the FLAT visualization nine participants first navigated to the last layer and then navigated to the front layer by layer. When they reached the front layer, all layers with all puzzle pieces were visible, hence there was no need to use the "Show all" button. The user feedback on the developed applications indicated that the prototypes were usable and enjoyable. Users could envision employing the prototypes in their work environments. We envision mobile VR to become an important tool for work, as it enables a large and private display on the go, regardless of the worker environment. However the physical environment of the user will not grow and the interactions should be designed accordingly. The presented applications are just a small probe of the complex design space encompassing all of the vast tasks information workers carry out, and this area is ripe for further research and development.

7.3.1 Use of Augmented Reality HMDs

The proposed interaction concepts, while being considered in VR, could be transferred to Augmented Reality (AR) and explored further. In AR, current generation optical see-through HMDs typically have a substantially smaller field of view compared to immersive VR HMDs. Hence, the proposed techniques might need adaptation. Potentially, off-screen visualization techniques could be integrated into an AR setting to compensate for this limited field of view [42]. Also, in AR displays without occlusion management, physical objects, such as the tablet or the user's hands are visible alongside the virtual objects, typically at a different focus distance, which can further lead to potential perceptual issues [26]. In contrast, combining an AR HMD with a physical tablet could make use of the potentially higher output resolution of the physical touch display and open up further possibilities for novel overview + detail techniques.

7.3.2 Limitations

Our work focused on a subset of a possibly large design space. While we foresee the combination of VR HMDs and tablets as one potentially promising direction on how mobile knowledge workers can be supported in the future, we are aware that it remains uncertain if and when such potential combinations will become products due to various factors such as technology, market, regulation, social acceptance and timing uncertainty [49]. For the CONTENT TRANSFER TASK, we tested both a small and a large number of virtual screens. However, other factors such as screen arrangement, screen size, or parameters of the interaction methods could impact the results. Although we foresee no substantial reversal of the effects indicated in the study, it may be interesting, as a future work, to test the limit of noisy eye gaze usage for very small screens and translation distances, following [116]. For the PUZZLE TASK we empirically determined the design parameters, such as distance between layers. Future work should investigate the effects of layer distance on various task types in more detail, as with increasing layer distance, the association between individual (x,y) coordinates on a layer could become harder for users to track. While we indicated the feasibility of mobile implementations using our prototype, we still opted for using an external tracking system for implementing our interaction techniques due to the limited accuracy of generation HMD-based hand tracking [100] at the time of the study. Newer technologies, such as Oculus Quest hand tracking offer higher accuracy and may enable sufficient performance. It remains to be evaluated how robust concurrent camera-based tracking of the hands, keyboard, the touchscreen, and, potentially, a pen is in real-world usage scenarios. An alternative approach may use a tablet to record the user's in air hand movement using a 45-degree mirror clipped above the tablet front camera [128].

8 CONCLUSIONS AND FUTURE WORK

Our work explored the opportunities arising out of a joint interaction space between VR HMDs and laptop/tablet capacitive touchscreens for supporting mobile information workers. We have demonstrated that such a joint interaction space is feasible to be implemented using consumer-oriented hardware by tracking the touchscreen and the user's finger, which could be useful in small constrained environments such as on an airplane or in touchdown spaces. We explored how to spatially arrange and interact with data within this joint interaction space. We further implemented a set of six applications (window manager, code version control, parallel coordinates exploration, map navigation, medical image viewer, and a virtual PowerPoint) making use of this design space. In future work, these interaction concepts should be studied in depth, such as recently done for spreadsheet interaction [34], to allow us to better understand their capabilities and limitations for supporting mobile knowledge workers in real-world working contexts. Also, additional study designs, e.g., for further investigations of the explosion diagram technique, comparisons with non-VR baseline techniques, effects of physical and virtual screen sizes, as well as the use of fully mobile implementations in real-world environments seem feasible. In particular, we see the potential for further multimodal interaction techniques that fit this shallow 3D space spanned by the hands resting on a 2D screen. One avenue of future work is to explore which particular document semantics are the most suitable for display above the documents (such as, for example, comments and other mark-up on a text document, notes on a code review), and which mappings fit different contexts. In addition, another avenue is to explore how to expand the interaction vocabulary with simultaneous pen and touch techniques.

REFERENCES

[1] D. Ahlström, K. Hasan, and P. Irani. Are you comfortable doing that?: acceptance studies of around-device gestures in and for public settings. In *Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services*, pp. 193–202. ACM, 2014.

[2] F. Argelaguet and C. Andujar. A survey of 3d object selection techniques for virtual environments. *Computers & Graphics*, 37(3):121–136, 2013.

[3] R. Arora, R. Habib Kazi, T. Grossman, G. Fitzmaurice, and K. Singh. Symbiosissketch: Combining 2d & 3d sketching for designing detailed 3d objects in situ. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, p. 185. ACM, 2018.

[4] M. Azmandian, M. Hancock, H. Benko, E. Ofek, and A. D. Wilson. Haptic retargeting: Dynamic repurposing of passive haptics for enhanced virtual reality experiences. In *Proceedings of the 2016 chi conference on human factors in computing systems*, pp. 1968–1979. ACM, 2016.

[5] T. Babic, H. Reiterer, and M. Haller. Pocket6: A 6dof controller based on a simple smartphone application. In *SUI'18: 6th ACM Symposium on Spatial User Interaction*, pp. 2–10, 2018.

[6] W. Barfield, C. Hendrix, and K.-E. Bystrom. Effects of stereopsis and head tracking on performance using desktop virtual environment displays. *Presence: Teleoperators & Virtual Environments*, 8(2):237–240, 1999.

[7] E. A. Bier, M. C. Stone, K. Pier, W. Buxton, and T. D. DeRose. Toolglass and magic lenses: the see-through interface. In *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, pp. 73–80. ACM, 1993.

[8] M. Billinghurst and T. Starner. Wearable devices: new ways to manage information. *Computer*, 32(1):57–64, 1999.

[9] D. Bowman, E. Kruijff, J. J. LaViola Jr, and I. P. Poupyrev. *3D User interfaces: theory and practice, CourseSmart eTextbook*. Addison-Wesley, 2004.

[10] D. A. Bowman and C. A. Wingrave. Design and evaluation of menu systems for immersive virtual environments. In *Proceedings IEEE Virtual Reality 2001*, pp. 149–156. IEEE, 2001.

[11] J. Brooke et al. Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189(194):4–7, 1996.

[12] L. D. Brown and H. Hua. Magic lenses for augmented virtual environments. *IEEE Computer Graphics and Applications*, 26(4):64–73, 2006.

[13] G. Bruder, F. Steinicke, and W. Sturzlinger. To touch or not to touch?: comparing 2d touch and 3d mid-air interaction on stereoscopic tabletop

surfaces. In *Proceedings of the 1st symposium on Spatial user interaction*, pp. 9–16. ACM, 2013.

[14] F. Brudy, C. Holz, R. Rädle, C.-J. Wu, S. Houben, C. N. Klokmose, and N. Marquardt. Cross-device taxonomy: Survey, opportunities and challenges of interactions spanning across multiple devices. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, p. 562. ACM, 2019.

[15] W. Büschel, J. Chen, R. Dachselt, S. Drucker, T. Dwyer, C. Görg, T. Isenberg, A. Kerren, C. North, and W. Stuerzlinger. Interaction for immersive analytics. In *Immersive Analytics*, pp. 95–138. Springer, 2018.

[16] S. Butscher, S. Hubenschmid, J. Müller, J. Fuchs, and H. Reiterer. Clusters, trends, and outliers: How immersive technologies can facilitate the collaborative analysis of multidimensional data. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–12, 2018.

[17] L. K. Chan and H. Y. Lau. Magicpad: the projection based 3d user interface. *International Journal on Interactive Design and Manufacturing (IJIDeM)*, 6(2):75–81, 2012.

[18] X. Chen, T. Grossman, D. J. Wigdor, and G. Fitzmaurice. Duet: exploring joint interactions on a smart phone and a smart watch. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 159–168. ACM, 2014.

[19] X. A. Chen, N. Marquardt, A. Tang, S. Boring, and S. Greenberg. Extending a mobile devices interaction space through body-centric interaction. In *Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services*, pp. 151–160. ACM, 2012.

[20] L.-P. Cheng, E. Ofek, C. Holz, H. Benko, and A. D. Wilson. Sparse haptic proxy: Touch feedback in virtual environments using a general passive prop. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 3718–3728. ACM, 2017.

[21] M. Cordeil, B. Bach, A. Cunningham, B. Montoya, R. T. Smith, B. H. Thomas, and T. Dwyer. Embodied axes: Tangible, actuated interaction for 3d augmented reality data spaces. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, 2020.

[22] M. Cordeil, A. Cunningham, T. Dwyer, B. H. Thomas, and K. Marriott. Imaxes: Immersive axes as embodied affordances for interactive multivariate data visualisation. In *In Proc. UIST'17*, pp. 71–83. ACM, 2017.

[23] T. Dorta, G. Kinayoglu, and M. Hoffmann. Hyve-3d and the 3d cursor: Architectural co-design with freedom in virtual reality. *International Journal of Architectural Computing*, 14(2):87–102, 2016.

[24] A. Drachen, P. Mirza-Babaei, and L. E. Nacke. *Games user research*. Oxford University Press, 2018.

[25] T. Drey, J. Gugenheimer, J. Karlbauer, M. Milo, and E. Rukzio. Vrsketchin: Exploring the design space of pen and tablet interaction for 3d sketching in virtual reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, 2020.

[26] A. Eiberger, P. O. Kristensson, S. Mayr, M. Kranz, and J. Grubert. Effects of depth layer switching between an optical see-through head-mounted display and a body-proximate display. In *Symposium on Spatial User Interaction*, pp. 1–9, 2019.

[27] B. Ens, J. D. Hincapié-Ramos, and P. Irani. Ethereal planes: a design framework for 2d information space in 3d mixed reality environments. In *Proceedings of the 2nd ACM symposium on Spatial user interaction*, pp. 2–12. ACM, 2014.

[28] B. Ens and P. Irani. Spatial analytic interfaces: Spatial user interfaces for in situ visual analytics. *IEEE computer graphics and applications*, 37(2):66–79, 2016.

[29] B. M. Ens, R. Finnegan, and P. P. Irani. The personal cockpit: a spatial interface for effective task switching on head-worn displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 3171–3180. ACM, 2014.

[30] A. O. S. Feiner. The flexible pointer: An interaction technique for selection in augmented and virtual reality. In *Proc. UIST'03*, pp. 81–82, 2003.

[31] S. Feiner, B. MacIntyre, M. Haupt, and E. Solomon. Windows on the world: 2 d windows for 3 d augmented reality. In *ACM Symposium on User Interface Software and Technology*, pp. 145–155, 1993.

[32] C. Forlines, D. Vogel, and R. Balakrishnan. Hybridpointing: fluid switching between absolute and relative pointing with a direct input device. In *Proceedings of the 19th annual ACM symposium on User interface software and technology*, pp. 211–220, 2006.

[33] D. Gasques, J. G. Johnson, T. Sharkey, and N. Weibel. Pintar: Sketching spatial experiences in augmented reality. In *Companion Publication of the 2019 on Designing Interactive Systems Conference 2019 Companion*, pp. 17–20. ACM, 2019.

[34] T. Gesslein, V. Biener, P. Gagel, D. Schneider, E. Ofek, M. Pahud, P. O. Kristensson, and J. Grubert. Pen-based interaction with spreadsheets in mobile virtual reality. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2020.

[35] J. G. Grandi, H. G. Debarba, L. Nedel, and A. Maciel. Design and evaluation of a handheld-based 3d user interface for collaborative object manipulation. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 5881–5891. ACM, 2017.

[36] R. Grasset, A. Duenser, H. Seichter, and M. Billinghurst. The mixed reality book: a new multimedia reading experience. In *CHI'07 extended abstracts on Human factors in computing systems*, pp. 1953–1958. ACM, 2007.

[37] J. Grubert, M. Heinisch, A. Quigley, and D. Schmalstieg. Multifi: Multi fidelity interaction with displays on and around the body. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pp. 3933–3942. ACM, 2015.

[38] J. Grubert, M. Kranz, and A. Quigley. Challenges in mobile multi-device ecosystems. *mUX: The Journal of Mobile User Experience*, 5(1):5, 2016.

[39] J. Grubert, E. Ofek, M. Pahud, P. O. Kristensson, F. Steinicke, and C. Sandor. The office of the future: Virtual, portable, and global. *IEEE computer graphics and applications*, 38(6):125–133, 2018.

[40] J. Grubert, M. Pahud, R. Grasset, D. Schmalstieg, and H. Seichter. The utility of magic lens interfaces on handheld devices for touristic map navigation. *Pervasive and Mobile Computing*, 18:88–103, 2015.

[41] J. Grubert, L. Witzani, E. Ofek, M. Pahud, M. Kranz, and P. O. Kristensson. Text entry in immersive head-mounted display-based virtual reality using standard keyboards. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 159–166. IEEE, 2018.

[42] U. Gruenefeld, A. E. Ali, W. Heuten, and S. Boll. Visualizing out-of-view objects in head-mounted augmented reality. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*, pp. 1–7, 2017.

[43] J. Gugenheimer, D. Dobbelstein, C. Winkler, G. Haas, and E. Rukzio. Facetouch: Enabling touch interaction in display fixed uis for mobile virtual reality. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, pp. 49–60. ACM, 2016.

[44] J. Guo, D. Weng, Z. Zhang, H. Jiang, Y. Liu, Y. Wang, and H. B.-L. Duh. Mixed reality office system based on maslow's hierarchy of needs: Towards the long-term immersion in virtual environments. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 224–235. IEEE, 2019.

[45] S. G. Hart and L. E. Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, vol. 52, pp. 139–183. Elsevier, 1988.

[46] O. Hilliges, S. Izadi, A. D. Wilson, S. Hodges, A. Garcia-Mendoza, and A. Butz. Interactions in the air: adding further depth to interactive tabletops. In *Proceedings of the 22nd annual ACM symposium on User interface software and technology*, pp. 139–148. ACM, 2009.

[47] T. Hirzle, J. Gugenheimer, F. Geiselhart, A. Bulling, and E. Rukzio. A design space for gaze interaction on head-mounted displays. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, p. 625. ACM, 2019.

[48] K. Huo, K. Ramani, et al. Window-shaping: 3d design ideation by creating on, borrowing from, and looking at the physical world. In *Proceedings of the Eleventh International Conference on Tangible, Embedded, and Embodied Interaction*, pp. 37–45. ACM, 2017.

[49] H. Jalonen. The uncertainty of innovation: a systematic review of the literature. *Journal of Management Research*, 4(1):1, 2012.

[50] H.-C. Jetter, R. Rädle, T. Feuchtner, C. Anthes, J. Friedl, and C. N. Klokmose. "in vr, everything is possible!": Sketching and simulating spatially-aware interactive spaces in virtual reality. In *CHI Conference on Human Factors in Computing SystemsACM Conference on Human Factors in Computing Systems*. Association for Computing Machinery, 2020.

[51] N. Katzakis, R. J. Teather, K. Kiyokawa, and H. Takemura. Inspect: extending plane-casting for 6-dof control. *Human-centric Computing and Information Sciences*, 5(1):22, 2015.

[52] 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):203–

220, 1993.

[53] P. Knierim, V. Schwind, A. M. Feit, F. Nieuwenhuizen, and N. Henze. Physical keyboards in virtual reality: Analysis of typing performance and effects of avatar hands. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, p. 345. ACM, 2018.

[54] M. Kobayashi and H. Koike. Enhanceddesk: integrating paper documents and digital documents. In *Proceedings. 3rd Asia Pacific Computer Human Interaction (Cat. No. 98EX110)*, pp. 57–62. IEEE, 1998.

[55] P. G. Kry, A. Pihuit, A. Bernhardt, and M.-P. Cani. Handnavigator: Hands-on interaction for desktop virtual reality. In *Proceedings of the 2008 ACM symposium on Virtual reality software and technology*, pp. 53–60. ACM, 2008.

[56] M. Kytö, B. Ens, T. Piumsomboon, G. A. Lee, and M. Billinghurst. Pinpointing: Precise head-and eye-based target selection for augmented reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, p. 81. ACM, 2018.

[57] J. J. LaViola Jr, E. Kruijff, R. P. McMahan, D. Bowman, and I. P. Poupyrev. *3D user interfaces: theory and practice*. Addison-Wesley Professional, 2017.

[58] J. H. Lee, S.-G. An, Y. Kim, and S.-H. Bae. Projective windows: Bringing windows in space to the fingertip. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, p. 218. ACM, 2018.

[59] S.-w. Leigh, P. Schoessler, F. Heibeck, P. Maes, and H. Ishii. Thaw: tangible interaction with see-through augmentation for smartphones on computer screens. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction*, pp. 89–96. ACM, 2015.

[60] J. R. Lewis. Psychometric evaluation of an after-scenario questionnaire for computer usability studies: the asq. *ACM Sigchi Bulletin*, 23(1):78–81, 1991.

[61] F. C. Y. Li, D. Dearman, and K. N. Truong. Virtual shelves: interactions with orientation aware devices. In *Proceedings of the 22nd annual ACM symposium on User interface software and technology*, pp. 125–128. ACM, 2009.

[62] Z. Li, M. Annett, K. Hinckley, K. Singh, and D. Wigdor. Holodoc: Enabling mixed reality workspaces that harness physical and digital content. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, p. 687. ACM, 2019.

[63] H.-N. Liang, C. Williams, M. Semegen, W. Stuerzlinger, and P. Irani. An investigation of suitable interactions for 3d manipulation of distant objects through a mobile device. *International Journal of Innovative Computing, Information and Control*, 9(12):4737–4752, 2013.

[64] J. Looser, R. Grasset, and M. Billinghurst. A 3d flexible and tangible magic lens in augmented reality. In *2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, pp. 51–54. IEEE, 2007.

[65] D. López, L. Oehlberg, C. Doger, and T. Isenberg. Towards an understanding of mobile touch navigation in a stereoscopic viewing environment for 3d data exploration. *IEEE Transactions on Visualization and Computer Graphics*, 22(5):1616–1629, 2015.

[66] Z. Lu, Y. Hu, and J. Dai. Watchar: 6-dof tracked watch for ar interaction. In *IEEE ISMAR 2019 Demonstrations*.

[67] K. Marriott, F. Schreiber, T. Dwyer, K. Klein, N. H. Riche, T. Itoh, W. Stuerzlinger, and B. H. Thomas. *Immersive Analytics*, vol. 11190. Springer, 2018.

[68] A. Marzo, B. Bossavit, and M. Hachet. Combining multi-touch input and device movement for 3d manipulations in mobile augmented reality environments. In *Proceedings of the 2nd ACM symposium on Spatial user interaction*, pp. 13–16. ACM, 2014.

[69] D. C. McCallum and P. Irani. Arc-pad: absolute+ relative cursor positioning for large displays with a mobile touchscreen. In *Proceedings of the 22nd annual ACM symposium on User interface software and technology*, pp. 153–156, 2009.

[70] M. McGill, D. Boland, R. Murray-Smith, and S. Brewster. A dose of reality: Overcoming usability challenges in vr head-mounted displays. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pp. 2143–2152. ACM, 2015.

[71] M. McGill, J. Williamson, A. Ng, F. Pollick, and S. Brewster. Challenges in passenger use of mixed reality headsets in cars and other transportation. *Virtual Reality*, pp. 1–21, 2019.

[72] D. Mendes, F. M. Caputo, A. Giachetti, A. Ferreira, and J. Jorge. A survey on 3d virtual object manipulation: From the desktop to immersive virtual environments. In *Computer Graphics Forum*, vol. 38, pp. 21–45. Wiley Online Library, 2019.

[73] P. Mohr, M. Tatzgern, T. Langlotz, A. Lang, D. Schmalstieg, and D. Kalkofen. Trackcap: Enabling smartphones for 3d interaction on mobile head-mounted displays. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, p. 585. ACM, 2019.

[74] A. Mossel, B. Venditti, and H. Kaufmann. 3dtouch and homer-s: intuitive manipulation techniques for one-handed handheld augmented reality. In *Proceedings of the Virtual Reality International Conference: Laval Virtual*, p. 12. ACM, 2013.

[75] J. Müller, G. Bailly, T. Bossuyt, and N. Hillgren. Mirrortouch: combining touch and mid-air gestures for public displays. In *Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services*, pp. 319–328. ACM, 2014.

[76] R. Munoz-Salinas. Aruco: a minimal library for augmented reality applications based on opencv. *Universidad de Córdoba*, 2012.

[77] M. A. Nacenta, R. L. Mandryk, and C. Gutwin. Targeting across displayless space. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 777–786, 2008.

[78] M. Nancel, O. Chapuis, E. Pietriga, X.-D. Yang, P. P. Irani, and M. Beaudouin-Lafon. High-precision pointing on large wall displays using small handheld devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 831–840, 2013.

[79] M. Nancel, E. Pietriga, and M. Beaudouin-Lafon. Precision pointing for ultra-high-resolution wall displays. 2011.

[80] E. Normand and M. J. McGuffin. Enlarging a smartphone with ar to create a handheld vesad (virtually extended screen-aligned display). In *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 123–133. IEEE, 2018.

[81] E. Ofek, J. Grubert, M. Pahud, M. Phillips, and P. O. Kristensson. Towards a practical virtual office for mobile knowledge workers. *Microsoft New Future of Work Symposium*, 2020.

[82] J.-y. Oh and H. Hua. User evaluations on form factors of tangible magic lenses. In *2006 IEEE/ACM International Symposium on Mixed and Augmented Reality*, pp. 23–32. IEEE, 2006.

[83] T. Pejsa, J. Kantor, H. Benko, E. Ofek, and A. Wilson. Room2room: Enabling life-size telepresence in a projected augmented reality environment. In *Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing*, pp. 1716–1725. ACM, 2016.

[84] J. Petford, M. A. Nacenta, and C. Gutwin. Pointing all around you: selection performance of mouse and ray-cast pointing in full-coverage displays. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, p. 533. ACM, 2018.

[85] K. Pfeuffer, J. Alexander, M. K. Chong, Y. Zhang, and H. Gellersen. Gaze-shifting: Direct-indirect input with pen and touch modulated by gaze. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, pp. 373–383, 2015.

[86] K. Pfeuffer and H. Gellersen. Gaze and touch interaction on tablets. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, pp. 301–311, 2016.

[87] K. Pfeuffer, B. Mayer, D. Mardanbegi, and H. Gellersen. Gaze+ pinch interaction in virtual reality. In *Proceedings of the 5th Symposium on Spatial User Interaction*, pp. 99–108. ACM, 2017.

[88] C. Pinhanez. The everywhere displays projector: A device to create ubiquitous graphical interfaces. In *International conference on ubiquitous computing*, pp. 315–331. Springer, 2001.

[89] E. D. Ragan, R. Kopper, P. Schuchardt, and D. A. Bowman. Studying the effects of stereo, head tracking, and field of regard on a small-scale spatial judgment task. *IEEE transactions on visualization and computer graphics*, 19(5):886–896, 2012.

[90] D. Raja, D. Bowman, J. Lucas, and C. North. Exploring the benefits of immersion in abstract information visualization. In *Proc. Immersive Projection Technology Workshop*, pp. 61–69, 2004.

[91] D. Ramanujan, C. Piya, K. Ramani, et al. Mobisweep: Exploring spatial design ideation using a smartphone as a hand-held reference plane. In *Proceedings of the TEI'16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction*, pp. 12–20. ACM, 2016.

[92] R. Raskar, G. Welch, M. Cutts, A. Lake, L. Stesin, and H. Fuchs. The office of the future: A unified approach to image-based modeling and spatially immersive displays. In *Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, pp. 179–188. ACM, 1998.

[93] P. Reipschläger and R. Dachselt. Designar: Immersive 3d-modeling

combining augmented reality with interactive displays. In *Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces*, pp. 29–41, 2019.

[94] J. Rekimoto and M. Saitoh. Augmented surfaces: a spatially continuous work space for hybrid computing environments. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pp. 378–385. ACM, 1999.

[95] J. Ren, Y. Weng, C. Zhou, C. Yu, and Y. Shi. Understanding window management interactions in ar headset+ smartphone interface. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems Extended Abstracts*, pp. 1–8, 2020.

[96] A. Ruvimova, J. Kim, T. Fritz, M. Hancock, and D. C. Shepherd. "transport me away": Fostering flow in open offices through virtual reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, 2020.

[97] K. Ryu, J.-J. Lee, and J.-M. Park. Gg interaction: a gaze–grasp pose interaction for 3d virtual object selection. *Journal on Multimodal User Interfaces*, pp. 1–11, 2019.

[98] D. Schmalstieg, L. M. Encarnação, and Z. Szalavári. Using transparent props for interaction with the virtual table. *SI3D*, 99:147–153, 1999.

[99] D. Schneider, A. Otte, T. Gesslein, P. Gagel, B. Kuth, M. S. Damlakhi, O. Dietz, E. Ofek, M. Pahud, P. O. Kristensson, et al. Reconviguration: Reconfiguring physical keyboards in virtual reality. *IEEE transactions on visualization and computer graphics*, 2019.

[100] D. Schneider, A. Otte, A. S. Kublin, P. O. Kristensson, E. Ofek, M. Pahud, A. Martschenko, and J. Grubert. Accuracy of commodity finger tracking systems for virtual reality head-mounted displays. In *IEEE VR 2020*. IEEE, IEEE, March 2020.

[101] R. Schweigert, V. Schwind, and S. Mayer. Eyepointing: A gaze-based selection technique. In *Proceedings of Mensch und Computer 2019*, pp. 719–723. ACM, 2019.

[102] M. Serrano, B. Ens, X.-D. Yang, and P. Irani. Gluey: Developing a head-worn display interface to unify the interaction experience in distributed display environments. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*, pp. 161–171, 2015.

[103] L. Sidenmark and H. Gellersen. Eye&head: Synergetic eye and head movement for gaze pointing and selection. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, pp. 1161–1174, 2019.

[104] L. Sidenmark, D. Mardanbegi, A. Ramirez Gomez, C. Clarke, and H. Gellersen. Bimodalgaze: Seamlessly refined pointing with gaze and filtered gestural head movement. In *Proceedings of Eye Tracking Research and Applications*, 2020.

[105] R. L. Sollenberger and P. Milgram. Effects of stereoscopic and rotational displays in a three-dimensional path-tracing task. *Human Factors*, 35(3):483–499, 1993.

[106] P. Song, W. B. Goh, C.-W. Fu, Q. Meng, and P.-A. Heng. Wysiwyf: exploring and annotating volume data with a tangible handheld device. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 1333–1342. ACM, 2011.

[107] M. Spindler and R. Dachselt. Paperlens: advanced magic lens interaction above the tabletop. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, p. 7. ACM, 2009.

[108] M. Spindler, C. Tominski, H. Schumann, and R. Dachselt. Tangible views for information visualization. In *ACM International Conference on Interactive Tabletops and Surfaces*, pp. 157–166. ACM, 2010.

[109] J. Steimle, A. Jordt, and P. Maes. Flexpad: highly flexible bending interactions for projected handheld displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 237–246. ACM, 2013.

[110] A. Strauss and J. Corbin. *Basics of qualitative research*. Sage publications, 1990.

[111] H. B. Surale, A. Gupta, M. Hancock, and D. Vogel. Tabletinvr: Exploring the design space for using a multi-touch tablet in virtual reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, p. 13. ACM, 2019.

[112] Z. Szalavári and M. Gervautz. The personal interaction panel–a two-handed interface for augmented reality. In *Computer graphics forum*, vol. 16, pp. C335–C346. Wiley Online Library, 1997.

[113] S. K. Tadeja, T. Kipouros, and P. O. Kristensson. Exploring parallel coordinates plots in virtual reality. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, p. LBW2617. ACM, 2019.

[114] C. Tominski, S. Gladisch, U. Kister, R. Dachselt, and H. Schumann. A survey on interactive lenses in visualization. In *EuroVis (STARs)*. Citeseer, 2014.

[115] C. Tominski, S. Gladisch, U. Kister, R. Dachselt, and H. Schumann. Interactive lenses for visualization: An extended survey. In *Computer Graphics Forum*, vol. 36, pp. 173–200. Wiley Online Library, 2017.

[116] J. Turner, J. Alexander, A. Bulling, and H. Gellersen. Gaze + rst: integrating gaze and multitouch for remote rotate-scale-translate tasks. In *Proceedings of 2015 CHI conference on Human Factors in Computing Systems*, pp. 4179–4188, 2015.

[117] B. Ullmer and H. Ishii. The metadesk: models and prototypes for tangible user interfaces. In *Proceedings of Symposium on User Interface Software and Technology (UIST??? 97), ACM*, 1997.

[118] R. Unger and C. Chandler. *A Project Guide to UX Design: For user experience designers in the field or in the making*. New Riders, 2012.

[119] J. Wagner, M. Nancel, S. G. Gustafson, S. Huot, and W. E. Mackay. Body-centric design space for multi-surface interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1299–1308. ACM, 2013.

[120] J. A. Wagner Filho, C. M. D. S. Freitas, and L. Nedel. Virtualdesk: a comfortable and efficient immersive information visualization approach. In *Computer Graphics Forum*, vol. 37, pp. 415–426. Wiley Online Library, 2018.

[121] M. Waldner, E. Kruijff, and D. Schmalstieg. Bridging gaps with pointer warping in multi-display environments. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries*, pp. 813–816, 2010.

[122] P. D. Wellner. Interacting with paper on the digitaldesk. Technical report, University of Cambridge, Computer Laboratory, 1994.

[123] D. Wenig, J. Schöning, A. Olwal, M. Oben, and R. Malaka. Watchthru: Expanding smartwatch displays with mid-air visuals and wrist-worn augmented reality. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 716–721, 2017.

[124] J. R. Williamson, M. McGill, and K. Outram. Planevr: social acceptability of virtual reality for aeroplane passengers. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–14, 2019.

[125] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 143–146, 2011.

[126] D. Wolf, J. J. Dudley, and P. O. Kristensson. Performance envelopes of in-air direct and smartwatch indirect control for head-mounted augmented reality. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 347–354. IEEE, 2018.

[127] R. Xiao, M. A. Nacenta, R. L. Mandryk, A. Cockburn, and C. Gutwin. Ubiquitous cursor: a comparison of direct and indirect pointing feedback in multi-display environments. In *Proceedings of Graphics Interface 2011*, pp. 135–142. 2011.

[128] C. Yoo, I. Hwang, E. Rozner, Y. Gu, and R. F. Dickerson. Symmetrisense: Enabling near-surface interactivity on glossy surfaces using a single commodity smartphone. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, p. 5126–5137. Association for Computing Machinery, New York, NY, USA, 2016. doi: 10.1145/2858036.2858286

[129] D. Yu, H.-N. Liang, X. Lu, T. Zhang, and W. Xu. Depthmove: Leveraging head motions in the depth dimension to interact with virtual reality head-worn displays. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 103–114. IEEE, 2019.

[130] S. Zhai, C. Morimoto, and S. Ihde. Manual and gaze input cascaded (magic) pointing. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pp. 246–253, 1999.

[131] F. Zhu and T. Grossman. Bishare: Exploring bidirectional interactions between smartphones and head-mounted augmented reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, 2020.

[132] D. Zielasko, M. Krüger, B. Weyers, and T. W. Kuhlen. Menus on the desk? system control in deskvr. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 1287–1288. IEEE, 2019.

[133] D. Zielasko, M. Krüger, B. Weyers, and T. W. Kuhlen. Passive haptic menus for desk-based and hmd-projected virtual reality. In *2019 IEEE 5th Workshop on Everyday Virtual Reality (WEVR)*, pp. 1–6. IEEE, 2019.

Chapter 5

PoVRPoint: Authoring Presentations in Mobile Virtual Reality

Summary

The previous chapter has shown, how displaying three-dimensional data and using multimodal interaction techniques involving eye-tracking and touch can increase performance. This chapter, shows that a similar approach can facilitate certain tasks when authoring presentations. To this end, a touchscreen was combined with a spatially tracked pen and eye-tracking to interact with content while editing presentation slides. Specifically, this article presents a technique that allows the manipulation of objects on a slide through the spatially tracked pen, which has the potential to make 3D rotations more intuitive. It also presents an animation editing technique that makes use of the 3D space in VR by displaying the timeline vertically above the slide. Similarly, a method for reordering overlapping objects on a slide is proposed, by showing all object layers slightly separated in 3D space. Finally, the large space in VR is used to display additional data, such as further presentation-files, PDFs or an image search from which the user can copy content while authoring the presentation.

A user study showed that users found these techniques to be enjoyable, useful and usable. The reordering technique was also compared to existing techniques, implemented in current authoring software, which showed that the VR technique is significantly faster and more usable. Another performance

study indicated that through the wide field of view in VR participants are significantly faster in identifying target slides than on a screen the size of a tablet for visually salient target slides.

Overall, this article provides examples of how knowledge work can benefit from using VR in terms of usability and performance. Also, the chosen input devices (tablet and pen) offer the possibility to use such a system also in mobile settings, as they are lightweight and do not require much space.

Contribution Statement

The presented techniques were designed and implemented by Verena Biener, Travis Gesslein, Daniel Schneider, Felix Kawala and Cuauhtli Campos in an iterative process and repeatedly discussed with Per Ola Kristensson, Michel Pahud, Eyal Ofek and Jens Grubert. Studies were conducted by Verena Biener, Travis Gesslein and Alexander Otte. The statistical analysis was performed by Verena Biener and she created the corresponding figures and tables. The images were created by Verena Biener with help of Eyal Ofek and Travis Gesslein. The first draft of the article was written by Verena Biener and subsequently refined by all authors.

Article [14]

PoVRPoint: Authoring Presentations in Mobile Virtual Reality

Verena Biener, Travis Gesslein, Daniel Schneider, Felix Kawala, Alexander Otte, Per Ola Kristensson,
Michel Pahud, Eyal Ofek, Cuauhtli Campos, Matjaž Kljun, Klen Čopič Pucihar and Jens Grubert

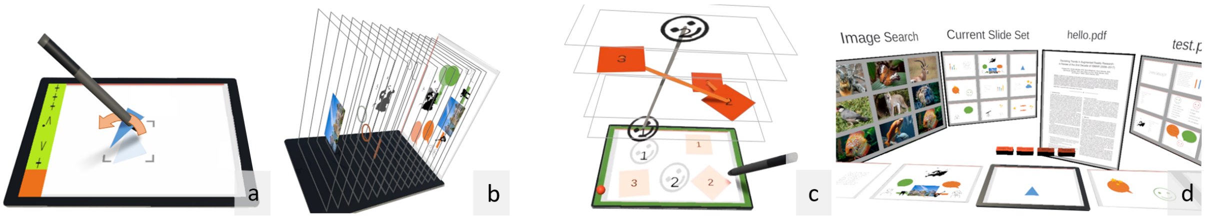

Fig. 1. Interaction techniques for authoring presentations in mobile VR. a) 3D object manipulation technique, b) occlusion handling, c) animations with time represented through height, d) working across slides and displaying different content resources

Abstract—Virtual Reality (VR) has the potential to support mobile knowledge workers by complementing traditional input devices with a large three-dimensional output space and spatial input. Previous research on supporting VR knowledge work explored domains such as text entry using physical keyboards and spreadsheet interaction using combined pen and touch input. Inspired by such work, this paper probes the VR design space for authoring presentations in mobile settings. We propose PoVRPoint—a set of tools coupling pen- and touch-based editing of presentations on mobile devices, such as tablets, with the interaction capabilities afforded by VR. We study the utility of extended display space to, for example, assist users in identifying target slides, supporting spatial manipulation of objects on a slide, creating animations, and facilitating arrangements of multiple, possibly occluded shapes or objects. Among other things, our results indicate that 1) the wide field of view afforded by VR results in significantly faster target slide identification times compared to a tablet-only interface for visually salient targets; and 2) the three-dimensional view in VR enables significantly faster object reordering in the presence of occlusion compared to two baseline interfaces. A user study further confirmed that the interaction techniques were found to be usable and enjoyable.

Index Terms—Virtual Reality, Presentation Authoring, Mobile Knowledge Work, Pen and Touch Interaction

◆

1 INTRODUCTION

Using slide-based presentation tools, such as Apple's Keynote and Microsoft's PowerPoint, has widespread use across many sectors, such as education, business and academia. In recent years, more people have had the need to work in less than perfect environments, at home, in makeshift offices and on the go. To enable workers to be productive everywhere, even the small space of a middle seat on a coach flight, we need to focus on using small, portable devices such as tablet computers and the small input space where the user may move their hands without interfering with their physical environment.

Virtual Reality (VR) as an instance of Extended Reality (XR), can benefit users immensely, as it provides a much larger display space than traditional mobile devices, independent of the physical environment and allows rendering of 3D elements above the screen, potentially mitigating effects of a small physical screen. [4, 15, 18]. HMDs can also increase privacy and reduce environmental clutter [18,57]. Using VR to support knowledge workers on the go, and therefore in limited spaces, has already been discussed in previous work [18,43,77] and interaction techniques for such scenarios have been developed for multi-screen environments [4] or spreadsheet applications [15]. This approach is also supported by the advancement of new hardware, as it is already possible to use inside-out tracking HMDs on the go and multiple manufacturers work on light HMDs as a replacement for physical screens. (While our prototype system relies on a stationary tracking system, we expect tracking systems to be available in mobile scenarios in the foreseeable future). Furthermore, new HMDs are designed as monitor replacements, increasing mobility of information workers in the near future [1,48], which we hope will increase the popularity of immersive displays for productivity.

In this work, we focus on exploring how to support the editing process of slide-based presentations using VR in a mobile scenario and present our prototype system called *PoVRPoint*. However, common interaction techniques for VR relying on in-air interaction with controllers or hands might not be well-suited for the limited interaction space in mobile settings, especially regarding practicality and social acceptability [43]. In addition to the restricted space, fatigue can limit the applicability of in-air interaction [23]. As hand and eye gaze tracking technologies are already being incorporated into commercial HMDs this opens up new interaction possibilities. Along with mobile devices such as tablets, touch-based interaction can be augmented with spatial interaction above and around the screen [4, 15] to facilitate knowledge workers' tasks and potentially improve the overall interaction experience in limited spaces. Therefore, we combine a VR HMD with eye-tracking and bimanual pen and touch techniques, where pen-based input can be sensed on and above mobile devices, requiring less interaction space [15].

We are concentrating on facilitating the authoring process of presentations by expanding typical 2D presentation editing interfaces, such as Microsoft PowerPoint, with an increased 3D output and input space, instead of inherently changing the nature of how presentations are authored today. This makes it possible to leverage familiarity of existing

---

- *V. Biener, T. Gesslein, D. Schneider, F. Kawala, A. Otte and J. Grubert are with Coburg University of Applied Sciences, Germany. E-mail: {verena.biener, travis.gesslein, daniel.schneider, felix.kawala, alexander.otte, jens.grubert}@hs-coburg.de.*
- *P.O. Kristensson is with University of Cambridge, United Kingdom. E-mail: pok21@cam.ac.uk.*
- *M. Pahud and E. Ofek are with Microsoft Research, United States. E-mail: {mpahud, eyalofek}@microsoft.com.*
- *C. Campos, M. Kljun and K.C. Pucihar are with University of Primorska, Slovenia. E-mail: cc.mijangos@gmail.com, matjaz.kljun@upr.si, klen.copic@famnit.upr.si.*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

tools and increases compatibility between VR and non-VR working modes.

In this paper we do not present a complete presentation authoring application, but we investigate common tools used in presentation editing that are concerned with both, editing *individual* slides and interacting with presentations across *multiple* elements (multiple slide-sets, PDFs, browsers) and how they can be enhanced using VR. Specifically, we explored if an extended display space can help users identify target slides (or images) faster, as this is a common task when editing a larger slide-set. In addition, we also investigated how to improve reordering of multiple shapes on a slide in presence of occlusions which can get problematic as the number of objects on a slide increases. We evaluated the performance of these techniques to understand how they can facilitate authoring presentations in VR. For additional concepts and techniques presented in this work, we report a usability study ($n = 18$), confirming the designed interaction techniques are usable and enjoyable.

In summary, this paper makes four central contributions. First, we present the design and implementation of four VR-based techniques for authoring presentations including manipulating objects, occlusion handling, animations, and working across slides. Second, we quantify the benefits of using the increased display space of a VR HMD compared to a tablet in a visual search task, e.g., when searching for slides or images. Our results indicate the superiority of VR when the matching task is easy (pre-attentive visual search). However, when the matching difficulty is high (attentive search), VR performs similarly to the tablet. Third, we show that the VR-based reordering technique supports significantly faster arranging of occluded objects (such as multiple overlapping shapes on a single slide) compared to the two existing baseline techniques used in PowerPoint. Fourth, we report that the VR-based techniques have been found usable and enjoyable in a usability study ($n = 18$).

2 RELATED WORK

Our work draws on the areas of supporting knowledge workers in XR, pen-based interaction, gaze-based interaction and in-air interaction as well as authoring and presenting in XR.

2.1 Knowledge Workers in VR

The use of XR for supporting knowledge work has attracted recent research interest, e.g., [18, 20, 42, 56]. Early research focused on projection systems to extend stationary physical office environments, e.g., [55, 72]). With the rise of affordable VR and AR HMDs these devices also have been explored as tools for assisting users when interacting with physical documents, e.g., [16, 37]. Further, Grubert et al. [18, 45] and McGill et al. [43] explored the positive and negative qualities that VR introduces in mobile knowledge work scenarios. Desktop-based environments have been studied for tasks such as text entry [19, 31, 41], system control [75, 76] and visual analytics [6, 70]. Research on productivity-oriented desktop-based VR has concentrated on the use of physical keyboards and mouse-based input along with HMDs [57, 71], controllers and hands [33, 76], and, recently, tablets [4, 15, 63].

Our work complements these prior studies by investigating the potential of editing slide-based presentations in VR using mobile devices such as tablets.

2.2 Pen-based, In-air and Gaze-based Interaction

Besides the commonly used single-point input with pens, enhanced interaction techniques have been explored. Examples include using touch input on the non-dominant hand, supporting pen input in bimanual interaction [26, 50], unimodal surface-based pen-postures [7], bending [14] or using sensors in or around the pen [24, 40] for gestures and postures, and examining pen-grips [61]. Our work was inspired by tilting [66] and hovering [17] the pen above interactive surfaces, which we use in a VR context.

The use of pens in AR and VR has also been investigated as a standard input device on physical props [64], as well as using grip-specific gestures for in-air interaction [36]. The accuracy of pen-based in-air pointing has also been studied [52]. In AR, pen-based interaction was specifically investigated for an object manipulation task [69].

In prior work on combining in-air with touch interaction, Marquardt et al. [39] investigated the use of on and above surface input on a tabletop. Chen et al. [10] explored in-air use of on and above surface input on a tabletop. They propose that interactions can be composed by interweaving in-air gestures before, between, and after touch on a prototype smartphone augmented with hover sensing. Hilliges et al. [22] have been using hover to allow more intuitive interaction with virtual objects that represent physical objects. More recently, Hinckley et al. [25] have been exploring a pre-touch modality on a smartphone including the approach to record trajectories of fingers to distinguish between different operations. Such technology can be used to connect 3D tracking and touchscreen digitizer for better accuracy of tracking.

Most VR in-air interaction typically aims at using unsupported hands. To enable reliable selection, targets are designed to be sufficiently large and spaced apart [62]. Our focus on mobile knowledge workers on the move implies small gestures to reduce working fatigue and to retain usability in potentially cramped environments, such as airplane seats or tiny work places such as touchdown spaces. We utilize gestures to be used by a hand, resting on the screen of a tablet and holding a pen.

In addition, the combination of eye-gaze with other modalities such as touch [49], in-air gestures [51, 58] and head-movements [34, 59] has been investigated for interaction in spatial user interfaces. For a recent survey on gaze-based interaction in AR and VR, see Hirzle et al. [27].

Specifically, our techniques were inspired by gaze-based interaction with virtual screens [4] as well as the combination of pen-based and touch-based interaction for mode switching [50] but adapted those techniques specifically for the use case of editing presentations.

2.3 Presenting and Authoring in XR

XR has been explored for complementing or substituting established methods for presenting materials, e.g., in the medical domain [46], general education [29] or training [9]. For example, Kockro et al. [32] compared VR and (2D) PowerPoint lectures in an anatomy context, and found no performance differences but found VR to be rated higher in domains such as spatial understanding and enjoyability. With the recent rise of online conferences, VR has also been explored for delivering oral presentations and poster sessions [35]. However, the benefits and drawbacks of presenting in VR compared to 2D conferencing tools such as Zoom, are yet to be explored in depth. XR has also been proposed as an aid for training public speaking [60] as well as in-situ support [47].

Besides using XR for presenting content to an audience, considerable work was invested in creating content for and in XR [3, 38, 44]. Specifically, XR was investigated for supporting modelling [11, 54], sketching [12] and creating animations [2, 8, 68].

Complementary to these previous approaches, our work focuses on utilizing VR as a tool for authoring 2D presentations.

3 INTERACTION TECHNIQUES FOR AUTHORING PRESENTATIONS IN MOBILE VR

We looked at several challenging aspects of using a 2D slide editing program, such as dealing with 3D orientation and ordering, dealing with temporal data, and retrieving information from a large corpus of graphics data. Then, we designed a set of interaction and visualization techniques using the advantages that VR provides, such as a large display space, a depth display and in-air interaction. Those techniques are just sample points in the entire space of tasks used for presentation authoring, but they can already show the advantages of using a VR environment. With our techniques, we want to support knowledge workers on the go or other confined spaces, limiting the choices of hardware. Therefore, our setup includes a tablet lying in front of the users, who hold a stylus with two buttons in their dominant hand, while their non-dominant hand is used for mode switches on the touchscreen in the area near the border. Both tablet and stylus are spatially tracked to represent them in VR. For designing these interaction techniques, we followed an iterative approach with multiple design iterations consisting of conceptualization, implementation and initial user tests (eating your own dog food) [67].

As the techniques are designed to support users when working in adverse conditions, such as confined spaces, lack of privacy and limited display size [4, 18], VR provides multiple advantages. First, users are no longer restricted to their available physical displays and can view information in the space around them, beyond the bounds of a mobile device. Second, the three-dimensional VR display enables depth visualization to allow utilization of the space above (or below) a 2D surface. Third, touch-based interaction can be complemented with further modalities such as in-air or gaze-based interaction. Fourth, the entire display is seen only by the users, maintaining their privacy and reducing visual disturbance from the environment.

Because we want to support small (mobile) work spaces, we use the tablet's surface as the main interaction space. This provides space for hand motions above it while the touchscreen supports easy and accurate input. However, in many cases it might not be comfortable to look at the tablet lying on a table. In such cases, the slide and also the pen could be re-projected to be in front of the user's head for indirect manipulations as suggested by prior work [19].

In the following subsections, we will present concepts for editing an individual slide, as well as concepts for working with multiple slides and other resources. Note: For video description of interaction techniques refer to the accompanying video.

3.1 Editing Slides

Preparing presentations requires the user to create slides that contain a collection of information, arranged both along the area of the slide, as well as in depth (layering of items) and in time (animation of the items). In the following, we propose techniques for such tasks using a pen and a tablet in VR.

3.1.1 Manipulation of Objects

Common presentation tools, such as Microsoft PowerPoint, let the user create slides that may contain a collection of items such as text, images, videos, three-dimensional objects and more. Such items can be selected, and dragged to change their position on the slide. They can be rotated and scaled using dedicated widgets and may even be rotated in three-dimensions using additional input fields or symbolic input.

We explore how to use a pen alongside touch input in VR to provide a unifying interface for 2D and 3D object manipulation of elements on the slide. We propose to use a pen that is spatially tracked which expands the interaction space to include not only the tablets surface but also the space around the user. This can potentially make object manipulations more intuitive.

Translating a selected object is supported by standard drag and drop using a stylus or the user's finger. An object can be rotated by rotating the pen (as if it is attached to the pen) and it can be scaled by moving the pen further away from or closer to the tablet (as if pulling to increase the size). To differentiate between translation, rotation and scaling, a finger of the non-dominant hand on the bezel of the touchscreen is used to control the modality of the manipulation, as shown in Figure 2. We chose a bezel-based technique, as they have already been successfully used for mode-switches in other scenarios (e.g. [4, 15]). When no finger touches the designated area on the bezel of the screen, the stylus is used to select an object and to drag it to a new position (Figure 2, a and e). Touching the bezel with the non-dominant hand, while still touching an object with the stylus, activates the 2D rotation mode. Rotating the stylus around its axis, while still touching the surface with the tip of the stylus, will rotate the object in the screen plane (Figure 2, b and f).

Lifting the pen away from the surface, while still touching the bezel in rotation mode, enables the user to perform 3D rotations by rotating the stylus in space. When performing a 3D rotation, two instances of the object are displayed - one as a flat projection on the slide, and a full 3D display of the object in the air above the screen, enabling the user to better see the three-dimensional pose (Figures 2, c and g and 1, a). The position of the 3D display is fixed above the objects position on the slide and its rotation is determined by the rotation of the stylus.

To scale an object, the non-dominant hand touches the bezel, while the tip of the stylus is in the air not touching the screen. Moving the stylus' tip up, away from the screen, increases the scale of the object uniformly, while bringing the tip closer to the screen reduces the its scale (Figure 2, d and h).

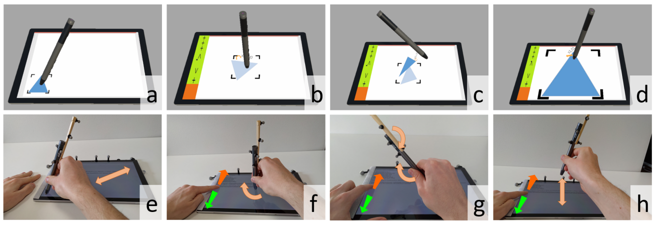

Fig. 2. Tracking of the stylus' six degrees of freedom in VR enables complex manipulation of an object on the slide. The dominant hand holding the stylus is used to drag, rotate, scale and 3D rotate the object, while the non-dominant hand's touch on the screen's bezel-area (green area on the left) modulates the interaction. Without any bezel touch (a, e) objects are dragged by the stylus. Bezel touch and rotation of the stylus along its axis while it touches the screen (b, f) rotates the object on the screen's plane. Rotating the stylus in the air while keeping the bezel touch (c, g) rotates the objects in 3D. Finally, touching the bezel while the stylus is in the air and moving the stylus towards or away from the screen (d, h) decreases or increases the scale of the object uniformly.

The sensitivity of these manipulations can be controlled by moving the finger's vertical location on the bezel (see Figure 2, f, g and h). Moving the touch point up increases the control-display gain of the rotation and the scaling, enabling larger changes with small pen movements. Moving the finger down enables better accuracy. With this form of object manipulation there is no need to select potentially small scale and rotation handles on already small display sizes because of the bimanual mode activation technique. For the case of 3D rotations in particular, the actual 3D preview of the rotation above the object combined with VR-based head movement and three-dimensional display is something not possible in classical 2D presentation tools and has the potential to make 3D rotations more intuitive for the user.

3.1.2 Occlusion Handling

Most presentation tools support explicit ordering of objects, where objects of a higher layer occlude objects of lower layers. As a result, objects may become partially or even completely hidden behind other objects, making visual identification difficult or impossible, and increasing selection difficulty. While some applications such as PowerPoint display a separate list of all objects on a slide, sorted by their levels, such a list requires the user to create a mental map, matching the location of each item on the list to the visual objects on the slide. An example can be seen in Figure 9, c which shows a simplified version of the PowerPoint interface used in our study.

The Mac version of PowerPoint uses a dynamic reordering mode[1] that displays objects with their layers slightly rotated, similar to Figure 9, b. In this mode, the layers can be grabbed and moved to rearrange their order. However, with increasing number of objects this can also become challenging as the layers can potentially overlap and make it harder to see which object belongs to which layer, impacting selection accuracy.

Inspired by this dynamic reordering mode, we propose to use the space above the tablet in VR to present the object-layers to the user in a way that facilitates assessing how the objects are ordered. Specifically, we propose to rotate the object-layers by 90 degrees, move them up to stand on top of the tablet and slightly separate them (see Figures 3 and 1, b). The tablet and the user's head can be repositioned in VR to resolve any potential occlusion issues that might arise from a fixed viewpoint. We also conducted informal experiments with further degrees of rotation. Zero degrees of rotation (layer parallel to the display) did not scale beyond a few layers due to the imprecise nature of in-air selection compared to touch-based selection. When comparing

---

[1] https://www.indezine.com/products/powerpoint/learn/shapes/dynamic-reorder-of-overlapping-shapes-in-ppt2011-mac.html Last access September 1, 2021

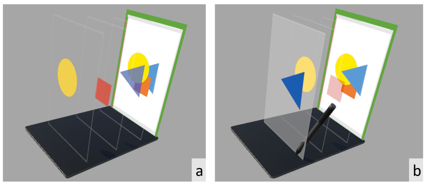

Fig. 3. Reordering layered objects. The layers are displayed with the top one on the far left of the tablet, and the bottom one on the right. On the right end of the tablet, we display the full composited slide for reference. a) An example of side view with a slide containing a yellow circle on top, than a red square followed by a blue triangle. b) A user dragging the blue triangle to place it on the top layer.

different amounts of rotations (0, 45 and 90 degrees), 90 degrees was perceived as most comfortable and efficient. Furthermore, the selection process can be supported by showing the intersections of the layers with the touchscreen (lines in Figures 3 and 1, b). Thus, instead of in-air selection, layers can be selected by touching their intersection lines, and precisely dragged to a new position (Figure 3, b). Using touch on the tablet display allows for precise interaction, even when the number of objects increases and the layers get closer to each other. There is also a projection of the complete slide to the right of the object layers which presents the current layout in 2D and facilitates the understanding of the ordering. In our implementation, the layer reordering mode is toggled by touching the lower left corner of the bezel area that is also used for mode switching when manipulating objects. An experiment showing superior performance of this technique compared to baseline Powerpoint implementations is presented in section 4.2.

3.1.3 Animations

Presentation applications such as PowerPoint show keyframes of object animations as locations on a slide. As the area of the slide is limited and may quickly be overloaded with information, PowerPoint displays only the start and end keyframes and only for the currently selected object. We facilitate visualizing and editing animations through a 3D visualization where the extra spatial dimension represents time.

By selecting an entry in an in-air menu opened by the second pen button, the user enters an explicit animation mode. As a base 2D visualization, we employ a similar approach to PowerPoint, visible in Figure 4, a, but showing all keyframes of all objects by default, instead of showing it only for the selected objects as in PowerPoint. However, the keyframes of certain objects can be hidden via a menu entry, which appears next to the pen after pressing the second pen button, as can be seen in Figure 4, b. The keyframes may be manipulated (dragged, rotated and scaled) like any other item on the slide. To represent time, the keyframes are numbered and connected by lines representing the sequence of the animation (Figure 4).

By using the space above the screen, we can also visualize the time of each keyframe by locating it at the corresponding height above the screen. The further up the keyframe is placed, the later the associated object state will be reached in the animation. This enables us to display all objects and the relations between their animations. All keyframes of an animation are connected with a colored curve that shows the path the animation takes in time (height), as visible in Figures 4, c and 1, c.

Similar to our reordering technique, the three-dimensional time view of the animation is toggled by touching the lower-left corner of the tablet's bezel area while the animation mode is active. A keyframe can be grabbed in-air using the first button on the stylus and moved up and down to change the corresponding time (Figure 4, d). For the animation interface, non-rotated layers were used (in contrast to occlusion handling), because object manipulation on the screen is possible while the 3D time view is active, overloading the touch inputs that would otherwise need to be used to move rotated layers. Also, animating depends, in contrast to pure reordering, much more on the actual position of the objects on the slide. Therefore, we prioritized spatial consistency over ease of selection.

Vertical timelines, displayed left of the tablet, help to indicate a keyframe's precise time. Applying a two finger pinch or move gesture in the bezel area can scale or scroll through the displayed animation along the time axis to better see parts of the animation. We display two timelines: one that displays the global animation timeline, and the other displays the currently viewed section of the timeline (Figure 4, c and d). In addition to toggling visibility of animations, the menu opened by the second pen button is also used to add and remove keyframes (Figure 4, b). A play button that can be touched by the stylus can play or pause the animation. The default display of the animation is on the tablet screen, but it is also possible to render it on a virtual screen placed away from the tablet for easier viewing.

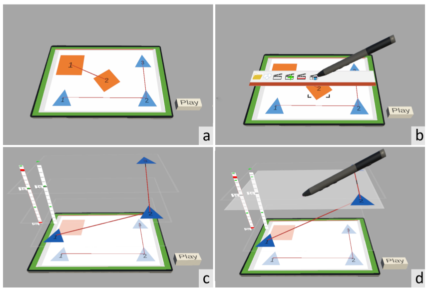

Fig. 4. a) Animation mode showing the keyframes of two objects on the slide. b) Menu that can be used to add or delete keyframes or make keyframes invisible. c) 3D animation mode which shows the keyframes as layers with their position in z-direction representing time. d) Moving the second keyframe up so the movement from the first to the second will be slower.

3.2 Working Across Slides

While the tablet screen is the default surface to display and edit the current slide, authoring presentations requires to also access further data sources, e.g., the slides of the currently edited slide-set, fetching resources from other presentations, or browsing external content such as images or PDF files. In the following, we discuss means for viewing further media concurrently to the active slide, and transferring content between these displays and the active slide screen.

3.2.1 Slide Overview

When authoring presentations, it is a common task to copy content between slides or to go back and forth between slides to review the content. Especially on a small display in presentations with several slides, it is often not possible to fit all slides on the screen with sizes that allow the user to recognize them, and the user has to scroll through the slide overview (or slide sorter view) to find the target slide. The large display space in VR can be used to mitigate this problem. We use two different ways to access other slides on the slideshow. First, while the current slide is displayed on the tablet screen, slides before and after the current slide are displayed to the left and right of the tablet, see Figure 1, d. This technique is inspired by similar visualizations of Gesslein et al. [15]. Second, a slide overview of the current slide-set is displayed in front of the user, as can also be seen in Figure 1, d. This overview can be scrolled and zoomed using a two-finger swipe or pinch gesture in the border area of the tablet while gazing at the overview area. While it would be possible to select a slide from the overview using in-air gestures, they may be exhaustive and may not fit limited work spaces. Instead, we use eye gaze to pre-select a slide of interest, indicated by a green frame, and while maintaining the gaze, swipe down on the touchscreen to confirm selection. The selected slide is then displayed as the active slide in the current slide set.

3.2.2 Multiple Content Sources

When creating a presentation, it is very common to use additional content such as text, images or videos. Just like displaying the slide overview, the space in front of the user can be used to display a myriad of content like images, videos or web pages that can be added to the presentation. The user can add such content areas through a menu

Fig. 5. a) The user is looking at an object (image) and sees the position of the pen above the tablet as a red dot. b) The user touches the tablet with the pen to select the image. c) The user looks back down and the selected object is copied to the currently edited slide.

Fig. 6. a) Study setup with 6 OptiTrack cameras and two VIVE lighthouse basestations (one is behind). OptiTrack cameras were used for the reordering task and the usability evaluation, while lighthouse basestations were used for the search task. b) Participant in the reordering task and the usability evaluation. c) Participant in the tablet search task condition.

opened, again, by a the second pen button. It is also possible to fade out some of these areas while they are not used by clicking on the corresponding toggle button with the pen (red buttons in front of the tablet in Figure 1, d).

3.2.3 Copying Content

The user can select objects from the external content displays described in Section 3.2.2 and copy them to the currently active slide. As the user gazes at an element such as a slide, a semi-transparent copy of the pen is visualized in the same relative pose to the gazed element as the physical pen is located above the physical tablet, and the touch location which it is hovering over is shown as a red dot (Figure 5, a). As the user moves the stylus above the tablet screen, the corresponding copy above the gazed element moves as well. The user can select an object from content displays in the same way as selecting an object on the tablet – by touching it as visualized in Figure 5, b. Looking back at the tablet while still touching the screen with the stylus will copy the selected object to the current slide (Figure 5, c). Similarly, the user can copy an entire slide from the current or another slide set by touching this slide at a position where no object is placed. Looking down will insert this slide after the currently edited slide. Both objects and slides can also be copied through the menu opened by the second pen button.

4 PERFORMANCE EVALUATION

The two main advantages of VR for knowledge worker tasks that we focus on in this paper are the larger display space and the three-dimensional viewing of 3D content. To evaluate their advantages for presentation authoring, we chose two tasks, each representing one of the above mentioned advantages: using the large field of view for displaying an overview of content and using the 3D space above the tablet for ordering slide objects that occlude each other. These two tasks represent issues that current users are often faced with: searching for a slide on a limited screen and ordering occluded objects. However, the presented concepts can also be generalized to other tasks. The evaluation was done by conducting two separate lab-based studies using a within-subjects design in each case.

4.1 Search Study

Virtual Reality HMDs enable the user to have a large display space around them, which can be used to show a large number of objects, such as slides in a slide sorter view or images during an image search, at the same time as shown in section 3.2.2. In contrast, using a small tablet screen may force the user to scroll to be able to view a similar amount of items with a comparative size. We selected a search task to quantify the possible advantages of VR HMDs compared to a tablet screen. The participants were presented with a target image and then had to find and select it among 63 other images. We chose a corpus of images of different animals. While this kind of visual stimuli are commonly used assets in presentations, they represent challenging content due to the amount of details in naturalistic images.

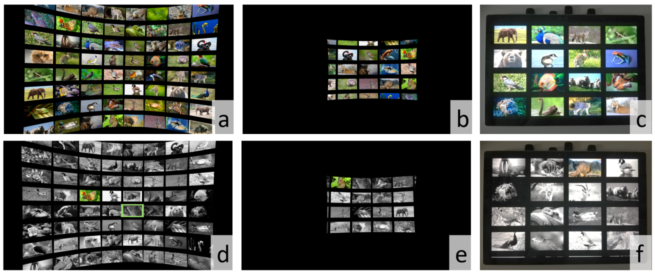

Fig. 7. Conditions for the visual search task: HARD conditions, with colored images only with a) VR-FULL, b) VR-LIMITED and c) TABLET INTERFACE; EASY conditions with only the target image colored with d) VR-FULL e) VR-LIMITED and f) TABLET INTERFACE. The green frame in d) indicates the image that is currently selected via gaze.

The experiment consisted of two independent variables: INTERFACE and DIFFICULTY. We used three types of INTERFACEs. The first, VR-FULL used the maximum field of view (FoV) provided by the HMD. The second, VR-LIMITED, artificially limited the users' FoV in VR to reflect a similar FoV to that of the tablet's display. And the third was TABLET, using the actual tablet display without VR. The second independent variable was DIFFICULTY. The levels were HARD, using the original colored images, leading to an attentive search, and EASY, where all images but the target image were reduced to grayscale making the search in this condition pre-attentive, so the target could be immediately spotted [74]. The combination of these two variables leads to six conditions which are depicted in Figure 7. The dependent variables measured in this experiment were task completion time, number of errors, usability (System Usability Scale, SUS) [5], workload (NASA TLX unweighted version) [21], and simulator sickness (SSQ) [30].

4.1.1 Participants

Twenty participants took part in this study (6 female, 14 male), with a mean age of 28.85 ($SD = 5.2$). Five participants wore glasses during the experiment and all but one had at least some prior VR experience.

4.1.2 Apparatus

In all conditions the participants were presented with a set of images. Each image was set to occupy approximately 8 degrees horizontally of the participant's field of view, both displayed on the tablet's screen or HMD. While it is known that image size impacts image search [28], we empirically determined this size to be both legible and selectable in VR and on the tablet. With an approximate distance to the tablet of 40 cm this resulted in an image width of 5.8 cm on the tablet. Sixty-four images were arranged in four columns, such that 16 images were fully visible without a need for scrolling.

In both VR conditions (VR-FULL and VR-LIMITED) the images were placed on a sphere around the user at a distance of 75 cm which resembles the focal distance of the HTC Vive HMD, resulting in an image width of 10 cm. In the VR conditions, the images were arranged in 8 columns covering 65 degrees horizontally and 45 degrees vertically to enable participants to comfortably reach all images with eye-gaze and only slight head movements. The field of view of the VR-LIMITED condition was artificially limited using black planes in all four directions, resulting in a field of view of about 36 x 24 degrees, resembling the field of view of the tablet (26 x 17.5 cm at a distance of 40 cm). Both the tablet and the VR applications were implemented using Unity 2019.4. The TABLET conditions were performed on a Microsoft Surface Pro 4 as shown in Figure 6, c. For the VR conditions, we used a HTC Vive Pro Eye, which provides built-in eye-tracking and two lighthouse base stations (Figure 6, a). We combined it with the Microsoft Surface to enable touch input. In contrast, to the further studies, the Optitrack system depicted in 6, a, was not used in this study.

4.1.3 Procedure

The study started by asking the participants to sign a consent form and fill out a demographic questionnaire. The order of the six conditions was balanced using a balanced Latin square. In each condition the participants were first presented with the target image, either on the tablet screen in the TABLET conditions or in the air in front of the user in VR. Upon touching the tablet's screen the target image vanished and the image search started. In the TABLET condition, the participants had to scroll through the images by touching and dragging on the display. When they found the target, they selected it by touching it. In the VR (VR-FULL and VR-LIMITED) conditions, the participants had to search for the target by moving their eyes and head and select the target by eye-gaze and confirm the selection by taping anywhere on the tablet. Prior to the start of the VR conditions, users conducted eye-gaze calibration using the built-in calibration routine of the HTC Vive Pro Eye. After selecting an image, the next target image was displayed. In each condition, participants had to find 30 images, which were always the same but in randomized order and positions while ensuring that targets are positioned in all regions. After completing a condition, participants answered the simulator sickness questionnaire [30], the system usability scale questionnaire [5] and the NASA task load index [21]. Also, we recorded the task completion times and errors for each task. On average, it took 45 minutes to complete this experiment.

4.1.4 Results

Repeated measures analysis of variance (RM-ANOVA) was used to analyze task completion times, which were non-normal and therefore log transformed. For multiple comparisons Bonferroni adjustments were used at an initial significance level of $\alpha = 0.05$. Aligned Rank Transform [73] was used for subjective data and errors that are not normally distributed (or could not be normalized using log transform). The main results are displayed in Table 1.

For each participant an average task completion was computed from the 30 tasks, including trials with errors. Analyzing the task completion times indicated significant simple main effects of INTERFACE and DIFFICULTY on task completion time and that there were also significant interaction effects. Specifically, the VR-FULL ($M = 4.7s$, $SD = 3.52$) conditions were significantly faster than the VR-LIMITED ($M = 5.47s$, $SD = 3.34$) and TABLET ($M = 5.75s$, $SD = 2.56$) conditions. Also, a significant difference could be found between VR-LIMITED and TABLET. As expected, the EASY ($M = 2.65s$, $SD = 1.02$) conditions were significantly faster than the HARD ($M = 7.97s$, $SD = 2.2$) conditions.

The interaction effect is visible in Figure 8. Post-hoc comparisons showed that for the EASY conditions, VR-FULL ($M = 1.71s$, $SD = 0.61$) is significantly faster than VR-LIMITED ($M = 2.7s$, $SD = 0.8$) and both VR methods are significantly faster than TABLET ($M = 3.53s$, $SD = 0.66$). This suggests that the wider FoV makes the search faster, as indicated by prior work [53]. The different input techniques (gaze vs. scrolling) are likely to contribute to difference between the VR-LIMITED and TABLET. However, for the HARD conditions, no significant differences between the INTERFACES were indicated. This could suggest that the larger FoV does not necessarily provide an advantage in a search were all possible targets have to be looked at in detail.

INTERFACE also significantly influenced the number of errors. Specifically, participants made significantly less errors in the TABLET ($M = 0.8$, $SD = 1.11$) conditions compared to both the VR-FULL ($M = 5.08$, $SD = 4.65$) and VR-LIMITED ($M = 3.58$, $SD = 3.64$) conditions. There was no significant difference between VR-LIMITED and VR-FULL. Also, the analysis showed that DIFFICULTY had a significant effect on the number of errors in such a way that the EASY conditions ($M = 2.63$, $SD = 3.81$) resulted in significantly less errors than the HARD conditions ($M = 3.67$, $SD = 3.9$). However, the error rate in all conditions was rather low and higher error rates in the VR conditions could be explained by the eye-gaze technique, relying on off-the shelf gaze-tracking in the HTC Vive Pro Eye. It is not always perfectly accurate and if the user is not fully concentrated, the gaze might lose the target in the moment of the confirmation, resulting in a wrong selection. However, there are possibilities for improvement like a delay before switching the selected images or a dwell time. More errors in the HARD conditions can be explained by user mistakes, for example when looking for an orange fish they chose the yellow one.

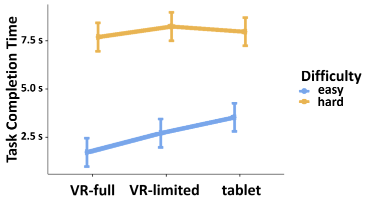

Fig. 8. Task completion times for the six conditions, which is significantly influenced by the interface in the easy conditions, but not in the hard.

Search Task - Subjective Ratings

	TS-SS					SUS				
	df_1	df_2	F	p	η_p^2	df_1	df_2	F	p	η_p^2
I	2	38	3.59	0.04	0.16	2	38	0.24	0.79	0.01
D	1	19	7.23	0.01	0.28	1	19	8.25	0.01	0.3
I × D	2	38	2.43	0.1	0.11	2	38	0.7	0.5	0.04

	Overall task load				
	df_1	df_2	F	p	η_p^2
I	2	38	0.42	0.66	0.02
D	1	19	40.51	< 0.001	0.68
I × D	2	38	1.98	0.15	0.09

Search Task - Performance Data

	Task Completion Time					Errors				
	df_1	df_2	F	p	η_p^2	df_1	df_2	F	p	η_p^2
I	2	38	32.3	$< .001$	.63	2	38	33.96	$< .001$	.64
D	1	19	782.5	$< .001$	.98	1	19	22.27	$< .001$	.54
I × D	2	38	33.3	$< .001$	.64	2	38	2.74	.08	.13

Table 1. RM-ANOVA results for the search task. Gray rows show significant findings. I = INTERFACE, D = DIFFICULTY. TS-SS: Total Severity Dimension of the Simulator Sickness Questionnaire. SUS: System Usability Scale. $df_1 = df_{effect}$ and $df_2 = df_{error}$.

Analyzing the results from the total severity dimension of the simulator sickness questionnaire indicated a significant influence of INTERFACE and DIFFICULTY. However, pairwise comparisons showed no significant differences. Unsurprisingly, a significant influence of DIFFICULTY on the usability score was detected in such a way that the HARD ($M = 85.88$, $SD = 13.07$) conditions had a significantly lower usability than the EASY ($M = 89.75$, $SD = 10.69$) conditions. However, no significant influence of INTERFACE could be detected, regarding usability. The DIFFICULTY also had a significant influence on the overall task load, such that the task load was significantly higher for the HARD conditions ($M = 29.65$, $SD = 16.54$) than for the EASY conditions ($M = 17.81$, $SD = 12.73$). This makes sense, because compared to the HARD conditions, in the EASY conditions less mental effort is required to find the target. Again, no significant influence of INTERFACE could be detected.

4.2 Reordering Study

To quantify the benefits of a three-dimensional visualization of layered information, we used the reordering task to compare the standard POWERPOINT tool for reordering object layers with the DYNAMIC REORDERING tool implemented in PowerPoint for Mac and our VR tool. The participants were presented with a number of objects on a slide. The displayed objects always included one yellow square and one red circle while the remaining objects were blue triangles. The task for the participants was to bring the red circle directly in front of the yellow square, by moving the corresponding layers to the front or to the back. To reproduce many different scenarios, the objects were placed with different amounts of overlap, the number of layers was varied and the target object was placed at different depths.

This experiment had one independent variable, INTERFACE, with three levels. First, the VR condition, a simplified version of the layer visualization and manipulation interface described in section 3.1.2 displays the layers in the air above the tablet (Figure 9, a). Second, the non-VR DYNAMIC REORDERING technique available in PowerPoint

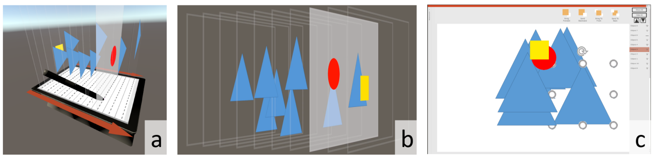

Fig. 9. Conditions in the reordering task: a) our VR technique. b) a re-implementation of the DYNAMIC REORDERING technique from PowerPoint for Mac c) a re-implementation of the standard POWERPOINT technique.

for Mac was used (Figure 9, b). Third, the non-VR baseline technique available in POWERPOINT was used. It presents the layers in a list next to the slide, where they can be dragged up and down, and buttons to bring a layer to the front or back (Figure 9, c). The dependent variables for this experiment were task completion time, usability (System Usability Scale, SUS) [5], workload by using NASA TLX (unweighted version) [21], simulator sickness (SSQ) [30] and user preferences.

4.2.1 Participants

Fourteen volunteers (2 female, 12 male, mean age 27.1, $SD = 4.2$) took part in the study. All but one had prior VR experience and normal or corrected to normal vision.

4.2.2 Apparatus

We implemented all three conditions using Unity 2019.4, in order to facilitate the task design, to get consistent logging data for all three techniques and to avoid potential confounds due to different input and output devices. In the VR condition, the object layers were rotated 90 degrees and lifted up above the tablet, so that the edge of each layer just touched the display. Therefore, it is possible to move a layer by touching the intersection on the tablet and moving it around. For the DYNAMIC REORDERING condition, the reordering feature from PowerPoint for Mac was re-implemented. The object layers were presented in 3D on the tablet and could be moved by touching a layer and dragging it to its new position. For the POWERPOINT condition, we implemented an application resembling the PowerPoint interface for rearranging object layers. Re-implementing this interface allowed us also to exclude potential confounds, such as visual noise added by extraneous buttons and submenus not relevant for the experiment. The two interfaces showed the same content as the original would, just leaving blank functionalities that are not used in the study.

For the VR condition we used a HTC Vive Pro Eye in combination with a Microsoft Surface Pro 4 tablet and an AZLink stylus pen. The pen, the tablet and the HMD were tracked via an OptiTrack motion-tracking system, using 4 Optitrack Prime 13 cameras placed above the user and two Prime 13W wide angle cameras placed closer to the user to support tracking for the stylus which was equiped with smaller markers. The VIVE-based HMD tracking was overriden to prevent interference between VIVE lighthouses and OptiTrack cameras. Optitrack was chosen for the study setup, because it provides very accurate tracking for the HMD, tablet and pen. Using Optitrack, we could utilize a normal lightweight pen in contrast to rather heavy trackable pens that are currently available. In future work this could also be implemented as a mobile prototype and be evaluated in-the-wild. Both the pen and the tablet were visualized in the virtual environment using 3D models of the real objects. The commodity PC running the VR application received touch inputs remotely from the Microsoft Surface via UDP. The two non-VR conditions (DYNAMIC REORDERING and POWERPOINT) were run directly on the Microsoft Surface Pro 4 tablet. The study setup can be seen in Figure 6, a and b.

4.2.3 Procedure

Upon arrival, the participants were first asked to sign a consent form and to fill out a demographic questionnaire. Then the participants started with one of the three conditions. The condition order was counterbalanced between the subjects to avoid effects due to fatigue or learning. With 14 participants, it was not possible to exactly counterbalance, but no significant order effects were detected. For each conditions the task was repeated 32 times. After each condition, participants completed the Simulator Sickness questionnaire [30], the System Usability Scale [5] and the NASA TLX questionnaire [21]. After all conditions were successfully completed, participants filled out a questionnaire about their preferred technique. We then conducted a semi-structured interview to understand their choice and give them the opportunity to give comments. In addition, we recorded the task completion times for all conditions. On average, this study took 30 minutes. Volunteers did not receive a compensation for participating.

4.2.4 Results

Repeated measures analysis of variance (RM-ANOVA) was used to analyze task completion times, which were non-normal and therefore log transformed. For multiple comparisons Bonferroni adjustments were used at an initial significance level of $\alpha = 0.05$. Aligned Rank Transform [73] was used for subjective data and errors that are not normally distributed (or could not be normalized using log transform).

The main results of the reordering task are shown in Table 2. Due to logging errors, we lost 12 samples for the task completion time from the first participant in the reordering task, so the mean task completion time for this participant only consisted of 20 samples. Statistical significance tests showed that the task completion time was significantly influenced by INTERFACE in such a way that the VR method ($M = 4.51$s, $SD = 2.46$) was significantly faster than both the DYNAMIC REORDERING ($M = 14.5$s, $SD = 7.4$) and POWERPOINT ($M = 16.1$s, $SD = 6.39$) methods. But no significant difference between DYNAMIC REORDERING and POWERPOINT was detected.

The NASA task load index also showed a significant influence of the INTERFACE on the overall task load. Pairwise comparisons showed that the VR ($M = 20.1$, $SD = 11.46$) interface resulted in a significantly lower task load than the DYNAMIC REORDERING ($M = 49.29$, $SD = 22.15$) interface. But no significant difference was detected between VR and POWERPOINT ($M = 34.23$, $SD = 21.56$) or between DYNAMIC REORDERING and POWERPOINT. No significant differences between the INTERFACEs regarding simulator sickness were detected.

The usability was also significantly influenced by the INTERFACE in such a way that the usability of the VR method ($M = 89.11$, $SD = 7.76$) was significantly higher than for the DYNAMIC REORDERING ($M = 53.39$, $SD = 21.63$) and POWERPOINT ($M = 69.64$, $SD = 18.76$) methods. Again, no significant difference between DYNAMIC REORDERING and POWERPOINT could be detected.

All but one participant preferred the VR method. One preferred the DYNAMIC REORDERING method. Six participants that preferred the VR techniques said that "it results in the fastest overview" (P1, P2, P3, P4, P7, P8) and "if something was occluded you could move your head" (P2, P7, P14). Five also mentioned that "it was easy to select the layers" (P1, P4, P6, P8, P12), three that "the interaction was more convenient" (P8, P9, P12,) and three that "it was easier and faster" (P1, P2, P11). Three participants also complained about the DYNAMIC REORDERING condition. One said "I was confused which slide is selected" (P1), one that "it was hard to identify the layer" (P2) and one that "I often selected the wrong layer" (P6). Another one mentioned that "it is a problem that it is only displayed in 2D, so it is hard to see where to tap" (P12). One participant proposed to highlight the slides when touching them and selecting them by for example pressing the pens button. This is in line with our observations that the two baseline techniques required a lot more trial and error to select the right objects, as they were partially occluded and it was hard to see which object belongs to which layer. In contrast, the three-dimensional view and head movement in VR helps with assigning objects to layers.

5 USABILITY EVALUATION

In addition to the performance evaluation that we presented in Section 4, we conducted a usability study on our prototype, which consists of the techniques presented in Section 3. Our objective was to find out if they are easy to understand and to use for regular users. Also, through participant's comments we gained valuable insights into how to improve our prototype. This study was divided into four parts representing the techniques from section 3—object manipulation, handling

Reordering Task

	df_1	df_2	F	p	η_p^2
TCT	2	26	71.8	< 0.001	0.85
TS-SS	2	26	0.68	0.52	0.05
SUS	2	26	22.8	< 0.001	0.64
Overall task load	2	26	18.98	< 0.001	0.59

Table 2. RM-ANOVA results for the reordering task. Gray rows show significant findings. TCT: Task Completion Time. TS-SS: Total Severity Dimension of the Simulator Sickness Questionnaire. SUS: System Usability Scale $\mathrm{d}f_1 = \mathrm{d}f_{effect}$ and $\mathrm{d}f_2 = \mathrm{d}f_{error}$.

occlusions, animations and working across slides. The three concepts (slide overview, multiple content sources, copying content) on working across slides, as presented in section 3, were presented to the participants jointly as one coherent workflow.

5.1 Participants

Eighteen participants (5 female, 13 male) took part in this study. Their mean age was 28.94 years ($SD = 5.3$). All had normal or corrected to normal vision and all but one had prior VR experience.

5.2 Apparatus

The VR applications described in this paper were implemented using the Unity 2019.4. We used a HTC Vive Pro Eye, which provides built-in eye tracking in combination with a Microsoft Surface Pro 4 tablet and an AZLink stylus pen. For the user study, the pen, tablet and HMD were tracked via an OptiTrack motion-tracking system with 6 Optitrack Prime 13 cameras (Figure 6, a and b), since pen tracking via VIVE-trackers or other VIVE-based tracking devices was unfeasible due to pen weight concerns. Therefore, the VIVE-based HMD tracking was overriden to prevent interference between VIVE lighthouses and OptiTrack cameras. Pen, tablet and the two fingers of the non-dominant hand, that were used for pinching, were visualized in the virtual environment, using 3D models of the real pen and tablet and spheres for the fingers. The tablet was connected to the PC running the VR application which received the touch inputs via UDP.

5.3 Procedure

First, participants were asked to sign a consent form and fill out a demographic questionnaire. Then the eye-tracking was calibrated using the built-in routine of the HTC Vive Pro Eye. All participants started with object manipulation, because it is a prerequisite for the animation techniques. The order of the remaining parts - handling occlusions, animations and working across slides - was counterbalanced. For each concept, the participants were walked through the possibilities and interaction techniques that are provided in the prototype. Then they had time to try out the technique as long as they liked. On average participants spent about 5 minutes exploring each technique. Following each interaction concept, the participants orally graded three statements (while wearing the HMD) by giving a score on a seven-item Likert scale, regarding ease of use ("I would find the application easy to use"), utility ("I would find the application to be useful") and enjoyment ("I would have fun interacting with the application") (1: totally disagree, 7: totally agree). Also they were encouraged to think out loud about their experience and make suggestions. At the end, when all concepts were explored, the participants completed the Simulator Sickness questionnaire [30], the System Usability Scale [5] and the NASA TLX questionnaire [21]. Also, they were asked to rank the four techniques by popularity and we conducted a semi-structured interview to give the participants a chance to further express their thoughts. The whole study took about 45 minutes on average.

5.4 Results

The results from the three questions on ease of use, utility and enjoyment that were asked after each technique are presented in Figure 10. It can be seen that the ratings for utility and enjoyment are high with more than 75% of the answers being at least a five on the seven-item Likert scale. Even tough the participants had to learn a lot in a short time, the ease of use rating was also high.

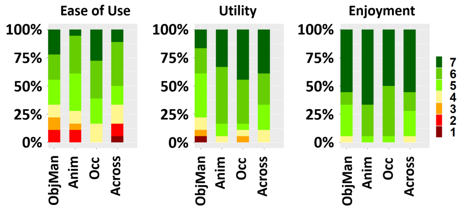

Fig. 10. Answers from questionnaire about Ease of use, utility and enjoyment for the four concept with 7 being the highest and 1 the lowest possible score. ObjMan = Object Manipulation, Anim = Animation, Occ = Occlusion Handling, Across = Across Slides

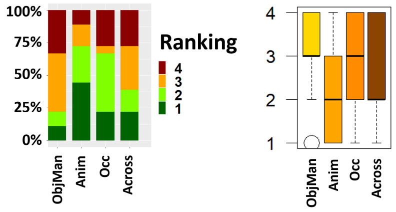

Fig. 11. Left: percentage of people ranking the technique either 1st, 2nd, 3rd or 4th. Right: boxplot of the ranking. ObjMan = Object Manipulation, Anim = Animation, Occ = Occlusion handling, Across = Across Slides

After each technique was presented, participants were asked to fill out three questionnaires (SUS, NASA TLX, SSQ). The system usability questionnaire reported an average usability score of 71.53 ($SD = 20.13$) which indicates that our prototype has an average usability. This is an acceptable result for a prototype. The average total severity dimension of the Simulator Sickness Questionnaire was 12.26 ($SD = 17.34$) (Nausea: $M = 5.83$, $SD = 13.15$, Oculo-motor: $M = 11.79$, $SD = 19.49$, Disorientation: $M = 15.47$, $SD = 15.76$) , which means the participants were not suffering from severe simulator sickness.

The average overall task load measured by the NASA task load index was 25.54 ($SD = 11.18$), indicating that they were not overwhelmed (due to technical problems, we lost the data for TLX of one participant and computed the average with one less value).

Participants were also asked to rank the techniques based on preference, see Figure 11. There is no clear trend visible, but animation and occlusion handling seem to be more popular than the object manipulation technique or working across slides.

Participants also had the chance to comment on the techniques. For the object manipulation technique, four participants mentioned that it requires some practice and P8 suggested an additional explicit mode switch. P2 would rather do the scaling in 2D, as it only affects two dimensions. P7 thought the 3D rotation was intuitive, and P15 said it was very useful, because "you can better see what happens".

Regarding the animation, two participants mentioned that it is easier and more intuitive than in a standard slide authoring application. They also suggested additional features such as the snapping of two frames (P5), other types of animation (P8) and the possibility to exactly enter time (P16). However, P2 and P12 also mentioned that a certain amount of practice time is needed.

P1 and P17 thought that the occlusion handling method makes it easier and more natural to find a specific layer. P8 suggested to add the possibility to see the final image not only on the right but also on the tablets surface or standing on the upper edge. To be more consistent with the animation mode, P4 suggested to move the layers in-air and P15 was concerned it could get confusing with a very large number of objects, but for standard slides this should not be an issue.

When working across slides, participants also mentioned that many new techniques need to be remembered (P1, P2, P12). Only one participant thought that too much content is displayed that confuses her. P2 liked the possibility to make the content in front of the user invisible and P16 would also like this feature for the slides presented to the sides of the tablet. P3 suggested to add the possibility to bring resources (like a PDF) closer, so that it is easier to read it. P15 found it challenging to hit the bezel while not looking at it. P10 mentioned to feel some motion-sickness when the content was moving around (when adding or closing content).

Participants were also asked to comment on the overall system and explained their choices for the rating. Four people said that they liked the system. P2, P10 and P14 said they had fun using the system. P6 and P17 could not imagine themselves using a VR HMD for work, but nevertheless, they liked the system. Five participants mentioned that they would need to practice to get used to all techniques. However, P11 mentioned that "with a little practice it would be faster than in standard PowerPoint". P8 suggested to display a cheat sheet somewhere, as VR provides a lot of space. Three participants mentioned that 3D view is useful and "makes things easier to see than in 2D" (P10). Four people mentioned the system seems viable and P14 said "I could imagine to use this system". P15 and P17 liked the display of additional information and the increased workspace. P14 was interested in seeing how these interaction techniques could be integrated in a standard desktop setup. P4, P17 and P18 especially liked the occlusion handling technique, because "it saves us a lot of work" (P4) and "it is what I always missed in PowerPoint" (P17). P14 and P15 especially liked the animation, because "it was fun" (P14) and "implemented well and easy to understand" (P15).

6 DISCUSSION

As people work more remotely, from touchdown offices, on the go or from home, the importance of mobility and privacy increases. This paper, and other prior works [4, 15] show that this mode of work can be more efficient than the current use of 2D displays, as HMDs can fill the user's full field of view with potentially very large virtual displays and show world grounded stereoscopic information, which the user can interact with directly.

To examine the effect of such a display on presentation authoring applications, we focused on four specific techniques in this paper: The use of the large field of view for selection and interactions with content catalogs, use of 3D visualization for animation, handling object ordering and occlusions, and tracking the user's stylus in 6 DOF, enabling complex manipulations of objects.

We evaluated the performance of two techniques utilizing the larger display space and the 3D view provided by VR. First, we examined the effect of the extended display space where we expected that a larger field of view would speed up a visual search task as indicated by prior work [53]. The results of the search study showed that this was true for the EASY conditions where the target was identifiable pre-attentively. Yet, contrary to our expectations, as matching required more mental effort per item, the field of view did not seem to significantly influence search time. Nevertheless, we argue that a wider field of view is desirable, as it performs at least as good or better when compared to a smaller field of view when identifying targets and has the potential to skip interactions for switching or toggling displayed information.

Our second performance study showed that VR-based 3D visualization for resolving occlusion outperformed two baseline techniques in terms of speed and usability. Follow-up interviews suggested that the technique was well received, as "easy", "fast" and "provided a good overview". Similar techniques for embedding 2D data in 3D could also be used in other applications from animation (representing time as the third dimension) or image editing (showing semantic layers and versioning in space) to displaying alignment constraints and relations between slide objects and more. We showed the benefits of VR as a work space, and hope to encourage more work in this direction.

Additionally, a usability study in a walk-up-and-use scenario was conducted to evaluate these techniques. Subjects were confronted with a large amount of new interaction techniques and input modalities that differ from traditional touch input techniques in several ways. In spite of this, participants gave positive ratings for ease of use, which is also reflected in the results of the system usability scale indicating an average usability of 71.53. It showed the feasibility of our approach and the techniques were rated usable and enjoyable by participants. Many participants also expressed the feeling that their level of comfort with the techniques would improve with more training, indicating their unfamiliarity due to the walk-up-and-use scenario but general level of comfort with the techniques. Future work could look at ways to further improve the techniques to be even more intuitive, discoverable, and require less setup and explanation to use or to determine the actual learning curve of the techniques. No severe levels of simulator sickness were measured among participants and their perceived task load was also not high, indicating that they had no major issues with the basic functionality of our prototype. We hope that these findings are precursors for supporting further tasks in presentation authoring beyond the tested ones. Our goal was to gain initial insights into the usability, yet it will be important for future work to evaluate them in a more extensive way.

In this paper we focus on the graphic organization of slides. Text entry is an important issue for presentation authoring, yet we did not address it in this paper, as prior works has already addressed typing in VR, e.g., [13, 19]. These techniques can be used in conjunction with the presented techniques, for example, after entering the text, it could be manipulated like any other object (translated, rotated, scaled).

One of our main objectives in designing interactions was to use small hand movements, to allow for longer interaction times without fatigue. One option to further extend the input space but still keep small hand movements would be to remap physical pen and finger movements using a C-D ratio and visulizing copies of the pen and fingers (similar to our technique used when working across slides), for example to allow users to reach higher times in the 3D time view of the animation mode without scrolling through the timeline.

Another outcome from this work is the use of eye-gaze along with retargeting the input of a stylus on the tablet. The use of such techniques allows the user to interact with a very large display space while working in a limited cluttered physical environment, as their hands are located on a small tablet screen. This skips the need to physically reach displayed content sources.

Finally, this paper has only looked at the authoring side of presentation applications. There is another aspect of these applications which is the presentation process. While being out of the scope of this work, there are many similar advancements that can come to play while presenting, from the use of the large display space to presenting information useful for the presenter such as upcoming slides or notes. Also the presented techniques could be used to control the presentations such as quickly switching between slides that are not neighbours.

7 CONCLUSIONS AND FUTURE WORK

In this work we prototyped an experience called PoVRPoint: a set of tools that couple pen-, touch- and gaze-based authoring of presentations on mobile devices with the interaction possibilities afforded by VR. We studied the utility of extended display space in VR for tasks such as visual search, spatial manipulation of shapes, animations and ordering overlapping shapes. The results showed that VR can improve usability and performance of common authoring tasks and are liked by participants.

We see multiple avenues of future work. First, we aim at investigating the knowledge worker experience within the office of the future in VR [18], with multiple applications in use at any given time. To achieve this, we plan to explore techniques for transferring content across applications. Second, we want to explore how to expand the knowledge worker experience in VR by opportunistically leveraging available physical objects *in-situ*, such as a tray surface or an armrest in an airplane. It has been shown that VR can improve the usability and performance of editing presentations during the limited time period of the study. Still, future work should evaluate the effects of working in VR for prolonged time periods. Finally, we would like to extend the work into a collaborative one, and see how we can further use the advantages of VR to create experiences with awareness of the context and the remote participants. For example, VR can enable remote collaboration that represents both the shared document (the task space), as well as a representation of collaborators with a reference space to show where they are pointing in relation to the shared document, and private spaces [65].

REFERENCES

[1] Microsoft research mt. rogers. `https://www.microsoft.com/en-us/research/project/mt-rogers-head-worn-virtual-displays/`, 2021.

[2] R. Arora, R. H. Kazi, D. M. Kaufman, W. Li, and K. Singh. Magicalhands: Mid-air hand gestures for animating in vr. In *Proc. of the 32nd Annual ACM Symp. on User Interface Software and Technology*, pp. 463–477, 2019.

[3] N. Ashtari, A. Bunt, J. McGrenere, M. Nebeling, and P. K. Chilana. Creating augmented and virtual reality applications: Current practices, challenges, and opportunities. In *Proc. of the 2020 CHI Conf. on Human Factors in Computing Systems*, pp. 1–13, 2020.

[4] V. Biener, D. Schneider, T. Gesslein, A. Otte, B. Kuth, P. O. Kristensson, E. Ofek, M. Pahud, and J. Grubert. Breaking the screen: Interaction across touchscreen boundaries in virtual reality for mobile knowledge workers. *IEEE Trans. on Visualization and Computer Graphics*, 26(12):3490–3502, 2020. doi: 10.1109/TVCG.2020.3023567

[5] J. Brooke. Sus: a "quick and dirty'usability. *Usability evaluation in industry*, p. 189, 1996.

[6] W. Büschel, J. Chen, R. Dachselt, S. Drucker, T. Dwyer, C. Görg, T. Isenberg, A. Kerren, C. North, and W. Stuerzlinger. Interaction for immersive analytics. In *Immersive Analytics*, pp. 95–138. Springer, 2018.

[7] D. Cami, F. Matulic, R. G. Calland, B. Vogel, and D. Vogel. Unimanual pen+ touch input using variations of precision grip postures. In *Proc. of the 31st Annual ACM Symp. on User Interface Software and Technology*, pp. 825–837, 2018.

[8] A. Cannavò, C. Demartini, L. Morra, and F. Lamberti. Immersive virtual reality-based interfaces for character animation. *IEEE Access*, 7:125463–125480, 2019.

[9] D. Checa and A. Bustillo. A review of immersive virtual reality serious games to enhance learning and training. *Multimedia Tools and Applications*, 79(9):5501–5527, 2020.

[10] X. Chen, J. Schwarz, C. Harrison, J. Mankoff, and S. E. Hudson. Air+ touch: interweaving touch & in-air gestures. In *Proc. of the 27th annual ACM symp. on User interface software and technology*, pp. 519–525, 2014.

[11] M. F. Deering. Holosketch: a virtual reality sketching/animation tool. *ACM Trans. on Computer-Human Interaction (TOCHI)*, 2(3):220–238, 1995.

[12] T. Drey, J. Gugenheimer, J. Karlbauer, M. Milo, and E. Rukzio. Vrsketchin: Exploring the design space of pen and tablet interaction for 3d sketching in virtual reality. In *Proc. of the 2020 CHI Conf. on Human Factors in Computing Systems*, pp. 1–14, 2020.

[13] T. J. Dube and A. S. Arif. Text entry in virtual reality: A comprehensive review of the literature. In M. Kurosu, ed., *Human-Computer Interaction. Recognition and Interaction Technologies*, pp. 419–437. Springer International Publishing, Cham, 2019.

[14] N. Fellion, T. Pietrzak, and A. Girouard. Flexstylus: Leveraging bend input for pen interaction. In *Proc. of the 30th Annual ACM Symp. on User Interface Software and Technology*, UIST '17, p. 375–385. Association for Computing Machinery, New York, NY, USA, 2017. doi: 10.1145/3126594.3126597

[15] T. Gesslein, V. Biener, P. Gagel, D. Schneider, E. Ofek, M. Pahud, P. O. Kristensson, and J. Grubert. Pen-based interaction with spreadsheets in mobile virtual reality. In *2020 IEEE Int. Symp. on Mixed and Augmented Reality (ISMAR)*, 2020.

[16] R. Grasset, A. Duenser, H. Seichter, and M. Billinghurst. The mixed reality book: a new multimedia reading experience. In *CHI'07 extended abstracts on Human factors in computing systems*, pp. 1953–1958. ACM, 2007.

[17] T. Grossman, K. Hinckley, P. Baudisch, M. Agrawala, and R. Balakrishnan. Hover widgets: Using the tracking state to extend the capabilities of pen-operated devices. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems*, CHI '06, p. 861–870. Association for Computing Machinery, New York, NY, USA, 2006. doi: 10.1145/1124772.1124898

[18] J. Grubert, E. Ofek, M. Pahud, P. O. Kristensson, F. Steinicke, and C. Sandor. The office of the future: Virtual, portable, and global. *IEEE computer graphics and applications*, 38(6):125–133, 2018.

[19] J. Grubert, L. Witzani, E. Ofek, M. Pahud, M. Kranz, and P. O. Kristensson. Text entry in immersive head-mounted display-based virtual reality using standard keyboards. In *2018 IEEE Conf. on Virtual Reality and 3D User Interfaces (VR)*, pp. 159–166. IEEE, 2018.

[20] J. Guo, D. Weng, Z. Zhang, H. Jiang, Y. Liu, Y. Wang, and H. B.-L. Duh. Mixed reality office system based on maslow's hierarchy of needs: Towards the long-term immersion in virtual environments. In *2019 IEEE Int. Symp. on Mixed and Augmented Reality (ISMAR)*, pp. 224–235. IEEE, 2019.

[21] S. G. Hart and L. E. Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, vol. 52, pp. 139–183. Elsevier, 1988.

[22] O. Hilliges, S. Izadi, A. Wilson, S. Hodges, A. Garcia-Mendoza, and A. Butz. Interactions in the air: Adding further depth to interactive tabletops. In *UIST '09 Proc. of the 22nd annual ACM symp. on User interface software and technology*, pp. 139–148. ACM, October 2009.

[23] J. D. Hincapié-Ramos, X. Guo, P. Moghadasian, and P. Irani. Consumed endurance: a metric to quantify arm fatigue of mid-air interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1063–1072, 2014.

[24] K. Hinckley, X. A. Chen, and H. Benko. Motion and context sensing techniques for pen computing. In *Proceedings of Graphics Interface 2013*, GI '13, p. 71–78. Canadian Information Processing Society, CAN, 2013.

[25] K. Hinckley, S. Heo, M. Pahud, C. Holz, H. Benko, A. Sellen, R. Banks, K. O'Hara, G. Smyth, and B. Buxton. Pre-touch sensing for mobile interaction. In *CHI '16 Proc. of the 2016 CHI Conference on Human Factors in Computing Systems*, pp. 2869–2881. ACM, May 2016.

[26] K. Hinckley, K. Yatani, M. Pahud, N. Coddington, J. Rodenhouse, A. Wilson, H. Benko, and B. Buxton. Pen+ touch= new tools. In *Proc. of the 23nd annual ACM symp. on User interface software and technology*, pp. 27–36, 2010.

[27] T. Hirzle, J. Gugenheimer, F. Geiselhart, A. Bulling, and E. Rukzio. A design space for gaze interaction on head-mounted displays. In *Proc. of the 2019 CHI Conf. on Human Factors in Computing Systems*, p. 625. ACM, 2019.

[28] W. Hürst, C. G. Snoek, W.-J. Spoel, and M. Tomin. Size matters! how thumbnail number, size, and motion influence mobile video retrieval. In *Int. Conf. on MultiMedia Modeling*, pp. 230–240. Springer, 2011.

[29] D. Kamińska, T. Sapiński, S. Wiak, T. Tikk, R. E. Haamer, E. Avots, A. Helmi, C. Ozcinar, and G. Anbarjafari. Virtual reality and its applications in education: Survey. *Information*, 10(10):318, 2019.

[30] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The int. journal of aviation psychology*, 3(3):203–220, 1993.

[31] P. Knierim, V. Schwind, A. M. Feit, F. Nieuwenhuizen, and N. Henze. Physical keyboards in virtual reality: Analysis of typing performance and effects of avatar hands. In *Proc. of the 2018 CHI Conference on Human Factors in Computing Systems*, p. 345. ACM, 2018.

[32] R. A. Kockro, C. Amaxopoulou, T. Killeen, W. Wagner, R. Reisch, E. Schwandt, A. Gutenberg, A. Giese, E. Stofft, and A. T. Stadie. Stereoscopic neuroanatomy lectures using a three-dimensional virtual reality environment. *Annals of Anatomy-Anatomischer Anzeiger*, 201:91–98, 2015.

[33] P. G. Kry, A. Pihuit, A. Bernhardt, and M.-P. Cani. Handnavigator: Hands-on interaction for desktop virtual reality. In *Proc. of the 2008 ACM symp. on Virtual reality software and technology*, pp. 53–60. ACM, 2008.

[34] M. Kytö, B. Ens, T. Piumsomboon, G. A. Lee, and M. Billinghurst. Pinpointing: Precise head-and eye-based target selection for augmented reality. In *Proc. of the 2018 CHI Conf. on Human Factors in Computing Systems*, p. 81. ACM, 2018.

[35] D. A. Le, B. MacIntyre, and J. Outlaw. Enhancing the experience of virtual conferences in social virtual environments. In *2020 IEEE Conf. on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 485–494. IEEE, 2020.

[36] N. Li, T. Han, F. Tian, Huang, S. Jin, P. Minhui, Irani, and J. Alexander. Get a grip: Evaluating grip gestures for vr input using a lightweight pen. In *Proc. of the 2020 CHI Conf. on Human Factors in Computing Systems*. ACM, 2020.

[37] Z. Li, M. Annett, K. Hinckley, K. Singh, and D. Wigdor. Holodoc: Enabling mixed reality workspaces that harness physical and digital content. In *Proc. of the 2019 CHI Conf. on Human Factors in Computing Systems*, p. 687. ACM, 2019.

[38] B. MacIntyre, M. Gandy, S. Dow, and J. D. Bolter. Dart: a toolkit for rapid design exploration of augmented reality experiences. In *Proc. of the 17th annual ACM symp. on User interface software and technology*, pp. 197–206, 2004.

[39] N. Marquardt, R. Jota, S. Greenberg, and J. A. Jorge. The continuous interaction space: interaction techniques unifying touch and gesture on and

above a digital surface. In *IFIP Conf. on Human-Computer Interaction*, pp. 461–476. Springer, 2011.

[40] F. Matulic, R. Arakawa, B. Vogel, and D. Vogel. Pensight: Enhanced interaction with a pen-top camera. In *Proc. of the 2020 CHI Conf. on Human Factors in Computing Systems*. ACM, 2020.

[41] 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*, pp. 2143–2152. ACM, 2015.

[42] M. Mcgill, A. Kehoe, E. Freeman, and S. Brewster. Expanding the bounds of seated virtual workspaces. *ACM Trans. on Computer-Human Interaction (TOCHI)*, 27(3):1–40, 2020.

[43] M. McGill, J. Williamson, A. Ng, F. Pollick, and S. Brewster. Challenges in passenger use of mixed reality headsets in cars and other transportation. *Virtual Reality*, pp. 1–21, 2019.

[44] M. Nebeling, K. Lewis, Y.-C. Chang, L. Zhu, M. Chung, P. Wang, and J. Nebeling. Xrdirector: A role-based collaborative immersive authoring system. In *Proc. of the 2020 CHI Conf. on Human Factors in Computing Systems*, pp. 1–12, 2020.

[45] E. Ofek, J. Grubert, M. Pahud, M. Phillips, and P. O. Kristensson. Towards a practical virtual office for mobile knowledge workers. *Microsoft New Future of Work Symposium*, 2020.

[46] P. Pantelidis, A. Chorti, I. Papagiouvanni, G. Paparoidamis, C. Drosos, T. Panagiotakopoulos, G. Lales, and M. Sideris. Virtual and augmented reality in medical education. In *Medical and Surgical Education-Past, Present and Future*, pp. 77–97. 2018.

[47] D. Parmar and T. Bickmore. Making it personal: Addressing individual audience members in oral presentations using augmented reality. *Proc. of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 4(2):1–22, 2020.

[48] L. Pavanatto, C. North, D. A. Bowman, C. Badea, and R. Stoakley. Do we still need physical monitors? an evaluation of the usability of ar virtual monitors for productivity work. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, pp. 759–767. IEEE, 2021. doi: 10.1109/VR50410.2021.00103

[49] K. Pfeuffer, J. Alexander, M. K. Chong, Y. Zhang, and H. Gellersen. Gaze-shifting: Direct-indirect input with pen and touch modulated by gaze. In *Proc. of the 28th Annual ACM Symp. on User Interface Software & Technology*, pp. 373–383, 2015.

[50] K. Pfeuffer, K. Hinckley, M. Pahud, and B. Buxton. Thumb+ pen interaction on tablets. In *CHI*, pp. 3254–3266, 2017.

[51] K. Pfeuffer, B. Mayer, D. Mardanbegi, and H. Gellersen. Gaze+ pinch interaction in virtual reality. In *Proc. of the 5th Symp. on Spatial User Interaction*, pp. 99–108. ACM, 2017.

[52] D.-M. Pham and W. Stuerzlinger. Is the pen mightier than the controller? a comparison of input devices for selection in virtual and augmented reality. In *25th ACM Symp. on Virtual Reality Software and Technology*, pp. 1–11, 2019.

[53] E. D. Ragan, D. A. Bowman, R. Kopper, C. Stinson, S. Scerbo, and R. P. McMahan. Effects of field of view and visual complexity on virtual reality training effectiveness for a visual scanning task. *IEEE trans. on visualization and computer graphics*, 21(7):794–807, 2015.

[54] P. Reipschläger and R. Dachselt. Designar: Immersive 3d-modeling combining augmented reality with interactive displays. In *Proc. of the 2019 ACM Int. Conf. on Interactive Surfaces and Spaces*, pp. 29–41, 2019.

[55] J. Rekimoto and M. Saitoh. Augmented surfaces: a spatially continuous work space for hybrid computing environments. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems*, pp. 378–385. ACM, 1999.

[56] A. Ruvimova, J. Kim, T. Fritz, M. Hancock, and D. C. Shepherd. ” transport me away”: Fostering flow in open offices through virtual reality. In *Proc. of the 2020 CHI Conf. on Human Factors in Computing Systems*, pp. 1–14, 2020.

[57] D. Schneider, A. Otte, T. Gesslein, P. Gagel, B. Kuth, M. S. Damlakhi, O. Dietz, E. Ofek, M. Pahud, P. O. Kristensson, et al. Reconviguration: Reconfiguring physical keyboards in virtual reality. *IEEE trans. on visualization and computer graphics*, 2019.

[58] R. Schweigert, V. Schwind, and S. Mayer. Eyepointing: A gaze-based selection technique. In *Proc. of Mensch und Computer 2019*, pp. 719–723. ACM, 2019.

[59] L. Sidenmark, D. Mardanbegi, A. Ramirez Gomez, C. Clarke, and H. Gellersen. Bimodalgaze: Seamlessly refined pointing with gaze and filtered gestural head movement. In *Proc. of Eye Tracking Research and Applications*, 2020.

[60] M. Slater, D.-P. Pertaub, C. Barker, and D. M. Clark. An experimental study on fear of public speaking using a virtual environment. *CyberPsychology & Behavior*, 9(5):627–633, 2006.

[61] H. Song, H. Benko, F. Guimbretiere, S. Izadi, X. Cao, and K. Hinckley. Grips and gestures on a multi-touch pen. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems*, CHI ’11, p. 1323–1332. Association for Computing Machinery, New York, NY, USA, 2011. doi: 10.1145/1978942.1979138

[62] 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, p. 1–13. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3173574.3174221

[63] H. B. Surale, A. Gupta, M. Hancock, and D. Vogel. Tabletinvr: Exploring the design space for using a multi-touch tablet in virtual reality. In *Proc. of the 2019 CHI Conf. on Human Factors in Computing Systems*, pp. 1–13, 2019.

[64] Z. Szalavári and M. Gervautz. The personal interaction panel–a two-handed interface for augmented reality. In *Computer graphics forum*, vol. 16, pp. C335–C346. Wiley Online Library, 1997.

[65] A. Tang, M. Pahud, K. Inkpen, H. Benko, J. C. Tang, and B. Buxton. Three’s company: Understanding communication channels in three-way distributed collaboration. In *Proc. of the 2010 ACM Conf. on Computer Supported Cooperative Work*, CSCW ’10, p. 271–280. Association for Computing Machinery, New York, NY, USA, 2010. doi: 10.1145/1718918.1718969

[66] F. Tian, L. Xu, H. Wang, X. Zhang, Y. Liu, V. Setlur, and G. Dai. Tilt menu: Using the 3d orientation information of pen devices to extend the selection capability of pen-based user interfaces. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems*, CHI ’08, p. 1371–1380. Association for Computing Machinery, New York, NY, USA, 2008. doi: 10.1145/1357054.1357269

[67] R. Unger and C. Chandler. *A Project Guide to UX Design: For user experience designers in the field or in the making*. New Riders, 2012.

[68] D. Vogel, P. Lubos, and F. Steinicke. Animationvr-interactive controller-based animating in virtual reality. pp. 1–6. IEEE, 2018.

[69] P. Wacker, O. Nowak, S. Voelker, and J. Borchers. Arpen: Mid-air object manipulation techniques for a bimanual ar system with pen & smartphone. In *Proc. of the 2019 CHI Conf. on Human Factors in Computing Systems*, pp. 1–12, 2019.

[70] J. A. Wagner Filho, C. M. D. S. Freitas, and L. Nedel. Virtualdesk: a comfortable and efficient immersive information visualization approach. In *Computer Graphics Forum*, vol. 37, pp. 415–426. Wiley Online Library, 2018.

[71] X. Wang, L. Besançon, D. Rousseau, M. Sereno, M. Ammi, and T. Isenberg. Towards an understanding of augmented reality extensions for existing 3d data analysis tools. In *Proc. of the 2020 CHI Conf. on Human Factors in Computing Systems*, pp. 1–13, 2020.

[72] P. D. Wellner. Interacting with paper on the digitaldesk. Technical report, University of Cambridge, Computer Laboratory, 1994.

[73] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proc. of the SIGCHI Conf. on human factors in computing systems*, pp. 143–146, 2011.

[74] J. M. Wolfe and T. S. Horowitz. Five factors that guide attention in visual search. *Nature Human Behaviour*, 1(3):1–8, 2017.

[75] D. Zielasko, M. Krüger, B. Weyers, and T. W. Kuhlen. Menus on the desk? system control in deskvr. In *2019 IEEE Conf. on Virtual Reality and 3D User Interfaces (VR)*, pp. 1287–1288. IEEE, 2019.

[76] D. Zielasko, M. Krüger, B. Weyers, and T. W. Kuhlen. Passive haptic menus for desk-based and hmd-projected virtual reality. In *2019 IEEE 5th Workshop on Everyday Virtual Reality (WEVR)*, pp. 1–6. IEEE, 2019.

[77] D. Zielasko, B. Weyers, and T. W. Kuhlen. A non-stationary office desk substitution for desk-based and hmd-projected virtual reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 1884–1889. IEEE, 2019.

Chapter 6

Quantifying the Effects of Working in VR for One Week

Summary

The previous chapters, especially Chapter 4 and Chapter 5 have looked into how VR can potentially support knowledge workers, especially in a mobile context. However, using VR for work on a regular basis also poses the question of how this could affect the users. As presented in Chapter 3, research on the long-term effects of using VR is very scarce. Therefore, this chapter presents a study in which participants worked in a VR environment for a whole workweek. For comparison participants were also observed working in a physical environment for one week. The aim of this study was to measure the baseline costs of VR and therefore, the VR and physical environment were designed to be as similar as possible.

The results show that VR resulted in significantly worse ratings for most measures, such as task load, frustration, negative affect, anxiety, eye strain, usability, flow, perceived productivity, wellbeing and simulator sickness. Therefore, this study shows the importance of conducting studies that last considerably longer than most current studies and provides a baseline for continuing research aiming at improving VR experiences.

Contribution Statement

The study design was discussed and agreed upon by Verena Biener, Jens Grubert and Stephan Streuber. Then, Verena Biener constructed and tested the study setup, including a literature review on related papers and useful questionnaires. Verena Biener conducted the study with around half the participants and instructed Snehanjali Kalamkar, John J. Dudley, Jinghui Hu and Maheshya Weerasinghe to conduct the study in remote locations. The statistical analysis was performed by Verena Biener and the results were discussed among all authors. Verena Biener also created all figures, with support from Negar Nouri. The first draft of the article was written by Verena Biener and then refined by all authors.

Article [16]

© 2022 IEEE. Reprinted, with permission, from Verena Biener, Snehanjali Kalamkar, Negar Nouri, Eyal Ofek, Michel Pahud, John J. Dudley, Jinghui Hu, Per Ola Kristensson, Maheshya Weerasinghe, Klen Čopič Pucihar, Matjaž Kljun, Stephan Streuber & Jens Grubert. Quantifying the Effects of Working in VR for One Week. In IEEE Transactions on Visualization and Computer Graphics (TVCG), vol. 28, no. 11, Nov. 2022, pp. 3810-3820. `https://doi.org/10.1109/TVCG.2022.3203103`

Quantifying the Effects of Working in VR for One Week

Verena Biener, Snehanjali Kalamkar, Negar Nouri, Eyal Ofek, Michel Pahud, John J. Dudley, Jinghui Hu,
Per Ola Kristensson, Maheshya Weerasinghe, Klen Čopič Pucihar, Matjaž Kljun, Stephan Streuber, Jens Grubert

Fig. 1. a) Participant working in a physical desktop setup in one of three physical locations (Germany), including a curved display and a keyboard with integrated touchpad. b) Participant working in the VR setup wearing an Oculus Quest 2, using the same keyboard but a virtual curved display. c) Participant's view within VR where the streamed content from a remote machine is visible. d) Participant's view of the keyboard and hands within VR.

Abstract—Virtual Reality (VR) provides new possibilities for modern knowledge work. However, the potential advantages of virtual work environments can only be used if it is feasible to work in them for an extended period of time. Until now, there are limited studies of long-term effects when working in VR. This paper addresses the need for understanding such long-term effects. Specifically, we report on a comparative study (n=16), in which participants were working in VR for an entire week—for five days, eight hours each day—as well as in a baseline physical desktop environment. This study aims to quantify the effects of exchanging a desktop-based work environment with a VR-based environment. Hence, during this study, we do not present the participants with the best possible VR system but rather a setup delivering a comparable experience to working in the physical desktop environment. The study reveals that, as expected, VR results in significantly worse ratings across most measures. Among other results, we found concerning levels of simulator sickness, below average usability ratings and two participants dropped out on the first day using VR, due to migraine, nausea and anxiety. Nevertheless, there is some indication that participants gradually overcame negative first impressions and initial discomfort. Overall, this study helps lay the groundwork for subsequent research, by clearly highlighting current shortcomings and identifying opportunities for improving the experience of working in VR.

Index Terms—virtual reality, long-term, knowledge work, user study

◆

1 INTRODUCTION

Virtual Reality (VR) has the potential to enhance physical working environments, for instance, by providing repeatable, location independent user experiences, or relieving physical world limitations such as screen sizes of physical screens [19]. For example, prior work studied transplanting or extending tasks performed by knowledge workers (as defined by [11]) from physical 2D displays to head-mounted displays (HMDs) (e.g. [6, 15, 45]). The use of a large display space around the user provided by the HMD, not limited by the size of physical monitors, supports visualization of information in multiple depth layers. The direct manipulation of data using natural hand motions, and the ability to map small physical motions to larger actions in such virtual environments can reduce fatigue and may open up the workspace for people with special needs. Remote collaboration in such environments brings people to the same virtual space, and, with a varying level of representation needed for a particular task at hand, can increase their sense of presence [46]. In addition, an office in VR can dynamically adapt to a user's work situation—it could transform into a calming beach when reading a paper or a formal office when writing an email. Virtual environments also provide privacy from the outside world and the removal of real environment disruptions may help users focus their attention on work. Further, collaborative work in virtual environments can save users the travel hassle, associated costs and reduce carbon footprint [26].

However, VR substitutes users' visual, audio and sometimes haptic sensations provided by the physical world with artificial inputs. With current VR technologies many of these inputs provide an inferior experience compared to the real world. For example, HMDs typically have a smaller field of view (FoV) compared to humans' visual field, and their resolution, while increasing over the years, is still lower than the retinal resolution. Most HMDs render the entire virtual world at a fixed focal distance and their dynamic range is smaller than what we can perceive as humans. Additionally, with their substantial weight and by blocking the air flow on the face, HMDs can reduce users' comfort. While many

---

- *V. Biener, S. Kalamkar, N. Nouri, S. Streuber and J. Grubert are with Coburg University of Applied Sciences and Arts, Germany. E-mail: {verena.biener, snehanjali.kalamkar, negar.nouri, stephan.streuber, jens.grubert}@hs-coburg.de.*
- *E. Ofek and M. Pahud are with Microsoft Research, United States. E-mail: {mpahud, eyalofek}@microsoft.com*
- *J. J. Dudley, J. Hu and P. O. Kristensson are with University of Cambridge, United Kingdom. E-mail: {jjd50,jh2265,pok21}@cam.ac.uk.*
- *M. Weerasinghe is with the University of Primorska, Slovenia and the University of St Andrews, United Kingdom. E-mail: amw31@st-andrews.ac.uk.*
- *K. Čopič Pucihar and M. Kljun are with University of Primorska, Slovenia. E-mail: {klen.copic, matjaz.kljun}@famnit.upr.si*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

of these limitations are likely to be addressed and improved in the future, visions of virtual knowledge work are not limited to research labs anymore. Startups such as Spatial or large corporations such as Meta and Microsoft are advancing products to facilitate virtual knowledge work.

Hence, we see it as beneficial to understand *how* VR technology available to end-users today influences knowledge work. In particular, we strive to quantify the effects of VR experiences that users can have today using commercial of-the-shelf hardware. Hence, we decided against designing an as good as possible virtual environment. Instead, we intentionally decided to study a specific operating point of VR hardware and software—the commercially available Oculus Quest 2 HMD and Logitech K830 physical keyboard with integrated touchpad. This specific combination enables integrated hand and keyboard tracking and allows virtual work for a large user group today.

To this end, we report on a study with users working for an entire workweek (five days, eight hours each day) in VR with the aforementioned setup (Fig. 1, b). To quantify the costs of this setup compared to physical workplaces, participants also worked another workweek in a corresponding physical work environment (Fig. 1, a). To be able to compare the two conditions, we made both the virtual and physical work environments as similar as possible: the same display size, shape, resolution and input device. At the same time, and in contrast to prior studies (e.g. [22, 50]), we decided against prescribing artificial tasks to participants but instead allowed participants to determine their own actual work. While this choice potentially impacts repeatability and replicability it also increases the ecological validity of the findings.

We are well aware that with the current state of VR technology, working in VR will be demanding on the user: the size, weight and quality of the HMD, its limited FoV, latency, and the authenticity of the representation of the world around users (drinks, keyboard, mouse, etc.) can affect the workflow. As a result, one can expect that the user experience in VR might be inferior to the one in the physical environment. Still, we see it as a worthwhile endeavor to *quantify* the effects of working in VR over five consecutive working days with eight hours in each day. This can serve as baseline for future optimized VR experiences that do not necessarily replicate a physical environment. Further, running a study over five consecutive days allows us to study gradual changes over time.

The main findings of the study are as follows: 1) self-rated task load was significantly higher in VR (approximately 35%), as was frustration (42%), negative affect (11%), anxiety (19%) and eye strain (48%); 2) VR resulted in significantly lower system usability scale scores (36%) with below average ratings, self-rated flow (14%), perceived productivity (16%) and wellbeing (20%); 3) VR resulted in (according to Stanney et al. [54]) poor ratings of simulator sickness. The values of some measures improved over the five days for both VR and PHYSICAL. However, we only found that the rate of change was significantly higher for VR regarding visual fatigue ratings compared to PHYSICAL which means it decreased significantly faster in VR.

In order to facilitate future work, we release an anonymized dataset along with this paper which can be found at `https://gitlab.com/mixedrealitylab/quantifying-the-effects-of-working-in-vr-for-one-week`. In summary, this paper presents the following contributions: 1) a study on the effects of working in VR for five working days, eight hours per day, 2) quantification of effects that occur between VR and a baseline physical work environment and 3) an accompanying dataset to aid replication and further analysis by the community.

2 RELATED WORK

There are several potential benefits of using VR as a working environment. Besides theoretical advantages of virtual offices (e.g., [19, 31, 32, 44] further work has empirically investigated specific benefits of VR for knowledge work, which are discussed in Sect. 2.1. Yet, the prolonged usage of VR can affect the usage and workflow. Therefore, Sect. 2.2 discusses previous work investigating long-term VR use. The potential benefits and understanding long-term usage motivated our work, which builds upon and extends the existing body of literature.

2.1 Benefits of VR for Knowledge Work

Prior research has examined the use of VR as an environment for knowledge work, showing a variety of effects of VR on the quality of work, the flexibility of the VR display, the direct interaction which can be used to increase productivity and the control of the environment around the worker which can be used to reduce stress. Ruvimova et al. [50] found that a VR office on a virtual beach successfully reduced distraction and simulated the workflow of a closed physical office. This indicates that VR can help users stay focused in open office environments. This capability is not restricted to VR. By displaying virtual separators in a physical office, augmented reality (AR) has also been shown to be useful against visual distractions [33]. In addition, users were allowed to personalize their work environment, which helped to increase their satisfaction and improved their experience of a shared workspace. Personalization of workspace has been shown beneficial in several studies and the literature emphasizes the importance of personalized design concepts for the so called non-territorial offices or shared office space [9, 29]. In such contexts, both VR and AR have the capability to allow users to design and decorate their virtual office space according to their preferences.

Prior research [2, 56] has demonstrated how virtual nature environments can reduce stress and improve mood during work. It is well known that spaces filled with greenery or even a view on greenery provide an opportunity for recovery from mental fatigue and are generally beneficial to human health [4, 27]. Further work indicates that VR can reduce stress more effectively than simply streaming a video of relaxing content [48] and that interactive VR environments are more effective than passively consuming VR content [57]. Also, Mostajeran et al. [41] found that showing a forest environment had positive effects on cognition. Despite these benefits, Li et al. [34] found that while users preferred nature environments, they were more productive when working in an office-like environment.

VR also provides new possibilities for interacting with and visualizing work-related content. Biener et al. [6] showed how multimodal interaction techniques, including eye tracking, can be used to efficiently navigate between a large number of virtual displays and how a three-dimensional visualization with depth-perception makes tasks involving multiple layers much easier. Pavanatto et al. [45] compared physical and virtual monitors in AR and concluded that while virtual monitors can be beneficial, they are still technically inferior. Therefore, they suggested to combine both physical and virtual monitors in the workflow, which is already put to practice in commercial products such as the Lenovo ThinkReality A3 Smart Glasses.

VR can also increase the usability and performance of knowledge worker tasks. For example, spreadsheet [15] and presentation authoring applications [5] can significantly benefit from depth perception, potentially large virtual display space, and new interaction possibilities such as a spatially tracked pen and eye-tracking, all provided by VR as well.

The benefits of extended virtual display spaces have been investigated in both stationary [12, 39] and mobile [40, 42] environments. For example, new interaction techniques for navigating large display spaces have been proposed using non-linear mapping of head-gaze [39] or above-surface interaction [40]. Personalized layouts of multiple virtual displays have also been studied [12, 42]. For example, Ens et al. [12] indicated that application switching times can be decreased by up to 40% using optimized layouts. Ng et al. [42] found that layout preferences depend on the perception of other passengers' physical presence, e.g., when sitting in an airplane seat next to another person. Using VR for knowledge work can also address privacy issues [19, 51] and enhance capabilities of existing devices, such as changing the keyboard on the fly to support other languages, symbols, and layouts [51].

However, these prior works have been studied in short term experiments. Also, while the proposed interaction techniques show the potential for supporting knowledge work in VR they often are restricted to specific lab-based setups. Instead, in our study, we focused on experiences that are accessible to potentially millions of users today, relying on commercially off-the-shelf hardware and software.

2.2 Working in VR for an Extended Period of Time

While several possibilities of using VR for knowledge work have been explored in recent years, showing advantages, the effects of long-term usage of such environments are still not sufficiently understood. Besides anecdotal reports (for example [10, 30]), researchers began to investigate long-term effects of AR and VR interfaces. Long-term usage of AR has been explored in manufacturing. In a study by Grubert et al. [18] participants used AR for four hours in an order picking task. The results showed a higher work efficiency in the AR condition without an increase in overall objective and subjective strain. However, some participants felt a higher eye discomfort compared to a non-AR baseline. Wille et al. [59] compared four hours of work on a monocular HMD, tablet and computer monitor, and did not find any objective physiological effect on participants' visual system. However, subjective ratings of strain were significantly higher for the HMD, which authors attributed to the unfamiliarity with the technology. In Funk et al.'s study [13], participants used a projector-based AR system for at least three full working days to perform an assembly task. The system was shown to be useful for untrained workers, yet it slowed down the performance and increased cognitive load for expert workers.

Lu et al. [37] conducted an in-the-wild study of glanceable AR for everyday use where participants used the prototype for three days. The authors concluded that participants liked the prototype and would use it daily if the HMD had a form factor of regular prescription glasses. In a 24-hour self-experiment [55] one participant worked, ate, slept and entertained himself in VR. In another case study [43], two participants used VR HMDs for 12 hours straight and reported only mild simulator sickness symptoms.

The prior work most similar to this paper has been conducted by Guo et al. [21–23] and Shen et al. [53]. They looked at prolonged use of VR for office work in a virtual environment. Their long-term study [21,22,53] had 27 participants working in a virtual and a physical office for eight hours, performing tasks such as document correction, keyword searching, text input and image classification. In the context of the Maslow's Hierarchy of Needs, the researchers found that emotional needs must be met in short and long-term use, while physiological, belongingness needs, temporal- and self-presence are only important for long-term use [22]. Evaluating the effects on visual discomfort, Guo et al. [21] found that signs of visual fatigue (subjective rating, pupil size, accommodation response) change with time in both physical and virtual conditions. They also indicated that female participants suffered more from visual fatigue and they speculated that this could be due to less experience with VR. They did not find significant differences in nausea and eye strain between the virtual and physical condition, but larger difficulty in focusing and physical discomfort (due to weight and form factor of HMD) were present in VR. In addition, Shen et al. [53] asked participants to perform a psychomotor vigilance task (PVT) six times during the day and found significantly less PVT lapses and higher reaction times in the physical environment, which indicates a higher mental fatigue in VR. They propose two explanations: either VR occupies more attentional resources or VR can increase attention of participants more effectively so that participants use more attentional resources. However, all these findings are based on artificial tasks designed by researchers, which might influence the results compared to real in-the-wild tasks. Also, prolonged use studies seem necessary to ensure participants obtain sufficient familiarity with the VR setup so that it is comparable with working in the physical environment. Hence, we here conduct a study that is five times as long in duration (five days, eight hours per day in each condition) compared to the prior studies [21, 22, 53]. Further, in our study we ask participants to work on their own everyday work tasks.

3 STUDY

Our goal was to quantify the effects of working in a virtual reality environment for extended periods of time. Specifically, we compared working in VR to working in a physical environment by analyzing their respective effects on a variety of measures as explained in the following sections. The study was approved by the ethics committee of Coburg University of Applied Sciences and Arts.

Table 1. Overview of measures taken during the study: before starting the daily work (START), and after two (2), four (4), six (6) and eight (8) hours of work. For the significance tests the average of each of the two values from the morning and afternoon are used.

Measure	Start	2h	4h	6h	8h
Task Load		✓	✓	✓	✓
System Usability		✓	✓	✓	✓
Flow		✓	✓	✓	✓
Perceived Productivity		✓	✓	✓	✓
Frustration		✓	✓	✓	✓
Presence		VR only	VR only	VR only	VR only
Pos. / Neg. Affect	✓				✓
Wellbeing	✓	✓	✓	✓	✓
Anxiety	✓	✓	✓	✓	✓
Simulator Sickness	✓	✓	✓	✓	✓
Visual Fatigue	✓	✓	✓	✓	✓
Heart Rate	continuous				
Break Times	continuous				
Typing Speed	first day				last day

3.1 Study Design

The experiment was designed as a within-subjects study. Each participant completed a full week of work (5 days, 8 hours per day) in a virtual and in a physical environment. The first independent variable is ENVIRONMENT, which is either VR or PHYSICAL. We counterbalanced ENVIRONMENT, with half the participants starting with the VR week and half of them with PHYSICAL. The second independent variable is DAY with five levels, namely DAY1, DAY2, DAY3, DAY4 and DAY5.

To get a better idea of how the measures evolve during the week, participants answered various questionnaires five times a day at fixed times: before starting the daily work, two before lunch break (after 2 and 4 hours of work) and two after a 45 minute lunch break after 6 and 8 hours of work. We subsequently merged these data-points into three values for each day, resulting in a third independent variable TIME with the three levels START, MORNING (mean of two data points before lunch break) and AFTERNOON (mean of two data points after lunch break).We merged the two data points from morning and afternoon to make measures more robust against the influence of different work tasks and to account for possible logging errors, as one missing value would eliminate all measures from this participant for ANOVA analysis.

Dependent variables included self-rated subjective as well as objective measures as explained in the next section. The number of data points, and, therefore, also the number of levels of the independent variables DAY and TIME varies between these measures, because depending on their purpose, some of them were recorded at different frequencies. The measures and the number of data-points obtained for each are summarized in Table 1.

3.2 Measures

In the course of the study we collected a range of subjective as well as objective measures which are presented in Table 1. Some of the subjective measures were assessed with questionnaires at four specified time intervals: after two, four, six and eight hours of work. These measures include: task load measured by the NASA TLX questionnaire [24], usability measured by the system usability scale [8], flow [49], presence using the IPQ questionnaire[1] [52], and two separate questions asking the participants to rate their perceived productivity and frustration on a 7-point-Likert-scale (from 1 to 7) ("I was very productive in the last two hours."; "I was very frustrated in the last two hours."). These measures where not taken at the START of the day, because they are not meaningful without referring to a prior period of work.

Further subjective measures were assessed five times a day. Specifically, data was collected before starting the daily work as well as every two hours as with the aforementioned questionnaires. These measures assessed different aspects of participants' general physical and mental wellbeing: anxiety using the short version of STAI [61] to examine if

---

[1] `http://www.igroup.org/pq/ipq`

VR has an effect on anxiety, the simulator sickness questionnaire [28], visual fatigue using six questions from [25] as was done by [3], and a separate question asking participants to rate their wellbeing on a 7-point-Likert-scale ("I was very comfortable in the last two hours."). Even though the simulator sickness questionnaire already includes one question about visual fatigue, we chose to use an additional set of questions to assess it more closely. In addition, participants answered the PANAS-SF questionnaire [58] to record a positive/negative affect of each condition on emotional state of users. This questionnaire was answered solely before starting the daily work and at the end of the working day to measure a change in emotional state induced by a the whole day of work. Sometimes, participants could not complete a particular questionnaire at the exact time because they had to participate in a meeting or lecture. Still the average duration of time blocks was very close to two hours ($m = 122$ *min*, $sd = 55.88$)

At the end of each week, further qualitative data was collected in a short interview. Participants were asked to talk about what they liked or disliked during the week, how they felt, what problems occurred and what they would improve in the particular condition used that week. After both weeks were completed, they were also asked about their preferences and if they could imagine using VR for work in the future.

Additionally, we also collected a set of objective measures. These included heart rate, which can be used as an indication for stress. It was continuously recorded using a Polar H10 chest strap. Using a webcam, we also recorded users' heads and we used these videos to track participants' break behaviors, that is, how much time participants spent away from the screen. Longer breaks were also recorded by ManicTime [1] software, which we used to additionally detect and confirm inactivity of more than 10 minutes as a break. At the beginning and end of each week the typing speed of each participant was assessed using a web-based typing speed test [16] to see if they were adapting to the unknown keyboard during the week. Note that we did not focus on text entry performance as a primary measure, as prior work already indicated that users can adapt well to physical keyboards in VR [20,47].

3.3 Apparatus

The experimental setup was designed to make the work environment in both conditions (VR and PHYSICAL) comparable, while still relying on commercial off-the-shelf hardware and standard system software. We used an Oculus Quest 2 HMD (Quest-Build 37.0) as it provides integrated tracking of user's hands and a physical keyboard, which was, at the time of the experiment, the Logitech K830. This keyboard has an integrated touchpad and was used as the main input device. Hence, we restrained from using an external mouse, which would also inhibit the repeatability of the experiment (as custom solutions would be needed for mouse tracking). In both conditions participants used a *work-computer* to work on during the whole experiment. Participants could either bring their own computer, or use the computer provided by us. In both conditions, a browser and Chrome Remote Desktop [35] were used to connect to the *work-computer*. This allowed participants to see the desktop environment through the Oculus Browser in VR HMD. To make the PHYSICAL condition as similar as possible to VR and to reduce confounding variables, we used a second computer (*display-computer*) in the non-VR condition as well and connected it to the *work-computer* via Chrome Remote Desktop. Therefore, using a personal laptop did not affect the study as users were accessing it remotely, solely using the curved monitor or the VR HMD and the keyboard. The language settings of the *work-computer* were set to English regardless of the geographical location of the participants, because at the time of the study, the Oculus Quest 2 only supported visualization of the English keyboard layout. A 24-inch curved display (AOC Gaming C24G1) was used as the *display-computer*, which was placed 60 cm from the participants, to resemble the field of view of the virtual browser window in VR (ca. 47° horizontally, 27.5° vertically) as closely as possible. The resolution of the *work-computer* was set to 1366×768 at 125% display scaling which allowed common user interface elements to still be legible in the virtual display. For example, the capital letter *A* rendered in the typeface Calibri at size 12 pt would result in a vertical FoV of 17.19 arcminutes at 60 cm viewing distance.

Fig. 2. a) VR condition in the UK, b) PHYSICAL condition in the UK, c) VR condition in Slovenia, d) PHYSICAL condition in Slovenia.

The physical curved display was present in both conditions, because a webcam was mounted on top of it to detect the participants' head movements. For the VR condition, an ArUco marker [14] was attached to the headset to allow analysis of break times. For the non-VR condition, we used the face-mesh algorithm of mediapipe [36]. In both conditions, the keyboard was connected directly to the *work-computer* via the Bluetooth, as otherwise the remote connection causes some keys to function incorrectly. This necessitated the use of a second keyboard in the VR condition which was connected to the Oculus Quest 2 via Bluetooth. When the Quest 2 has a Bluetooth connection to a K830 keyboard, it will detect and display any K830 in view. This allowed the keyboard connected to the *work-computer* and used by the participant to be displayed in VR. The keyboard actually connected to the Quest 2 was hidden out of sight.

In conjunction with the keyboard tracking in the Oculus Quest 2, the hand tracking was also enabled. We added a virtual desk using the 'Bring Your Desk Into VR' option of the Oculus Quest 2. To mitigate distractions in the virtual environment, we selected the 'Bubbles' home environment. We did not replicate the rest of the physical environment (e.g., walls surrounding the user). While this might create a potential confound, we decided on such design in favor of repeatability and ecological validity of the experiment (i.e. experiences that actual users can have outside of lab environments without the need to install custom hardware or software). Also, while we did not quantify the illuminance at the user's eye in the physical workplace we ensured that the perceived lighting conditions were comparable.

In addition, participants could double-tap the side of the HMD to toggle the pass-through-mode, so that they could more easily drink, eat or pick up their phone. We also reduced distractions in the physical environment as much as possible, however, the setup slightly differed between the three physical locations in which the experiment was carried out (Germany, UK, Slovenia). In Germany the participants (ten in total) were partially shielded from the rest of the room using mobile walls, as can be seen in Fig. 1, a and Fig. 1, b. In the UK, participants (2) were sitting on their own in the corner of a vacant open office, as displayed in Fig. 2, a and Fig. 2, b. In Slovenia, participants (4) were sitting in a corner of a small office, as shown in Fig. 2, c and Fig. 2, d. In general, participants were sitting on their own, but due to the length of the experiment we could not completely control the occasional presence of other people. In these situations, other people present were asked to be as quiet as possible.

In both conditions participants were wearing a Polar H10 chest strap, which was used for collecting heart rate data. The strap sent data to an Android Phone with the Polar App installed via Bluetooth.

3.4 Participants

In total, 16 participants (*mean age* = 29.31, *sd* = 5.52, 10 male, 6 female) participated in the study and completed both weeks. All participants were employees or researchers at a university. Two additional participants (age 32 and 33, one male, one female) dropped out on the first day of the VR condition due to a migraine, nausea and anxiety. Eight participants started the study with the VR week and the other eight with the PHYSICAL week. Three participants were left handed, but all participants used their right hand to operate a mouse which was consistent with the touch pad of the K830 being on the right-hand side of the keyboard. All participants had normal, or corrected to normal eyesight and they saw everything clearly in the virtual environment. Two (2) participant had no previous experience with VR, six (6) only slight experience, two (2) had moderate experience, four (4) substantial experience, and two (2) extensive experience. Participants were also asked to indicate how often they usually look at the keyboard while typing on a scale from 1 (never) to 7 (all the time), which resulted in a mean rating of 3.19 ($sd = 1.38$). When asked about how often they use a touchpad the mean rating (on the same scale) was 2.88 ($sd = 1.67$).

3.5 Procedure

All participants were informed about the procedure and the content of the study, signed a consent form and filled out a demographic questionnaire. Next, the participant attached the Polar H10 heart rate sensor. On the first day of VR, participants received a short introduction using the HMD, how to activate the pass-through mode, and how to reconnect via Chrome Remote Desktop.

Next, the camera recording of the participant's face and ManicTime software were started and the participant filled out the first set of questionnaires (the further ones followed after 2, 4, 6 and 8 hours according to Table 1). The questionnaires were filled out on the same screens that the participants worked on (a physical screen in PHYSICAL and virtual one in VR). Each participant conducted a typing speed test before the first day and following the last day of each condition as explained in Sect. 3.2 and visible in Table 1. All participants were required to take a 45 minute break after four hours. All together, the duration of the whole workday was 8 hours 45 minutes. At the end of each week participants were interviewed and all data collected during the week was secured.

3.6 Results

We used a three-way repeated measures analysis of variance (ANOVA) to analyze the collected data. Non-normal data was log-transformed (heart rate, break time, typing speed) and for the subjective feedback from questionnaires we used the Aligned Rank Transform [60] before conducting ANOVA (task load, system usability, flow, perceived productivity, frustration, presence, positive/negative affect, wellbeing, anxiety, simulator sickness, visual fatigue). For multiple comparisons in post hoc tests, we used Wilcoxon signed-rank test for ART data and t-test otherwise, both with Bonferroni adjustments at an initial significance level of $\alpha = 0.05$. As already mentioned, the number of independent variables and the number of levels of each independent variable differ between the measures. The results for the measures are presented in the following sections. We only display main effects of the ENVIRONMENT and interaction effects involving the ENVIRONMENT, as our focus was on exploring the differences between VR and PHYSICAL and not general variances over time. Therefore, no main effects of DAY or TIME are reported in the paper. We provide a more extensive analysis of other main and interaction effects in the supplementary material. In addition to examining interaction effects of DAY and ENVIRONMENT, we compared the slopes of a fitted line through all days between VR and PHYSICAL by using a one-sided t-test. We chose a one-sided t-test, because we wanted to know if the slopes for VR are significantly higher, as we hypothesize that VR changes are greater, as participants are getting used to a relatively new system while the PHYSICAL environment is familiar from the start. We only report this, for the measures with interaction effects between ENVIRONMENT and DAY (negative affect, anxiety, simulator sickness, visual fatigue). Due to data logging errors, we had to remove some participants from the analysis of several measures (as indicated below). Data logging errors occurred due to the following reasons: 1) the website providing the questionnaires was once not reachable; 2) participants answered the wrong questionnaire-set four times (after 8 hours they did the questionnaire meant for after 6 hours, which did not include positive/negative affect); 3) one participant skipped the first questionnaire once; 4) Polar's heart rate logging application failed to properly sync data for 6 participants on at least one day. As data logging errors rarely occurred for questionnaires (5 times), we do not believe it affected the results.

Task Load: Over the whole week, VR induced a significantly higher taskload ($m = 46.48$, $sd = 2.64$) compared to PHYSICAL ($m = 34.37$, $sd = 1.55$). The mean task load for each DAY and TIME in both weeks is displayed in Fig. 3. The ANOVA results are displayed in Table 2. Due to logging errors, one participant had missing data and was therefore excluded from the analysis. There were no interaction effects between DAY and ENVIRONMENT. This result indicates that participants experienced a significantly higher perceived workload when working in VR than in the comparable physical setup.

System Usability: Over the whole week, PHYSICAL resulted in a significantly higher system usability ($m = 73.88$, $sd = 1.49$) compared to VR ($m = 54.71$, $sd = 1.32$). The mean system usability for each DAY and TIME in both weeks is displayed in figure Fig. 3. The ANOVA results are displayed in Table 2. There were no interaction effects between DAY and ENVIRONMENT. We can conclude from this result that participants found the VR working arrangement far less usable than the comparable physical setup.

Flow: Over the whole week, PHYSICAL resulted in a significantly higher flow ($m = 4.76$, $sd = 0.15$) compared to VR ($m = 4.11$, $sd = 0.18$). The mean flow score for each DAY and TIME in both weeks is displayed in Fig. 3. The ANOVA results are displayed in Table 2. Due to logging errors, we lost the data of three participants for this measure. There were no interaction effects between DAY and ENVIRONMENT. This result suggests that working in VR did not support participants' focus and sense of active engagement in their work activity in a better way compared to PHYSICAL.

Perceived Productivity: Over the whole week, PHYSICAL induced a higher level of perceived productivity ($m = 4.89$, $sd = 0.23$) compared to VR ($m = 4.11$, $sd = 0.28$). The mean productivity scores for each DAY and TIME in both weeks are displayed in Fig. 3. The ANOVA results are displayed in Table 2. In addition, an interaction effect between TIME and ENVIRONMENT was detected, but post hoc tests, comparing VR with PHYSICAL indicated that VR resulted in lower perceived productivity for both times (MORNING, AFTERNOON). This result suggests that working in VR did lead to a significant decrease in perceived productivity compared to PHYSICAL.

Frustration: Over the whole week, VR resulted in a significantly higher score for frustration ($m = 3.49$, $sd = 0.34$) compared to PHYSICAL ($m = 2.45$, $sd = 0.26$). The mean frustration scores for each DAY and TIME in both weeks are displayed in Fig. 3. The ANOVA results are displayed in Table 2. There were no interaction effects between DAY and ENVIRONMENT. This result suggests that working in VR did lead to a significant increase in frustration compared to PHYSICAL.

Presence: Questions about presence only make sense in the VR condition. Therefore, this measure is merely descriptive. The mean presence scores for each DAY and TIME are displayed in Fig. 3. The presence score ranges from 0 to 6, with 6 being the maximal amount of presence perceived by participants. Among all participants over the whole week the mean total presence score was 3.06 ($sd = 1.15$). The sub-scores for spatial presence was 3.66 ($sd = 1.4$), for involvement 2.39 ($sd = 1.01$) and for experienced realism 2.49 ($sd = 1.21$).

Positive/Negative Affect: Over the whole week, VR resulted in a significantly higher negative affect ($m = 11.97$, $sd = 1.04$) compared to PHYSICAL ($m = 11.11$, $sd = 0.52$). No such effect could be detected for positive affect. The analysis also indicated a significant interaction effect between ENVIRONMENT and DAY on negative affect, which was not confirmed in post hoc tests. The mean scores for positive and negative affect for each DAY and TIME for both weeks are displayed

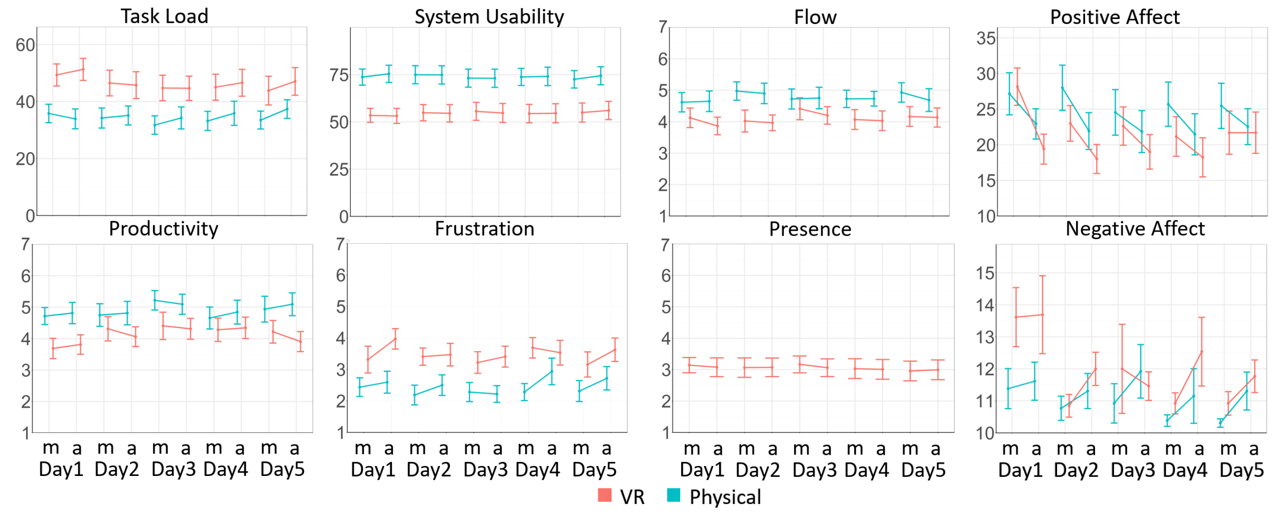

Fig. 3. Average values and standard error for subjective measures in the morning (m) and afternoon (a).

Table 2. RM-ANOVA results for subjective measures. df_1 = df_{effect} and df_2 = df_{error}.

	Task Load					System Usability					Flow				
	df_1	df_2	F	p	η_p^2	df_1	df_2	F	p	η_p^2	df_1	df_2	F	p	η_p^2
Environment	1	14	12.03	.003	.46	1	15	21.14	$<.001$	.58	1	12	7.72	.02	.39
Environment*Day	4	56	2.04	.10	.13	4	60	.75	.57	.05	4	48	1.37	.26	.10
Environment*Time	1	14	.08	.78	.01	1	15	.01	.91	$<.001$	1	12	.33	.57	.03

	Perceived Productivity					Frustration					Positive Affect					Negative Affect				
	df_1	df_2	F	p	η_p^2	df_1	df_2	F	p	η_p^2	df_1	df_2	F	p	η_p^2	df_1	df_2	F	p	η_p^2
Environment	1	15	1.01	.01	.46	1	15	11.70	.003	.44	1	12	2.14	.17	.15	1	12	14.44	.003	.55
Environment*Day	4	60	1.16	.34	.07	4	60	.19	.94	.01	4	48	.83	.52	.06	4	48	4.11	.006	.25
Environment*Time	1	15	6.96	.02	.32	1	15	.02	.88	.001	1	12	.05	.82	.004	1	12	3.15	.10	.21

in Fig. 3. Due to data collection errors 4 data points of 3 different participants are missing, so these participants had to be removed to conduct the ANOVA. The ANOVA results are displayed in Table 2. There were no significant differences in the trendline slopes for negative affect over days ($p = 0.082$). These findings indicate that working in VR was more detrimental to participants' moods than working in the physical setup.

Wellbeing: Over the whole week, PHYSICAL resulted in a significantly higher wellbeing ($m = 5.31$, $sd = 0.34$) compared to VR ($m = 4.25$, $sd = 0.59$). The mean scores for wellbeing for each DAY and TIME in both weeks are displayed in Fig. 4. The ANOVA results are displayed in Table 3. Due to logging errors, data of one participant is missing. There were no interaction effects between DAY and ENVIRONMENT. However, an interaction effect between the ENVIRONMENT and TIME was detected. Post hoc tests, comparing the VR to PHYSICAL condition for each of the three times (START, MORNING, AFTERNOON) indicated that the VR condition results in a significantly lower wellbeing in the MORNING ($V = 10$, $p = 0.015$, $r = 0.71$, $mean_{VR} = 4.07$, $sd_{VR} = 1.4$, $mean_{Physical} = 5.26$, $sd_{Physical} = 1.42$) and AFTERNOON ($V = 11$, $p = 0.01$, $r = 0.74$, $mean_{VR} = 3.91$, $sd_{VR} = 1.2$, $mean_{Physical} = 5.15$, $sd_{Physical} = 1.27$), but not at START.

Anxiety: Over the whole week, VR resulted in a significantly higher anxiety ($m = 5.3$, $sd = 5.84$) compared to PHYSICAL ($m = 2.42$, $sd = 5.34$). The mean scores for anxiety for each DAY and TIME in both weeks are displayed in Fig. 4. The ANOVA results are displayed in Table 3. There were also interaction effects between ENVIRONMENT and both DAY and TIME, which were not confirmed in post-hoc tests. Also, there were no significant differences in the trendline slopes over days ($p = 0.21$). The findings suggest that working in VR elevated participants' feelings of anxiety.

Simulator Sickness: Over the whole week, simulator sickness scores were significantly higher in VR ($m = 34.3$, $sd = 10.16$) compared to PHYSICAL ($m = 9.21$, $sd = 4.47$). According to Stanney et al. [54] these symptoms in VR can be considered bad. The mean scores for each DAY and TIME for both weeks are displayed in Fig. 4. The ANOVA results are displayed in Table 3. Interaction effects between the ENVIRONMENT and both DAY and TIME were detected. Post hoc tests revealed a significant difference between VR and PHYSICAL condition on all days and for all three time periods. Also, there were no significant differences in the trendline slopes over days ($p = 0.078$). These findings suggest that VR leads to substantial simulator sickness symptoms over the course of the week.

Visual Fatigue: Over the whole week, VR ($m = 1.61$, $sd = 0.22$) resulted in a significantly higher visual fatigue than PHYSICAL ($m = 1.09$, $sd = 0.05$). The mean visual fatigue scores for each DAY and TIME in both weeks are displayed in Fig. 4. The ANOVA results are displayed in Table 3. Additionally, interaction effects between all variables were found. Post hoc tests revealed a significant difference between VR and PHYSICAL on every day and for all time periods (START, MORNING and AFTERNOON). Comparing the slopes of lines fitted through the mean ratings for each day, we found that in VR ($m = -0.069$, $sd = 0.13$) visual fatigue decreased at a significantly higher rate than in PHYSICAL ($m = -0.01$, $sd = 0.03$) during the course of the week ($p = 0.04$, $Cohen's\ d = -0.47$). This suggests that while visual fatigue was substantially higher in VR, it also decreased significantly faster that PHYSICAL.

Heart Rate: Statistical tests showed no significant influence of ENVIRONMENT on heart rate. The mean heart rate among all participants for each DAY and TIME for both weeks are displayed in Fig. 5. The ANOVA results are displayed in Table 4. There was a significant interaction effect between DAY and ENVIRONMENT. It can be seen in Fig. 5 that on DAY4 and DAY5 the average heart rate is higher in

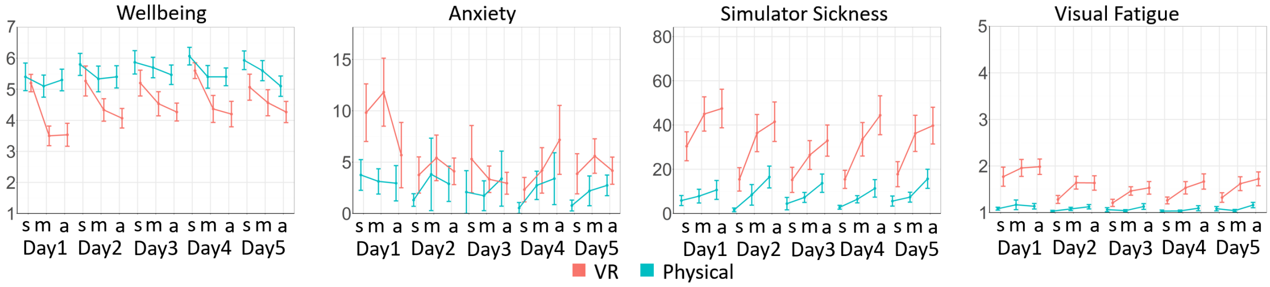

Fig. 4. Average values and standard error for subjective measures at the start of the day (s), in the morning (m) and afternoon (a).

Table 3. RM-ANOVA results for Wellbeing, Anxiety, Simulator Sickness and Visual Fatigue. $df_1 = df_{effect}$ and $df_2 = df_{error}$.

	Wellbeing					Anxiety					Simulator Sickness					Visual Fatigue				
	df_1	df_2	F	p	η_p^2	df_1	df_2	F	p	η_p^2	df_1	df_2	F	p	η_p^2	df_1	df_2	F	p	η_p^2
Environment	1	14	13.34	.002	.49	1	15	20.35	$< .001$	.58	1	15	24.34	$< .001$	.62	1	15	26.30	$< .001$	.64
Environment*Day	4	56	.70	.59	.05	4	60	5.98	$< .001$	.28	4	60	10.32	$< .001$	.41	4	60	12.98	$< .001$	.46
Environment*Time	2	28	5.70	.008	.29	2	30	4.07	.03	.21	2	30	19.06	$< .001$	.56	2	30	27.10	$< .001$	.64

PHYSICAL condition while it is higher in VR for the first three days. Post hoc tests, however, could not identify significant differences. Due to logging errors we lost the data of 6 participants.

Break Times: No significant effect of ENVIRONMENT could be detected on the number of breaks. However, the average duration of a break was significantly higher in VR ($m = 617.05s$, $sd = 137.42$) compared to PHYSICAL ($m = 442.121s$, $sd = 98.1$). The mean number of breaks and the average duration are displayed in Fig. 5. The ANOVA results are displayed in Table 4.

For acquiring break times, we used the videos and head-/marker-tracking algorithms, and considered it a break if no head/marker could be detected. We were interested only in actual breaks that participants used to rest. Hence, in this analysis, we only consider a break if in VR the participant took off the headset for more than 30 seconds and in PHYSICAL if the participant turned away from the screen for more than 30 seconds. We do not consider shorter breaks, since the videos indicated, that these are mainly due to participants quickly turning around, adjusting the headset, picking something up or behaving in a way that results in tracking being lost. For all break times generated by the tracking algorithms, we manually verified them. Additionally, we compared the resulting break times with the ManicTime data, which logged all time frames where the user was inactive for more than 10 minutes, so we do not miss any major breaks, which might have not been detected by the tracking algorithms (e.g., a user facing the camera while using a smartphone instead of the work PC).

Typing Speed: Both ENVIRONMENT and DAY had an influence on typing speed such that PHYSICAL condition resulted in a significantly faster typing speed ($m = 46.88$, $sd = 20.92$) than VR ($m = 43.09$, $sd = 23.98$). On DAY5 ($m = 46.97$, $sd = 21.42$) participants were typing significantly faster than on DAY1 ($m = 43.0$, $sd = 23.52$) in both conditions. The RM-ANOVA results are displayed in Table 4 and the means among all participants are displayed in Fig. 5. There were no interaction effects between DAY and ENVIRONMENT. These results are in line with prior work suggesting a mild performance drop when typing with physical keyboards in VR [20].

Interviews: At the end of each week, we gained additional information from the participants through an interview in which we asked them about how they felt during the week, what they liked or disliked, which problems occurred, and what they would improve. At the end of the second week, they were also asked which ENVIRONMENT they preferred and if they could imagine using VR for work in the future.

For the VR condition, 11 participants disliked the comfort or proposed to increase it (P01-P03, P05, P07-P11, P13, P15). Major issues mentioned include the weight of the HMD and its pressure against the face. P03 and P05 mentioned that the peripheral view was not satisfactory, so they had to move their heads more often. P03 and P05 also pointed out that they needed to take off their headsets for drinking or eating because they were afraid to spill something, and P04 said that such tasks were harder with the HMD on. P05 and P13 missed the ability to write something down on paper. Participants also mentioned technical details that could be improved. For example, removing the headset sometimes made it necessary to reset the position of the virtual screen (using a menu in the Oculus system software) or the remote connection (P03, P04, P13). In addition, hand gestures were sometimes falsely recognized while typing (P06, P02, P12), resulting in an involuntary selection action. Also, four participants mentioned the tracking of the keyboard could be improved (P06, P09, P13, P11).

Three participants mentioned that the study was too long (P07, P09, P10). Seven participant said they felt tired during VR condition (P07, P09, P10, P13 - P16) and only two (P09, P13) during PHYSICAL condition. However, the main reason for this seemed to be deviation from their normal schedules. For example, P13 said "Maybe I was more tired than usual, because I usually do not work for that much time continuously". Five participants (P04, P06, P12, P13, P16) mentioned that they got used to wearing the HMD during the VR condition; however, P06 and P09 also mentioned that the second half of the day was usually harder. P13 mentioned "I did more work than I usually do," while P12 felt that he "was not as productive, because of the low resolution and keyboard". P03 revealed that he "had a blurry vision when driving home on the first day". Seven participants (P03, P04, P06 - P08, P13, P16) mentioned that they felt "alright" during the PHYSICAL condition, while only P08 explicitly stated that she felt alright in the VR condition.

Nine participants (P01, P03, P05, P06, P09, P12 - P15) liked that the isolation in the VR condition allowed them to concentrate more on the tasks at hand, because they were not distracted, especially in combination with music from their private headphones. However, this could also have drawbacks, and as P01, P06 and P08 mentioned, the VR condition was "a bit scary," because they could not see the presence of other people in the real world (see also [38]).

P12 also mentioned that "without [private] music turned on, I was trying to guess what was happening around me". P01 and P13 said that they even forgot that they were wearing HMDs when concentrating hard on their work and P11 mentioned that the experience and movement in VR felt natural. Four participants (P04, P07, P08, P12) specifically mentioned that they liked to try out and experience VR in a work context. P09 liked the privacy that VR offered, as "nobody can see what you are doing". While P10 and P13 liked to relax in the virtual environment, P13 also "liked to look around when taking a break and just looking at empty space". On the other hand, P06, P08 and P12 mentioned that they felt more comfortable seeing the real surrounding and P09 liked to look somewhere else and not on a display when resting.

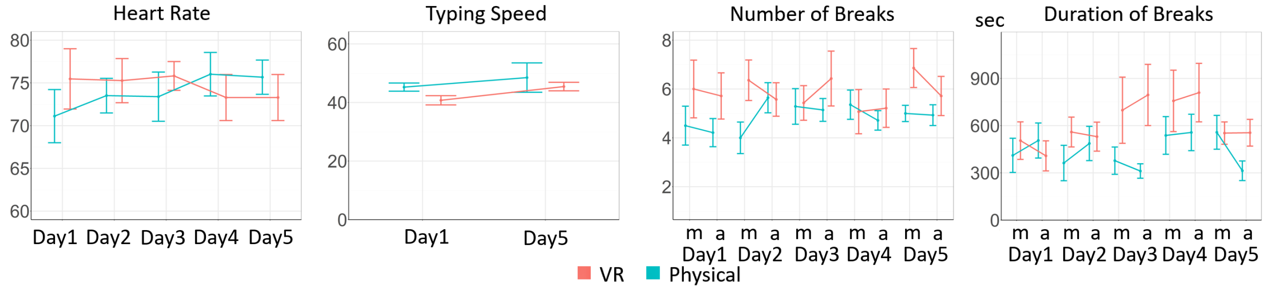

Fig. 5. Average values and standard error for objective measures in the morning (m) and afternoon (a).

Table 4. RM-ANOVA results for breaks. df_1 = df_{effect} and df_2 = df_{error}.

	Number of Breaks					Break Duration					Heart Rate					Typing Speed				
	df_1	df_2	F	p	η_p^2	df_1	df_2	F	p	η_p^2	df_1	df_2	F	p	η_p^2	df_1	df_2	F	p	η_p^2
Environment	1	13	0.87	0.37	0.06	1	13	12.27	.004	.49	1	9	.08	.78	.009	1	15	8.46	.01	.36
Environment*Day	4	52	1.57	0.20	0.11	4	52	2.01	0.11	0.13	4	36	1.77	.16	.17	1	15	.31	.59	.02
Environment*Time	1	13	0.71	0.41	0.05	1	13	< 0.001	0.98	.0										

Only three participants (P01, P11, P13) preferred the VR condition. P01 felt more relaxed in VR, P13 liked the isolation and was able to do more work, while P11 liked it because he already knew the system/study from the PHYSICAL condition week. P02, P05 and P12 preferred the PHYSICAL condition because it was the same as the VR, but without downsides such as the heavy HMD. Others preferred the PHYSICAL condition because it was more familiar (P04, P08, P09), it felt less limiting (P03), it did not require wearing the HMD (P07), it was easier (P10, P16) and it allowed to focus more on work as she "did not have to tackle problems with the system" (P15). Nevertheless, P06 and P05 added that VR was more exciting.

All participants could imagine using VR for work in the future if some conditions are met, such as having lighter HMDs with higher resolution and being able to have multiple displays. Also, all participants mentioned that they could imagine using VR for a limited amount of time (on train rides or for certain tasks). Regarding time, P12 mentioned that "in VR I had 45 minutes of high performance and then 3 hours of headache". P06 suggested that using VR could improve ergonomics, because displays are adjustable, as well as that VR would be good for working at home to separate work from personal life. P03 and P07 mentioned that they prefer not to sit in one place all day and they usually like to walk around and talk to coworkers. P04 would also like to have the possibility to play games with other users in VR during breaks and to have their phone integrated in VR. P16 mentioned that the isolation in VR could hinder collaboration with colleagues. For both conditions, participants mentioned concerns about the keyboard, touchpad, screen resolution and delay induced by the remote desktop. Please note that this is expected and a result of making the two interfaces comparable with available hardware.

A week after the experiment, participants were also asked if they observed any effects after completing the VR week. P04 mentioned that sometimes during the weekend, she felt as if she was still wearing the HMD. P16 and P15 still had a feeling of dry eyes after finishing the use of VR and P15 felt sleepy and dizzy for about 2 hours afterwards. Also, P01 mentioned she felt a swelling of the face around the eyes and her neck and shoulders were stiff. P16 and P06 were "amazed by how detailed the real world is after removing HMD". P02 and P06 felt their skin suffered after wearing the headset for one week. All other participants did not report major effects.

3.7 Dropout

In total, two participants decided to drop out. The first participant who dropped out experienced regular migraines and mentioned that the weight of the headset triggered them. Therefore, this participant dropped out after two blocks of work (four hours). The second participant who dropped out explained that due to the weight of the headset it was not possible to sit in a relaxed position. In addition, this participant mentioned that not having access to the usual setup reduced motivation and productivity. After approximately two hours this participant experienced anxiety and felt nauseated and disoriented, resulting in the participant dropping out from the study.

4 DISCUSSION

In the presented study, we have examined the experience of users working for one work week in VR and a comparable physical setup. To control for various factors, such as screen size, input device and working conditions, our experimental protocol enforced as similar configurations in both conditions as possible, while considering work experiences that are accessible to a wide number of users using today's commercial off-the-shelf VR solutions (an Oculus Quest2 with accompanying Logitech keyboard). Given the limitations of current technology and the fact that VR provides a virtual approximation of the real environment, we did not expect the VR condition to outperform the PHYSICAL condition which is also confirmed by the results. However, the quantified results of the studied VR experience that is comparable to a physical one, can serve as a baseline for future optimized VR systems. In fact, the extent to which some of our measures diverged between VR and PHYSICAL is notable. For example, VR clearly resulted in below average system usability scale ratings, while PHYSICAL resulted in above average ratings, even though both systems used the same input devices and had comparable screen real estates for carrying out work. Similarly, while it is expected that a VR system induces higher simulator sickness ratings than a non-VR system, the absolute values of the SSQ ratings indicate that VR corresponded to the worst category of simulator sickness [54][2]. We find this surprising given that we utilized a setup (Oculus Quest2, Logitech Keyboard), which can be considered widespread among consumers and professionals alike. These high ratings of simulator sickness could be observed throughout the week, even though we note that the SSQ ratings decreased slightly over the week. On the other hand, not all differences are as concerning. While the self-rated task load in VR was approximately 35% higher compared to PHYSICAL condition, it is still within the 50th percentile of ratings of computer activities [17].

We also examined, if there are any differences between touch-typists and non-touch-typists. We divided the participants into two groups based on their need to look at the keyboard while typing which they reported in the demographic questionnaire on a seven-point likert scale. Six participants answered either 1 (never) or 2 and were therefore considered touch-typists. An analysis with TOUCHTYPIST as a between-subjects factor revealed no significant main effects of this variable. This

---

[2] Please note that this categorization is based mostly on military simulators and that researchers discuss about the associated challenges [7].

suggests that the need to look at the keyboard while typing does not have a significantly negative influence on the results.

When fitting a linear model to the data, we observed that any improvements across most measures in both conditions are not significant. Still, examining the development of the scores over the week can serve as an indication of a possible emerging trend in the future. When inspecting the graphs, it is possible to observe a rapid adaptation of users to the VR condition. Within a day or two, many of the scores for VR improved. At this stage, we do not know if this improvement was a result of participants' brains adapting to the new condition or if they overcame the initial expectations that people had previously about VR. Another effect observable in the results was a gradual accumulation of some exhaustion across the week. We observed this effect in some of the measures of VR (specifically, regarding task load, simulator sickness). We also observed such an effect in the PHYSICAL condition, although with slower growth, which may hint that this factor might be independent of the environment and instead more related to the duration of the experiment.

It is clear that there is still a long way to go for the development of more comfortable hardware. There are already HMDs that offer higher resolution displays, faster refresh rates, variable focal distance displays, and wider field of view. We anticipate future HMDs will be available in a form factor similar to conventional glasses, and be lighter and allow the flow of air around users' face. We expect that such hardware will further reduce the gap between VR and the PHYSICAL conditions. Some of the more mundane issues encountered by the participants in the VR condition are relatively straightforward to address. For example, multiple participants complained about the keyboard periodically vanishing from the VR environment, the relative position of the home environment shifting, and hand movements on the keyboard being inadvertently recognized as input gestures. To address this, a dedicated 'work' mode may be appropriate, allowing device tracking and gesture recognition subsystems to operate in a more persistent manner when the user is known to be engaged in seated work.

The comments from participants in the interviews highlight the challenge of implementing an effective and enjoyable VR working experience given the potential influence of personal preference. We note that several participants appreciated how working in VR helped isolate them from their physical workspace and enabled periods of greater focus, separation and privacy. Conversely, other participants had negative experiences due to this isolation, whether due to a feeling of unease produced by an inability to perceive who is nearby or the obstacles the setup presents for face-to-face collaboration. Nevertheless, despite the generally negative experience reported by the majority of participants, all commented that they could imagine using VR for at least some work tasks or at least a portion of the day. This hints to the future when knowledge workers will combine two modes depending on the needs for the work at hand.

4.1 Limitations and Future Work

Conducting a complex in-situ study carried out across an entire working week is inherently intertwined with variables that are outside of our control. Therefore, many of the measures depend on factors that we cannot fully control. For example, reflections on frustration, perceived productivity or ability to concentrate may be influenced by the type of work being performed. There are also a number of other aspects of the study that should be considered when interpreting the findings. First, we discovered that although the overall task load was higher in VR, task load was at a relatively high level in both conditions (but still within the bound of comparable computer work [54]). This is likely a consequence of the limitations of the setup common to both conditions, such as no use of a mouse, and a display set at a relatively low resolution. Second, many of our measures are based on participants' subjective responses. However, as observed by Wille et al. [59], there can be a disconnect between objective physiological effects and subjective user ratings. In terms of eye strain, Wille et al. [59] suggest that this disconnect may be influenced by the level of familiarity a user has with the technology. If true, this may serve to explain some of the reduction seen in some measures over the first few days of the study. Third, our understanding of other factors related to working in VR, and how they impact the user experience, is still emerging. Shen et al. [53] suggest that VR allows for the use of more attention resources. This may help explain the experience of P12, who commented that he experienced a high efficiency for 45 minutes and then a headache for three hours. If VR does indeed allow for the use of more attention resources, steps should be taken to avoid overloading users, particularly when they are still acclimating to working in VR. Finally, we only presented selected analyses of the collected data. Yet, we are planning to investigate some aspects in more detail, such as micro-breaks and and other behavioral patterns that could be detected, by closely analyzing the recorded videos. As we release the collected anonymized data, future work can also investigate other aspects of the data further. Since we employed a widely available commercial hardware and software solution, we would also hope that further researchers could add to the data set by replicating the study.

This paper has studied the effects of working in VR compared to a regular working environment at one particular operating point with experimental parameters set to align as much as possible between both environments, given the constrains of using a commodity off-the-shelf VR system. We hope this work will stimulate further work at different operating points, investigating how some of the quantified effects we observed in this study may possibly change if the VR condition is allowed to deviate from the operating point of a regular working environment, for instance, by providing flexible solutions to allow VR users making maximum use of the available VR space and novel VR interaction techniques to make VR interaction more comfortable.

Additionally, the duration of the study, the need to exercise control over the work environment, and the fact that participants were required to perform their standard work tasks largely restricted feasible recruitment to individuals already embedded within the three different university sites. Such individuals may inadvertently be more forgiving of the deficiencies of the setup. Future work is required to look at how a broader population may experience working in VR.

We also see interesting further work in examining stress and heart rate more closely when working in different VR environments. There are also open issues around the social acceptability of working in VR for a prolonged amount of time, as well as users possibly feeling isolated or having difficulties in collaborating with their colleagues.

5 CONCLUSIONS

In this paper, we have studied the effects of working in VR for an entire workweek. While VR has been repeatedly pitched as providing new exciting possibilities for modern knowledge work, in practice the potential advantages of virtual work environments can only be used if it is feasible to work in a virtual environment for an extended period of time. Prior to this work, there were only limited studies of long-term effects of working in VR. We reported the results from a comparative study with 16 participants working for an entire workweek in both VR and in a baseline physical desktop environment. As a first study of this kind and scale, we deliberately opted to design these conditions to be as similar as possible to allow as many quantitative comparisons as possible. Therefore, the study did not present the participants with the best possible VR system but instead a setup that delivered a comparable experience to working in a physical desktop environment. The study revealed that, as expected, VR resulted in significantly worse ratings across most measures. For example, VR resulted in below average system usability scale ratings while the physical environment resulted in above average ratings. We also found that VR resulted in the worst category of simulator sickness although the severity decreased slightly across the week. However two participants even dropped out on the first VR day, due to migraine, nausea and anxiety. Nevertheless, there was some indication that participants gradually overcame negative first impressions and initial discomfort. Overall, this study helps laying the groundwork for subsequent research, highlighting current shortcomings and identifying opportunities for improving the experience of working in VR. We hope this work will stimulate further research investigating longer-term productive work in-situ in VR.

REFERENCES

[1] Time tracker management tracking software. `https://www.manictime.com/`. Accessed: 2022-05-31.

[2] A. P. Anderson, M. D. Mayer, A. M. Fellows, D. R. Cowan, M. T. Hegel, and J. C. Buckey. Relaxation with immersive natural scenes presented using virtual reality. *Aerospace medicine and human performance*, 88(6):520–526, 2017.

[3] S. Benedetto, V. Drai-Zerbib, M. Pedrotti, G. Tissier, and T. Baccino. E-readers and visual fatigue. *PloS one*, 8(12):e83676, 2013.

[4] M. G. Berman, J. Jonides, and S. Kaplan. The cognitive benefits of interacting with nature. *Psychological science*, 19(12):1207–1212, 2008.

[5] V. Biener, T. Gesslein, D. Schneider, F. Kawala, A. Otte, P. O. Kristensson, M. Pahud, E. Ofek, C. Campos, M. Kljun, K. Č. Pucihar, and G. Jens. Povrpoint: Authoring presentations in mobile virtual reality. *IEEE Trans. on Visualization and Computer Graphics*, 2022. doi: 10.1109/TVCG.2022.3150474

[6] V. Biener, D. Schneider, T. Gesslein, A. Otte, B. Kuth, P. O. Kristensson, E. Ofek, M. Pahud, and J. Grubert. Breaking the screen: Interaction across touchscreen boundaries in virtual reality for mobile knowledge workers. *IEEE Trans. on Visualization and Computer Graphics*, 26(12):3490–3502, 2020. doi: 10.1109/TVCG.2020.3023567

[7] P. Bimberg, T. Weissker, and A. Kulik. On the usage of the simulator sickness questionnaire for virtual reality research. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 464–467. IEEE, 2020.

[8] J. Brooke. Sus: a "quick and dirty'usability. *Usability evaluation in industry*, p. 189, 1996.

[9] S. Brunia and A. Hartjes-Gosselink. Personalization in non-territorial offices: a study of a human need. *Journal of Corporate Real Estate*, 2009.

[10] S. Buckley. Two people spent 48 hours in nonstop virtual reality. `https://www.engadget.com/2017-06-07-two-people-spent-48-hours-in-nonstop-virtual-reality.html`, 2017. Accessed: 2022-05-24.

[11] P. Drucker. The effective executive harper & row. *New York, NY*, 1966.

[12] B. M. Ens, R. Finnegan, and P. P. Irani. The personal cockpit: a spatial interface for effective task switching on head-worn displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 3171–3180, 2014.

[13] M. Funk, A. Bächler, L. Bächler, T. Kosch, T. Heidenreich, and A. Schmidt. Working with augmented reality? a long-term analysis of in-situ instructions at the assembly workplace. In *Proceedings of the 10th International Conference on PErvasive Technologies Related to Assistive Environments*, pp. 222–229, 2017.

[14] S. Garrido-Jurado, R. Muñoz-Salinas, F. J. Madrid-Cuevas, and M. J. Marín-Jiménez. Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition*, 47(6):2280–2292, 2014.

[15] T. Gesslein, V. Biener, P. Gagel, D. Schneider, P. O. Kristensson, E. Ofek, M. Pahud, and J. Grubert. Pen-based interaction with spreadsheets in mobile virtual reality. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 361–373. IEEE, 2020.

[16] C. V. C. GmbH. 10fastfingers.com - typing test, competitions, practice & typing games. `https://10fastfingers.com/`. Accessed: 2022-05-31.

[17] R. A. Grier. How high is high? a meta-analysis of nasa-tlx global workload scores. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 59, pp. 1727–1731. SAGE Publications Sage CA: Los Angeles, CA, 2015.

[18] J. Grubert, D. Hamacher, R. Mecke, I. Böckelmann, L. Schega, A. Huckauf, M. Urbina, M. Schenk, F. Doil, and J. Tümler. Extended investigations of user-related issues in mobile industrial ar. In *2010 IEEE International Symposium on Mixed and Augmented Reality*, pp. 229–230. IEEE, 2010.

[19] J. Grubert, E. Ofek, M. Pahud, and P. O. Kristensson. The office of the future: Virtual, portable, and global. *IEEE computer graphics and applications*, 38(6):125–133, 2018.

[20] J. Grubert, L. Witzani, E. Ofek, M. Pahud, M. Kranz, and P. O. Kristensson. Text entry in immersive head-mounted display-based virtual reality using standard keyboards. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 159–166. IEEE, 2018.

[21] J. Guo, D. Weng, H. Fang, Z. Zhang, J. Ping, Y. Liu, and Y. Wang. Exploring the differences of visual discomfort caused by long-term immersion between virtual environments and physical environments. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 443–452. IEEE, 2020.

[22] J. Guo, D. Weng, Z. Zhang, H. Jiang, Y. Liu, Y. Wang, and H. B.-L. Duh. Mixed reality office system based on maslow's hierarchy of needs: Towards the long-term immersion in virtual environments. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 224–235. IEEE, 2019.

[23] J. Guo, D. Weng, Z. Zhang, Y. Liu, and Y. Wang. Evaluation of maslows hierarchy of needs on long-term use of hmds–a case study of office environment. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 948–949. IEEE, 2019.

[24] S. G. Hart and L. E. Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, vol. 52, pp. 139–183. Elsevier, 1988.

[25] H. Heuer, G. Hollendiek, H. Kröger, and T. Römer. Rest position of the eyes and its effect on viewing distance and visual fatigue in computer display work. *Zeitschrift fur experimentelle und angewandte Psychologie*, 36(4):538–566, 1989.

[26] C. J. Jordan and A. A. Palmer. Virtual meetings: A critical step to address climate change, 2020.

[27] S. Kaplan. The restorative benefits of nature: Toward an integrative framework. *Journal of environmental psychology*, 15(3):169–182, 1995. doi: 10.1016/0272-4944(95)90001-2

[28] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The int. journal of aviation psychology*, 3(3):203–220, 1993.

[29] J. Kim, C. Candido, L. Thomas, and R. de Dear. Desk ownership in the workplace: The effect of non-territorial working on employee workplace satisfaction, perceived productivity and health. *Building and Environment*, 103:203–214, 2016.

[30] B. Klemens. I spent hundreds of hours working in vr. here's what i learned. `https://www.wired.com/story/hours-working-vr-tips/`, 2022. Accessed: 2022-05-24.

[31] P. Knierim, T. Kosch, and A. Schmidt. The nomadic office: A location independent workspace through mixed reality. *IEEE Pervasive Computing*, 20(4):71–78, 2021.

[32] P. Knierim and A. Schmidt. The virtual office of the future: Are centralized workplaces obsolete? 2020.

[33] H. Lee, S. Je, R. Kim, H. Verma, H. Alavi, and A. Bianchi. Partitioning open-plan workspaces via augmented reality. *Personal and Ubiquitous Computing*, pp. 1–16, 2019.

[34] J. Li, C. George, A. Ngao, K. Holländer, S. Mayer, and A. Butz. Rear-seat productivity in virtual reality: Investigating vr interaction in the confined space of a car. *Multimodal Technologies and Interaction*, 5(4):15, 2021.

[35] G. I. Limited. Chrome remote desktop. `https://remotedesktop.google.com/`. Accessed: 2022-05-31.

[36] G. LLC. Face mesh - mediapipe. `https://google.github.io/mediapipe/solutions/face_mesh.html`. Accessed: 2022-05-31.

[37] F. Lu and D. A. Bowman. Evaluating the potential of glanceable ar interfaces for authentic everyday uses. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, pp. 768–777. IEEE, 2021.

[38] M. McGill, D. Boland, R. Murray-Smith, and S. Brewster. A dose of reality: Overcoming usability challenges in vr head-mounted displays. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pp. 2143–2152, 2015.

[39] M. Mcgill, A. Kehoe, E. Freeman, and S. Brewster. Expanding the bounds of seated virtual workspaces. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 27(3):1–40, 2020.

[40] T. Menzner, T. Gesslein, A. Otte, and J. Grubert. Above surface interaction for multiscale navigation in mobile virtual reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 372–381. IEEE, 2020.

[41] F. Mostajeran, J. Krzikawski, F. Steinicke, and S. Kühn. Effects of exposure to immersive videos and photo slideshows of forest and urban environments. *Scientific Reports*, 11(1):1–14, 2021.

[42] A. Ng, D. Medeiros, M. McGill, J. Williamson, and S. Brewster. The passenger experience of mixed reality virtual display layouts in airplane environments. In *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 265–274. IEEE, 2021.

[43] R. Nordahl, N. C. Nilsson, A. Adjorlu, E. Magalhaes, S. Willemsen, N. S. Andersson, J. Wang, and S. Serafin. 12 hours in virtual reality: Two cases of long-term exposure to consumer-grade virtual reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 2019.

[44] E. Ofek, J. Grubert, M. Pahud, M. Phillips, and P. O. Kristensson. Towards a practical virtual office for mobile knowledge workers. *arXiv preprint*

arXiv:2009.02947, 2020.

[45] L. Pavanatto, C. North, D. A. Bowman, C. Badea, and R. Stoakley. Do we still need physical monitors? an evaluation of the usability of ar virtual monitors for productivity work. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, pp. 759–767. IEEE, 2021.

[46] T. Pejsa, J. Kantor, H. Benko, E. Ofek, and A. Wilson. Room2room: Enabling life-size telepresence in a projected augmented reality environment. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW '16)*, 2016.

[47] D.-M. Pham and W. Stuerzlinger. Hawkey: Efficient and versatile text entry for virtual reality. In *25th ACM Symposium on Virtual Reality Software and Technology*, pp. 1–11, 2019.

[48] J. Pretsch, E. Pretsch, J. Saretzki, H. Kraus, and G. Grossmann. Improving employee well-being by means of virtual reality–realex: an empirical case study. *European Journal of Economics and Business Studies*, 6(1):95–105, 2020.

[49] F. Rheinberg, R. Vollmeyer, and S. Engeser. Die erfassung des flowerlebens. 2003.

[50] A. Ruvimova, J. Kim, T. Fritz, M. Hancock, and D. C. Shepherd. "transport me away": Fostering flow in open offices through virtual reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–14, 2020.

[51] D. Schneider, A. Otte, T. Gesslein, P. Gagel, B. Kuth, M. S. Damlakhi, O. Dietz, E. Ofek, M. Pahud, P. O. Kristensson, et al. Reconviguration: Reconfiguring physical keyboards in virtual reality. *IEEE transactions on visualization and computer graphics*, 25(11):3190–3201, 2019.

[52] T. W. Schubert. The sense of presence in virtual environments: A three-component scale measuring spatial presence, involvement, and realness. *Z. für Medienpsychologie*, 15(2):69–71, 2003.

[53] R. Shen, D. Weng, S. Chen, J. Guo, and H. Fang. Mental fatigue of long-term office tasks in virtual environment. In *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 124–127. IEEE, 2019.

[54] K. M. Stanney, R. S. Kennedy, and J. M. Drexler. Cybersickness is not simulator sickness. In *Proceedings of the Human Factors and Ergonomics Society annual meeting*, vol. 41, pp. 1138–1142. SAGE Publications Sage CA: Los Angeles, CA, 1997.

[55] F. Steinicke and G. Bruder. A self-experimentation report about long-term use of fully-immersive technology. In *Proceedings of the 2nd ACM symposium on Spatial user interaction*, pp. 66–69, 2014.

[56] K. D. Thoondee and A. Oikonomou. Using virtual reality to reduce stress at work. In *2017 Computing Conference*, pp. 492–499. IEEE, 2017.

[57] D. Valtchanov, K. R. Barton, and C. Ellard. Restorative effects of virtual nature settings. *Cyberpsychology, Behavior, and Social Networking*, 13(5):503–512, 2010.

[58] D. Watson, L. A. Clark, and A. Tellegen. Development and validation of brief measures of positive and negative affect: the panas scales. *Journal of personality and social psychology*, 54(6):1063, 1988.

[59] M. Wille, L. Adolph, B. Grauel, S. Wischniewski, S. Theis, and T. Alexander. Prolonged work with head mounted displays. In *Proceedings of the 2014 ACM International Symposium on Wearable Computers: Adjunct Program*, pp. 221–224, 2014.

[60] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proc. of the SIGCHI Conf. on human factors in computing systems*, pp. 143–146, 2011.

[61] A. N. Zsido, S. A. Teleki, K. Csokasi, S. Rozsa, and S. A. Bandi. Development of the short version of the spielberger state—trait anxiety inventory. *Psychiatry research*, 291:113223, 2020.

Chapter 7

Hold Tight: Identifying Behavioral Patterns During Prolonged Work in VR through Video Analysis

Summary

The previous chapter reported on a study in which participants were working in VR and a comparable physical setup for a whole workweek. The measures reported in that chapter showed that VR performed generally worse than the physical setup. However, detailed insights into how participants behaved during that week and how they might have coped with the difficulties of VR were missing. Therefore, this article provides a closer look at the video recordings from that study, which were obtained through a webcam filming the participants upper body during the whole study. Six annotators watched four entire work days for each of the 16 participants (the first, third and fifth day of VR and the first day of physical) and annotated their behavior.

The results indicate that participants adapted to certain restrictions of the HWD because they mostly did not take off the HWD for talking or using their phone. Also, throughout the week the frequency of adjusting or supporting the HWD reduced, which could indicate that participants get used to it. On the other hand, they took it off for longer periods towards the end of the week and actions such as eating, drinking, or interacting with physical objects were less frequent in VR compared to the physical setup, indicating that VR can

be disruptive to normal behavioral patterns. These unique results can inform the design of less restricting and possibly more ergonomic XR systems that can also be used for extended periods of time.

Contribution Statement

In an iterative process Verena Biener and Forouzan Farzinnejad created a codebook for labeling the videos. Verena Biener and Forouzan Farzinnejad together with Rinaldo Schuster, Seyedmasih Tabaei, Leon Lindlein and Jinghui Hu labelled all the videos. Verena Biener processed the data gained through the labels, executed the statistical analysis and created visualizations of the results. Negar Nouri created visualizations of the labeled actions. Verena Biener, John J. Dudley, Per Ola Kristensson and Jens Grubert discussed and interpreted the results. An initial draft of the paper was written by Verena Biener which was subsequently refined by all authors.

Article [13]

Hold Tight: Identifying Behavioral Patterns During Prolonged Work in VR through Video Analysis

Verena Biener[1]* Forouzan Farzinnejad[1] Rinaldo Schuster[1] Seyedmasih Tabaei[1] Leon Lindlein[1]
Jinghui Hu[2] Negar Nouri[1] John J. Dudley[2] Per Ola Kristensson[2] Jörg Müller[3] Jens Grubert[1]†

[1]Coburg University of Applied Sciences and Arts, Germany
[2]University of Cambridge, United Kingdom
[3]University of Bayreuth, Germany

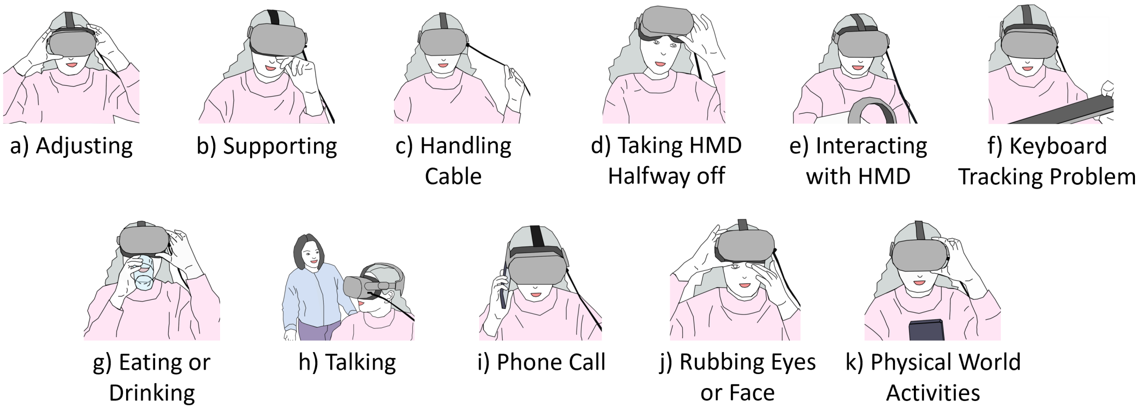
a) Adjusting
b) Supporting
c) Handling Cable
d) Taking HMD Halfway off
e) Interacting with HMD
f) Keyboard Tracking Problem
g) Eating or Drinking
h) Talking
i) Phone Call
j) Rubbing Eyes or Face
k) Physical World Activities

Fig. 1: Categories of actions observed in the video-analysis: (a) adjusting the HMD; (b) supporting the HMD; (c) handling the HMD's power cable; (d) taking the HMD halfway off, by lifting it up a bit; (e) interacting wit the HMD through controllers or gestures; (f) encountering keyboard tracking problems; (g) eating or drinking something; (h) talking to another physically present person; (i) having a phone call; (j) rubbing eyes or face; (k) doing physical world activities such as using a phone or reading from paper.

Abstract—VR devices have recently been actively promoted as tools for knowledge workers and prior work has demonstrated that VR can support some knowledge worker tasks. However, only a few studies have explored the effects of prolonged use of VR such as a study observing 16 participants working in VR and a physical environment for one work-week each and reporting mainly on subjective feedback. As a nuanced understanding of participants' behavior in VR and how it evolves over time is still missing, we report on the results from an analysis of 559 hours of video material obtained in this prior study. Among other findings, we report that (1) the frequency of actions related to adjusting the headset reduced by 46% and the frequency of actions related to supporting the headset reduced by 42% over the five days; (2) the HMD was removed 31% less frequently over the five days but for 41% longer periods; (3) wearing an HMD is disruptive to normal patterns of eating and drinking, but not to social interactions, such as talking. The combined findings in this work demonstrate the value of long-term studies of deployed VR systems and can be used to inform the design of better, more ergonomic VR systems as tools for knowledge workers.

Index Terms—virtual reality, video-analysis, productivity work, long-term, prolonged use, office work, future of work

◆

1 INTRODUCTION

Virtual Reality (VR) has already gained popularity in the entertainment domain, but it has also been explored in recent years as a tool for knowledge work [4,6,46]. VR can provide various advantages for improving work experiences, such as enhancing interactivity [4], adapting work-environments [46] or utilizing large virtual displays [6,35]. Still, the prospect of wearing current-generation VR headsets for a prolonged period of time, such as a full workday or even a whole workweek, could be off-putting due to the current state of VR systems. When using VR applications intended for entertainment this may be less of an issue as users might be distracted from the less-than-optimal hardware by an engaging virtual experience. However, this might not be the case for knowledge workers who need to wear HMDs for prolonged periods of time. This motivates research with the objective of understanding how knowledge workers respond to the prolonged use of VR.

To this end, Biener et al. [5] conducted a study in which participants completed a full workweek in VR and compared it to a week in a comparable physical setup. They found that the VR condition resulted in significantly worse ratings for measures of task load, frustration, negative affect, anxiety, eye strain, system usability, flow, productivity, well-being and simulator sickness. Nevertheless, some of the reported measures improved slightly in the course of the five days. While conducting this study, Biener et al. [5] generated a dataset of over 1,400 hours of video material, capturing the participants' behavior in both conditions throughout the duration of both workweeks. However, this data was not analyzed in the original paper. We have obtained access to this dataset of the original study and present an extensive analysis of the user behaviors exhibited in this rich video data.

We report on the behavior of 16 participants as they respond to the

---

*e-mail: verena.biener@hs-coburg.de
†e-mail: jens.grubert@hs-coburg.de

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

experience of working in VR over a workweek, by sampling the first, third, and last day. Where possible, we also compare this behavior with the observed behavior of participants in the non-VR condition. This resulted in a total of 559 hours of analyzed video data. To our knowledge, this is the first paper reporting on such a comprehensive video study of VR work and it provides substantial and detailed insights into the behavior of users in a VR work setting. It is covering much longer time periods in comparison to typical VR studies, and, therefore, also allows us to describe how participants' behavior changes over time.

Common behaviors that we observed include: adjusting, supporting, or removing the headset, standing, eating and drinking, using a phone, or interacting with other people. Key insights are: over the five days (1) the frequency of actions related to adjusting the headset was reduced by 46% and the duration by 50%; (2) the frequency of actions related to supporting the headset was reduced by 42%; (3) the HMD was removed 31% less frequently but for 41% longer periods; (4) wearing an HMD is disruptive to normal patterns of eating and drinking, but not to social interactions, such as talking.

These observations reveal a pattern of accommodation, adaptation, and appropriation to a novel working setup. Viewed in concert with the findings of Biener et al. [5], however, these adaptations were insufficient to bridge the gap between the VR and physical work setups. Our findings thereby provide evidence and motivation for a broader consideration of usability and ergonomic issues encountered by wearers of HMDs under extended use, which so far has been insufficiently explored in the literature. The insights presented in this paper can thus inform both: (1) the physical design of HMDs to improve comfort and simplify adjustment; and (2) the development of ergonomic guidance for workers tasked to wear an HMD for any extended period of time.

2 RELATED WORK

VR has several possible advantages and challenges as a tool for knowledge work which are discussed in Section 2.1. However, the studies investigating the benefits of VR are usually quite short. Therefore, Section 2.2 reviews previous work about using VR for longer periods of time. Then, Section 2.3 reports on prior studies using video-analysis to uncover the behavior of participants.

2.1 Knowledge Work in VR

For more than a decade VR has been used in industry to support work in various contexts [3]. Recent VR HMDs have been promoted specifically as a tool for knowledge workers (e.g. [42]) and previous studies have also shown that VR has the potential to support knowledge work. Biener et al. [4, 6] have shown that multimodal interaction, combining touch and eye-tracking can be used to efficiently navigate multiple screens and that three-dimensional visualizations can facilitate tasks involving multiple layers of information. In combination with a spatially tracked pen, VR has been shown to facilitate certain tasks in spreadsheet applications [12], or when authoring presentations [4]. Other work has investigated how to use a 2D mouse to interact with 3D content in VR [59]. Using eye-gaze and blink, Meng et al. [37] explored techniques for hands-free text-selection in VR, and Lee et al. [29] proposed a technique that uses gaze to make reading in VR more efficient and less demanding.

In addition to novel interaction techniques, VR makes it possible to have any number of displays which could be especially helpful in a mobile context where the screen space of conventional devices is limited. However, after comparing physical and virtual monitors in AR, Pavanatto et al. [44] suggested to use a mix of both, due to the current limitations of AR devices such as low resolution. McGill et al. [35] suggested to manipulate the virtual display position using the users gaze-direction to use a large display space with less head movement. Even though virtual displays can be beneficial in settings with limited space, Ng et al. [39] found that passengers in an airplane preferred to limit virtual displays to their personal seating area. Similarly, Medeiros et al. [36] found that users avoid placing displays at the location of other passengers in a public transportation scenario, but that they use them to shield themselves from others.

VR can also address privacy issues when used in a public space, for example by randomizing the layout of a physical keyboard when entering a password [47], or by the fact that only the person wearing the HMD can see the content on the virtual screens [16].

In addition, VR can be used to reduce distractions and stress. Ruvimova et al. [46] reported that a virtual beach environment when sitting in an open office environment can reduce distractions and induce flow. Similarly, Lee et al. [30] showed that using AR to add visual separators in an open office reduces distractions and allows to easily personalize the work environment. Thoondee et al. [54] reported that participants who were experiencing a VR relaxation environment were more relaxed. It has also been indicated that experiencing nature in VR can reduce stress better than watching a video on a regular display [45], and that interactive nature environments in VR have more positive effects on stress than passive VR experiences [55]. However, even though users prefer natural environments, they perform better in familiar work environments [31].

This prior work demonstrates a range of positive aspects of working in VR, yet the results have been gathered through short-term studies. In addition, a review of previous work about office-like tasks in VR [50] found that VR could induce increased visual fatigue, muscle fatigue, acute stress and mental overload. Therefore, further research is needed on the effects of working in VR, especially for extended periods of time.

VR can reduce distractions [46] or help users relax [54], by shielding them from the physical world. On the other hand, there are also aspects of the physical surroundings that the user should be aware of while working in VR. To overcome this problem, McGill et al. [34] suggested to integrate relevant parts of the physical environment into the virtual, so that users are aware of other people or relevant objects, such as a nearby cup of tea on the desk. To this end, Wang et al. [56] present a customizable physical world view and Hartmann et al. [22] include real-time 3D reconstructions of the physical environment into the VR application. Tao and Lopes [53] show that potential real-world distractions can be integrated in VR to improve presence. Therefore, Simeone et al. [49] used a depth camera to detect and display bystanders' positions. Displaying bystanders is especially helpful, as O'Hagen et al. [43] found that users can feel uncomfortable when knowing that bystanders are present without being aware of their position. Other research focuses on introducing smartphones into the virtual environment by showing a video pass-through of the smartphone and the users hands [1], using a camera stream and screenshots sent from the phone to the VR device [11], or replacing the phone and hands by a virtual representation in VR [2].

In the video recordings that we obtained from Biener et al. [5], the participants were using an off-the-shelf HMD without any of the previously mentioned advanced techniques. Besides seeing a virtual representation of the keyboard and a video pass-through of their hands, they could enable video pass-through to see the physical world. Therefore, it gives us the opportunity to observe how participants behave in a baseline VR environment and how they handled situations in which users had to interact with bystanders or physical objects without the above-mentioned advanced techniques.

2.2 Prolonged Studies in AR and VR

To better understand and observe all effects of using VR, it is advisable to conduct studies with a longer duration. There are several studies that looked at the prolonged use of VR or AR devices. For example, Steinicke and Bruder [52] conducted a self experiment where one person spend 24 hours working, eating, sleeping and entertaining himself in VR. Nordahl et al. [41] reported on two participants who used VR for 12 hours. Lu et al. [32] reported on an in-the-wild study in which participants used their glanceable AR prototype for three days concluding that the main issue that would keep the participants from using this system daily is the form factor of the glasses. Grubert et al. [14] observed user of AR HMDs in an order processing task for four hours. While the work efficiency of users increased, they also reported higher visual discomfort.

Guo et al. [19–21] and Shen et al. [48] report on a study where

27 participants worked in a virtual and a physical office environment for eight hours each. They focused on emotional and physiological needs [20] during short and long-term use of VR, but also on visual fatigue and physical discomfort [19]. They found that the weight and form factor of the HMD resulted in higher physical discomfort in VR compared to the physical condition. Yet, there are no further insights into how participants coped with wearing the HMD.

So far, the longest VR study was reported by Biener et al. [5] who compared working in VR to working in a regular physical environment for five days each. They reported a wide range of measures (task load, usability, flow, productivity, frustration, positive/negative affect, wellbeing, anxiety, simulator sickness, visual fatigue, heart rate, break times, typing speed) taken at regular time intervals. Yet, this paper did not closely examine the video data recorded during the study, which includes valuable and interesting insights in how participants use the HMD, how they interact with it and what problems they encounter. We have been granted access to this data and analyzed around 40% of the video data to study participant behavior in more detail. This can also contribute towards a better understanding of the ergonomics of current VR devices, as these aspects should get more attention and consideration [9, 10].

2.3 Video Analysis Studies

Analyzing videos recorded during studies have been used before to get a better understanding of participant behavior. Southgate [51] analyzed types of learning behavior in virtual environments by analyzing screen capture videos, concluding that this can be a good process to understand learning behavior within an immersive VR environment. Kang et al. [26] conducted a study where children used a mobile AR system to learn about mathematical concepts, and analyzed the videos recorded during the study to see how children used the system. Segura et al. [33] used video analysis to describe the behavior of participants while playing games in an immersive exergame platform. For handheld AR, Grubert et al. used video analysis for a series of studies investigating the behavior of players and bystanders in public gaming [15, 18] and tourism [17] scenarios.

In the workplace context, Hindmarsh and Heath [23] provide an overview of several video-based studies that were conducted in different workplaces such as call centers, mobile offices or control centers. They argue that videos provide access to details of social interactions that might not be available otherwise.

3 METHODOLOGY

This work aims at closely analyzing the behavior of 16 participants during an experiment (6 female, 10 male, $m = 29.31$ years, $sd = 5.52$, ranging from 22–38 years). Throughout this paper we abbreviate arithmetic mean as m and standard deviation as sd. All participants were university employees or researchers. Two participants had no previous experience with VR, six only slight, two moderate, four substantial, and two extensive experience. The participants were recorded wearing an off-the-shelf Oculus Quest 2 while working for five consecutive days (VR condition) and also for another five without an HMD (PHYSICAL condition), as specified in Biener et al. [5]. The ethics committee of Coburg University approved this study, which took place in quiet lab areas. The VR setup provided the possibility to exchange the virtual environment with a video pass-through of the physical world and participants could see a virtual representation of the physical keyboard and a video pass-through of their hands. Otherwise both conditions were kept as similar as possible. Each day, participants worked for eight hours, with a mandatory 45-minute lunch break after four hours. Instead of a predefined study task, participants carried out their own everyday work tasks. The videos we use in this study were recorded during this previously mentioned study [5] using a webcam facing the participants and recording their face and parts of their upper body. Similar to prior research involving video analysis [26, 51] we used open and axial coding. Six people were involved in watching and annotating interesting behavior in the videos (Annotator A, B, C, D, E, and F). All annotators had a computer science background, either as students or as employees in our lab. As a first step, annotator A and B were skimming through one VR video for each participant, taking notes on possible codes for relevant behavior. Thereafter they discussed and concluded on a first joint codebook. Both annotator A and B used this codebook to code four hours of one video during which there were multiple iterations of discussion to extend and refine the codebook.

The codebook for the PHYSICAL videos was derived directly from the VR codebook, to allow us to compare participants' behavior in VR with their behavior in the PHYSICAL condition. Therefore, annotator A and B discussed which codes from VR would also be applicable for PHYSICAL, as well as which additional codes were needed for PHYSICAL. Codes involving the HMD were substituted with regular glasses, if participants were wearing any. However, actions that were inapplicable to normal glasses were removed, such as managing cables or using controllers. All codes were grouped into categories as described in section 3.3.

The study by Biener et al. [5] provided us with around 698 hours of video material for the VR condition and around 702 hours for the PHYSICAL condition. The time demand for labeling one video was very high as annotators required about one hour for processing one hour of video material. Therefore, we decided to annotate all VR videos for day 1, 3, and 5 and the PHYSICAL videos only for day 1. As participants were already familiar with using a standard desktop setup for work, we were not expecting a change of behavior over time for the PHYSICAL videos. This means we annotated 48 VR videos (420 hours) and 16 PHYSICAL videos (139 hours) for a total of 559 hours. In the end, annotator A and B finished eight, annotator C and D 16, annotator E 12, and annotator F four videos. The amount of completed videos varied between annotators, depending on their available time and the number of events in each video.

3.1 Inter-Coder Reliability

To ensure a high consistency among annotations, annotator A and B first explained the codebook to the other annotators and thereafter each annotator annotated a 30-minute training video, extracted from one of the VR videos. The results from each annotator were compared to the results from annotator A, by computing the F-score [25]:

$$F = \frac{(1+\beta^2) \cdot recall \cdot precision}{(\beta^2 \cdot precision) + recall}. \tag{1}$$

As there was no indication that we should weight recall over precision or vice versa, the weight β was set to 1. Recall is the number of overlap between two annotators divided by all annotations of the first annotator, and precision is the number of overlap between two annotators divided by all annotations of the second annotator.

The F-score was first calculated for each code and then averaged among all codes to arrive at a single value comparing the two annotators. The average F-score for the training video was 0.77 (sd=0.09). The annotations for the training video were also checked manually to identify any problems, which were then discussed among annotators, as some actions could not be unambiguously assigned to one code.

After discussing the training video, each annotator was assigned videos of several participants. Upon starting annotating videos of a new participant, a random 30-minute-video-section of that participant was annotated by two different annotators, one of which was always annotator A, and then checked for inter-coder reliability. The average F-score for these tests was 0.81 (sd=0.09). After discussing potential issues, the corresponding annotator proceeded to annotate all videos of this participant, including short discussions with other annotators about unclear events. Therefore, all four videos (three VR and one PHYSICAL) of one participant were annotated by a single annotator to allow for the highest practically possible consistency in the annotation process.

3.2 Labeling Process

We used ELAN [8, 40] to code the videos. We set it up to have multiple tiers to add codes, as it was possible that labels would overlap. For example, participants could be talking while rubbing their eyes. Annotators usually watched the video at twice the original speed and stopped for adding annotations by dragging the mouse in the timeline to mark the corresponding time-span and then selecting the appropriate annotation from a predefined dictionary with all the codes. HMD-related events were generally annotated from the moment the participant was touching the HMD to the moment when the hands stopped touching it. Taking off the headset was annotated from the moment the participant touched the HMD until the HMD no longer touched the head of the participant and the other way round for putting the HMD on. For other events, the annotation was added from the point where it was apparent to the annotator that the event started until the point when the event concluded. This included observing arm movements to estimate the start and end of actions, such as using the controller while it is not visible in the video.

3.3 Statistical Analysis

We grouped all codes into 17 categories. We analyzed the occurrence of each category in VR using a repeated measures ANOVA. As in prior work [5] we analyzed DAY (DAY 1, DAY 3, DAY 5) and TIME (MORNING, AFTERNOON) as independent variables. As the average number and the total duration of the events per hour are the dependent variables, we divided the total number and duration during each time period (MORNING, AFTERNOON) by the length of this time period, which was usually around four hours and 22 minutes (four hours of work plus half of the 45-minute break). To ensure the robustness of the ANOVA, even with data that is not normally distributed [7], we used Greenhouse-Geisser correction whenever the sphericity assumption was violated. We applied Bonferroni-correction to all post-hoc tests involving multiple comparisons.

For categories that are also sensible in PHYSICAL, we compared DAY 5 of VR, in which participants were already more familiar with the HMD, to DAY 1 of PHYSICAL using a repeated measures ANOVA with the independent variables INTERFACE (VR, PHYSICAL) and TIME (MORNING, AFTERNOON).

In addition, we also ran a separate repeated measures ANOVAs for each measure with gender as a between-subjects factor to test for gender differences. We did not run separate analysis for each gender, as the two groups are very small. However, the descriptive data and trends of both individual groups (male, female) is in line with the significant differences reported in the results.

4 RESULTS

The results are clustered into 17 categories, including 'Other'. All significant ANOVA results can be found in table 1 and 2.

Adjusting This category describes all events where the participants adjusted their HMD, by moving it or making any part tighter or looser. This includes: adjusting the HMD from below by pushing it upwards; adjusting from the side with one hand; adjusting with both hands; adjusting with one hand or two hands at the back of the head; adjusting, moving or touching the fitting strap on top of the head; making it tighter or looser by rotating the fitting wheel at the back of the head; and fixing the face pad by moving the finger between the face and the headset. A representative visualization of such events is depicted in Fig. 1 (a).

Analyzing how often the participants adjusted the HMD per hour showed that this number was significantly influenced by the DAY. Post-hoc tests showed that there were significantly less adjusting actions on DAY 5 ($m = 8.06$, $sd = 7.01$) compared to DAY 1 ($m = 15.07$, $sd = 12.64$, $p = 0.010$). Similarly, the average time per hour spent on adjusting the HMD was significantly influenced by the DAY and post-hoc tests showed it was significantly higher on DAY 1 ($m = 44.88$ s, $sd = 29.24$) compared to DAY 3 ($m = 25.7$ s, $sd = 21.91$, $p = 0.008$) and DAY 5 ($m = 22.3$ s, $sd = 17.87$, $p < 0.001$). The number and duration of adjusting events can be seen in Fig. 2. We did not find significant differences between genders for this measure. Examination of the frequency of the individual codes used to label the videos revealed that 34.4% of them was using two hands, which was the most common, followed by one hand from the side at 25.3%.

There were five participants who wore glasses during the study. Comparing their adjusting behavior in VR with PHYSICAL revealed that two of them adjusted their regular glasses more often than the HMD (one participant around 80% more; another more than ten times as much), while the other participants did it less (96%, 67% and 30% less). For all except one participant, the time spent on adjusting per hour on average was less in PHYSICAL.

Table 1: Significant RM-ANVOA results describing changes within VR.

Number of Events per Hour

Ind. Variable	df_1	df_2	F	p	η_p^2
Adjusting					
DAY	1.42	21.3	8.06	0.005	0.35
Supporting					
DAY	2	30	4.8	0.016	0.24
GENDER	1	14	11.4	0.008	0.45
Taking HMD Off					
DAY	2	30	5.05	0.013	0.25
Touching HMD					
GENDER	1	14	9.63	0.008	0.41
Keyboard Tracking Problem					
TIME	1	15	5.08	0.04	0.25
Rubbing Eyes or Face					
GENDER	1	14	5.28	0.057	0.27
Colleague					
GENDER	1	14	4.69	0.048	0.25

Total Duration per Hour

Ind. Variable	df_1	df_2	F	p	η_p^2
Adjusting					
DAY	2	30	14.93	< 0.001	0.5
Supporting					
GENDER	1	14	5.06	0.041	0.27
Taking HMD Off					
DAY	2	30	8.87	< 0.001	0.37
Screen Time					
DAY	2	30	8.31	0.001	0.36
Standing up					
DAY	2	30	6.15	0.006	0.29
TIME	1	15	136.79	< 0.001	0.9

Supporting This category describes all events where participants held the HMD in a way that suggests the purpose was to hold the HMD in a more comfortable or correct position. In contrast to the adjust-category, the HMD is barely moving and supporting generally lasts longer than adjusting. This category includes: supporting with one hand or both hands from below; and supporting it with one hand or both hands from the side. A representative visualization of such events is depicted in Fig. 1 (b).

Analyzing how often the participants supported the HMD per hour showed that this number was significantly influenced by the DAY. Post hoc-tests showed there were significantly less supporting actions on DAY 5 ($m = 4.38$, $sd = 9.07$) compared to DAY 3 ($m = 6.67$, $sd = 10.85$) ($p = 0.041$) as displayed in Fig. 2. However, we could not find a significant influence of DAY or TIME on the total duration of supporting events even though it was decreasing from DAY 1 ($m = 385.76$ s, $sd = 695.3$) to DAY 5 ($m = 308.78$ s, $sd = 730.70$). An analysis of gender as a between-subjects factor revealed a significant difference in the number of supporting events, such that it occurred around ten times more often for female ($m = 14.24$, $sd = 12.69$) than male ($m = 1.34$, $sd = 2.48$) participants. We also found a significant difference between the duration of supporting events, which was around 12 times longer for females ($m = 790.21$ s, $sd = 995.95$) than males ($m = 62.9$ s, $sd = 174.42$). This is also displayed in Fig. 2. This indicates that in general female participants were supporting the HMD a lot more than male participants. However, it is also notable, that the standard deviation of these measures for females is much higher than for males.

We also observed than in the beginning of the study, the number and the total duration of supporting events was much higher (average around 30 times higher) for six participants compared to the others. For all of them except P15 and P10, the number and duration of supporting events decreased slightly during the week. When asked about the

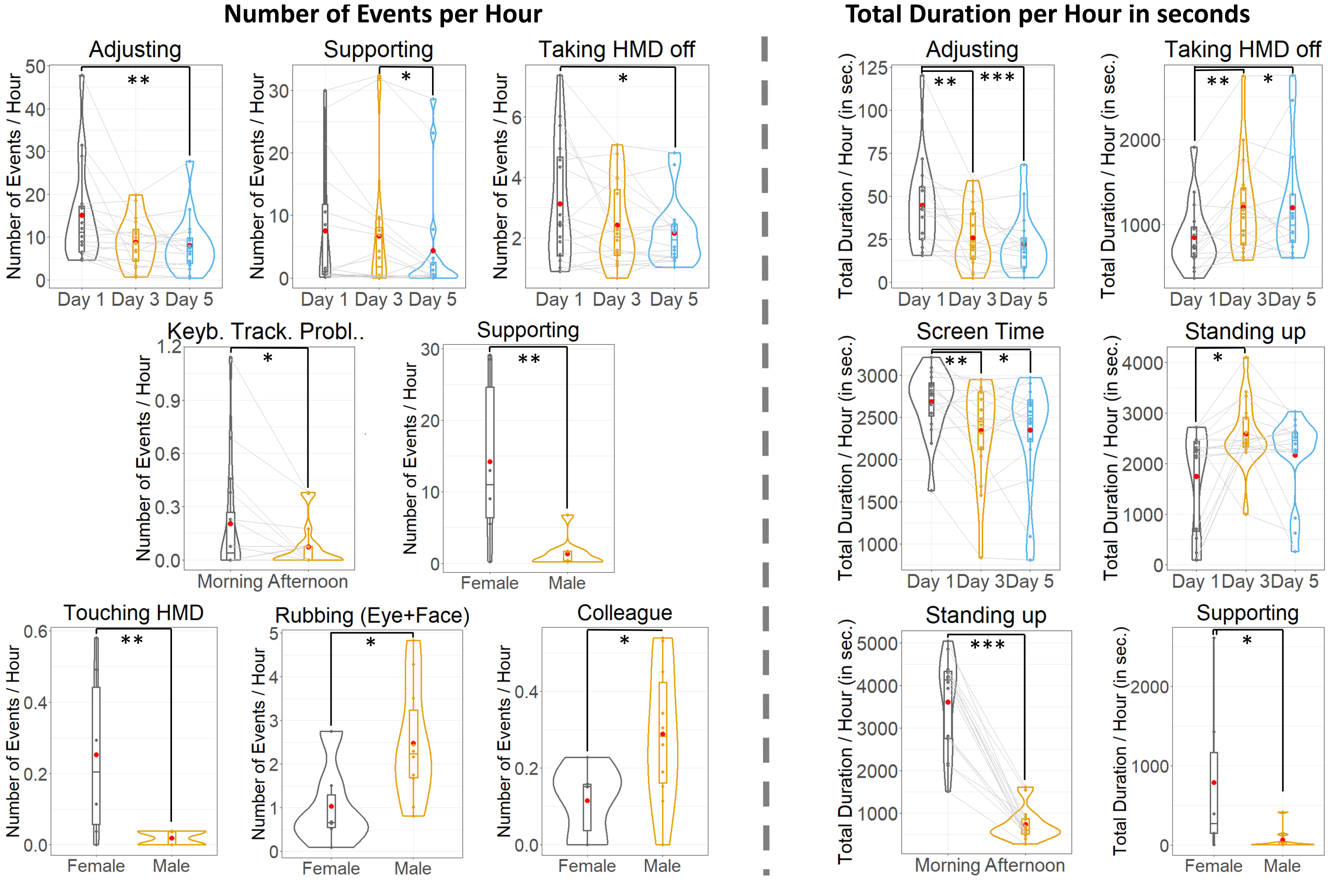

Fig. 2: Violin plots, showing the number and total duration of events per hour in VR separated by DAY, TIME and gender. The red dot displays the arithmetic mean. Stars indicate the level of significance in post-hoc tests: $* < 0.05$, $** < 0.01$, $*** < 0.001$.

reasons for supporting the HMD, P04, P08, P09 and P10 mentioned that this helped them to fix the HMD in a position where they get the sharpest image. P04 mentioned this was especially difficult for her as she wore regular glasses underneath the HMD. P08 and P15 also mentioned that without using their hands, they would need to use their cheek muscles to lift the HMD. P15 said that making the head-strap of the HMD tighter caused her to feel dizzy and P14 and P10 did not like the pressure on their face, so they kept it rather loose and supported the HMD with their hands instead. P09 mentioned the HMD bothered him because he is not used to wearing something on his head. However, it was visible in the video that the participants had to stop supporting the HMD at times to be able to do things such as typing. Analyzing the frequency of individual codes also revealed that for more than 91% of supporting events participants used one hand only. We could not find indications that wearing glasses or a lack of VR experience would lead to a high number of supporting events. However, matching the significant gender differences, five out of these six participants were female, while only six out of the 16 participants in the study were female.

Taking HMD Off To analyze how often and for how long participants took the HMD off, we annotated parts in the video where participants take the HMD off and put it back on. In the process of these events adjusting the HMD was not labeled separately, but considered as being part of putting the headset on, or taking it off.

The time durations for not wearing a HMD are the time difference between take-off and put-on events. Analyzing how often the participants took off the HMD per hour showed that this was significantly influenced by the DAY. Post-hoc tests showed that participants took off the HMD significantly less on DAY 5 ($m = 2.17$, $sd = 1.16$) compared to DAY 1 ($m = 3.14$, $sd = 2.26$) ($p = 0.04$). For the average time per hour that participants were not wearing the HMD, we also found that this was significantly influenced by DAY. In contrast to the number of events, post-hoc tests showed that the duration was significantly lower on DAY 1 ($m = 849.42\ s$, $sd = 492.85$) compared to DAY 3 ($m = 1202.78\ s$, $sd = 687.28$, $p = 0.002$) and DAY 5 ($m = 1199.20\ s$, $sd = 643.40$, $p = 0.015$). The number and duration of taking-off events are displayed in Fig. 2. We did not find significant differences between genders for this measure. Also, we could not find any significant correlations between the average number and duration of supporting actions and the average number and duration of taking off the HMD for each participant.

To gain a better understanding of why participants took the HMD off, we checked which other labels occurred while the HMD was not worn. The most frequent ones were talking (14.6%), using the phone (12.4%), standing up (12.0%) or sitting down (11.8%), rubbing the face (11.4%) and drinking (10.5%).

For the five participants who wore regular glasses in the PHYSICAL condition, we noticed that one participant took off the regular glasses more often than the HMD while all other participants did it less and one participant never did it. The average duration per hour that they did not wear the glasses was much lower for all participants in PHYSICAL ($m = 103.84\ s$, $sd = 201.12$) compared to VR ($m = 1277.58\ s$, $sd = 692.87$).

Screen Time We calculated screen time in VR by including all the times when the participants were wearing the HMD, not including all the times in which they were taking the HMD halfway off, using the controller or tapping on the HMD, encountering keyboard tracking problems or were reading, writing or using their phone. We found a significant main effect of DAY on the screen time per hour in VR, such that the screen time was higher on DAY 1 ($m = 2684.62\ s$, $sd = 494.58$)

Table 2: Significant RM-ANOVA results comparing VR and PHYSICAL.

Number of Events per Hour

Ind. Variable	df_1	df_2	F	p	η_p^2
Eating or Drinking					
INTERFACE	1	15	7.97	0.013	0.45
Phone Call					
TIME	1	15	7.2	0.02	0.33
Rubbing Eyes or Face					
INTERFACE	1	15	5.6	0.032	0.27
TIME	1	15	7.64	0.014	0.34
Rubbing Eyes					
INTERFACE	1	15	11.75	0.004	0.44
TIME	1	15	9.03	0.009	0.38
Physical World Activity					
INTERFACE	1	15	7.669	0.014	0.34

Total Duration per Hour

Ind. Variable	df_1	df_2	F	p	η_p^2
Standing up					
TIME	1	15	66.97	< 0.001	0.82
Rubbing Eye					
TIME	1	15	10.33	0.006	0.41

compared to DAY 3 ($m = 2343.18$ s, $sd = 686.0$, $p = 0.004$) and DAY 5 ($m = 2348.58$ s, $sd = 643.88$, $p = 0.017$). We did not find significant differences between genders. The screen time for all three VR days is shown in Fig. 2.

In PHYSICAL we included two additional codes that mark events where the participant turned away from the screen and turned back to look at the screen to compare the screen time in VR and PHYSICAL. This code could not be reliably used for VR, because it is not clear from the videos where the participants had placed their virtual monitor and therefore if they are looking at it or not. Screen time in PHYSICAL was then calculated similarly to VR, as the time that participants were looking at the screen not including all times in which they were reading, writing or using the phone. Comparing VR with PHYSICAL showed an interaction effect of DAY and TIME but post-hoc tests did not show significant differences.

Touching HMD We introduced one code which labels events where participants are touching the HMD without adjusting or supporting it or doing any other purposeful action.

Overall, this is a very rare event which did not even occur for 6 participants. Therefore, we could also not find any significant influence of TIME or DAY on the number or duration of such events. However, looking at gender as a between-subjects factor revealed a significant difference in number of touching events such that it occurred more often for females ($m = 0.25$, $sd = 0.57$) than males ($m = 0.02$, $sd = 0.06$) as is displayed in Fig. 2. No difference between genders was detected for the total time spent on such events.

From what the annotators can estimate from the videos, four participants touched the HMD to check its position, five were just touching it for no apparent reason, three moved their fingers along the HMD, one was tapping on both sides, and one participant was briefly touching it with a second hand while adjusting it.

Handling Cable This category describes all events related to the power-cable that was usually plugged in to the HMD during the study. It can either be that the participants are purposefully touching the cable, for example to move it out of the way or that the participants take out or plug in the cable. A representative visualization of such events is depicted in Fig. 1 (c).

This is also a rather rare event that only occurs in around 67% of the analyzed videos. We did not find any significant influence of TIME or DAY on the number or duration of such events nor any differences between genders. Looking at the labeled sections, we noticed that six participants were moving the cable behind their arm, six participants checked that it is properly plugged in and six participants simply tried to move it to the side a little. P01 seemed to struggle with the weight of the cable pulling the head down on one side, so she tried to find a more comfortable position by laying it across her head. Four participants unplugged it during their workday: one of them to be able to turn around and redraw the boundary; one to stand up while wearing the HMD; and two participants seemed to take it out because it was annoying. In six videos the cable was coming from the top hanging over a movable wall to avoid extra weight. Even though the sample size is very small, this could be a strategy to avoid cable events, as the average number of events per hour with the cable hanging down was 0.46 ($sd = 0.85$) and with the cable coming from the top only 0.10 ($sd = 0.18$).

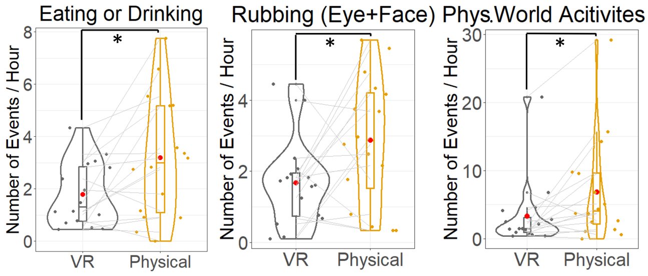

Fig. 3: Violin plots showing the number of events per hour for VR and PHYSICAL. The red dot displays the arithmetic mean. Stars indicate the level of significance in post-hoc tests: $* < 0.05$, $** < 0.01$, $*** < 0.001$.

Taking HMD Halfway off This category describes actions in which participants lift either the front of the HMD above eye-level with one or two hands, or lift the strap up at the back without completely removing the HMD. A representative visualization of such events is depicted in Fig. 1 (d).

We could not find any main effects of DAY or TIME on the number or duration of such actions, nor significant gender differences. Analyzing the frequency of the individual codes revealed that 74% of these actions were performed with one hand and 25% with two hands in the front and 1% in the back. To get an idea of why participants took the HMD halfway off, we checked which other labels overlapped with it. The most common ones were drinking (38.0%), rubbing the face (8.5%), talking (8.0%), rubbing the eyes (6.9%), using the controller (6.1%) and using the phone (5.2%). By examination of the overlaps we cannot be sure that they caused the participants to take the HMD halfway off, but they can be indications. The back of the HMD was mainly lifted to scratch the head.

Interacting with HMD We captured how often the participants had to interact with the interface of the HMD by using the controller or hand gestures, by tapping on the HMD to enable or disable passthrough mode, by using the volume-buttons on the HMD or by redrawing the boundary of the guardian area, which always coincides with controller usage which is depicted in Fig. 1 (e).

We could not find any main effects of DAY or TIME on the number or duration of those actions nor did we find significant differences between genders. This is also not expected since these actions are mainly mandatory to operate the system. Looking at the individual codes, we saw that 83.9% of the occurrences was about using the controller (or hand-gestures), 12.6% about tapping for pass-through and 3.6% about adjusting the volume. It seemed like only four participants had to redraw the boundary during the study. We observed that participants regularly used the controllers after putting the HMD on. This is probably due to the fact that after putting it back on the virtual desktop often moved a little, so they had to reposition it.

Keyboard Tracking Problem To get an indication on how often participants had problems with the keyboard tracking, we annotated parts of the video in which participants were obviously trying to get the keyboard tracked.

This is also a rather rare event, as it only occurs in around 42% of the analyzed VR videos. In these situations, participants generally lifted the keyboard and moved it around to try to track it again as visualized in Fig. 1 (f). We found a significant main effect of TIME on the number of keyboard tracking problems, such that there were significantly more in the MORNING ($m = 0.2$, $sd = 0.52$) compared to the AFTERNOON ($m = 0.07$, $sd = 0.19$) as visualized in Fig. 2. We did not find significant differences between genders.

Standing up To analyze how often and for how long participants stood up from their chair, we annotated parts in the videos in which they got up and sat down.

We could not detect a significant influence of DAY or TIME on the number of such events. However, we found a significant main effect of both DAY and TIME on the duration of standing per hour. Post-hoc tests showed that participants were standing longer on DAY 3 ($m = 2590.74\ s$, $sd = 1862.65$) compared to DAY 1 ($m = 1752.45\ s$, $sd = 1855.36$, $p = 0.027$). In addition, participants were standing significantly less during the AFTERNOON ($m = 729.83\ s$, $sd = 522.68$) compared to the MORNING ($m = 3612.36\ s$, $sd = 1558.06$). These results are visualized in Fig. 2.

Comparing VR to PHYSICAL revealed a significant influence of TIME on the duration of standing per hour, yet no interaction effect between INTERFACE and TIME. This indicates that both in VR and PHYSICAL participants stood less during the AFTERNOON ($m = 742\ s$, $sd = 513.7$) than during the MORNING ($m = 3425\ s$, $sd = 1498.5$) (see Fig. 3).

No significant gender differences were found for such events.

Eating or Drinking This category summarizes the two codes for drinking and eating, labeling the parts in which participants put drinks or food into their mouth, ignoring longer periods of chewing. Fig. 1 (g) depicts a representative visualization of such events.

We combined the occurrences of both codes to see how often and for how long participants were engaged with food. In VR, we could not find a significant influence of DAY or TIME on the number or duration of such events.

When comparing PHYSICAL to VR, we found a significant influence of INTERFACE on the number of such events, such that in PHYSICAL ($m = 3.19$, $sd = 2.67$) participants drank and ate more frequently than in VR ($m = 1.79$, $sd = 1.31$) (Fig. 3). However, we did not find a significant influence on the total time spent eating or drinking per hour. We did not find any significant differences in gender. Looking at the frequency of drinking and eating individually, we found that participants were mostly drinking, accounting for 73.7% of the occurrences in VR and 76.0% in PHYSICAL.

Generally, participants used the same containers for drinking in VR and PHYSICAL which included cups, glasses, and bottles. In VR, for the majority (92.2%) of drinking and eating events and the majority (94.1%) of time spend eating, participants did not take off the HMD and of those in only 16% did they lift the HMD (take halfway off), and in only 10% of the cases did they support the HMD. We noticed that P06 often did not support the HMD when drinking out of a bottle but supported or took it halfway off when using a cup. Instead of supporting the HMD with the hands, P04 sometimes used the cup itself to push the HMD slightly upwards. P02, however, at one point attempts to drink out of a glass but gives up because the HMD blocks access to the mouth.

Talking To detect how often, and for how long, participants were talking to others in the same room, visualized in Fig. 1 (h), we labelled such parts in the videos. As the videos were recorded without audio, we included all events in which the participants move their lips in a way that suggests they are having a conversation, not including phone calls or video calls, if this is apparent in the video.

We could not find a significant influence of DAY or TIME on the number or duration of talking events per hour, and no significant difference between VR and PHYSICAL or between genders. This is unsurprising, as talking is probably dependent on many external factors such as availability of another person. However, we found that 78.2% of all talking occurrences and 79.5% of the total talking duration was done while wearing the HMD and only in 1.9% of these did participants lift the front of the HMD (take halfway off).

Phone Call We introduced one code for annotating all parts of the video in which the participant is talking on the phone or in a video call such as visualized in Fig. 1 i).

This is a rare event, as it only occurs in around 40% of the analyzed VR videos. We did not find any significant influence of DAY or TIME on the number or duration of phone calls and also no significant influence of INTERFACE or gender. However, when comparing VR and PHYSICAL there was a significant main effect of TIME, such that there were more phone calls in the MORNING ($m = 0.25$, $sd = 0.42$) compared to the AFTERNOON ($m = 0.15$, $sd = 0.33$). However, we have to be cautious here, as the frequency of phone calls are probably much more influenced by other people than the participants experience during the VR condition.

We could observe that 90.9% of all phone call occurrences and 94.4% of the total phone call duration occurred while wearing the HMD and in only 7% of these instances did the participant lift the front of the HMD (take it halfway off). P04 once removed the HMD after answering the call. Yet, at another time she put the HMD back on after starting a call. P11 removed the HMD to answer the call and then afterwards immediately put it back on. From the videos it appears participants were both answering and starting calls.

Rubbing Eyes or Face This category combines instances in which participants are rubbing their eyes or face whenever these actions are very obvious, not including events in which the participants very shortly, and probably involuntarily, touch their faces. A representative visualization of such events is depicted in Fig. 1 (j).

We did not observe any significant influence of DAY or TIME on the number or duration of such actions in VR. An analysis of gender as a between-subjects factor revealed a significant difference in the number of instances of rubbing eyes or faces, which was about twice as high for males ($m = 2.48$, $sd = 1.82$) than for females ($m = 1.03$, $sd = 1.14$), yet no difference was detected regarding the total duration. Significant differences are visualized in Fig. 2.

Also, when comparing VR and PHYSICAL, we found a significant main effect of INTERFACE and TIME on the number of such actions , but no effect on the duration. These effects indicate that participants rubbed their eyes and faces more in PHYSICAL ($m = 2.87$, $sd = 2.25$) than VR ($m = 1.68$, $sd = 1.44$) (see Fig. 3) and that they did this more often in the MORNING ($m = 2.75$, $sd = 2.36$) compared to the AFTERNOON ($m = 1.8$, $sd = 1.36$).

An analysis of the frequency of individual codes shows that in VR 67.1% were about rubbing the face and 32.9% about rubbing eyes, while, in PHYSICAL rubbing the eyes had a portion of 59.0%.

Therefore, we compared VR and PHYSICAL also for both individual codes. There were significantly more eye rubbing events in PHYSICAL ($m = 1.7$, $sd = 1.68$) compared to VR ($m = 0.58$, $sd = 0.613$), but no significant difference for the duration or between genders. Yet, there were significantly more eye rubbing events in the MORNING ($m = 1.45$, $sd = 1.66$) compared to the AFTERNOON ($m = 0.83$, $sd = 0.94$) and significantly more time spent on eye rubbing in the MORNING ($m = 4.82\ s$, $sd = 5.68$) compared to the AFTERNOON ($m = 2.74\ s$, $sd = 3.06$). No significant differences were found regarding rubbing the face.

Physical World Activities To explore how often and for how long participants concern themselves with things outside of the virtual world, this category combines codes that describe events where the annotators believe the participants were reading something outside of VR, writing something outside of VR, using a smartphone (Fig. 1 (k)) or otherwise peeking under the HMD to see something. These events can occur together with previously mentioned events such as taking the HMD (halfway) off.

We found no significant influence of DAY or TIME on the number or duration of events in this category. Comparing VR and PHYSICAL indicated a significant main effect of INTERFACE such that physical world activities where significantly more frequent in PHYSICAL ($m = 6.88$, $sd = 8.40$) than in VR ($m = 3.28$, $sd = 4.97$) (see Fig. 3). However, no significant differences have been found regarding the total time spent on such actions. We also did not find significant differences between gender.

An analysis of the frequency of individual events in VR showed that in 64.0% of the cases participant were using their phone. In 20.1% of the cases they were probably reading and in 10.0% of the cases they were writing something. However, in PHYSICAL only 46.6% of the actions were related to the phone and 35.6% to writing and 17.7% to reading. We also found that 92.3% of phone usage in VR happened while wearing the HMD and only in 7% of these cases did participants lift the HMD in the front (taking it halfway off).

Colleague To detect if an event was likely triggered by a colleague or another person, we used this code in combination with any other code to label such events. For example, the participant could be taking off the HMD when a colleague approaches.

We could not find any statistically significant influence of DAY or TIME. An analysis of gender as a between-subjects factor revealed a significant difference in number of events triggered by colleagues which was higher for males ($m = 0.29$, $sd = 0.34$) than females ($m = 0.11$, $sd = 0.19$) as depicted in Fig. 2. However, this measure heavily depends on the coworkers, their availability and relationship with the participant.

Checking which actions were triggered by colleagues in VR, showed that 37.4% were taking off the HMD, 26.6% were talking, 8.6% were taking the HMD halfway off, and 5.7% were taking the headphones off. In PHYSICAL, most common actions were to turn away from (81%) or towards (6%) the screen, taking the headphones off (5%) and talking (3%). These values can give some indications on how users react when approached by a colleague.

Other We also added a code to label any behavior that the annotators might find interesting that do not match any of the previously mentioned codes. For example, sometimes the washable face-pad that we used in the study got out of place, so we saw five participants adjusting the face pad while not wearing the HMD. Additionally, we could observe three participants cleaning the HMD during the study.

5 DISCUSSION

The results provide insights into the behavioral patterns of the participants as they acclimated to the experience of working in VR. Exposure to unfamiliar technology is well known to demand an initial period of familiarization and adaptation, as also seen in studies on user experience in VR and AR [27, 28]. In this work, however, we investigated behavioral changes over a much longer timescale than typical VR user studies so far, and our observations reveal rich patterns of user behavior not previously documented.

The observed decline of adjusting events suggests that users need less time to fix the HMD in a comfortable position as they become more familiar with the device which highlights the need for improved tutorials for initial fitting. In addition, the number of supporting events but not the total time spent supporting declined significantly over the week. This could indicate that users put up with the need to support instead of repeatedly trying to wear the HMD without support. The main reasons for support, as mentioned by the participants, were to obtain a clear image and to avoid pressure on the head by having the head-strap too tight. This emphasizes the importance of ongoing efforts to improve the form factor of HMDs. In 95% of the cases, the participants only supported the HMD with one hand, which made it possible to use the keyboard or mousepad with the other hand. Still, from an ergonomic point of view, this is clearly not desirable, especially for prolonged use of VR or AR HMDs.

Despite the limited number of participants, we observed some significant differences between genders, which were especially prominent in the number and duration of supporting events. Similarly, in their eight-hour-long study, Guo et al. [19] also noticed significant gender differences regarding visual fatigue, which was much higher for females. Yet in both cases, the number of participants was rather small, and other factors such as experience with VR devices or games, as mentioned by Guo et al. [19], cannot be ruled out. Therefore, gender differences should be investigated more closely in future research and these findings emphasize the importance of having a diverse group of participants.

The frequency of taking the HMD off decreased over the five days. However, the break durations and time spent standing increased and consequently, screen time decreased. Yet, we did not detect any significant difference in screen time between VR and PHYSICAL. Also, our results suggest that supporting the HMD does neither significantly increase, nor decrease the need for taking the HMD off.

Reviewing why participants might lift the HMD in the front, or even take it off completely, indicated that this was done to drink, rub the face and eyes, talk, or to use a phone. These are either basic needs or common actions in work environments that should not be restricted by the VR device. Future HMDs could reduce the need for lifting the HMD when drinking or rubbing the face, by reducing the form factor or allowing easier access to the face and especially the mouth while wearing the HMD. One solution could be a raisable visor as offered by the HoloLens 2 [38]. When talking or using the phone users could benefit from a clearer pass-through mode or other techniques to include outside information, such as smartphones [1, 2, 11] or parts of the surrounding environment [22, 56].

Differences between VR and PHYSICAL were found in the frequency of participants rubbing their eyes, which was significantly more frequent in PHYSICAL and could be explained by the eyes being more accessible. The HMD could prevent participants from involuntarily rubbing their eyes as it requires more actions, such as lifting the HMD. Also, in PHYSICAL, participants interacted significantly more with physical objects, such as reading, writing, or using a phone which could indicate that they were more accessible for spontaneous actions. This matches the reports of Biener et al [5], who mentioned that participants missed the ability to write things down in VR. Similarly, the frequency of eating and drinking was higher in PHYSICAL, indicating that the HMD was too obstructive to support such activities. This was also visible in the videos as participants often supported or lifted the HMD to drink and it is also in line with participants' comments reported by Biener et al. [5] about being afraid to spill something, as well as prior observations by McGill et al. [34]. On the other hand, no significant differences were found in the total duration of drinking or eating events, physical world activities, or in rubbing of eyes or faces. This indicates that participants did not neglect these needs, but instead did them more carefully and for longer periods, as it was not convenient to do them as often in VR. Also, the HMD did not substantially restrict participants from performing actions outside VR. For example, 92% of phone usage in the VR condition was done while wearing the HMD and participants only rarely lifted the HMD to do so. Additionally, during 80% of the time spent talking, 94% of the time spent on phone calls, and 94% of the time spent drinking or eating, participants wore the HMD. Still, further efforts to make such interactions more comfortable should not be neglected, such as including the phone, as mentioned above, or hinting at and visualizing bystanders [49] or objects. This finding sheds further light on how people adopt and appropriate technology in their daily routines, even though those routines might initially be hampered by that very technology.

Another problem could be caused by the power cable, which can pull the HMD down and become distracting when obstructing users' movements. When the HMD should be used for a time span that requires charging while working, one should think about how to comfortably adjust the cable, for example by hanging the cable from the ceiling or running it down at the back of the head rather than the side, which has already been done for newer HMDs such as Varjo Aero or Apple's Vision Pro, or include enough breaks to charge the HMD in between uses. Participants also seemed to have issues with the keyboard tracking. This happened more often in the MORNING and we can speculate that it could be caused by different lighting conditions or because participants simply coped with not seeing the keyboard in the AFTERNOON, albeit window blinds were used to control lighting conditions. Taking into account the significantly fewer standing events that tended to occur in the afternoon it could also hint at participants being generally less active in the afternoon than earlier in the day. Still, the observed problemsunderline the need to further improve all components of VR HMDs, including robust object and environment tracking algorithms.

Overall, these varied observations and findings can inform the design of more comfortable HMDs for work applications, and serve as guidance for extended use of VR in work settings and risk assessments of such activities. Hardware solutions are likely required to address the trade-off between setting straps tightly to ensure a stable image and the associated discomfort arising from this tightness after extended use. Current generation HMDs, such as the Oculus Quest Pro, do offer different ergonomics compared to the Quest 2. Yet, on an anecdotal level only, we experienced that even with these changes prolonged use beyond 30–45 minutes can still lead to increased pressure on the forehead, and, subsequently, headache. Hence, it should be carefully studied if these new headset generations deliver a better long-term usage experience or not. Prior work has analyzed and designed workplace ergonomics by using VR as a tool [13, 58]. However, less focus has been on the ergonomics of VR itself [9, 10], or on possible solutions, such as Wentzel et al. [57], who amplified hand movements for more comfortable interactions in VR, or McGill et al. [35], who reduced the need for head movement in large-display setups. Software solutions may potentially be able to address other comfort issues, such as streamlining interactions to minimize user effort, prompting users towards more ergonomic virtual display configurations, and facilitating common tasks currently not easy to perform while wearing a HMD, such as locating items on the desk, using a phone, and eating or drinking. Software-based interventions may also allow for promoting compliance with ergonomic guidance, for example by encouraging workers to take breaks after predefined time periods, using eye exercises to alleviate digital eye strain [24], performing some form of exercise after set periods, or incorporating relaxation methods designed for VR [54, 55]. Although many of these design implications are intuitive, there has been limited evidence to support and motivate such remedial efforts, and, hence, we call for increased community efforts to work towards VR (and AR) experiences that can support prolonged use in various contexts.

Currently, research is often focused on brief usability evaluations. Yet, we argue that long-term usability issues require further attention. In addition, the issues of working in VR for a prolonged time, as described in this study, which have not been formally documented and described before, provide a reference against which future hardware and software improvements can be measured.

6 LIMITATIONS AND FUTURE WORK

As mentioned in the study by Biener et al. [5], a limitation is that tasks are chosen by participants. This means we cannot fully eliminate the possibility that the type of tasks on different days influenced the behavior of the participants. Also, no VR-specific benefits were employed, as the goal of the previous study was to quantify baseline usability issues. Therefore, the behavior of participants could be different when using an optimized version of VR that, for example, facilitates certain tasks or specifically aims at increasing the well-being of the users. Also, only one type of HMD, the Quest 2, was investigated. Future work should also consider HMDs that vary in important aspects such as weight, weight-distribution and resolution. Additionally, this study, as well as the prior study by Biener et al. [5] is reporting the effects of using VR in the context of knowledge workers, which does not necessarily imply a generalizability to other application areas which should therefore be evaluated in future work.

Another limitation is that we do not have a complete view of the participants' surroundingsso we can only report what was within the camera's view. For P11 and P14 the camera orientation and participant positioning were such that it was impossible to see some supporting actions or when they were using their phone.

Our analysis also hinted at some gender differences, however, female participants were underrepresented in this study (six out of 16 participants) and future work should therefore have a closer look at such gender-related behavioral differences.

In addition, due to the high time demand for annotating the videos, six people were involved in the annotation process. Through training, calculation of the inter-coder reliability and repeated discussions we were trying to make the annotations as consistent as possible. Still, the annotations depend on subjective decisions by the annotators, because certain actions can not always be unambiguously assigned to one of the categories. However, as all videos of one participant were only coded by one annotator, we maximized the coding reliability within each participant as far as possible.

7 CONCLUSIONS

We presented unique insights into how users behave during prolonged use of VR HMDs, obtained through reviewing and coding 559 hours of video material from a study in which participants worked in VR for an entire workweek. We found indications that participants are getting used to the HMD during the week, for example, the frequency of adjusting and supporting actions decreased over the five days of working in VR. Also, participants seemed to be able to adapt to certain restrictions, as they mostly did not take off the HMD while talking or using a phone. On the other hand, by the end of the week participants removed the HMDs for longer periods of time which also resulted in less screen time. In addition, we found that the HMD can be disruptive to normal patterns of eating and drinking, as well as interacting with objects in the physical environment, as they were less frequent in VR. These insights can be used to inform the design of less restricted, more ergonomic VR systems as tools for knowledge workers. Overall, this paper presents detailed insights into user behavior while working in VR, which can only be detected through in-depth video analysis of data gathered over an extended period of time. Therefore, we hope this work inspires further research on the long-term use of VR in various contexts.

REFERENCES

[1] G. Alaee, A. P. Deasi, L. Pena-Castillo, E. Brown, and O. Meruvia-Pastor. A user study on augmented virtuality using depth sensing cameras for near-range awareness in immersive vr. In *IEEE VR's 4th workshop on everyday virtual reality (WEVR 2018)*, vol. 10, p. 3, 2018. 2, 8

[2] H. Bai, L. Zhang, J. Yang, and M. Billinghurst. Bringing full-featured mobile phone interaction into virtual reality. *Computers & Graphics*, 97:42–53, 2021. 2, 8

[3] L. P. Berg and J. M. Vance. Industry use of virtual reality in product design and manufacturing: a survey. *Virtual reality*, 21:1–17, 2017. 2

[4] V. Biener, T. Gesslein, D. Schneider, F. Kawala, A. Otte, P. O. Kristensson, M. Pahud, E. Ofek, C. Campos, M. Kljun, et al. Povrpoint: Authoring presentations in mobile virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, 28(5):2069–2079, 2022. 1, 2

[5] V. Biener, S. Kalamkar, N. Nouri, E. Ofek, M. Pahud, J. J. Dudley, J. Hu, P. O. Kristensson, M. Weerasinghe, K. Č. Pucihar, et al. Quantifying the effects of working in vr for one week. *IEEE Transactions on Visualization and Computer Graphics*, 28(11):3810–3820, 2022. 1, 2, 3, 4, 8, 9

[6] V. Biener, D. Schneider, T. Gesslein, A. Otte, B. Kuth, P. O. Kristensson, E. Ofek, M. Pahud, and J. Grubert. Breaking the screen: Interaction across touchscreen boundaries in virtual reality for mobile knowledge workers. *arXiv preprint arXiv:2008.04559*, 2020. 1, 2

[7] M. J. Blanca, J. Arnau, F. J. García-Castro, R. Alarcón, and R. Bono. Non-normal data in repeated measures anova: Impact on type i error and power. *Psicothema*, 35(1):21–29, 2023. 4

[8] H. Brugman, A. Russel, and X. Nijmegen. Annotating multi-media/multi-modal resources with elan. In *LREC*, pp. 2065–2068, 2004. 3

[9] Y. Chen, X. Wang, and H. Xu. Human factors/ergonomics evaluation for virtual reality headsets: a review. *CCF Transactions on Pervasive Computing and Interaction*, 3(2):99–111, 2021. 3, 9

[10] B. A. Ciccone, S. K. Bailey, and J. E. Lewis. The next generation of virtual reality: recommendations for accessible and ergonomic design. *Ergonomics in Design*, 31(2):24–27, 2023. 3, 9

[11] A. P. Desai, L. Pena-Castillo, and O. Meruvia-Pastor. A window to your smartphone: exploring interaction and communication in immersive vr with augmented virtuality. In *2017 14th conference on computer and robot vision (CRV)*, pp. 217–224. IEEE, 2017. 2, 8

[12] T. Gesslein, V. Biener, P. Gagel, D. Schneider, P. O. Kristensson, E. Ofek, M. Pahud, and J. Grubert. Pen-based interaction with spreadsheets in mobile virtual reality. In *2020 IEEE International symposium on mixed and augmented reality (ISMAR)*, pp. 361–373. IEEE, 2020. 2

[13] D. Grajewski, F. Górski, P. Zawadzki, and A. Hamrol. Application of virtual reality techniques in design of ergonomic manufacturing workplaces. *Procedia Computer Science*, 25:289–301, 2013. 9

[14] J. Grubert, D. Hamacher, R. Mecke, I. Böckelmann, L. Schega, A. Huckauf, M. Urbina, M. Schenk, F. Doil, and J. Tümler. Extended investigations

of user-related issues in mobile industrial ar. In *2010 IEEE International Symposium on Mixed and Augmented Reality*, pp. 229–230. IEEE, 2010. 2

[15] J. Grubert, A. Morrison, H. Munz, and G. Reitmayr. Playing it real: magic lens and static peephole interfaces for games in a public space. In *Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services*, pp. 231–240, 2012. 3

[16] J. Grubert, E. Ofek, M. Pahud, and P. O. Kristensson. The office of the future: Virtual, portable, and global. *IEEE computer graphics and applications*, 38(6):125–133, 2018. 2

[17] J. Grubert, M. Pahud, R. Grasset, D. Schmalstieg, and H. Seichter. The utility of magic lens interfaces on handheld devices for touristic map navigation. *Pervasive and Mobile Computing*, 18:88–103, 2015. 3

[18] J. Grubert and D. Schmalstieg. Playing it real again: a repeated evaluation of magic lens and static peephole interfaces in public space. In *Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services*, pp. 99–102, 2013. 3

[19] J. Guo, D. Weng, H. Fang, Z. Zhang, J. Ping, Y. Liu, and Y. Wang. Exploring the differences of visual discomfort caused by long-term immersion between virtual environments and physical environments. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 443–452. IEEE, 2020. 2, 3, 8

[20] J. Guo, D. Weng, Z. Zhang, H. Jiang, Y. Liu, Y. Wang, and H. B.-L. Duh. Mixed reality office system based on maslow's hierarchy of needs: Towards the long-term immersion in virtual environments. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 224–235. IEEE, 2019. 2, 3

[21] J. Guo, D. Weng, Z. Zhang, Y. Liu, and Y. Wang. Evaluation of maslows hierarchy of needs on long-term use of hmds–a case study of office environment. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 948–949. IEEE, 2019. 2

[22] J. Hartmann, C. Holz, E. Ofek, and A. D. Wilson. Realitycheck: Blending virtual environments with situated physical reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–12, 2019. 2, 8

[23] J. Hindmarsh and C. Heath. Video-based studies of work practice. *Sociology Compass*, 1(1):156–173, 2007. 3

[24] T. Hirzle, F. Fischbach, J. Karlbauer, P. Jansen, J. Gugenheimer, E. Rukzio, and A. Bulling. Understanding, addressing, and analysing digital eye strain in virtual reality head-mounted displays. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 29(4):1–80, 2022. 9

[25] G. Hripcsak and A. S. Rothschild. Agreement, the f-measure, and reliability in information retrieval. *Journal of the American medical informatics association*, 12(3):296–298, 2005. 3

[26] S. Kang, E. Shokeen, V. L. Byrne, L. Norooz, E. Bonsignore, C. Williams-Pierce, and J. E. Froehlich. Armath: Augmenting everyday life with math learning. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, p. 1–15. Association for Computing Machinery, New York, NY, USA, 2020. doi: 10.1145/3313831.3376252 3

[27] K. Kim, M. Billinghurst, G. Bruder, H. B.-L. Duh, and G. F. Welch. Revisiting trends in augmented reality research: A review of the 2nd decade of ismar (2008–2017). *IEEE transactions on visualization and computer graphics*, 24(11):2947–2962, 2018. 8

[28] Y. M. Kim, I. Rhiu, and M. H. Yun. A systematic review of a virtual reality system from the perspective of user experience. *International Journal of Human–Computer Interaction*, 36(10):893–910, 2020. 8

[29] G. Lee, J. Healey, and D. Manocha. Vrdoc: Gaze-based interactions for vr reading experience. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 787–796, 2022. doi: 10.1109/ISMAR55827.2022.00097 2

[30] H. Lee, S. Je, R. Kim, H. Verma, H. Alavi, and A. Bianchi. Partitioning open-plan workspaces via augmented reality. *Personal and Ubiquitous Computing*, pp. 1–16, 2019. 2

[31] J. Li, C. George, A. Ngao, K. Holländer, S. Mayer, and A. Butz. Rear-seat productivity in virtual reality: Investigating vr interaction in the confined space of a car. *Multimodal Technologies and Interaction*, 5(4):15, 2021. 2

[32] F. Lu and D. A. Bowman. Evaluating the potential of glanceable ar interfaces for authentic everyday uses. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, pp. 768–777. IEEE, 2021. 2

[33] E. Márquez Segura, K. Rogers, A. L. Martin-Niedecken, S. Niedecken, and L. T. Vidal. Exploring the design space of immersive social fitness games: The imsofit games model. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3411764.3445592 3

[34] M. McGill, D. Boland, R. Murray-Smith, and S. Brewster. A dose of reality: Overcoming usability challenges in vr head-mounted displays. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pp. 2143–2152, 2015. 2, 8

[35] M. Mcgill, A. Kehoe, E. Freeman, and S. Brewster. Expanding the bounds of seated virtual workspaces. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 27(3):1–40, 2020. 1, 2, 9

[36] D. Medeiros, M. McGill, A. Ng, R. McDermid, N. Pantidi, J. Williamson, and S. Brewster. From shielding to avoidance: Passenger augmented reality and the layout of virtual displays for productivity in shared transit. *IEEE Transactions on Visualization and Computer Graphics*, 28(11):3640–3650, 2022. doi: 10.1109/TVCG.2022.3203002 2

[37] X. Meng, W. Xu, and H.-N. Liang. An exploration of hands-free text selection for virtual reality head-mounted displays. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 74–81, 2022. doi: 10.1109/ISMAR55827.2022.00021 2

[38] Microsoft. Hololens 2 - overview, features and specs. 2023. Accessed: 2023-12-19. 8

[39] A. Ng, D. Medeiros, M. McGill, J. Williamson, and S. Brewster. The passenger experience of mixed reality virtual display layouts in airplane environments. In *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 265–274. IEEE, 2021. 2

[40] Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. Elan (version 6.4) [computer software]. 2022. Retrieved from `https://archive.mpi.nl/tla/elan`. 3

[41] R. Nordahl, N. C. Nilsson, A. Adjorlu, E. Magalhaes, S. Willemsen, N. S. Andersson, J. Wang, and S. Serafin. 12 hours in virtual reality: Two cases of long-term exposure to consumer-grade virtual reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 2019. 2

[42] E. Ofek, J. Grubert, M. Pahud, M. Phillips, and P. O. Kristensson. Towards a practical virtual office for mobile knowledge workers. *arXiv preprint arXiv:2009.02947*, 2020. 2

[43] J. O'Hagan and J. R. Williamson. Reality aware vr headsets. In *Proceedings of the 9TH ACM International Symposium on Pervasive Displays*, pp. 9–17, 2020. 2

[44] L. Pavanatto, C. North, D. A. Bowman, C. Badea, and R. Stoakley. Do we still need physical monitors? an evaluation of the usability of ar virtual monitors for productivity work. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, pp. 759–767. IEEE, 2021. 2

[45] J. Pretsch, E. Pretsch, J. Saretzki, H. Kraus, and G. Grossmann. Improving employee well-being by means of virtual reality–realex: an empirical case study. 2021. 2

[46] A. Ruvimova, J. Kim, T. Fritz, M. Hancock, and D. C. Shepherd. " transport me away": Fostering flow in open offices through virtual reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–14, 2020. 1, 2

[47] D. Schneider, A. Otte, T. Gesslein, P. Gagel, B. Kuth, M. S. Damlakhi, O. Dietz, E. Ofek, M. Pahud, P. O. Kristensson, et al. Reconviguration: Reconfiguring physical keyboards in virtual reality. *IEEE transactions on visualization and computer graphics*, 25(11):3190–3201, 2019. 2

[48] R. Shen, D. Weng, S. Chen, J. Guo, and H. Fang. Mental fatigue of long-term office tasks in virtual environment. In *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 124–127. IEEE, 2019. 2

[49] A. L. Simeone. The vr motion tracker: visualising movement of non-participants in desktop virtual reality experiences. In *2016 IEEE 2nd Workshop on Everyday Virtual Reality (WEVR)*, pp. 1–4. IEEE, 2016. 2, 8

[50] A. D. Souchet, D. Lourdeaux, A. Pagani, and L. Rebenitsch. A narrative review of immersive virtual reality's ergonomics and risks at the workplace: cybersickness, visual fatigue, muscular fatigue, acute stress, and mental overload. *Virtual Reality*, pp. 1–32, 2022. 2

[51] E. Southgate. Using screen capture video to understand learning in virtual reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 418–421, 2020. doi: 10.1109/VRW50115.2020.00089 3

[52] F. Steinicke and G. Bruder. A self-experimentation report about long-term use of fully-immersive technology. In *Proceedings of the 2nd ACM symposium on Spatial user interaction*, pp. 66–69, 2014. 2

[53] Y. Tao and P. Lopes. Integrating real-world distractions into virtual reality. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*, pp. 1–16, 2022. 2

[54] K. D. Thoondee and A. Oikonomou. Using virtual reality to reduce stress

at work. In *2017 Computing Conference*, pp. 492–499, 2017. doi: 10.1109/SAI.2017.8252142 2, 9

[55] D. Valtchanov, K. R. Barton, and C. Ellard. Restorative effects of virtual nature settings. *Cyberpsychology, Behavior, and Social Networking*, 13(5):503–512, 2010. 2, 9

[56] C.-H. Wang, B.-Y. Chen, and L. Chan. Realitylens: A user interface for blending customized physical world view into virtual reality. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*, pp. 1–11, 2022. 2, 8

[57] J. Wentzel, G. d'Eon, and D. Vogel. Improving virtual reality ergonomics through reach-bounded non-linear input amplification. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–12, 2020. 9

[58] L. E. Whitman, M. Jorgensen, K. Hathiyari, and D. Malzahn. Virtual reality: its usefulness for ergonomic analysis. In *Proceedings of the 2004 Winter Simulation Conference, 2004.*, vol. 2, pp. 1740–1745. IEEE, 2004. 9

[59] Q. Zhou, G. Fitzmaurice, and F. Anderson. In-depth mouse: Integrating desktop mouse into virtual reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pp. 1–17, 2022. 2

Chapter 8

Working with XR in Public: Effects on Users and Bystanders

Summary

As already mentioned before, XR could be especially useful to support mobile knowledge work. Because many mobile spaces are also public, it is important to know how using XR for work in public is currently perceived by both the user and also by people around them, called bystanders. Therefore, this article reports on a study conducted in a public university cafeteria in which participants were performing a task once using only a laptop, once using a combination of a laptop with an augmented reality (AR) HWD, and once using a combination of a laptop and a VR HWD.

The results indicate that XR users feel more safe if they can see their physical environment, which was the case in the laptop condition and partially in the AR condition. Also, using XR in public still makes you stand out. However, these findings could also change in the future as XR becomes more common.

Contribution Statement

The basic study design was discussed between all authors and subsequently refined by Verena Biener. Then, Verena Biener also constructed and tested the study setup. Verena Biener and Snehanjali Kalamkar each conducted around half of the study sessions. The collected data was processed by Verena Biener with the help of Snehanjali Kalamkar. The results were discussed by all authors. Verena Biener wrote the first draft of the article and created the figures with help of Snehanjali Kalamkar. All authors subsequently refined the article.

Article [15]

© 2024 IEEE. Reprinted, with permission, from Verena Biener, Snehanjali Kalamkar, John J. Dudley, Jinghui Hu, Per Ola Kristensson, Jörg Müller & Jens Grubert. Working with XR in Public: Effects on Users and Bystanders. In IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), 2024, pp. 779-780. `https://doi.org/10.1109/VRW62533.2024.00186`.

Working with XR in Public: Effects on Users and Bystanders

Verena Biener[1]* Snehanjali Kalamkar[1] John J. Dudley[2] Jinghui Hu[2] Per Ola Kristensson[2]
Jörg Müller[3] Jens Grubert[1]†

[1]Coburg University of Applied Sciences and Arts, Germany
[2]University of Cambridge, United Kingdom
[3]University of Bayreuth, Germany

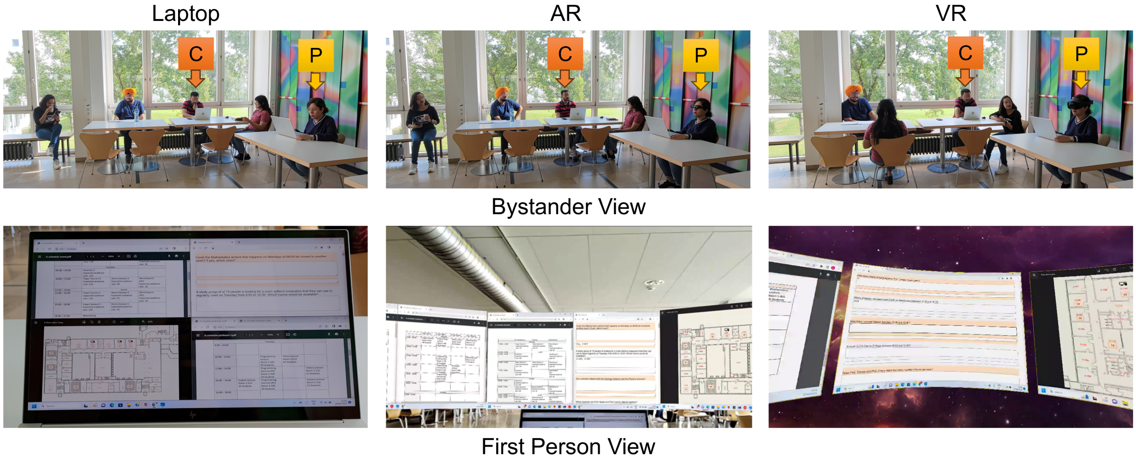

Figure 1: The top row shows the view of a bystander passing the area of the study for all three conditions (LAPTOP, AR, VR). C indicates the coordinator of the study and P indicates the participant. All other people are bystanders. The illustrations are reenactments of the actual situations. The bottom row shows the participant's views during the task for each condition.

ABSTRACT

Recent commercial virtual and augmented reality (VR / AR) devices have been promoted as tools for knowledge work and research findings show that they can be beneficial. One benefit is the possibility to display virtual screens which could be especially helpful in mobile contexts, in which users might not have access to an optimal physical work setup. Such situations often occur in public settings, for example when working on a train while traveling to a business meeting. However, using such devices in public is still uncommon, which motivates our study to better understand the implications of using AR and VR for work in public on the user itself, and also on bystanders. Therefore, we present initial results of a study in a university cafeteria comparing three different systems: a laptop with a single screen; a laptop combined with an optical see-through AR headset; a laptop combined with an immersive VR headset.

Keywords: Virtual Reality, Augmented Reality, Extended Reality, Knowledge Work, Public Environment

1 INTRODUCTION

Using extended reality (XR) for knowledge work has recently gained popularity through the promotion of current commercial off-the-shelf devices as tools for work, such as the Meta Quest Pro and the Lenovo Think Reality glasses. With XR we refer to both augmented reality (AR) and virtual reality (VR) technologies. Prior research has also demonstrated the potential benefits of using XR for knowledge work, such as how VR can be used in open office spaces to reduce distractions [7] and how new interaction possibilities provided by VR can improve the performance for certain tasks [3, 2]. In addition, researchers have evaluated how the large display space provided by AR and VR can be used in the context of knowledge work [5, 6]. These properties of XR seem especially advantageous in mobile scenarios, where work environments can be less optimal, such as in crowded spaces with many distractions, or in confined spaces that limit the size of physical hardware [4], for example, in public transportation or in a café. However, prior studies on supporting knowledge work have focused on new technologies evaluated mainly through laboratory-based experiments which are typically not representative of public spaces. Therefore we see the need to explore how users of XR devices *experience* working with these devices in public spaces, and also to investigate how people around them react to the currently relatively rare sight of XR-users in public.

2 METHODOLOGY

The study was conducted as a within-subjects design with one independent variable INTERFACE having three levels, LAPTOP, AR and VR. LAPTOP was chosen as the baseline, as it is commonly used for working in public spaces. For AR and VR, we chose commercially available headsets with current software that can connect to a laptop as an input device while displaying multiple virtual screens. Specifically, we used the Meta Quest Pro in combination with the application "Immersed" [1] for VR and the Lenovo Think Reality glasses A3 with its built-in display manager in AR. Both devices were designed and promoted by their respective manufacturers as

---

*e-mail: verena.biener@hs-coburg.de
†e-mail: jens.grubert@hs-coburg.de

tools for knowledge work. For all conditions we used a 16-inch HP Envy laptop and an external mouse. While participants could only use the laptop screen in the LAPTOP condition, they could use three virtual monitors in AR and VR (see Fig. 1).

The study was conducted in three sessions, one for each condition. All sessions started in a lab in which participants were introduced to the device and completed a 20 minute training task. Then, participant and coordinator moved to different places in the public cafeteria and the participant completed a task for 35 minutes while the coordinator was observing the situation (see Fig. 1). Afterwards participants answered various questionnaires and also several bystanders were asked to fill out a short online questionnaire. Then, participant and coordinator returned to the lab for a short interview. Three similar but different tasks were used for the three conditions which were counterbalanced along with the order of the conditions. Eighteen participants (all students or employees at our university) successfully completed the study (5 female). Their mean age was 25.67 years ($sd = 5.2$). The number of bystanders was counted at the beginning and end of each session resulting in a minimum of 7 and a maximum of 37. In total, we collected 231 responses from bystanders, given by 209 different people (104 female, 103 male, 2 others).

3 RESULTS

In the following, we will present initial findings of our study. We used a repeated measures analysis of variance (RM-ANOVA) to analyze the data obtained through the study. For subjective data, we used aligned rank transform (ART) [8] and for post-hoc comparisons we applied Bonferroni adjustments at an initial significance level of $\alpha = 0.05$. In the following we use m to abbreviate the arithmetic mean and sd for standard deviation. We used axial coding to structure statements from free-text-fields and interviews.

Safety and Isolation We found a significant main effect of INTERFACE ($F(2,34) = 16.824, p < 0.001, \eta_p^2 = 0.497$) on participants' feeling of safety rated on an 11-point Likert scale (1 = threatened, 11 = safe). Post-hoc tests indicate that they felt significantly more safe in LAPTOP ($m = 8.44,\ sd = 2.2$) as compared to both AR ($m = 7.61,\ sd = 2.03$) ($p = 0.04$) and VR ($m = 5.94,\ sd = 2.67$) ($p =< 0.001$). In addition, participants felt significantly safer in AR than in VR ($p = 0.009$). Participants also stated that they liked that they could see the real environment in AR (P5, P15, P21, P22, P24) which made them feel more secure (P5) and some participants (P16, P18, P24) did not like being unaware of their physical environment in VR. These combined findings might indicate that participants' feeling of safety is related to how well they can perceive their physical environment.

Only P5 used full passthrough throughout the VR condition, and P9 switched to passthrough occasionally. All other participants chose one of the virtual environments, in which they felt comfortable or which they thought looked nice. Fourteen participants did not even use a passthrough window in VR, except for the keyboard, as they thought it was distracting (P4, P14, P16), was not needed (P7, P11, P17, P18, P19, P23), felt more focused and immersed without passthrough (P23, P24) or felt like they had less space for the task (P12). However, P6 regretted not using passthrough, as she wanted to see who was talking. Of the three participants who used a passthrough window, P8 moved it out of sight as it interfered with the task, and P9 and P22 had one on both sides but made little use of it during the task.

Standing out We also found a significant main effect of INTERFACE ($F(2,34) = 6.214, p = 0.005, \eta_p^2 = 0.268$) on participants' rating of feeling more observed (1) or unobserved (11). Post-hoc tests indicate that they felt significantly less observed in LAPTOP ($m = 7.22,\ sd = 2.76$) as compared to VR ($m = 4.78,\ sd = 2.67$) ($p = 0.004$). P23 also stated that he felt people were looking at him in VR. From observing the bystanders, we know that this feeling is valid. We observed that no bystander seemed to have any interest in the participants during the LAPTOP condition, as they were blending in well with other people working on their laptops. During the AR and VR conditions, individual bystanders stared at the participants for multiple seconds, some also stared repeatedly and some were obviously talking about the participant. However, the majority of bystanders either did not, or pretended not to notice the participants or only looked at them in a very subtle way. Some seemed to find excuses to look such as when they walked past the participant. In addition, while in the LAPTOP condition only 65% indicated that they noticed the participant, compared to 93% and 95% for AR and VR respectively. Fisher's exact test revealed that there is a significant effect of INTERFACE on whether bystanders noticed the participant or not ($p < 0.001$), with post-hoc tests showing that in the LAPTOP condition they noticed the participant significantly less often than in AR ($p < 0.001$) and VR ($p < 0.001$). These findings indicate that XR-users stand out, and that users are sensitive to that, especially in VR. However, the reactions of the bystanders remain mostly subtle and we did not observe any actions that could have been dangerous for the participants.

4 CONCLUSION

With the goal of observing how using XR devices in public affects the users and also the bystanders, we conducted a study in which participants used an AR and VR HMD as well as a LAPTOP in a university cafeteria. Initial findings indicate that, at the time of the study, using XR in public still makes users stand out, yet we observed mainly subtle reactions from bystanders and this behavior might change if XR becomes more widespread in the future. Our findings also suggest that seeing the physical surrounding can increase users' feeling of safety, yet, 14 out of 16 participants stated that passthrough windows in VR were unnecessary or distracting. Therefore, more research is required to integrate passthrough in a sensible way.

In future work, we will analyze the full data-set collected during our study, including a range of questionnaires and interviews from the XR users and more detailed descriptions of how bystanders experienced the situations. This will allow us to present a more complete picture of the current perception of XR use in public as well as possible challenges.

REFERENCES

[1] Immersed. `https://www.immersed.com/`. Accessed: 2023-10-05. 1

[2] V. Biener, T. Gesslein, D. Schneider, F. Kawala, A. Otte, P. O. Kristensson, M. Pahud, E. Ofek, C. Campos, M. Kljun, K. Č. Pucihar, and G. Jens. Povrpoint: Authoring presentations in mobile virtual reality. *IEEE Trans. on Visualization and Computer Graphics*, 2022. doi: 10.1109/TVCG.2022.3150474 1

[3] T. Gesslein, V. Biener, P. Gagel, D. Schneider, P. O. Kristensson, E. Ofek, M. Pahud, and J. Grubert. Pen-based interaction with spreadsheets in mobile virtual reality. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 361–373. IEEE, 2020. 1

[4] J. Grubert, E. Ofek, M. Pahud, and P. O. Kristensson. The office of the future: Virtual, portable, and global. *IEEE computer graphics and applications*, 38(6):125–133, 2018. 1

[5] M. Mcgill, A. Kehoe, E. Freeman, and S. Brewster. Expanding the bounds of seated virtual workspaces. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 27(3):1–40, 2020. 1

[6] L. Pavanatto, C. North, D. A. Bowman, C. Badea, and R. Stoakley. Do we still need physical monitors? an evaluation of the usability of ar virtual monitors for productivity work. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, pp. 759–767. IEEE, 2021. 1

[7] A. Ruvimova, J. Kim, T. Fritz, M. Hancock, and D. C. Shepherd. "transport me away": Fostering flow in open offices through virtual reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–14, 2020. 1

[8] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 143–146, 2011. 2

Chapter 9

Discussion

In the following sections, I will summarize and discuss the results presented in chapter 3 [17] to 8 [15]. First, I will discuss the results related to the first area of research questions about facilitating mobile knowledge work using the advantages of VR, such as a large 3D display space and multimodal interaction opportunities. Then, I will discuss how current off-the-shelf VR and XR devices affect knowledge workers during prolonged use and in public environments. Finally, I will also reflect on the limitations of the work presented in this thesis and future research directions.

9.1 Facilitating Mobile Knowledge Work Using VR

Chapter 4 [18] and 5 [14], had a closer look at how mobile knowledge work can be facilitated using certain affordances provided by VR. We have seen two studies that quantify the benefits of utilizing 3D visualizations in knowledge worker tasks and saw that they can outperform existing 2D techniques. In addition, chapter 5 [14] also quantified the effects of the increased display space in VR in the context of a visual search task, showing that an increased field of view reduces the search time significantly if the target is preattentive and, therefore, easy to find among other images. Chapter 4 [18] also presented a study that shows how a multimodal interaction approach that combines touch and eye-tracking has benefits, such as a higher usability and faster task completion times, over a touch-only interaction. Subsequently, chapter 4 [18] described several example applications that utilize the large 3D space and multimodal interaction techniques. Using similar techniques, chapter 5

[14] proposed several VR based techniques for certain presentation authoring tools. In the following, I will discuss these results in more detail.

9.1.1 Advantages of a Large 3D Display

We have seen in two studies, presented in chapter 4 [18] and 5 [14], that the 3D display in VR and, therefore, the three-dimensional representation of information can make certain tasks more efficient. In the puzzle task study in chapter 4 [18], we have seen that participants were around 15% faster when using the depth visualization compared to using a flat 2D visualization. Very similar results were presented in the reordering study of chapter 5 [14]. The 3D VR technique for reordering layers of information was significantly faster than both baseline techniques (69% and 72%), which were established in current PowerPoint and PowerPoint for Mac versions.

In both studies, participants also reported a significantly higher usability for the VR techniques. For the puzzle task, we could also observe a lower mental demand and lower simulator sickness for the condition that used the depth visualization. For the reordering task, a significantly higher task load was reported for one of the baseline techniques, namely the dynamic reordering. No significant difference regarding simulator sickness was found for the reordering task, which can be seen as a good sign when comparing to non-VR baselines. In both studies, I received feedback from the participants on why they liked the VR techniques with depth visualization. They found it “provides more information” [18] (chapter 4) and it is easier to identify mistakes [18] (chapter 4), “if something was occluded you could move your head” [14] (chapter 5), and it provides a "faster overview" [14] (chapter 5).

These results show how displaying information in 3D using VR devices can be advantageous. Of course, the kind of 3D visualizations presented in this thesis are directly applicable to scenarios in which there is layered or overlapping information or information that has some sort of three dimensional relation. Such scenarios could, for example, be found in the medical area with data obtained from CT scans, in architecture when looking at layouts of multiple floors of a building, or when creating digital pictures, which often involves layers containing different elements of the picture. It remains to be seen if and how other kind of information can benefit from such a representation. In chapter 5 [14], we have seen that a 3D representation in VR can also outperform current well-established 2D techniques implemented in current office applications. However, there is also a possibility that these established 2D techniques will be advanced in the future to catch up with the 3D version. Yet, considering that this 3D technique is an early-stage prototype, it is likely that it can also be further optimized. Amongst others, possible improvements could be made by refining the distance between information layers or by incorporating more visualization techniques, such as transparency, colors, or varying sizes. Such improvements will not be trivial and will most likely depend on the type of information and the number of layers. The goal, however, should not be to promote 3D over 2D applications. Therefore, combining their strengths regarding different aspects of knowledge worker tasks could be a promising avenue.

Even though this work focused on mobile work, the results indicate that VR can not only support knowledge work in a mobile context with limited space and hardware but also in standard office settings. This means that certain tasks that benefit a lot from a three-dimensional representation could be completed more efficiently in VR, also outside of the mobile context. This could not only be the case for 2D data with three-dimensional relations but also for a range of tasks involving inherently three-dimensional data, such as modelling [91], sketching [28] or creating animations [7, 113, 21]. This is also supported by a survey from Berg and Vance [12] that shows how VR is successfully used as a tool in various industries for use-cases such as exploring ergonomics, evaluating aesthetics, or visualizing abstract data.

In addition to the benefits of a 3D representation, chapter 5 [14] also quantified the benefits of the larger screen space in VR on a search task. For this purpose, the large screen space in VR was compared to a smaller screen space in VR and a standard mobile display - in this case a tablet. Prior work has already suggested that a larger field of view can speed up target detection [88]. The search task study in chapter 5 [14] showed that for an easy search task, a so-called preattentive search, VR with it's full field of view makes the search significantly faster than VR with only a limited field of view or a tablet. A preattentive search means that the target image stood out from the rest of the images because it was the only colored one. For such a search, VR with using the full field of view was around 37% faster than VR with only a limited field of view and around 52% faster than the tablet. However, for harder tasks, in which all images were colored, the results did not indicate significant differences between the conditions. For hard tasks, the target did not stand out from the rest, and therefore, participants needed to look at each image carefully, which seems to strongly diminish the advantage of a large field of view. Still, it can be concluded that the larger field of view is at least as good as a smaller one and in certain situations definitely outperforms it, such as when searching for a specific image or, more generally, a specific piece of information.

In addition, there might be usability benefits provided by the larger field of view, as there is no need for additional interactions to scroll through the information. The study also indicated potential usability benefits of VR through the built-in eye-tracking. We saw that for the easy tasks VR with the same field of view as the tablet was still faster than the tablet. This can be explained by the use of eye-tracking in combination with touch input. In contrast, the tablet condition used only touch input, which required a lot of scrolling. The benefits of combining touch and eye-gaze interaction have also been studied in detail in chapter 4 [18] and will be discussed in the next subsection. In addition to shorter search times, the large display space provided by VR can also provide the benefit of displaying more information at once. This has not been explored in-depth by the presented studies, but the concept was implemented in the prototype of chapter 5 [14] by visualizing multiple sources of information, such as a PDF, an image search, or multiple presentation-slides, and participants rated it to be useful. Pavanatto et al. [82] have investigated this more closely in AR and conclude that, due to technical constraints of the hardware, the performance and usability of physical screens is currently higher than that of virtual screens. Therefore, they recommend to combine both physical and virtual screens. This would allow for a larger display space in situations where multiple physical monitors are not available, while also having a familiar high resolution physical screen.

Overall, both the large screen space as well as the 3D display have shown to be able to facilitate certain knowledge worker tasks, particularly tasks that naturally consist of layers of information or benefit from a presentation in multiple layers. The large display space is of course especially helpful in restricted spaces or wherever additional hardware such as displays is not feasible. Similarly, it has been shown that the 3D display can make certain tasks more efficient, even with only a small interaction device such as a tablet available. In addition, the 3D display also has the potential to support knowledge work in non-mobile situations due to visualization options not possible on a standard display.

9.1.2 Multimodal Techniques that Combine VR and Established Devices

When creating new interaction techniques and applications as described in chapter 4 and 5, I focused on a small input area, as the main focus was on supporting mobile knowledge work. As mentioned before, in mobile contexts the physical space is often restricted and interacting with novel devices also poses the question of social acceptability [65]. Therefore, common VR interaction techniques such as in-air gestures might not be well-suited, as they require a certain amount of free space, and additionally, they might induce fatigue [47]. For these reasons, I decided to combine VR with touch interfaces, namely tablets, and therefore, the input area is restricted to the small space on the tablet. The first study, presented in chapter 4 [18], was conducted to evaluate how a combination of touch and eye-gaze could enhance the interaction across multiple screens compared to a bimanual, touch-only technique. We saw that the gaze-based technique was about 30% faster than the bimanual technique and all participants also perceived it as being faster. In addition, the gaze-based technique resulted in a significantly higher usability rating and a lower perceived task load. Therefore, all but one participant preferred the gaze-based approach. This shows that eye-gaze can be an important addition to the set of input techniques for doing knowledge work in VR, especially when making use of the large screen space. The value of gaze-based techniques in VR has also been indicated in other studies such as by McGill et al. [64], who manipulated the virtual display position depending on the users gaze-direction to allow them to use a large display space with less head movement. Other examples are provided by Meng et al. [70] who explored techniques for hands-free text selection or Lee et al. [59], who proposed a technique that uses gaze to make reading in VR more efficient and less demanding.

Both in chapter 4 [18] and chapter 5 [14], I have presented a range of examples on how multimodal interaction involving a VR HWD, touch on a tablet surface, gaze, and spatial interaction can be realized. For all of them, I conducted a usability evaluation in which participants could try the techniques and then rate their ease of use, utility, and enjoyment. Overall, these applications received positive ratings. Similar results were also reported by Gesslein et al. [37] through a video-based survey of multimodal techniques involving VR, gaze, a tablet, and a spatially tracked pen in the context of spreadsheets. Interestingly, I observed that the ease of use of the presentation authoring tools in chapter 5 [14] were also rated pretty high, considering that the presented applications and the interaction with it was totally new to the participants, and therefore, they had to learn and remember a lot in a very short time. These ratings acknowledge that combining a variety of interaction modalities is feasible in the context of VR for knowledge work. In chapter 5 [14], the participants also rated the usability of the combined feature set for authoring presentations to be average (71.53 on the system usability scale [19]), which is promising for a first prototype. While participants found the presented techniques for object manipulation and animation quite complex and they said it requires some practice, they also mentioned that it can be useful and intuitive. One participant even mentioned that "with a little practice it would be faster than in standard PowerPoint" [14] (chapter 5). This is also in line with the results of the reordering task, discussed in chapter 9.1.1, which showed that the VR technique outperforms current state-of-the-art techniques. Still, these comments highlight that users need time to learn a new technique. This process could be supported by well-designed tutorials that help users to get the most out of a new technique as fast as possible. Overall, participants see value in the presented applications and many could also imagine using such a system in the future.

In chapter 5 [14], we have seen a selection of tools for presentation authoring and chapter 4 [18] proposed a variety of possible applications, including a window manager, code version control, a parallel coordinate plot, a map, and medical imaging. However, these kinds of techniques can be also transferred to many more application areas, such as already shown by Gesslein et al. [37] for interacting with spreadsheets.

Even though I have made the design choice to evaluate the presented techniques using VR HWDs, I believe that they are also viable in other XR devices. For example, the space around the tablet could also be augmented with additional visualizations in AR. Using VR HWDs had the advantage of a much larger field of view than that of optical see-through AR devices and, at the time of the studies, the passthrough mode of mixed reality HWDs was rather poor (such as Meta Quest 2). When using the techniques for current AR HWDs with a much smaller field of view, some adaptation might be needed. To avoid large head movements to see information distributed on the large virtual screen, techniques such as proposed by McGill et al. [64] could be employed that change the display position depending on the gaze angle. Also, the size of virtual screens or elements could be adapted to fit the field of view of a certain device. Additionally, instead of fully replacing the physical screens, a combination of physical and virtual screens could be used in AR, as suggested by Pavanatto et al. [82]. Even though displaying virtual along with physical content could lead to perceptual issues [31], directly involving the physical screen can provide a higher resolution than is possible on current XR devices.

Overall, we have seen that it is indeed feasible to combine touch-input with VR, which can be especially useful in small constraint environments. This can then be enhanced by combining the touch input with a spatially tracked pen or eye-gaze to retarget the touch input on the tablets surface, which enables the user to interact with the large display space in VR while staying within the limited interaction space on or above the tablet. For other use cases, the tablet could of course also be replaced with other portable devices such as the touchpad of a laptop, which would then also allow the use of a physical keyboard for efficient text-entry. Using any kind of small portable device as the basis for interacting, solves the problem of the limited interaction space mentioned in the beginning of this section. It makes the interaction more subtle and, therefore, potentially more socially acceptable, and also reduces the risk for fatigue induced by large gestures.

9.2 Effects of Current VR Devices on Knowledge Workers in Realistic Use-Cases

The previous sections have looked at possible advantages that VR can provide, such as utilizing the large 3D display space and novel multimodal interaction concepts. However, the studies presented in Chapters 4 [18] and 5 [14], as well as the majority of other studies in VR or XR, are rather short in duration. For example, the studies from Chapters 4 [18] and 5 [14] lasted for less than two hours. In addition, they are often conducted in lab environments and engage the participants in made-up tasks and, therefore, lack ecological validity. Hence, Chapters 6 [16] and 7 [13] considered the effects of working in VR for a much longer time-period, namely a whole workweek. In addition, the participants were able to continue their actual work tasks, which increases the ecological validity of the study. And finally, Chapter 8 [15], investigated the effects of using both VR and AR devices in an actual public space instead of a controlled lab environment.

The study presented in Chapter 6 [16] showed that the baseline costs of VR are higher than for the physical setup for nearly all measures. However, there were also indications that participants got used to VR, which could also be concluded by the video analysis presented in Chapter 7 [13]. In general, both Chapters 6 [16] and 7 [13] provide unique insights into how users are affected by using VR for prolonged times.

As already mentioned, Chapter 8 [15] provided insights into the use of XR in a mobile context, by observing XR users and bystanders in a public space. The results showed that using XR can influence the participants' feeling of safety with significant differences between AR and VR and that currently XR still makes users stand out.

A detailed discussion of the results obtained from the one-week-study and the study conducted in a public space is presented in the following subsections.

9.2.1 Knowledge Workers Response to Prolonged Use of VR

The results presented in Chapter 6 [16] showed that the usability of the VR system was rated around 36% lower than that of the physical system. The average usability score for VR was around 54, which is below average but can still be considered ok [11]. In addition, the reported frustration was around 42% higher in VR and task load was rated 35% higher in VR compared to the physical system. However, the average task load of around 46 was still within the 50th percentile of computer activities [39], which can be interpreted such that the task load was still on a level that represents normal computer activities. In a study that compared VR and AR HWDs with a tablet, Li et al. [60] also found increased perceived workload when using HWDs and in a study lasting for one day (eight hours), Shen et al. [99] also reported a higher mental fatigue for VR than for a physical system. Shen et al. speculated that participants achieve a higher engagement level in VR and, therefore, spend more attention resources on a task. This theory could be supported by a statement of P12, as reported in Chapter 6 [16], who said that he could work really focused for 45 minutes, but then suffered from a headache during the next 3 hours. More research on this topic is needed, but if VR could enable users to spend more attention resources on a task, it could be a valuable tool for knowledge workers in situations that require them to be highly concentrated or that are highly time critical. Yet, these studies indicate that such intense work can also lead to higher mental fatigue and should therefore be carefully dosed. In contrast to this theory, the participants perceived their productivity in VR as lower in the study presented in Chapter 6 [16] and also report a significantly lower flow in VR, which indicates that participants did not feel more focused while using VR. However, the full potential of VR, which includes shielding the user from distractions in the physical environment and potentially make them work more focused, could not be explored in this study, as in both conditions the participants were sitting in a calm environment and, therefore, there were barely any distractions present. Still, indications that VR could increase flow and performance by shielding the user from a busy environment were found by Ruvimova et al. [94].

For other subjective measures, Chapter 6 [16] also reported significantly worse ratings for VR compared to the physical system. This includes higher ratings for negative affect, anxiety, simulator sickness and visual fatigue, and lower ratings for wellbeing. The exact cause for these bad ratings need to be explored in dedicated future studies. From the observations made in the presented study and specifically the interview answers, I can only infer that participants generally were less happy and felt less comfortable in VR.

The average simulator sickness score of around 34 reported in Chapter 6 [16] can be considered bad [103] which is alarming because a widespread consumer device was used. Also, due to the predominantly static VR scene (neutral background, a screen, and desk with a keyboard), I expected rather low simulator sickness ratings. Therefore, this needs to be further investigated to make out the main factors contributing to these ratings and propose changes in hardware and software accordingly. As the average simulator sickness ratings also slightly increased during the day for the physical condition, it is likely that the high values in VR could at least partially be explained by factors that are not directly related to VR. However, at this point, it is important to note that two additional participants dropped out of the study during the first hours of the VR condition due to nausea, anxiety, and a migraine. This might be due to the known drawbacks of current HWDs or about these participants general susceptibility for VR sickness [22]. Such people should also be kept in mind when thinking of a future in which VR could replace offices.

In addition to the higher visual fatigue in VR compared to the physical setup, the results indicate that the average visual fatigue increased during the day. Similarly, Guo et al. [42] found an increase over time regarding subjective eye-strain and difficulty in focusing during their eight hour long study. Also, Souchet et al. [102], who reviewed previous work regarding office-like tasks in VR, report that using VR can lead to higher visual fatigue as well as muscle fatigue, acute stress, and mental overload. On the other hand, Wille et al. [116], who conducted a four hour study in VR, found that subjective ratings suggest a higher strain and no habituation, yet their objective strain measures indicated habituation effects. Therefore, they hypothesize that the unfamiliarity with a technology could influence subjective ratings. Nevertheless, when working in VR, it would be important to consider including eye exercises to mitigate eye strain [48] or relaxation methods to reduce stress [108, 112] to increase users overall wellbeing.

The video-analysis in Chapter 7 [13] also revealed that participants took the HWD off less often at the end of the week, yet the duration of not wearing the HWD significantly increased during the week, indicating that participants took fewer but longer breaks. Taking the HWD off could often be associated with participants standing up, using their phone, rubbing their face, or drinking. Rubbing the face or standing up are actions that are only possible after removing the HWD. However, this is not true for other actions.

For example, 95% of the time spent drinking or eating and 92% of the time spent using the phone participants were wearing the HWD. Therefore, these can not be considered main reasons for taking it off. Also other actions that are connected with the physical environment were mostly done while wearing the HWD, such as talking or phone calls. I speculate that participants also took off the HWD because it got heavy or they felt tired and needed a break. Especially at the end of the week, they might also have been less motivated to wear the HWD, as it was not new and exciting anymore.

In Chapter 7 [13], we also saw that supporting the HWD was less frequent on day 5, yet no difference was found regarding the time spent supporting. Therefore, I speculate that participants got more used to supporting and did not repeatedly try not to support the HWD. Participants who supported the HWD a lot explained this through the need of getting a sharp image without the pain induced by fixing the HWD in a very tight position. However, regularly supporting the HWD is not optimal because one of the users hands is always occupied holding the HWD. Therefore, this reinforces the need to further improve the form factor of current HWDs.

Interestingly, I found a very distinct difference between male and female participants, such that females were supporting the HWD significantly more often and for longer times. On the other hand, I found that male participants rubbed their eyes and faces more frequently than female participants. Prior work by Guo et al. [42] also reported more severe visual and physical discomfort for female participants. They hypothesize that this difference could be explained by male participants having more VR experience and generally spending more time in front of a display. However, in the studies presented here, no significant effect of VR experience on the supporting behavior could be detected. Also, it is important to note that both Guo et al.'s study and the one presented in Chapter 6 [16] have a rather small sample size, which means these results could be due to differences between individuals. Therefore, future work should more closely inspect gender differences when evaluating the effects of VR and also other XR systems, especially for prolonged usage durations.

In the scope of Chapter 6 [16], participants also reported on problems with drinking or when they wanted to write something on paper while wearing the HWD. These problems were explored more closely through the video analysis in Chapter 7 [13]. We saw that participants were eating and drinking less frequently in VR, yet there was no significant difference regarding the total time spent on these actions. Therefore, I speculate that overall, participants did not neglect these needs but were more careful when eating or drinking in VR. This is also supported by the fact that for 94% of the time spent drinking or eating participants were actually wearing the HWD and some participants reported they were afraid to spill something. The video analysis also showed that the frequency of the participants rubbing their faces or eyes was less in VR which can be explained by it being less convenient, as the face is less accessible. In addition, there is less of a chance that participants do such actions unconsciously. Furthermore, as already mentioned, participants had problems to write something on paper and, therefore, such interactions with physical objects, also including the usage of phones, were less frequent in VR. Because of these observations, I believe that future HWDs should provide easier access to the users face to facilitate actions such as drinking, rubbing one's eyes, or scratching the nose. An example for a HWD that already allows easier access to the face is the HoloLens 2 [73] with its raisable visor. Additionally, interactions such as writing on a paper or using a phone could be supported through an integration of the physical world into VR. This will be discussed in more detail in subsection 9.2.2.

Even though Chapter 6 [16] generally showed worse ratings for VR than the physical system, participants also saw advantages of using VR for work. Some reported that they liked the isolation in VR, specifically in combination with earphones, but admitted it could also be scary to be totally cut off from the physical world. Others also valued the privacy. All participants said they could imagine working in VR in the future, yet only if certain improvements are realized, such as more lightweight HWDs, similar to prescription glasses. A similar finding was reported by Lu et al. [61], who conducted an in-the-wild study of glanceable AR. Such responses indicate that XR is a promising tool for knowledge work, yet there are still technical limitations that restrict a more widespread use.

In general, Chapter 6 [16] quantified the effects of using an off-the-shelf VR environment for work during a whole week and can serve as a baseline for future optimized systems. It showed that the baseline costs for doing knowledge work in VR are higher for most measures as compared to a physical setup. In addition, Chapter 7 [13] uncovered a range of behavioral patterns during such a prolonged VR use that were not previously documented. It showed that VR restricts normal behavior such as eating or rubbing the face, yet participants also seem to cope with the HWD as adjustments decrease over time and participants do not take the HWD off for certain actions such as talking, drinking, or using a phone. Nevertheless, the presented results show that the hardware limitations of current HWDs still implicate some of the main concerns when using VR for knowledge work. Therefore, HWDs should be improved further to allow a more comfortable experience and be less restrictive when it comes to personal needs.

9.2.2 Implications on Users and Bystanders when Using XR for Work in Public

The implications of using XR in public, on both the users and the bystanders, were examined in Chapter 8 [15]. We saw that participants felt safest in the laptop condition, followed by AR, and felt least safe in VR. Looking at this in concert with participants comments saying that they liked to see the physical environment in AR and that they disliked not seeing the physical environment in VR, I hypothesize that being able to see the physical surroundings can increase the feeling of safety. These findings are related to findings of Li et al. [60], who reported that participants using HWDs had a lower awareness of their physical environment in comparison to using a tablet and O'Hagen et al. [79], who found that VR users feel uncomfortable if they are not aware of the position of bystanders. Also Eghbali et al. [30] reported that, amongst others, a sense of safety and the freedom to switch between realities is an important factor influencing the experience of using VR in public. It was also already mentioned by McGill et al. [63] that it is important to include certain real world things to VR. Nevertheless, most participants in the study described in Chapter 8 [15] did not use passthrough mode or passthrough windows in VR. They thought it was not needed, it was distracting, or it reduced the space available for the work task. Yet, there are also participants who regretted not using it. The experience from this study suggests that users should either get more guidelines on how to efficiently use currently available possibilities of displaying real-world content in VR or other techniques should be offered to integrate the physical surrounding in a better way. There are various approaches to that, such as real-time 3D reconstructions of the physical environment [45], a customizable physical world view [114], adding cues from reality such as people and furniture to the virtual world [10], notifying the user that bystanders are present [79, 66], or displaying bystanders by using depth cameras [100]. Other approaches try to include certain objects or devices, such as smartphones [3, 8, 26], physical keyboards [63, 54, 57], other people [63, 38], or they even substitute the whole physical environment with a virtual one while preventing the users to hit obstacles [119, 24]. George et al. [36] presented a design space that covers a range of possibilities for combining the real and virtual world.

The goal for VR systems should be to allow the user to integrate as much of the physical world into the virtual one as is needed to make the user aware of safety relevant objects and actions from the real world and also other aspects that would help the user to feel comfortable and productive. This could include food or work items, as mentioned in subsection 9.2.1, or relevant personal belongings.

We saw that bystanders noticed the participants less during the laptop condition compared to the AR and VR conditions. This is also perceived by the users as they reported to feel more observed when using VR than when using only a laptop. Even though many bystanders where not too familiar with the XR devices, they did generally not behave in a rude or negative way. Most bystanders were very subtle when looking at the XR users, even though some of them were obviously staring or talking to each other about what they see. This means XR currently still makes users stand out. However, these findings are only representative of the time of the study. If XR becomes more wide-spread in the future, the observed behavior might change.

Overall, Chapter 8 [15] revealed that further improvements of XR systems are required to make users feel safe in public environments and to provide them with the necessary information about their environment. On another note, it is encouraging that while XR users still stand out, bystanders mostly reacted in a very subtle and positive way.

9.3 Limitations and Future Work

When designing the studies presented in this thesis, certain design decisions had to be made that can influence the results and impose certain limitations. In the following paragraphs I will reflect on these limitations and propose future research directions to overcome them. In addition, I will point out further research opportunities that evolve from the results that were discussed throughout this thesis.

9.3.1 Choice of Hardware

In the presented studies, generally only one type of HWD was used, except for the last study in Chapter 8 [15] in which an AR and a VR device were compared. As at the time of these studies, there were no vast differences between HWDs and I generally used state-of-the-art devices, I would not expect significantly different results when repeating the studies with other current devices. Still, it would be valuable to replicate the studies with different devices to validate the findings. Also, as new devices and software

are released regularly, new studies should be conducted to explore if current issues have been solved and which problems remain.

Even though I focused on a mobile context when implementing interaction techniques in Chapters 4 [18] and 5 [14], I relied on an external tracking system (OptiTrack motion-tracking system [81]) that cannot be considered mobile. This was done to achieve the best possible accuracy. However, newer HWDs might provide better tracking of hands, and also already include tracking of other objects such as keyboards [109]. Therefore, the proposed systems will also be viable in a fully mobile implementation. In future work they could be translated into mobile systems and evaluated in different contexts.

When designing the prototypes, I decided to focus on a physical tablet to provide a basis for the interaction concepts. However, these techniques could also be translated to other mobile devices, such as smartphones or laptops. Laptops might provide a smaller touch-input area (touchpad) but on the other hand would facilitate text-entry through its full-size physical keyboard. Different physical devices could therefore be valuable in different situations or for different tasks. That is why future work should also explore how the presented techniques can be translated to other devices, if and how they would need to be adapted, or which new techniques could be realized. It is also thinkable to adapt the techniques to situations without any physical devices, for example, by opportunistically utilizing objects in the environment such as a table or, as suggested by Medeiros et al. [69], the back of a seat. However, for all interaction techniques that are proposed for use in public scenarios it is important to also consider their social acceptability, which can be especially critical when sitting face-to-face with other people [67].

Furthermore, multiple of the presented studies have revealed the form factor and weight of the HWDs to be one of the main challenges. Therefore, research should continue making the devices smaller which could also increase social acceptability and lower barriers for more widespread use. In addition, the design of HWDs could be adapted to facilitate the fulfillment of personal needs such as eating or drinking or rubbing one's eyes. As mentioned before, a raisable visor such as provided by the HoloLens 2 [73] could solve such problems. For VR devices, this could also be achieved by using less restrictive light blockers such as provided for the Meta Quest Pro [72]. However, this can then reduce immersion because larger parts of the real world become visible. Therefore, clever systems need to be developed that support the users in fulfilling their personal needs while also maintaining properties of VR such as immersion in a virtual world.

9.3.2 Advancing the Presented Techniques

In the scope of this thesis, I have explored the advantages of VR for knowledge worker tasks, which was done most detailed for authoring presentations in Chapter 5 [14]. Moreover, Chapter 4 [18] presented examples of several more applications on how to enhance knowledge work through VR covering a wide spectrum of different applications. However, the examples from Chapter 4 [18] were not evaluated in detail, which would be an interesting avenue for future work. Also, the techniques presented in these studies could be advanced, such as making the reordering technique [14] (Chapter 5) scalable also to very large numbers of layers or adding further interaction possibilities to the object manipulation technique that is controlled through a spatially tracked pen. Of course, there are also further common knowledge worker tasks not yet mentioned in this work that could be supported in VR and that should be explored, such as authoring text or searching for and organizing information [90]. Also, an important concept that should be investigated is the combination of multiple VR applications and a seamless interaction between them, just as we are used to doing when we are working on our standard devices such as a laptop. This would also allow an investigation of more realistic use cases, as users could then efficiently use VR for a variety of tasks.

So far, I have only looked at single users. As knowledge workers also commonly collaborate with each other, it should be investigated how the proposed techniques can be translated to collaborative contexts. This would also include the compatibility to non-VR devices, as some collaborators might not always have access to such devices. Also, workers might not always want to or might not be able to use VR, therefore, non-VR alternatives should always be maintained.

9.3.3 Participants and Social Aspects

Another limitation of the presented work is that the participants were mostly university students and employees. Therefore, it would be of interest to rerun the studies with a much more diverse group of participants that is equally representing different ages, genders, and professions. This is especially true for the public study (Chapter 8 [15]) in which most bystanders were university students and might be more open-minded towards new technology. Also, the participants using XR in public could have been influenced by the location, which was known to all participants, as they were affiliated with the university, and therefore, might have been less unsettling to them. Repeating the study in different locations would be very insightful to gain a more complete understanding of how XR is perceived in public today. Of course, many of the results are likely to also change in the future as HWDs and also serious use cases such as for work or school might become more common. This is already in progress, as many new devices are advertised in a work context, such as the Quest Pro [72], Lenovo Think Reality glasses [106], or Apple Vision Pro [6]. Therefore, it could be very insightful to regularly monitor user and bystander reactions in the future in order to inform the design of future devices. However, there are also privacy concerns that need to be addressed if XR use in public becomes more common. For example, O'Hagan et al. [80] point out the conflict between bystanders privacy and the need for obtaining data through sensors on the XR device such as a camera, to ensure the functionality of the device.

While observing the behavior of participants during prolonged use of VR [13] (Chapter 7), I found indications of gender differences and also other authors reported on such findings [42]. Therefore, it would be important to investigate them in a more systematic way to work out design guidelines to increase usability for all users. The same would also apply for other groups of people that are often underrepresented in current user studies, such as much older people. If certain differences between groups of users can be confirmed, then these findings should be adopted in the design of new HWDs. This could include adaptions in the hardware such as different sizes of HWDs or a widespread support for different diopters.

9.3.4 Further Investigations of Long-term Effects

Especially the long-term study of Chapter 6 [16] could be influenced by a row of confounding variables that could not be controlled. For example, the type of work that the participants chose to do could have had an influence on the measures, such as frustration or perceived productivity. However, I chose to let participants do their own tasks to increase ecological validity and because doing an artificial task for a whole week could be unpleasant for the participants.

As I only captured the baseline costs of using VR, not optimizing it with any potential benefits, the natural next step would be to create an optimized setup and quantify the effects of working with such a system for a longer period of time. This could also include using newer and more ergonomic devices, which will hopefully be available in the near future. Such investigations could also be done in other contexts, such as public spaces or only for certain tasks, but on a regular basis.

In addition, various measures such as negative affect, anxiety, simulator sickness, visual fatigue, and wellbeing have shown to be negatively impacted by VR compared to a physical setup [16] (Chapter 6). To address these issues appropriately, further research is required that systematically investigates the effect of VR on these measures. This can then lead to insights on how to overcome these issues. First steps to tackle such problems have already been made, such as looking at exercises to counteract eyestrain [48] or helping the users to relax [108, 112]. As mentioned in Chapter 7 [13], VR software could be enhanced to promote a more ergonomic work position or, in a simple case, just suggest breaks. To optimize this, the system could also recognize the users' current condition and propose actions accordingly. Yet these ideas should be further studied to provide guidelines on how to integrate them in a work setting.

Moreover, it has been suggested that VR allows the user to access more attention resources [99], which has also been indicated in Chapter 6 [16]. However, further research is required to confirm this. If it holds true, additional research would be required to determine how to efficiently make use of this property without overburdening the users. In general, I hope that this work inspires more research in the direction of knowledge work, the effects of using VR for extended periods or on a regular basis, and how to optimally support the users.

9.3.5 Integrating the Physical Environment in VR

In addition, we have seen that integrating parts of the physical environment into VR is crucial as this has also been expressed by prior work [63], even though we saw in Chapter 7 [13] that users could cope with many situations while wearing the HWD and did not bother to take it off, for example, to talk or drink. Yet, these results were obtained in the secure location of a lab with only known people around. It is possible that users would not behave that way in public scenarios, as they might not know who they are even talking to or they might be more concerned to spill something or looking foolish while drinking.

However, in Chapter 8 [15] we have seen that participants were reluctant to use passthrough in VR, even though they were in a public environment and felt safer with systems in which they could see their surroundings. Therefore, it is important to further investigate how to integrate parts of the real environment into VR in a way that is not distracting or limiting the advantages of VR but still makes the users aware of everything important, increase their feeling of safety, and allow them to use real-world objects if required, such as a glass of water, a sheet of paper, or a phone. To increase safety while still keeping the user as immersed as possible, relevant real-world objects could also be replaced with virtual content as proposed by Yang et al. [119] for walking while wearing a VR HWD.

9.3.6 Translating Findings to AR

Lastly, the techniques presented in this thesis were implemented using VR HWDs. It was already mentioned in section 9.1.2 that it should be possible to translate them to AR, yet the smaller field of view in AR could impose some challenges. This should be investigated in future work to see if some adaptations would be required and which issues might open up. Also, new techniques could be developed specifically for the use with AR devices.

In addition, it would be interesting to investigate if the long-term use of AR devices leads to similar issues as the use of VR devices. Problems such as how to include the physical world in VR is not applicable to AR, in which the physical world is naturally visible, and therefore, other aspects are more prominent such as the placement of virtual screens, which has been explored in other experiments [68, 75]. For example, Medeiros et al. [68] have shown that users avoid placing screens at the position of other people. A useful feature for such scenarios could be a system that automatically adapts to a new environment, as already proposed for different scenarios [35, 58, 34], finding the optimal compromise between a user's preferences and the environmental conditions.

Chapter 10

Conclusion

Overall, this thesis has quantified the benefits that certain properties of VR provide for knowledge work, such as the increased screen space, the possibility to display information three-dimensionally and the enhanced interaction possibilities. In addition, a variety of prototypes, such as a display manager or presentation authoring techniques, were proposed that can support mobile knowledge work through VR. Even though the long-term study, which lasted for a full work week, indicated that users can adapt to the VR HWD, the study also showed certain limitations of current devices when compared to a physical setup. These limitations were indicated through a range of measures such as task load, usability, frustration, anxiety or visual fatigue.

In Chapter 8 [15], we have seen that using VR in public is still not well known. This will likely change as VR becomes overall more widespread and therefore, also using VR in public could become more common in the next years. However, there are still important aspects to consider when designing VR experiences for use in public. Main concerns are the feeling of safety and being aware of the physical surrounding. Therefore, systems should be designed in such a way that users can focus on their work tasks in VR while still being totally aware of all kinds of critical events that might go on around them, such as pickpockets or medical emergencies. In addition, participants of all studies mentioned that the hardware, specifically the weight, resolution and form factor of the VR devices is a major drawback that would keep them from using the presented systems regularly.

All these findings combined show that common knowledge worker devices such as laptops cannot simply be replaced by current VR devices. On the software side, it is important to carefully consider the use cases and design

VR-applications that support the knowledge worker in certain tasks or situations in an optimal way. If this cannot be provided, the drawbacks of current VR systems will most certainly outweigh the benefits. In addition to supporting knowledge workers in their tasks, VR knowledge worker applications should also be designed in a way that further increases the user's wellbeing, beyond optimizing the work task. This can, for example, be done by improving the ergonomics, such as placing virtual screens in good positions, or by designing the virtual environment such that the user feels comfortable. In addition, VR applications can also actively encourage users to take breaks or do eye exercises. Even if future VR devices can overcome the above-mentioned hardware limitations, for example by providing an extremely high resolution display in a device with the form factor of normal prescription glasses, the advantages that can be achieved by designing good software should not be neglected to make VR a valuable tool for knowledge work.

In conclusion, this thesis has shown that for a range of tasks, VR can provide real benefits for mobile knowledge workers. Yet, there are still drawbacks which are often caused by the hardware properties of current HWDs and which will hopefully be resolved in the future. Still, even with current hardware there are many possibilities to enhance VR systems from a software perspective to, for example, increase safety and wellbeing or make certain knowledge worker tasks more efficient. With such carefully designed systems VR can be a viable option for mobile knowledge work.

References

[1] William C Adams: Conducting semi-structured interviews. In: *Handbook of practical program evaluation* (2015), pp. 492–505.

[2] Ammar Ahmad, Cyrille Migniot, and Albert Dipanda: Hand pose estimation and tracking in real and virtual interaction: A review. In: *Image and Vision Computing* 89 (2019), pp. 35–49.

[3] Ghassem Alaee, Amit P Deasi, Lourdes Pena-Castillo, Edward Brown, and Oscar Meruvia-Pastor: A user study on augmented virtuality using depth sensing cameras for near-range awareness in immersive VR. In: *IEEE VR's 4th workshop on everyday virtual reality (WEVR 2018)*. Vol. 10. (2018), p. 3.

[4] Antonio de Lucas Ancillo, María Teresa del Val Núñez, and Sorin Gavrila Gavrila: Workplace change within the COVID-19 context: a grounded theory approach. In: *Economic Research-Ekonomska Istraživanja* 34.1 (2021), pp. 2297–2316.

[5] Allison P Anderson, Michael D Mayer, Abigail M Fellows, Devin R Cowan, Mark T Hegel, and Jay C Buckey: Relaxation with immersive natural scenes presented using virtual reality. In: *Aerospace medicine and human performance* 88.6 (2017), pp. 520–526.

[6] *Apple Vision Pro - Apple*. `https://www.apple.com/apple-vision-pro`. Accessed: 2023-12-03. (2023).

[7] Rahul Arora, Rubaiat Habib Kazi, Danny M Kaufman, Wilmot Li, and Karan Singh: Magicalhands: Mid-air hand gestures for animating in vr. In: *Proceedings of the 32nd annual ACM symposium on user interface software and technology.* (2019), pp. 463–477.

[8] Huidong Bai, Li Zhang, Jing Yang, and Mark Billinghurst: Bringing full-featured mobile phone interaction into virtual reality. In: *Computers & Graphics* 97 (2021), pp. 42–53.

[9] NBKDE Bailey and Nancy B Kurland: The advantages and challenges of working here, there, anywhere, and anytime. In: *Organizational dynamics* 28.2 (1999), pp. 53–68.

[10] Laura Bajorunaite, Stephen Brewster, and Julie R. Williamson: Reality Anchors: Bringing Cues from Reality to Increase Acceptance of Immersive Technologies in Transit. In: *Proc. ACM Hum.-Comput. Interact.* 7.MHCI (2023). DOI: `10.1145/3604266`. URL: `https://doi.org/10.1145/3604266`.

[11] Aaron Bangor, Philip Kortum, and James Miller: Determining what individual SUS scores mean: Adding an adjective rating scale. In: *Journal of usability studies* 4.3 (2009), pp. 114–123.

[12] Leif P Berg and Judy M Vance: Industry use of virtual reality in product design and manufacturing: a survey. In: *Virtual reality* 21 (2017), pp. 1–17.

[13] Verena Biener, Forouzan Farzinnejad, Rinaldo Schuster, Seyedmasih Tabaei, Leon Lindlein, Jinghui Hu, Negar Nouri, John J. Dudley, Per Ola Krlstensson, Jörg Müller, and Jens Grubert: Hold Tight: Identifying Behavioral Patterns During Prolonged Work in VR Through Video Analysis. In: *IEEE Transactions on Visualization and Computer Graphics* 30.5 (2024), pp. 2796–2806. DOI: `10.1109/TVCG.2024.3372048`.

[14] Verena Biener, Travis Gesslein, Daniel Schneider, Felix Kawala, Alexander Otte, Per Ola Kristensson, Michel Pahud, Eyal Ofek, Cuauhtli Campos, Matjaž Kljun, et al.: Povrpoint: Authoring presentations in mobile virtual reality. In: *IEEE Transactions on Visualization and Computer Graphics* 28.5 (2022), pp. 2069–2079.

[15] Verena Biener, Snehanjali Kalamkar, John J. Dudley, Jinghui Hu, Per Ola Kristensson, Jörg Müller, and Jens Grubert: Working with XR in Public: Effects on Users and Bystanders. In: *IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW).* (2024), pp. 779–780. DOI: `10.1109/VRW62533.2024.00186`.

[16] Verena Biener, Snehanjali Kalamkar, Negar Nouri, Eyal Ofek, Michel Pahud, John J Dudley, Jinghui Hu, Per Ola Kristensson, Maheshya Weerasinghe, Klen Čopič Pucihar, et al.: Quantifying the effects of working in VR for one week. In: *IEEE Transactions on Visualization and Computer Graphics* 28.11 (2022), pp. 3810–3820.

[17] Verena Biener, Eyal Ofek, Michel Pahud, Per Ola Kristensson, and Jens Grubert: Extended reality for knowledge work in everyday environments. In: *Everyday Virtual and Augmented Reality.* Springer, (2023), pp. 21–56.

[18] Verena Biener, Daniel Schneider, Travis Gesslein, Alexander Otte, Bastian Kuth, Per Ola Kristensson, Eyal Ofek, Michel Pahud, and Jens Grubert: Breaking the screen: Interaction across touchscreen boundaries in virtual reality for mobile knowledge workers. In: *IEEE*

transactions on visualization and computer graphics 26.12 (2020), pp. 3490–3502.

[19] John Brooke: Sus: a “quick and dirty’usability. In: *Usability evaluation in industry* 189.3 (1996), pp. 189–194.

[20] S. Buckley: *Two people spent 48 hours in nonstop virtual reality.* `https://www.engadget.com/2017-06-07-two-people-spent-48-hours-in-nonstop-virtual-reality.html`. Accessed: 2022-05-24. (2017).

[21] Alberto Cannavò, Claudio Demartini, Lia Morra, and Fabrizio Lamberti: Immersive virtual reality-based interfaces for character animation. In: *IEEE Access* 7 (2019), pp. 125463–125480.

[22] Eunhee Chang, Hyun Taek Kim, and Byounghyun Yoo: Virtual reality sickness: a review of causes and measurements. In: *International Journal of Human–Computer Interaction* 36.17 (2020), pp. 1658–1682.

[23] Xiang’Anthony’ Chen, Tovi Grossman, Daniel J Wigdor, and George Fitzmaurice: Duet: exploring joint interactions on a smart phone and a smart watch. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* (2014), pp. 159–168.

[24] Lung-Pan Cheng, Eyal Ofek, Christian Holz, and Andrew D Wilson: Vroamer: generating on-the-fly VR experiences while walking inside large, unknown real-world building environments. In: *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR).* IEEE. (2019), pp. 359–366.

[25] Andreas Dengel, Stefan Agne, Bertin Klein, Achim Ebert, and Matthias Deller: Human-centered interaction with documents. In: *Proceedings of the 1st ACM international workshop on human-centered multimedia.* (2006), pp. 35–44.

[26] Amit P Desai, Lourdes Pena-Castillo, and Oscar Meruvia-Pastor: A window to your smartphone: exploring interaction and communication in immersive vr with augmented virtuality. In: *2017 14th conference on computer and robot vision (CRV).* IEEE. (2017), pp. 217–224.

[27] James Dominic, Brock Tubre, Charles Ritter, Jada Houser, Colton Smith, and Paige Rodeghero: Remote pair programming in virtual reality. In: *2020 IEEE International Conference on Software Maintenance and Evolution (ICSME).* IEEE. (2020), pp. 406–417.

[28] Tobias Drey, Jan Gugenheimer, Julian Karlbauer, Maximilian Milo, and Enrico Rukzio: Vrsketchin: Exploring the design space of pen and tablet interaction for 3d sketching in virtual reality. In: *Proceedings of the 2020 CHI conference on human factors in computing systems.* (2020), pp. 1–14.

[29] P Drucker: The effective executive. In: *Harper & row, New York, NY* (1966).

[30] Pouya Eghbali, Kaisa Väänänen, and Tero Jokela: Social acceptability of virtual reality in public spaces: Experiential factors and design recommendations. In: *Proceedings of the 18th international conference on mobile and ubiquitous multimedia.* (2019), pp. 1–11.

[31] Anna Eiberger, Per Ola Kristensson, Susanne Mayr, Matthias Kranz, and Jens Grubert: Effects of depth layer switching between an optical see-through head-mounted display and a body-proximate display. In: *Symposium on Spatial User Interaction.* (2019), pp. 1–9.

[32] Katja Einsfeld, Stefan Agne, Matthias Deller, Achim Ebert, Bertin Klein, and Christian Reuschling: Dynamic visualization and navigation of semantic virtual environments. In: *Tenth International Conference on Information Visualisation (IV'06).* IEEE. (2006), pp. 569–574.

[33] Anthony Elliott, Brian Peiris, and Chris Parnin: Virtual reality in software engineering: Affordances, applications, and challenges. In: *2015 IEEE/ACM 37th IEEE International Conference on Software Engineering.* Vol. 2. IEEE. (2015), pp. 547–550.

[34] Andreas Fender and Jörg Müller: SpaceState: Ad-Hoc definition and recognition of hierarchical room states for smart environments. In: *Proceedings of the 2019 ACM international conference on interactive surfaces and spaces.* (2019), pp. 303–314.

[35] Ran Gal, Lior Shapira, Eyal Ofek, and Pushmeet Kohli: FLARE: Fast layout for augmented reality applications. In: *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR).* (2014), pp. 207–212. DOI: `10.1109/ISMAR.2014.6948429`.

[36] Ceenu George, An Ngo Tien, and Heinrich Hussmann: Seamless, bidirectional transitions along the reality-virtuality continuum: A conceptualization and prototype exploration. In: *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR).* IEEE. (2020), pp. 412–424.

[37] Travis Gesslein, Verena Biener, Philipp Gagel, Daniel Schneider, Per Ola Kristensson, Eyal Ofek, Michel Pahud, and Jens Grubert: Pen-based interaction with spreadsheets in mobile virtual reality. In: *2020 IEEE International symposium on mixed and augmented reality (ISMAR).* IEEE. (2020), pp. 361–373.

[38] Matt Gottsacker, Nahal Norouzi, Kangsoo Kim, Gerd Bruder, and Greg Welch: Diegetic representations for seamless cross-reality interruptions. In: *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR).* IEEE. (2021), pp. 310–319.

[39] Rebecca A Grier: How high is high? A meta-analysis of NASA-TLX global workload scores. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting.* Vol. 59. 1. SAGE Publications Sage CA: Los Angeles, CA. (2015), pp. 1727–1731.

[40] Jens Grubert, Matthias Heinisch, Aaron Quigley, and Dieter Schmalstieg: Multifi: Multi fidelity interaction with displays on and around the body. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems.* (2015), pp. 3933–3942.

[41] Jens Grubert, Eyal Ofek, Michel Pahud, and Per Ola Kristensson: The office of the future: Virtual, portable, and global. In: *IEEE computer graphics and applications* 38.6 (2018), pp. 125–133.

[42] Jie Guo, Dongdong Weng, Hui Fang, Zhenliang Zhang, Jiamin Ping, Yue Liu, and Yongtian Wang: Exploring the differences of visual discomfort caused by long-term immersion between virtual environments and physical environments. In: *2020 IEEE Conference on virtual reality and 3D user interfaces (VR).* IEEE. (2020), pp. 443–452.

[43] Jie Guo, Dongdong Weng, Zhenliang Zhang, Haiyan Jiang, Yue Liu, Yongtian Wang, and Henry Been-Lirn Duh: Mixed reality office system based on maslow's hierarchy of needs: Towards the long-term immersion in virtual environments. In: *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR).* IEEE. (2019), pp. 224–235.

[44] Sandra G Hart and Lowell E Staveland: Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In: *Advances in psychology.* Vol. 52. Elsevier, (1988), pp. 139–183.

[45] Jeremy Hartmann, Christian Holz, Eyal Ofek, and Andrew D Wilson: Realitycheck: Blending virtual environments with situated physical reality. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems.* (2019), pp. 1–12.

[46] H Heuer, G Hollendiek, H Kröger, and T Römer: Rest position of the eyes and its effect on viewing distance and visual fatigue in computer display work. In: *Zeitschrift fur experimentelle und angewandte Psychologie* 36.4 (1989), pp. 538–566.

[47] Juan David Hincapié-Ramos, Xiang Guo, Paymahn Moghadasian, and Pourang Irani: Consumed endurance: a metric to quantify arm fatigue of mid-air interactions. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* (2014), pp. 1063–1072.

[48] Teresa Hirzle, Fabian Fischbach, Julian Karlbauer, Pascal Jansen, Jan Gugenheimer, Enrico Rukzio, and Andreas Bulling: Understanding, addressing, and analysing digital eye strain in virtual reality head-

mounted displays. In: *ACM Transactions on Computer-Human Interaction (TOCHI)* 29.4 (2022), pp. 1–80.

[49] Teresa Hirzle, Jan Gugenheimer, Florian Geiselhart, Andreas Bulling, and Enrico Rukzio: A Design Space for Gaze Interaction on Head-Mounted Displays. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems.* CHI '19. Glasgow, Scotland Uk: Association for Computing Machinery, (2019), pp. 1–12. ISBN: 9781450359702. DOI: `10.1145/3290605.3300855`.

[50] Deborah Hix and H. Rex Hartson: *Formative evaluation: ensuring usability in user interfaces.* (1993). URL: `https://api.semanticscholar.org/CorpusID:107367884`.

[51] *Horizon Workrooms virtual office and meetings.* `https://forwork.meta.com/de/en/horizon-workrooms/`. Accessed: 2023-12-01. (2023).

[52] *Hubs - Private virtual 3D worlds in your browser.* `https://hubs.mozilla.com/`. Accessed: 2023-12-01. (2023).

[53] *Immersed.* `https://www.immersed.com/`. Accessed: 2024-01-10. (2023).

[54] Haiyan Jiang, Dongdong Weng, Zhenliang Zhang, Yihua Bao, Yufei Jia, and Mengman Nie: Hikeyb: High-efficiency mixed reality system for text entry. In: *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct).* IEEE. (2018), pp. 132–137.

[55] Robert S Kennedy, Norman E Lane, Kevin S Berbaum, and Michael G Lilienthal: Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. In: *The int. journal of aviation psychology* 3.3 (1993), pp. 203–220.

[56] B. Klemens: *I Spent Hundreds of Hours Working in VR. Here's What I Learned.* `https://www.wired.com/story/hours-working-vr-tips/`. Accessed: 2022-05-24. (2022).

[57] Pascal Knierim, Thomas Kosch, Johannes Groschopp, and Albrecht Schmidt: Opportunities and challenges of text input in portable virtual reality. In: *Extended abstracts of the 2020 CHI conference on human factors in computing systems.* (2020), pp. 1–8.

[58] Wallace S Lages and Doug A Bowman: Walking with adaptive augmented reality workspaces: design and usage patterns. In: *Proceedings of the 24th International Conference on Intelligent User Interfaces.* (2019), pp. 356–366.

[59] Geonsun Lee, Jennifer Healey, and Dinesh Manocha: VRDoc: Gaze-based Interactions for VR Reading Experience. In: *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR).* IEEE. (2022), pp. 787–796.

[60] Jingyi Li, Luca Woik, and Andreas Butz: Designing mobile mr workspaces: Effects of reality degree and spatial configuration during passenger productivity in hmds. In: *Proceedings of the ACM on Human-Computer Interaction* 6.MHCI (2022), pp. 1–17.

[61] Feiyu Lu and Doug A Bowman: Evaluating the potential of glanceable ar interfaces for authentic everyday uses. In: *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. IEEE. (2021), pp. 768–777.

[62] Andrés Lucero and Akos Vetek: NotifEye: using interactive glasses to deal with notifications while walking in public. In: *Proceedings of the 11th conference on advances in computer entertainment technology.* (2014), pp. 1–10.

[63] Mark McGill, Daniel Boland, Roderick Murray-Smith, and Stephen Brewster: A dose of reality: Overcoming usability challenges in vr head-mounted displays. In: *Proceedings of the 33rd annual ACM conference on human factors in computing systems.* (2015), pp. 2143–2152.

[64] Mark Mcgill, Aidan Kehoe, Euan Freeman, and Stephen Brewster: Expanding the bounds of seated virtual workspaces. In: *ACM Transactions on Computer-Human Interaction (TOCHI)* 27.3 (2020), pp. 1–40.

[65] Mark McGill, Julie Williamson, Alexander Ng, Frank Pollick, and Stephen Brewster: Challenges in passenger use of mixed reality headsets in cars and other transportation. In: *Virtual Reality* 24 (2020), pp. 583–603.

[66] Daniel Medeiros, Rafael dos Anjos, Nadia Pantidi, Kun Huang, Maurício Sousa, Craig Anslow, and Joaquim Jorge: Promoting Reality Awareness in Virtual Reality through Proxemics. In: *2021 IEEE Virtual Reality and 3D User Interfaces (VR).* (2021), pp. 21–30. DOI: `10.1109/VR50410.2021.00022`.

[67] Daniel Medeiros, Romane Dubus, Julie Williamson, Graham Wilson, Katharina Pöhlmann, and Mark McGill: Surveying the Social Comfort of Body, Device, and Environment-Based Augmented Reality Interactions in Confined Passenger Spaces Using Mixed Reality Composite Videos. In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7.3 (2023), pp. 1–25.

[68] Daniel Medeiros, Mark McGill, Alexander Ng, Robert McDermid, Nadia Pantidi, Julie Williamson, and Stephen Brewster: From Shielding to Avoidance: Passenger Augmented Reality and the Layout of Virtual Displays for Productivity in Shared Transit. In: *IEEE Transactions on Visualization and Computer Graphics* 28.11 (2022), pp. 3640–3650. DOI: `10.1109/TVCG.2022.3203002`.

[69] Daniel Medeiros, Graham Wilson, Mark Mcgill, and Stephen Anthony Brewster: The Benefits of Passive Haptics and Perceptual Manipulation for Extended Reality Interactions in Constrained Passenger Spaces. In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems.* (2023), pp. 1–19.

[70] Xuanru Meng, Wenge Xu, and Hai-Ning Liang: An exploration of hands-free text selection for virtual reality head-mounted displays. In: *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR).* IEEE. (2022), pp. 74–81.

[71] *Meta Quest 2: Immersive all-in-one VR headset.* `https://www.meta.com/de/en/quest/products/quest-2/`. Accessed: 2024-01-18. (2024).

[72] *Meta Quest Pro: Premium mixed reality.* `https://www.meta.com/de/en/quest/quest-pro/`. Accessed: 2023-12-03. (2023).

[73] Microsoft: *HoloLens 2 - Overview, Features and Specs.* `https://www.microsoft.com/en-gb/hololens/hardware`. Accessed: 2023-12-19. (2023).

[74] Paul Milgram, Haruo Takemura, Akira Utsumi, and Fumio Kishino: Augmented reality: A class of displays on the reality-virtuality continuum. In: *Telemanipulator and telepresence technologies.* Vol. 2351. Spie. (1995), pp. 282–292.

[75] Alexander Ng, Daniel Medeiros, Mark McGill, Julie Williamson, and Stephen Brewster: The passenger experience of mixed reality virtual display layouts in airplane environments. In: *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR).* IEEE. (2021), pp. 265–274.

[76] Cheuk Fan Ng: Public spaces as workplace for mobile knowledge workers. In: *Journal of Corporate Real Estate* 18.3 (2016), pp. 209–223.

[77] Cuong Nguyen, Stephen DiVerdi, Aaron Hertzmann, and Feng Liu: Vremiere: In-headset virtual reality video editing. In: *Proceedings of the 2017 CHI conference on human factors in computing systems.* (2017), pp. 5428–5438.

[78] Erwan Normand and Michael J McGuffin: Enlarging a smartphone with ar to create a handheld vesad (virtually extended screen-aligned display). In: *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR).* IEEE. (2018), pp. 123–133.

[79] Joseph O Hagan and Julie R Williamson: Reality aware vr headsets. In: *Proceedings of the 9th ACM international symposium on pervasive displays.* (2020), pp. 9–17.

[80] Joseph O'Hagan, Pejman Saeghe, Jan Gugenheimer, Daniel Medeiros, Karola Marky, Mohamed Khamis, and Mark McGill: Privacy-Enhancing

Technology and Everyday Augmented Reality: Understanding Bystanders' Varying Needs for Awareness and Consent. In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6.4 (2023), pp. 1–35.

[81] *OptiTrack - Motion Capture for Virtual Reality.* `https://optitrack.com/applications/virtual-reality/`. Accessed: 2024-01-18. (2024).

[82] Leonardo Pavanatto, Chris North, Doug A Bowman, Carmen Badea, and Richard Stoakley: Do we still need physical monitors? an evaluation of the usability of ar virtual monitors for productivity work. In: *2021 IEEE Virtual Reality and 3D User Interfaces (VR).* IEEE. (2021), pp. 759–767.

[83] Aravind Gandhi Periyasamy, Amarjeet Singh, and Khaiwal Ravindra: Carbon emissions from virtual and physical modes of conference and prospects for carbon neutrality: an analysis from India. In: *Air, Soil and Water Research* 15 (2022). DOI: `10.1177/11786221221093298`.

[84] Ken Pfeuffer, Jason Alexander, Ming Ki Chong, Yanxia Zhang, and Hans Gellersen: Gaze-shifting: Direct-indirect input with pen and touch modulated by gaze. In: *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology.* (2015), pp. 373–383.

[85] Ken Pfeuffer and Hans Gellersen: Gaze and touch interaction on tablets. In: *Proceedings of the 29th Annual Symposium on User Interface Software and Technology.* (2016), pp. 301–311.

[86] Alexander Plopski, Teresa Hirzle, Nahal Norouzi, Long Qian, Gerd Bruder, and Tobias Langlotz: The eye in extended reality: A survey on gaze interaction and eye tracking in head-worn extended reality. In: *ACM Computing Surveys (CSUR)* 55.3 (2022), pp. 1–39.

[87] Halley Profita, Reem Albaghli, Leah Findlater, Paul Jaeger, and Shaun K Kane: The AT effect: how disability affects the perceived social acceptability of head-mounted display use. In: *proceedings of the 2016 CHI conference on human factors in computing systems.* (2016), pp. 4884–4895.

[88] Eric D Ragan, Doug A Bowman, Regis Kopper, Cheryl Stinson, Siroberto Scerbo, and Ryan P McMahan: Effects of field of view and visual complexity on virtual reality training effectiveness for a visual scanning task. In: *IEEE transactions on visualization and computer graphics* 21.7 (2015), pp. 794–807.

[89] Ramesh Raskar, Greg Welch, Matt Cutts, Adam Lake, Lev Stesin, and Henry Fuchs: The office of the future: A unified approach to image-based modeling and spatially immersive displays. In: *Proceedings of*

the 25th annual conference on Computer graphics and interactive techniques. (1998), pp. 179–188.

[90] Wolfgang Reinhardt, Benedikt Schmidt, Peter Sloep, and Hendrik Drachsler: Knowledge worker roles and actions—results of two empirical studies. In: *Knowledge and process management* 18.3 (2011), pp. 150–174.

[91] Patrick Reipschläger and Raimund Dachselt: Designar: Immersive 3d-modeling combining augmented reality with interactive displays. In: *Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces.* (2019), pp. 29–41.

[92] Jun Rekimoto and Masanori Saitoh: Augmented surfaces: a spatially continuous work space for hybrid computing environments. In: *Proceedings of the SIGCHI conference on Human Factors in Computing Systems.* (1999), pp. 378–385.

[93] Falko Rheinberg, Regina Vollmeyer, and Stefan Engeser: *Die Erfassung des Flow-Erlebens.* postprint. (2006).

[94] Anastasia Ruvimova, Junhyeok Kim, Thomas Fritz, Mark Hancock, and David C Shepherd: " Transport Me Away": Fostering Flow in Open Offices through Virtual Reality. In: *Proceedings of the 2020 CHI conference on human factors in computing systems.* (2020), pp. 1–14.

[95] Daniel Schneider, Verena Biener, Alexander Otte, Travis Gesslein, Philipp Gagel, Cuauhtli Campos, Klen Čopič Pucihar, Matjazz Kljun, Eyal Ofek, Michel Pahud, et al.: Accuracy evaluation of touch tasks in commodity virtual and augmented reality head-mounted displays. In: *Proceedings of the 2021 ACM Symposium on Spatial User Interaction.* (2021), pp. 1–11.

[96] Daniel Schneider, Alexander Otte, Travis Gesslein, Philipp Gagel, Bastian Kuth, Mohamad Shahm Damlakhi, Oliver Dietz, Eyal Ofek, Michel Pahud, Per Ola Kristensson, et al.: Reconviguration: Reconfiguring physical keyboards in virtual reality. In: *IEEE transactions on visualization and computer graphics* 25.11 (2019), pp. 3190–3201.

[97] Thomas W Schubert: The sense of presence in virtual environments: A three-component scale measuring spatial presence, involvement, and realness. In: *Z. für Medienpsychologie* 15.2 (2003), pp. 69–71.

[98] Valentin Schwind, Jens Reinhardt, Rufat Rzayev, Niels Henze, and Katrin Wolf: Virtual reality on the go? a study on social acceptance of vr glasses. In: *Proceedings of the 20th international conference on human-computer interaction with mobile devices and services adjunct.* (2018), pp. 111–118.

[99] Ruiying Shen, Dongdong Weng, Shanshan Chen, Jie Guo, and Hui Fang: Mental fatigue of long-term office tasks in virtual environment.

In: *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE. (2019), pp. 124–127.

[100] Adalberto L Simeone: The VR motion tracker: visualising movement of non-participants in desktop virtual reality experiences. In: *2016 IEEE 2nd Workshop on Everyday Virtual Reality (WEVR)*. IEEE. (2016), pp. 1–4.

[101] Richard Skarbez, Missie Smith, and Mary C Whitton: Revisiting Milgram and Kishino's reality-virtuality continuum. In: *Frontiers in Virtual Reality* 2 (2021). DOI: `10.3389/frvir.2021.647997`.

[102] Alexis D Souchet, Domitile Lourdeaux, Alain Pagani, and Lisa Rebenitsch: A narrative review of immersive virtual reality's ergonomics and risks at the workplace: cybersickness, visual fatigue, muscular fatigue, acute stress, and mental overload. In: *Virtual Reality* 27.1 (2023), pp. 19–50.

[103] Kay M Stanney, Robert S Kennedy, and Julie M Drexler: Cybersickness is not simulator sickness. In: *Proceedings of the Human Factors and Ergonomics Society annual meeting*. Vol. 41. 2. SAGE Publications Sage CA: Los Angeles, CA. (1997), pp. 1138–1142.

[104] Frank Steinicke and Gerd Bruder: A self-experimentation report about long-term use of fully-immersive technology. In: *Proceedings of the 2nd ACM symposium on Spatial user interaction*. (2014), pp. 66–69.

[105] Hemant Bhaskar Surale, Aakar Gupta, Mark Hancock, and Daniel Vogel: Tabletinvr: Exploring the design space for using a multi-touch tablet in virtual reality. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. (2019), pp. 1–13.

[106] *ThinkReality A3 Smart Glasses*. `https://www.lenovo.com/de/de/thinkrealitya3/`. Accessed: 2023-12-03. (2023).

[107] David R Thomas: *A general inductive approach for qualitative data analysis*. (2003).

[108] Krsna Das Thoondee and Andreas Oikonomou: Using virtual reality to reduce stress at work. In: *2017 computing conference*. IEEE. (2017), pp. 492–499.

[109] *Tracked keyboards for Meta Quest*. `https://www.meta.com/en-gb/help/quest/articles/headsets-and-accessories/meta-quest-accessories/tracked-keyboards-meta-quest/`. Accessed: 2023-12-03. (2023).

[110] Ying-Chao Tung, Chun-Yen Hsu, Han-Yu Wang, Silvia Chyou, Jhe-Wei Lin, Pei-Jung Wu, Andries Valstar, and Mike Y Chen: User-defined game input for smart glasses in public space. In: *Proceedings of the 33rd Annual ACM conference on human factors in computing systems*. (2015), pp. 3327–3336.

[111] Matthew Turk: Multimodal interaction: A review. In: *Pattern recognition letters* 36 (2014), pp. 189–195.

[112] Deltcho Valtchanov, Kevin R Barton, and Colin Ellard: Restorative effects of virtual nature settings. In: *Cyberpsychology, Behavior, and Social Networking* 13.5 (2010), pp. 503–512.

[113] Daniel Vogel, Paul Lubos, and Frank Steinicke: Animationvr-interactive controller-based animating in virtual reality. In: *2018 IEEE 1st Workshop on Animation in Virtual and Augmented Environments (ANIVAE)*. IEEE. (2018), pp. 1–6.

[114] Chiu-Hsuan Wang, Bing-Yu Chen, and Liwei Chan: Realitylens: A user interface for blending customized physical world view into virtual reality. In: *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology.* (2022), pp. 1–11.

[115] David Watson, Lee Anna Clark, and Auke Tellegen: Development and validation of brief measures of positive and negative affect: the PANAS scales. In: *Journal of personality and social psychology* 54.6 (1988), p. 1063.

[116] Matthias Wille, Lars Adolph, Britta Grauel, Sascha Wischniewski, Sabine Theis, and Thomas Alexander: Prolonged work with head mounted displays. In: *Proceedings of the 2014 ACM international symposium on wearable computers: adjunct program.* (2014), pp. 221–224.

[117] Michael Williams and Tami Moser: The art of coding and thematic exploration in qualitative research. In: *International Management Review* 15.1 (2019), pp. 45–55.

[118] Jacob O Wobbrock, Leah Findlater, Darren Gergle, and James J Higgins: The aligned rank transform for nonparametric factorial analyses using only anova procedures. In: *Proceedings of the SIGCHI conference on human factors in computing systems.* (2011), pp. 143–146.

[119] Jackie Yang, Christian Holz, Eyal Ofek, and Andrew D Wilson: Dreamwalker: Substituting real-world walking experiences with a virtual reality. In: *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology.* (2019), pp. 1093–1107.

[120] Fengyuan Zhu and Tovi Grossman: Bishare: Exploring bidirectional interactions between smartphones and head-mounted augmented reality. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems.* (2020), pp. 1–14.

[121] Daniel Zielasko, Benjamin Weyers, Martin Bellgardt, Sebastian Pick, Alexander Meibner, Tom Vierjahn, and Torsten W Kuhlen: Remain seated: towards fully-immersive desktop VR. In: *2017 IEEE 3rd workshop on everyday virtual reality (WEVR).* IEEE. (2017), pp. 1–6.

[122] Andras N Zsido, Szidalisz A Teleki, Krisztina Csokasi, Sandor Rozsa, and Szabolcs A Bandi: Development of the short version of the spielberger state—trait anxiety inventory. In: *Psychiatry research* 291 (2020). DOI: `https://doi.org/10.1016/j.psychres.2020.113223`.

Eigene Publikationen

Hier findet sich eine Liste der Publikationen, die im Rahmen meiner Promotion entstanden sind und die in dieser Arbeit in den Kapiteln 3 bis 8 vorgestellt werden:

- Verena Biener, Daniel Schneider, Travis Gesslein, Alexander Otte, Bastian Kuth, Per Ola Kristensson, Eyal Ofek, Michel Pahud & Jens Grubert. Breaking the Screen: Interaction Across Touchscreen Boundaries in Virtual Reality for Mobile Knowledge Workers. In IEEE Transactions on Visualization and Computer Graphics (TVCG), vol. 26, no. 12, Dec. 2020, pp. 3490-3502. `https://doi.org/10.1109/TVCG.2020.3023567`.
- Verena Biener, Travis Gesslein, Daniel Schneider, Felix Kawala, Alexander Otte, Per Ola Kristensson, Michel Pahud, Eyal Ofek, Cuauhtli Campos, Matjaž Kljun, Klen Čopič Pucihar & Jens Grubert. PoVRPoint: Authoring Presentations in Mobile Virtual Reality. In IEEE Transactions on Visualization and Computer Graphics (TVCG), vol. 28, no. 5, May 2022, pp. 2069-2079. `https://doi.org/10.1109/TVCG.2022.3150474`.
- Verena Biener, Snehanjali Kalamkar, Negar Nouri, Eyal Ofek, Michel Pahud, John J. Dudley, Jinghui Hu, Per Ola Kristensson, Maheshya Weerasinghe, Klen Čopič Pucihar, Matjaž Kljun, Stephan Streuber & Jens Grubert. Quantifying the Effects of Working in VR for One Week. In IEEE Transactions on Visualization and Computer Graphics (TVCG), vol. 28, no. 11, Nov. 2022, pp. 3810-3820. `https://doi.org/10.1109/TVCG.2022.3203103`
- Verena Biener, Eyal Ofek, Michel Pahud, Per Ola Kristensson, Jens Grubert. Extended Reality for Knowledge Work in Everyday Environments. In: Simeone, A., Weyers, B., Bialkova, S., Lindeman, R.W. (eds) Everyday Virtual and Augmented Reality. Human–Computer In-

teraction Series. Springer, Cham. 2023. `https://doi.org/10.1007/978-3-031-05804-2\_2`

- Verena Biener, Forouzan Farzinnejad, Rinaldo Schuster, Seyedmasih Tabaei, Leon Lindlein, Jinghui Hu, Negar Nouri, John J. Dudley, Per Ola Kristensson, Jörg Müller & Jens Grubert. Hold Tight: Identifying Behavioral Patterns During Prolonged Work in VR through Video Analysis. In IEEE Transactions on Visualization and Computer Graphics (TVCG), vol. 30, no. 5, May 2024, pp. 2796-2806. `https://doi.org/10.1109/TVCG.2024.3372048`
- Verena Biener, Snehanjali Kalamkar, John J. Dudley, Jinghui Hu, Per Ola Kristensson, Jörg Müller & Jens Grubert. Working with XR in Public: Effects on Users and Bystanders. In IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), 2024, pp. 779-780. `https://doi.org/10.1109/VRW62533.2024.00186`

Im folgenden findet sich eine Liste weiterer Publikationen an denen ich bis zur Einreichung dieser Arbeit mitgewirkt habe:

- Christoph Müller, Michael Krone, Markus Huber, Verena Biener, Dominik Herr, Steffen Koch, Guido Reina, Daniel Weiskopf, & Thomas Ertl. Interactive Molecular Graphics for Augmented Reality Using HoloLens. Journal of Integrative Bioinformatics, vol. 15, no. 2, 2018, pp. 20180005. `https://doi.org/10.1515/jib-2018-0005`
- Travis Gesslein, Verena Biener, Philipp Gagel, Daniel Schneider, Per Ola Kristensson, Eyal Ofek, Michel Pahud & Jens Grubert. Pen-based Interaction with Spreadsheets in Mobile Virtual Reality. In Proceedings of the 2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR 2020), 2020, pp. 361-373. `https://doi.org/10.1109/ISMAR50242.2020.00063`
- Francisco Kiss, Paweł W. Woźniak, Verena Biener, Pascal Knierim, & Albrecht Schmidt. VUM: Understanding Requirements for a Virtual Ubiquitous Microscope. In Proceedings of the 19th International Conference on Mobile and Ubiquitous Multimedia (MUM '20), 2020, pp. 259–266. `https://doi.org/10.1145/3428361.3428386`
- Daniel Schneider, Verena Biener, Alexander Otte, Travis Gesslein, Philipp Gagel, Cuauhtli Campos, Klen Čopič Pucihar, Matjaž Kljun, Eyal Ofek, Michel Pahud, Per Ola Kristensson & Jens Grubert. Accuracy Evaluation of Touch Tasks in Commodity Virtual and Augmented

Reality Head-Mounted Displays. In Proceedings of the 2021 Symposium on Spatial User Interaction (SUI '21), 2021, Article 7, pp. 1–11. `https://doi.org/10.1145/3485279.3485283`

- Verena Biener. [DC] Mixed Reality Interaction for Mobile Knowledge Work. In IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), 2022, pp. 928-929. `https://doi.org/10.1109/VRW55335.2022.00315`
- Maheshya Weerasinghe, Verena Biener, Jens Grubert, Aaron J. Quigley, Alice Toniolo, Klen Čopič Pucihar & Matjaž Kljun. VocabulARy: Learning Vocabulary in AR Supported by Keyword Visualisations. In IEEE Transactions on Visualization and Computer Graphics (TVCG), vol. 28, no. 11, Nov. 2022, pp. 3748-3758. `https://doi.org/10.1109/TVCG.2022.3203116`
- Negar Nouri, Snehanjali Kalamkar, Forouzan Farzinnejad, Verena Biener, Fabian Schick, Stefan Kalkhof & Jens Grubert. Improving Understanding of Biocide Availability in Facades through Immersive Analytics. In Adjunct Proceedings of IEEE International Symposium on Mixed and Augmented Reality 2022 (ISMAR 2022), pp.98-103. `https://doi.org/10.1109/ISMAR-Adjunct57072.2022.00028`
- Verena Biener & Jens Grubert. Content Transfer Across Multiple Screens with Combined Eye-Gaze and Touch Interaction - A Replication Study. In Adjunct Proceedings of IEEE International Symposium on Mixed and Augmented Reality 2022 (ISMAR 2022), pp.270-273. `https://doi.org/10.1109/ISMAR-Adjunct57072.2022.00060`
- Maheshya Weerasinghe, Verena Biener, Jens Grubert, Aaron J. Quigley, Alice Toniolo, Klen Čopič Pucihar & Matjaž Kljun. VocabulARy replicated: comparing teenagers to young adults. In Adjunct Proceedings of IEEE International Symposium on Mixed and Augmented Reality 2022 (ISMAR 2022), pp. 283-285. `https://doi.org/10.1109/ISMAR-Adjunct57072.2022.00064`
- Snehanjali Kalamkar, Verena Biener, Fabian Beck and Jens Grubert. Remote Monitoring and Teleoperation of Autonomous Vehicles—Is Virtual Reality an Option? In IEEE International Symposium on Mixed and Augmented Reality 2023 (ISMAR 2023), pp. 463-472. `https://doi.org/10.1109/ISMAR59233.2023.00061`
- Snehanjali Kalamkar, Verena Biener & Jens Grubert. VR for Remote Monitoring of Automated Vehicles: Replications Across Follow-

Up Evaluations. In Adjunct Proceedings of IEEE International Symposium on Mixed and Augmented Reality Adjunct 2023 (ISMAR 2023), pp. 3-4. `https://doi.org/10.1109/ISMAR-Adjunct60411.2023.00008`

- Verena Biener, Forouzan Farzinnejad, Rinaldo Schuster, Seyedmasih Tabaei, Leon Lindlein, Jinghui Hu, Negar Nouri, John J. Dudley, Per Ola Kristensson, Jörg Müller & Jens Grubert. Video Analysis of Behavioral Patterns During Prolonged Work in VR. In IEEE International Symposium on Mixed and Augmented Reality Adjunct 2023 (ISMAR-Adjunct), pp. 466-467. `https://doi.org/10.1109/ISMAR-Adjunct60411.2023.00101`

Eidesstattliche Versicherung

Hiermit versichere ich an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet habe.

Weiterhin erkläre ich, dass ich die Hilfe von gewerblichen Promotionsberatern bzw. – vermittlern oder ähnlichen Dienstleistern weder bisher in Anspruch genommen habe, noch künftig in Anspruch nehmen werde.

Zusätzlich erkläre ich hiermit, dass ich keinerlei frühere Promotionsversuche unternommen habe.

Bayreuth, den

Unterschrift