Augmented Everything: Engineering Compelling Ubiquitous AR Experiences

Zur Erlangung des akademischen Grades eines

Doktors der Wirtschaftswissenschaften
(Dr. rer. pol.)

von der KIT-Fakultät für Wirtschaftswissenschaften
des Karlsruher Instituts für Technologie (KIT)

genehmigte

Dissertation

von

Marius Franz Schenkluhn (M.Sc.)

Tag der mündlichen Prüfung: 01.07.2024

Referent: Prof. Dr. Christof Weinhardt

Korreferentin: Prof. Dr. Dr.-Ing. Dr.h.c. Jivka Ovtcharova

Karlsruhe, 2024

Abstract

Augmented Reality (AR) is expected to become the next leap in the interaction with information technology. By projecting digital information directly into the user's field of view, AR offers continuous access to knowledge and communication, everywhere and in a context-aware manner. Thereby, the interaction with digital content and applications fundamentally changes user experience. Multi-modal interaction with hands-free options through voice, gestures, and gaze allows for both natural and magical experiences. However, the promises of AR follow a repeated trajectory of overhyped expectations and underwhelming results. While significant efforts are put into continuously improving the maturity of the hardware, many questions regarding the user experience of the software for ubiquitous applications remain unanswered. Yet, as low entry barriers are crucial, this thesis sets out to investigate relevant aspects of the user experience in AR that keep the technology from reaching its full potential.

One fundamental interaction problem is the input of text, which is particularly challenging in AR due to the lack of a physical keyboard and the need for hands-free options. The human-computer interaction discipline has proposed various text entry methods for Virtual Reality (VR) and AR, yet, there is no convergence towards a standard method for AR due to the lack of comparability between the methods and a comprehensive understanding of the user's needs and preferences. Thus, key contributions of this dissertation include identifying design requirements, discussing human interaction factors and limitations, and evaluating the performance of three promising text entry methods for AR in a controlled laboratory study, highlighting the need for tailored and user-adaptive text input.

In addition, the application of AR poses specific challenges in the respective domain and requires a thorough understanding to craft tailored solutions. While it remains unclear if the single "killer application" of AR exists, providing a seamless user experience is a key factor for the success of AR applications and its platform. Applied to the domain of Smart Home control, which is a well-suited testbed, this thesis further investigates the user's needs and preferences for interaction with Internet of Things (IoT) devices through AR. After discussing design requirements with a focus group by applying universal design principles, the thesis focuses on automation within Smart Homes. Two novel concepts for automated Smart Home control via

AR-based indoor positioning and automated spatial setup are presented and the tradeoff between automation and perceived autonomy is evaluated in a second laboratory study.

For both perspectives, the design of text entry methods and the application of AR to the Smart Home domain, the results demonstrate the need for adaptive and user-centered solutions that are tailored to the specific context of use and user preferences. One-size-fits-all approaches that were applied to prior platforms are not able to sufficiently live up to the expectations of the successor of the smartphone and the acceptance of this technology.

Overall, this thesis contributes to research and practice by studying both grounded AR user experience aspects and exploring novel applications through various methods and mixed-methods approaches such as controlled laboratory studies, prototypical implementations on state-of-the-art hardware, user interviews, and a focus group.

Acknowledgement

I love languages. This does not primarily refer to foreign languages in the traditional sense, but rather to the languages of different disciplines, roles, and trades found in science, industry, or simply everyday life. The babelesque differences between the languages of programmers, designers, engineers, managers, creatives, students, and scientists necessitate interfaces. This, in my view, makes interdisciplinary fields both exciting and challenging, while also invigorating cross-disciplinary collaboration. Communication is key.

In the field of Human-Computer Interaction (HCI), the situation is similar. This interdisciplinary field aims to overcome communicative barriers between humans and computers. Communicating with often seemingly stubborn computers on how they should behave is an art in itself. They follow clear rules, which are not always obvious to us humans. Designing communication that works in both directions is a fascinating challenge. I hope that with this work, I have been able to make a sustainable contribution to the future human-computer interface of AR.

I am grateful for the support and accompaniment of many individuals along this journey, through active support, constructive criticism, stimulating discussions, and inspiring encounters.

My special thanks go to Prof. Christof Weinhardt for giving me the opportunity to pursue my PhD at the Institute of Information Systems and Marketing (IISM). Thank you for your trust and the freedom to set my own research priorities and develop teaching concepts. I thank Christian Peukert for his excellent supervision, integration into the DXP team while the world was still barricaded in home offices, and the many stimulating discussions about the talk of the town which we then sent into the ocean. Likewise, I thank Michael Knierim for seamlessly taking over supervision and sharing some of the most intense and satisfying brainstorming sessions with me, which were ultimately rewarded with a trip to Hawaii. I am grateful to Anke Greif-Winzrieth for her guidance on all administrative questions in the jungle of lab experiments, but also for always having an open ear. My thanks go to Thimo Schulz, for sharing (sometimes involuntarily) my enthusiasm for everything new in AR and for jointly designing and conducting seminars that we would have liked to attend as students ourselves. Overall, I thank the entire chair team for all the shared experiences and friendships.

I also thank Robert Bosch GmbH for the opportunity to carry out this industrial doctorate together, the trust in my work, and the opportunity to present the results worldwide. I am particularly grateful to Jens Müller for initiating the program and selecting this exciting topic area. My thanks also go to Andreas Klein, who always ensured that every obstacle was removed from our path, and to the entire group for their close integration in various development projects. I also thank the team of the Bosch Smart Life Lab for opening their doors, allowing me to significantly professionalize my research. Special gratitude goes to Francisco Kiss, who not only provided wisdom and HCI insights but also saved my vacation.

Thanks also to numerous anonymous reviewers who provided valuable and often challenging critiques of published papers. I remain fascinated by science's ability to create a global forum where researchers take so much time to review and improve their colleagues' work. In this regard, I also thank my examiners for their time and interest in my work.

Moreover, I am grateful for my friends and especially my small group for their companionship and support over the past years. Thank you for never tiring of having the differences between Augmented Reality and Virtual Reality explained to you. And thanks to those who shared my (too many) new hobbies acquired during my PhD time with me and helped me kickstart them. I thank my parents for their support, appreciation, and patience when they had to wait a few days longer for a sign of life.

A huge thank you goes to my wife Laura, who tolerated the transformation of our apartment into a nerdy Smart Home lab, supported me at every step with advice and action, and could now probably defend my dissertation better than I could. Thank you for your excitement, support, and love.

I thank everyone who has accompanied and supported me on this exciting, challenging, educational, frustrating, fascinating, enlightening, and valuable journey. And lastly, I thank my God for the many open doors and gifts that have enabled me to this path.

Contents

List of Figures

List of Tables

List of Abbreviations

AAL Ambient Assisted Living

ADL Activities of Daily Living

AI Artificial Intelligence

ANOVA Analysis of Variance

AoI Area of Interest

API Application Programming Interface

AR Augmented Reality

ARKB Augmented Reality Keyboard

ASR Automatic Speech Recognition

ATI Affinity for Technology Interaction scale questionnaire

AV Augmented Virtuality

BCI Brain-Computer Interface

CAGR Compound Annual Growth Rate

CER Corrected Error Rate

CPS Characters Per Second

DPx Design Principle

DSR Design Science Research

FDA Food and Drug Administration

FoV Field of View

FPS Frames Per Second

GUI Graphical User Interface

HCI Human-Computer Interaction

HMD Head-Mounted Display

HVAC Heating, Ventilation, and Air Conditioning

IS Information Systems

IVA Intelligent Virtual Assistant

IoMT Internet of Medical Things

IoT Internet of Things

IPS Indoor Positioning System

JSON JavaScript Object Notation

KIT Karlsruhe Institute of Technology

LLM Large Language Model

LiDAR Light Detection and Ranging

MRx Meta Requirement

MR Mixed Reality

MRTK Microsoft Mixed Reality Toolkit

NASA-TLX NASA Task Load Index

NFT Non-Fungible Token

PIR Passive Infrared

PARCS Prototypical Augmented Reality Configuration System

RQ Research Question

SD Standard Deviation

SDT Self-Determination Theory

SoTA State of the Art

SPA Single Page Application

SLAM Simultaneous Localization and Mapping

TAM Technology Acceptance Model

TENS Technology-based Experience of Need Satisfaction

TER Total Error Rate

ToL Transfer of Learning

UD Universal Design

UEQ User Experience Questionnaire

UER Uncorrected Error Rate

UI User Interface

UX User Experience

VR Virtual Reality

WPM Words Per Minute

XR Extended Reality

6DoF Six Degrees of Freedom

Part I

Fundamentals

Introduction 1

> *The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it.*
>
> — **Mark Weiser (1991)**,
> Former Chief Technologist at Xerox PARC

1.1 Motivation

In the rapidly evolving landscape of digital technology, the advancements in Artificial Intelligence (AI) and Large Language Models (LLMs) have significantly altered our understanding of machines. These technologies simplify complex tasks ranging from language understanding and image creation to copiloting data analysis and software development (Betker et al., 2023; Chowdhary, 2020; Moradi Dakhel et al., 2023). So far, their interfaces – usually simplistic chat windows – mask the complexity and potential inherent in these tools.

Somewhat hidden behind the latest advances in AI research, another technology is constantly evolving that may become the future interface not only to AI but to digital technology in general: Augmented Reality (AR). The unique affordances of AR – ranging from three-dimensional interactions to context-sensitive experiences – herald a new era of digital engagement, where the physical and virtual worlds blend seamlessly (Azuma, 1997). In this sense, AR could complement AI by transforming the frontend of applications, just as AI redefines their backend (Longo et al., 2021; Sahu et al., 2021).

In the discipline of Human-Computer Interaction (HCI), the long-standing vision has been to bridge the communication gap between humans and machines (Quigley et al., 2013). The concept of ubiquitous computing, as described by Weiser (1991), envisions technology seamlessly integrated into daily life, becoming an indistinguishable part of our existence rather than a distinct entity. The continuous miniaturization and enhancement of devices bring us closer to realizing these visions, with AR representing the next evolutionary step towards lightweight and pervasive digital integration (Yin et al., 2021).

AR's potential lies in its ability to create immersive, intuitive interactions by extending interfaces beyond screens into the user's surroundings as one digital canvas. Instead of using a mouse and keyboard or touchscreens, users control AR systems naturally with gestures, natural language, or gaze direction (Billinghurst et al., 2015). In addition, the context-sensitive nature of AR enables the processing of environmental data through various embedded sensors to personalize and contextualize applications and create believable, immersive experiences (Kim et al., 2021a). Controlling a Smart Home (Binary Banana, 2018), shopping in a supermarket (Nott, 2022; Peukert, 2020), or navigating a city (Google, 2019) no longer have to be done in a smartphone app but can take place directly in the user's real environment. Thanks to eye tracking, many of these interactions can be carried out hands-free (Lu et al., 2021). As with the smartphone, the wealth of information on the Internet, the performance of AI models, and communication platforms are available, however, directly in the user's field of view at all times.

Scientific research on AR has already produced significant results, ranging from optics, computer vision, and sensor technology to interaction design and human factors (Billinghurst et al., 2015). While pancake optics reduce the weight and size of AR glasses (Cakmakci et al., 2021), advancements in sensor technology, such as eye-tracking systems, enable increasingly precise and faster user detection, leading to more natural interactions (Carter & Luke, 2020). Simultaneously, advancements in computer vision facilitate the precise recognition of objects and real-time interaction with them (Guo et al., 2022). In interaction design, research has already yielded numerous innovations that, for example, allow for the natural reaching of distant objects (Poupyrev et al., 1996) or enhance interaction with virtual objects through haptic feedback (Zhu et al., 2020).

In addition, the transformative potential of AR has attracted significant investment expecting the emergence of a new device platform that might replace the smartphone and potentially other personal computing devices (Gallagher, 2023; Hatmaker, 2023; Porter & Heppelmann, 2017; Rauschnabel, 2021). Forecasts for the Compound Annual Growth Rate (CAGR) of AR include 31.5% (Skyquest, 2024), 39.8% (Grand View Research, 2024), or even 45.4% (Fortune Business Insights, 2024) until the early 2030s, with an estimated market size of $1.2 trillion for the Extended Reality (XR) market by 2035 (McKinsey, 2022). Moreover, the smartphone platforms iOS and Android represent a relevant market power, mainly through their app stores. For instance, Apple charges up to 30% commission on app sales and in-app purchases while dictating the terms of apps on its platform, including contents and design (Apple, 2024c). Recent action of the European Union further underlines the market power of digital platforms by taking regulatory action to address it through legislation such as the Digital Services Act and the Digital Markets Act (European Commission, 2024; European Union, 2022; Webster, 2024). The potential succession in the form of the detachment from physical screens thus represents the vast economic and social impact that AR could have. Recognizing these implications, companies like Meta and Apple are investing large sums in AR and Virtual Reality (VR) as the potential next-generation platform, aiming to replicate and perhaps surpass the utility and versatility offered by current devices and establishing their position as platform providers (Hatmaker, 2023; Heath, 2023; Leswing, 2020).

However, AR's success hinges not only on technological advancements but also on its ability to offer a diverse, useful, and innovative range of applications from its inception (Rauschnabel & Ro, 2016). Assuming the emergence of this technology as the next-generation platform, its implications extend far beyond significant economic impacts. The advent of previous technological innovations, such as smartphones and the internet, has already established prerequisites that disproportionately disadvantage non-digital natives and individuals with disabilities, at the very least amplifying the effort necessary to access services and information (Chadwick & Wesson, 2016; Hargittai, 2010). For proficient users, mastering complex operational paradigms and interactions tailored to a specific application may be feasible. However, this expectation is not realistic for a wide range of novice users (Thies, 2015; Ziefle & Bay, 2005). Consequently, the intricacies of AR technology should not be exacerbated by further complicating user interactions or escalating the demands placed on users. Therefore, crafting a well-designed User Experience (UX) is paramount for the technology's acceptance and inclusive and barrier-free digital society and is therefore imperative for the success of the technology for companies and legislators alike.

Previous attempts to penetrate the end-consumer market with AR glasses such as Google Glass in 2013 and Magic Leap One in 2018 have not met with success (Metz, 2022; Weidner, 2023). Only in the past couple of months have companies like Meta and Apple begun to address this market again with Mixed Reality (MR) Head-Mounted Displays (HMDs) that enable AR via video-see-through displays (Apple, 2024a; Meta, 2024b). This raises the critical question regarding the prerequisites for achieving widespread acceptance of such technology: What makes this technology ready enough?

Research within the domain of technology acceptance suggests that factors including functional benefits, technology affinity, and social norms significantly influence the adoption of AR technologies (Rauschnabel & Ro, 2016). At the same time, transitioning from smartphones to AR glasses does not present a marked reduction in device size as observed in the transition from mainframes to personal computers to smartphones. Instead, it introduces the necessity for consumers to wear relatively bulky glasses, at least in the initial stages.

Thus, the crux of fostering acceptance for AR glasses lies in delivering an initial UX that is not only compelling but also capable of competing with established technological platforms (Billinghurst, 2021). Achieving this objective necessitates the development and consistent application of new UX principles specifically tailored to the unique interaction paradigms of AR (Dey et al., 2018). These principles must transcend traditional interactions bound by screens and physical input devices, instead embracing and leveraging the potential of three-dimensional space for a truly immersive AR experience.

UX design encompasses both foundational research into basic interaction mechanisms and applied research focusing on real-world applications (Hassenzahl & Tractinsky, 2006). Despite significant advancements in the field, numerous challenges and unexplored opportunities remain (Billinghurst, 2021; Dey et al., 2018).

Thus, this dissertation focuses on two parts, fundamental and applied aspects of UX research, aiming to enrich the field's understanding from multiple perspectives and contribute to making technology more accessible to the general public.

The first, fundamental part of the dissertation examines the intricacies of text input in AR. Text entry is a ubiquitous interaction across digital platforms, yet it presents a unique challenge within the AR context. However, the development of the iPhone, for example, had to be paused and developer resources reallocated because no satisfactory way of entering text had been found at that time (Balakrishnan & CNBC, 2017). The prevalent QWERTY keyboard layout, although originally designed for typewriters, persists across digital devices despite recurrent scrutiny regarding its efficacy (Noyes, 1983). While voice input and physical keyboards are valid and useful means for text entry in AR, they are not always suitable or desirable (Schenkluhn et al., 2023a). An optimal method for text input in AR remains undetermined, with the risk of early successful products inadvertently setting a de facto standard (Dube & Arif, 2019; Xu et al., 2019a).

Text input research in VR and AR is already well-established, offering various proposals for implementation. However, the evaluations of individual text input candidates often cannot be compared with each other and research is not yet converging on a convincing standard method (Dube & Arif, 2019; Speicher et al., 2018).

The second part of the dissertation explores AR-based control of Internet of Things (IoT) devices. In everyday use, not only the interactions with the device itself but also the interactions with other devices through AR are an interesting subject of research. IoT is playing an increasingly important role in numerous industries and will become even more important in a connected world in the future (Gubbi et al., 2013).

One prominent application of IoT is the Smart Home, where various devices are interconnected to create a more comfortable and efficient living environment. The Smart Home environment serves as an exemplary context for applied AR research, particularly due to the natural integration of AR interactions within the user's physical space as opposed to abstract digital interfaces such as entity lists on smartphones. This contextual advantage, coupled with the higher likelihood of technology adoption among Smart Home users, renders the Smart Home a compelling case for investigating AR's potential in enhancing IoT device control. When conducting scientific studies, the home environment is also expected to be more comprehensible for random participants without prior domain experience than imagining themselves in a manufacturing, surgery, or workshop environment. Additionally, resulting research implications could be generalized to other use cases, such as Industry 4.0, smart cities, or the Internet of Medical Things (IoMT).

Current Smart Home solutions often suffer from complexity and lack of interoperability among different devices and systems (Purdy, 2022), presenting a challenge that AR could address by enabling direct interaction with IoT devices and rendering IoT controls in a unified spatial user interface. Although research in this domain is nascent, its exploratory nature complements the established field of text input research in this thesis. Consequently, merging AR with Smart Home technologies not only enriches scientific inquiry but also harbors significant practical application potential.

1.2 Research Agenda and Research Questions

The general objective of this dissertation is to critically examine and enhance the UX in future AR applications, to uncover untapped potentials, and to make the technology more accessible and useful to the general public. This goal is motivated by the fact that AR technology exhibits greater technical complexity than previous end-user platforms, thereby amplifying the significance of UX design in making the technology accessible to a wider audience (Billinghurst, 2021; Wagner et al., 2019). Furthermore, AR technology has not yet reached a level of widespread adoption where established standards for UX have been developed, indicating that research in this area can still influence essential design decisions and make valuable contributions to the field.

To provide both substantial and broad contributions to research, the objective is focused on two main areas, whose relevance for AR UX research was motivated above. Text entry in AR is an established research field that has yet to produce a convergent solution. Therefore, this work aims to comprehensively investigate the requirements and comparatively evaluate existing, promising proposals for text entry to contribute to the convergence of the field. Simultaneously, research in the area of text entry can make an important contribution to foundational research in AR UX. On the other hand, AR-based control of IoT devices is a nascent research field that holds significant potential for AR technology, especially due to the rising importance of IoT and Smart Home technologies in industry and daily life, offering substantial opportunities for mediated Human-Computer Interaction in general.

From this general research agenda, a series of research questions are derived to be addressed in this work. All research questions tackled in this work are introduced and contextualized within the research framework in this section.

A significant portion of text entry research focuses on VR systems (Dube & Arif, 2019). While some aspects are transferable, issues such as occlusion do not exist in AR, as the environment, including hands and potentially a physical keyboard, remains visible (Grubert et al., 2018; Walker et al., 2017). Therefore, the first research question primarily addresses the requirements specific to text entry in AR systems. Additionally, stationary use of AR HMDs can employ regular physical keyboards as with desktop PCs or laptops. Hence, focusing on mobile use of AR systems is intriguing, where text entry performance still faces significant trade-offs (Dube & Arif, 2019). This leads to the formulation of the initial research question:

> **Research Question 1.1 (RQ1.1)**
> How to design a mobile virtual keyboard for AR systems to increase text entry performance?

Since text entry is only one of many new interaction types in AR systems, ensuring quick learnability of text entry modes is also crucial to avoid hindering ubiquitous use of AR systems:

> **Research Question 1.2 (RQ1.2)**
> How to design a mobile virtual keyboard for AR systems to increase learnability?

After highlighting the technical limitations of existing systems and deriving requirements and design principles for text entry in AR systems (Schenkluhn et al., 2022b), human factors are considered as the counterpart following Norman (2013). Various studies have shown that gaze-based text entry is a promising option for writing text (Kristensson & Vertanen, 2012; Kurauchi et al., 2016, 2020; Xu et al., 2019a). Dwell-based typing, a form of gaze-based text entry, is particularly interesting as it allows for hands-free input. This involves focusing gaze on a letter and selecting it after a defined duration without further action (Majaranta et al., 2004, 2009; Yu et al., 2017). The dwell time until letter selection is a crucial factor affecting the performance of gaze-based text entry. Therefore, the following research question aims to investigate the human limit for dwell-based gaze-typing depending on text length:

> **Research Question 2 (RQ2)**
> What is the influence of dwell time and text length on gaze typing performance?

Based on insights from previous research questions that primarily illuminated the problem space of text entry in AR (Schenkluhn et al., 2022b), the goal is to eventually determine which text entry method is most suitable for mobile use of AR systems. Since the solution space has been well researched individually but the various studies are often not comparable (Dube & Arif, 2019; Speicher et al., 2018), a comparative evaluation is conducted to investigate different methods. This evaluation will not only consider the performance of methods but also users' subjective preferences. As with previous research questions, the focus is on mobile use of AR systems since established methods like physical keyboards or voice input are more suitable for stationary systems (Dube & Arif, 2019). Thus, the following research question concludes the text entry research in this work:

Research Question 3 (RQ3)
Which text entry method is most suitable for mobile AR devices in terms of performance and user preference?

In the domain of AR-based control of Smart Home devices, an initial investigation is conducted to ascertain users' requirements and needs for AR interfaces in Smart Home applications. This precedes the development and evaluation of concrete AR interfaces for such applications. Within the so-called Universal Design approach, particular attention is given to the needs of people with disabilities and older adults to ensure high usability and accessibility when designing things (Steinfeld & Maisel, 2012). Moreover, employing Universal Design is hoped to provide a more comprehensive understanding of user needs since these groups often have specific requirements whose practical implementation can be beneficial for a broad user base (Steinfeld & Maisel, 2012). Thus, the first step in exploring the Smart Home context leads to the following research question:

Research Question 4 (RQ4)
What are possible use cases and requirements for AR-based applications that are accessible and inclusive for people with different abilities and disabilities?

The results of the study addressing RQ4 underscore the necessity for adaptive automation in Smart Home applications. Automations within the Smart Home context can be triggered by various events, such as the user's presence in a specific room or time of day. From a UX perspective, maximizing automations' capabilities and scope could reduce interaction frequency, workload, learning needs, and potential frustration. If a Smart Home system is aware of the user's position within the home, it can support the user with context-aware automations. Moreover, as each user's needs may vary, identifying specific users allows for the personalization of automations. AR HMDs are equipped with sensors capable of both identification and precise indoor positioning of the user, thus offering potential for adaptive Smart Home interactions and automations. To explore this potential, the following research question is formulated:

Research Question 5 (RQ5)
How can sensor technology of AR glasses be leveraged for precise in-

door user positioning, and what implications does this have for Smart Home automations?

Since not all interactions in a Smart Home can be automated, active engagement with IoT devices in a Smart Home remains an important aspect of Smart Home interaction. Through AR HMDs, this interaction can occur not only via lists or voice commands but directly in the spatial context of the user. For the system to map each virtual entity to its physical location in space, an initial spatial configuration of Smart Home devices is necessary. However, with a large number of devices, this process can quickly become cumbersome and time-consuming. Thus, in addition to using automations for interacting with the Smart Home, automating the spatial setup process also presents an opportunity. During the development of a prototype for spatial configuration of Smart Home devices using AR HMDs, it became clear that maximizing automation is not necessarily desirable. Although UX dimensions such as usability, cognitive load, and frustration are improved through a high degree of automation, the psychological needs of users must also be considered to promote technology acceptance. These include needs for competence and autonomy, which could be undermined by high automation (Ryan & Deci, 2000). This trade-off between classical UX dimensions and psychological needs in the context of the spatial configuration of Smart Home systems is therefore examined in the following research questions:

Research Question 6.1 (RQ6.1)
Does a manual Smart Home spatial setup design maximize psychological need recognition, and do classical usability dimensions undermine the benefit of this characteristic?

Research Question 6.2 (RQ6.2)
Does a fully automated spatial setup maximize classical UX dimensions like ease of use, mental workload, and frustration, but reduce the attractiveness of the interaction design by thwarting psychological needs?

Research Question 6.3 (RQ6.3)
Does a combination of manual and automated features strike an effec-

tive balance between psychological needs recognition and classical UX dimensions, effectively enhancing technology acceptance?

To address these research questions, this thesis employs a variety of methods from HCI research. Requirements are analyzed, prototypes are developed accordingly, and evaluated through controlled laboratory studies. These nine research questions lay the foundation for the subsequent chapters, which will be introduced in detail in the following section.

1.3 Structure of Dissertation

This dissertation is structured into four cohesive parts: Foundations (I), the primary subjects of Text Input in AR (II) and AR-based Smart Home Control (III), and concluding with the final considerations (IV). It spans 13 chapters, each contributing to a comprehensive analysis and enrichment of the field concerning UX in AR.

Part I establishes the groundwork by presenting the objectives and state-of-the-art concepts guiding this research. **Chapter 1** delves into the thesis's motivation and research questions, while **Chapter 2** explores the theoretical underpinnings and current state of AR and UX research.

Part II addresses a critical area of UX research: text input within AR environments. Introduced in **Chapter 3**, this section initiates with a comprehensive requirements analysis in **Chapter 4** to ascertain the prerequisites for an effective AR text input method. This analysis evaluates user, application, and contextual needs based on extant literature and theories. **Chapter 5** complements the application side by focusing on the human limitations associated with gaze-based text input, proposing a study design for its empirical examination. **Chapter 6** introduces Xperisight, a software tool designed for the remote control and parallelization of AR experiments, facilitating the comprehensive evaluation in **Chapter 7** of various text input methods. This study distinguishes itself by not merely expanding the repertoire of exploratory ideas but by implementing and comparing existing methods for optimal comparability. The results indicate that no singular text input method prevails; rather, the selection hinges on application specifics, context, and user preferences. Recommendations for system designers emphasize accommodating diverse text input methods, granting users the autonomy to choose. Further research implications suggest that text input performance is not solely determinant. Subjective perceptions and distortions concerning keyboard characteristics equally influence user preferences.

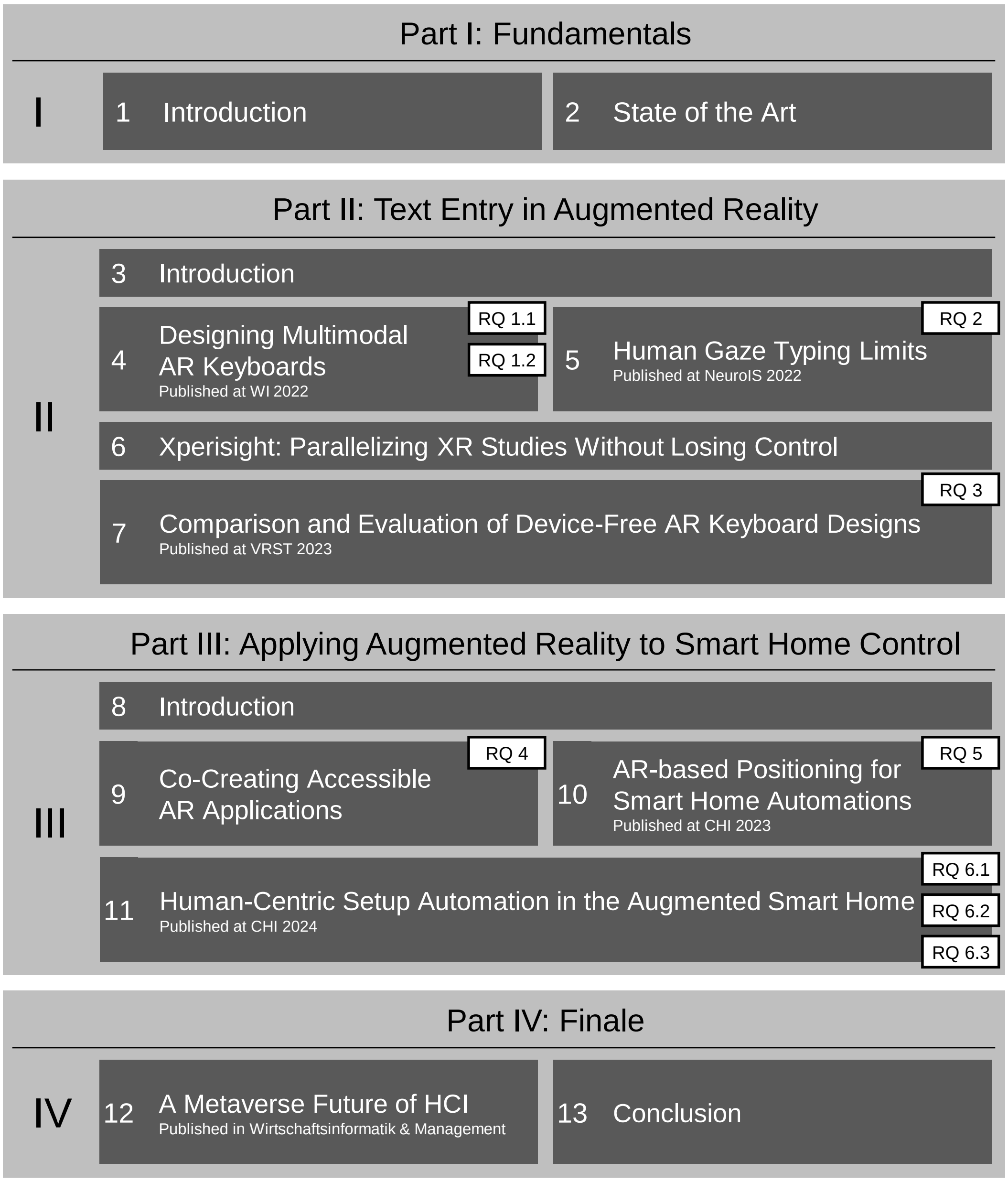

Figure 1.1.: Structure of the dissertation with the four parts and their respective chapters.

Part III ventures into applied UX research through the lens of designing AR interfaces for Smart Home applications. Following an introduction in **Chapter 8**, **Chapter 9** outlines a focus group study exploring the needs and requirements of individuals with disabilities and older adults concerning AR Smart Home interfaces, adopting a universal design framework. This approach endeavors to create accessible and inclusive products and environments, benefiting a wide range of users. The study underscores how automating daily interactions can significantly enhance the quality of life for these groups.

Subsequently, **Chapter 10** explores the innovative use of AR glasses' sensor technology for precise indoor user positioning, highlighting its superior advantages over traditional sensors without compromising system latency. **Chapter 11** then delves into the development and evaluation of an AR interface for Smart Home device control, critically assessing the desirability of automation from the user perspective. Findings suggest that striking a balance between automation and user autonomy is crucial for technology acceptance.

In **Part IV**, **Chapter 12** offers forward-looking perspectives on integrating AR within the metaverse, examining its potential implications across various hypothetical scenarios. The conclusive **Chapter 13** summarizes the dissertation's core findings, discussing their general implications for both academia and industry practice, and outlining potential avenues for future research. An overview of the dissertation's structure and Research Questions (RQs) is depicted in Figure 1.1.

State of the Art 2

This chapter will provide an overview of the current state of the art of AR within the HCI domain to identify research gaps that will be addressed in the course of the dissertation.

The terminology surrounding AR is important to differentiate, even if the boundaries are becoming increasingly blurred and usage is often inconsistent, as they are used both in science and corporate marketing and are subject to constant technological change (Rauschnabel et al., 2022). For instance, Microsoft coined their AR and VR platform *Windows Mixed Reality*, using MR as an umbrella term that includes both AR and VR (Janssen, 2017). At the same time, the company differentiates AR from MR by the amount of interaction between virtual objects and the real environment in that AR has limited interactions as an overlay (Microsoft, 2024). Similarly, Meta (2024a) views MR as an advanced form that is capable of blending virtual objects into the physical environment, while AR is only an overlay that is interactive but mostly separated from the real environment. However, Meta contradicts the prior statement by referring to the IKEA Kreativ app[1] as an example of AR, which overlays IKEA furniture over the real environment and, thus, registers in the real environment (Meta, 2024a). Intel (2024) describe VR as the umbrella term and differentiates AR from MR by the degree of user interaction. In contrast, Apple (2024a) uses the term *Spatial Computing* for its new Apple Vision Pro headset, moving away from the prior usage of AR, e.g., for their ARKit platform (Apple, 2024b) and even requesting third-party developers to use the term Spatial Computing over AR, VR, MR, or XR (Apple, 2024d).

These discrepancies reveal that the terms and definitions of AR and related technologies are not uniform and are in a state of flux. In science, the terms are more clearly defined, but here too different views are reflected in the literature. Depending on the definition, different perspectives are adopted that relate to the technology, the interplay between the real and virtual worlds, or the user experience. For example, Azuma (1997) describes AR systems as those that combine real and virtual elements (1), are interactive in real-time (2), and register in 3D (3). The definition is not bound to specific technologies, but rather to the user experience. Therefore, several

[1] https://www.ikea.com/us/en/home-design/learnmore/ (Accessed: 12.04.2024)

form factors can be used to realize AR, such as HMDs, smartphones, or projectors (Azuma et al., 2001). To differentiate AR from related concepts, Milgram and Kishino (1994) describe the virtuality continuum, which ranges from the real world to the virtual world. In between are AR and Augmented Virtuality (AV), which describe states in which real and virtual elements are combined. AR is understood as the insertion of virtual elements into the real environment, while AV describes the insertion of real elements into the virtual environment. However, Milgram and Kishino (1994) admit that the distinction between AR and AV can be difficult as future technologies are increasingly converging and the boundaries are blurring. The factors influencing this distinction are understood differently in literature (Klopfer & Squire, 2008; Wu et al., 2013). MR is understood by Milgram and Kishino (1994) as an umbrella term for AR and AV, which covers the entire range of the spectrum, but not the extrema of pure reality and VR. Milgram et al. (1995) describe AR as a subclass of MR and "a form of virtual reality where the participant's HMD is transparent, allowing a clear view of the real world", binding the definition to the HMD technology. Other sources define AR as a real-time view of the physical real-world environment that is augmented by virtual computer-generated information (Carmigniani et al., 2011) or by the differentiation of superimposing real, existing objects (AR) and possible but non-existing objects (MR) onto the real world (Farshid et al., 2018). Rauschnabel et al. (2022) argues that AR and VR should be considered as separate concepts, as they are technically related but completely different experiences from the user's point of view.

Whether a one-dimensional continuum is sufficient to describe and differentiate AR and related technologies is also controversial. Augmentation does not necessarily have to be visual, but can also be auditory, haptic, olfactory, or a combination (Azuma, 1997; Geronazzo & Serafin, 2023; Skarbez et al., 2021; Speicher et al., 2019). Additional dimensions could be the extent of world knowledge, reproduction fidelity, the extent of presence (Milgram & Kishino, 1994), immersion, coherence (Skarbez et al., 2021), or additional senses (e.g., sonar, radar) (Mann et al., 2018). The concept of the continuum is also called into question (Skarbez et al., 2021). If one follows Sutherland (1965), the term virtual world alone can be graded from an environment of digital elements such as a VR game to a "The Matrix"-like world in which "a bullet [...] would be fatal" (Sutherland, 1965), separating purely visual sensation from comprehensive immersion (Speicher et al., 2019).

For this dissertation, it is sufficient to understand AR based on the definition of Azuma (1997). The Microsoft HoloLens 2, which is used as the reference device in this paper, is therefore understood as an AR device. If required, the term MR is used to describe systems that are able to switch between pure virtuality and pure reality, i.e. that cover the entire spectrum, following one possible interpretation of MR (Speicher et al., 2019). XR is used as an umbrella term for AR, VR, and MR (Rauschnabel et al., 2022) While until recently there was a clearer distinction between AR hardware and VR hardware, numerous devices can now enable a smooth transition by projecting external camera images onto the displays of headsets with low latency and high image quality. The limitation of these video-see-through devices is the recording and display quality, which cannot yet match that of the human eye (Rolland et al., 1995). Current examples include the Apple Vision Pro (Apple, 2024a), Varjo XR-4 (Varjo, 2023), and the Meta Quest 3 (Meta, 2024b) in various price segments.

AR has evolved significantly since its inception, with advancements in tracking, display, development tools, and interaction technologies (Billinghurst et al., 2015). Despite these advancements, the field experienced a period of stagnation in the past couple of years, described as the "VR Winter," due to limitations in portability, computational power, and optics (Evans, 2020; Steffen et al., 2019). However, interest in AR has been rekindled, as evidenced by strategic acquisitions by tech giants like Apple, Alphabet, and Meta, indicating a belief in AR as the next major computing platform (Bastian, 2017, 2020a, 2020b, 2020c). Meta's CEO Mark Zuckerberg supposedly wrote an e-mail to COO Sheryl Sandberg and others in 2015 about VR / AR stating: "VR / AR will be the next major computing platform after mobile in about 10 years. [...] Once you have a good VR / AR system, you no longer need to buy phones or TVs or many other physical objects - they can just become apps in a digital store" (Zuckerberg, 2015).

The scientific community has also shown growing interest in AR, with a significant increase in publications across various domains such as manufacturing, education, retail, construction, and surgery (Kohn & Harborth, 2018; Oesterreich & Teuteberg, 2017; Sommerauer & Müller, 2018; Tepper et al., 2017). A search on the Web of Science database[2] reveals that the number of publications is rising each year; from 161 publications in 2010, the number grew to at least 2212 in 2023 as measured by the Web of Science. The research highlights the benefits of AR, including enhanced learning outcomes, improved customer satisfaction, and increased efficiency in tasks (Billinghurst et al., 2015; Kammler et al., 2019; Tarafdar et al., 2019).

AR applications typically utilize devices like see-through smart glasses or head-mounted displays (HMDs), handhelds (smartphones or tablets), and projectors (Kohn & Harborth, 2018). Among these, HMDs are preferred for proof-of-concept applications due to their advanced capabilities (Kohn & Harborth, 2018). While some

[2]https://www.webofknowledge.com (Accessed: 12.04.2024)

current devices compromise on features and performance to achieve a lightweight and slim form factor such as the Xreal Air 2 (XReal, 2024) or Meta Ray-Ban Stories (Ray-Ban, 2024), others prioritize performance, sensors, and image quality at the expense of weight and bulk like the Microsoft HoloLens 2 (Microsoft, 2020) or the Magic Leap 2 (Leap, 2024). The Microsoft HoloLens 2 has seen increased popularity for application in scientific studies (Kortekamp et al., 2019).

AR devices have seen significant advancements in recent years with technological convergence along distinct dimensions toward the goal of ubiquitous and unobtrusive yet powerful devices. Current examples for the application of AR are that the US Food and Drug Administration (FDA) has approved the first AR-based application for surgical navigation (Dador, 2023), NASA has developed AR-based applications for the maintenance and repair of spacecraft (NASA, 2021), and the US Army is testing AR for training and combat (Shakir, 2023).

Despite the potential of AR, widespread adoption has been hindered by several challenges. Challenges in the development of AR hardware include the miniaturization of components, the optimization of power consumption and heat dissipation, and the improvement of display technology to achieve high resolution, high refresh rates, and a wide field of view (Billinghurst, 2021). Moreover, cognitive overload due to new interaction techniques, occlusion problems, and misalignment between AR affordances and industry requirements are challenges for AR (Azuma, 2017; Kohn & Harborth, 2018; Oesterreich & Teuteberg, 2017; Sommerauer & Müller, 2018; Steffen et al., 2019; Tarafdar et al., 2019; Tian et al., 2010). Furthermore, the design of comprehensive AR applications is constrained by a lack of knowledge about AR systems design (Duenser et al., 2007).

In UX research, challenges for AR research include the high effort and complexity of conducting user studies as the expensive equipment and space requirements limit the scalability of studies, the lack of standardized evaluation methods, and the need for interdisciplinary skills and knowledge from various domains such as computer science, psychology, and design (Billinghurst et al., 2015). To implement prototypes, researchers often rely on (modified) off-the-shelf hardware and software frameworks from the industry and either have to create custom software themselves or commission external service providers (Kortekamp et al., 2019; Peukert et al., 2019). However, there are also successful efforts from scientists to create open-source software frameworks for VR and AR such as PolyVR (Häfner, 2019).

To address these challenges, applying HCI principles to AR design is crucial. This includes using direct 3D manipulation for more intuitive interactions, minimizing cognitive overload through user-friendly UIs, reducing physical effort and enhancing learnability for better user acceptance, ensuring user satisfaction through immersive experiences, and incorporating flexibility and error tolerance into design (Billinghurst et al., 2015; Dey et al., 2018; Duenser et al., 2007; Evans et al., 2017). Moreover, UX design principles such as simplicity, consistency, and depth should be emphasized to decrease cognitive load and enhance realism (UXDA, 2020; Vi et al., 2019). UX researchers also advocate for understanding hardware capabilities, ensuring comfort and safety, immersing users with convincing illusions and audio, and lowering the entry barrier by guiding users and reusing familiar User Interface (UI) patterns (Babich, 2020; Prilla et al., 2019).

To further investigate interaction design in AR and create an overview, a structured literature review was conducted for the time period from 01.01.2011 to 31.12.2021 following Webster and Watson (2002)[3]. The search query consisted of a combination of the technology (AR or MR, HMDs, Smart Glasses, HoloLens, Varjo, Magic Leap), the research area (Interaction Design, HCI), UX (UX, Design Principles and Elements, Usability). Starting with 241 papers from Scopus[4], Web of Science[5], and the Association for Information Systems (AIS) eLibrary[6], the selection was narrowed down to 34 papers through scanning abstracts and papers and forward/backward searches. The selection was based on the relevance to future interactions with AR, design principles for user-friendly interaction, and future user profiles.

A resulting concept matrix categorizes the 34 papers into interaction techniques, design principles, UX, and user profiles, further divided into 8 main and 30 subcategories. Literature suggests various supplementary devices to enhance interactions with AR headsets, aiming to facilitate complex visual tasks and improve daily micro-interactions. These include wearables (smart rings, bracelets, watches, and belts) (Lee & Hui, 2018), mice/controllers with varying degrees of freedom (Hoppe et al., 2017; Özacar et al., 2017; Saidi et al., 2019), on-body interaction ("Body-as-a-Device") (Lee & Hui, 2018), smartphones for precise 2D interactions (Lee & Chu, 2018; Sun et al., 2020), keyboards for virtual text entry (Jiang et al., 2018), and additional cameras for improved remote collaboration (Teo et al., 2018).

Freeform control in AR encompasses gesture and motion control (Bertarini, 2014), touch control (Lee & Hui, 2018), voice control (Wang et al., 2021), and eye-tracking

[3]The literature review was conducted as part of a supervised seminar by David Kappelmann. This section is based on the results and the resulting term paper.
[4]https://www.scopus.com (Accessed: 12.04.2024)
[5]https://www.webofknowledge.com (Accessed: 12.04.2024)
[6]https://aisel.aisnet.org (Accessed: 12.04.2024)

(Lee & Hui, 2018; Wang et al., 2021). Gesture control is highlighted as particularly relevant, with eye-tracking recognized for its potential to speed up input tasks.

Tasks in AR interaction are categorized into manipulation/change, selection, navigation, and application control (Messaci et al., 2022). Specific tasks of interest include text entry (Lu et al., 2020), interaction with external devices (Knierim et al., 2019; Lin et al., 2016), interaction with tags for quick device pairing (Sorgalla et al., 2018), authentication through picture passwords (Hadjidemetriou et al., 2019), and interaction with virtual agents (Wang et al., 2019).

The modality of interaction – whether singlemodal or multimodal – is discussed, with multimodal interactions seen as a way to enhance user experience by combining gestures with voice commands for more efficient and accurate control (Wang et al., 2021).

Design principles and guidelines from the literature focus on improving immersive environments, UX design for AR applications, and specific guidelines for industrial service scenarios (Greenfeld et al., 2018; Vi et al., 2019; Xu et al., 2020). Notably, comprehensive guidelines by Vi et al. (2019) and principles for awareness of interaction boundaries by (Xu et al., 2020) are highlighted.

Future user profiles are categorized into work/industrial, private/domestic, and public/everyday settings. Each category outlines potential applications and developments in AR technology use (Knierim et al., 2019; Kymäläinen, 2016; Lee et al., 2021b).

Overall, supplementary devices enhance interaction precision but often require additional equipment (Lee & Hui, 2018). Gesture control is prevalent in current AR devices, with eye-tracking emerging as a promising technology (Bertarini, 2014). Specific tasks like text entry remain relevant, with innovative approaches to intuitive interaction (Lu et al., 2020). Additionally, multimodal interactions could improve user experience significantly (Wang et al., 2021). The design principles and guidelines offer valuable insights but must be contextualized within specific studies (Vi et al., 2019).

The concept matrix provides a detailed overview of the current state of AR interactions and is depicted in Table 2.1.

Table 2.1.: Concept matrix of the reviewed literature in the context of Augmented Reality interactions.

	Interaction Free Form Control				Interaction Cross-Control						Interaction Tasks				Specific Interaction Tasks					Modality of Interaction		Guidelines / Principles		User Profiles			Concerns			
	Gesture / Movement	Touch	Voice	Eye/Gaze-Tracking	Mouse / Controller	Smartphone	Wearables	On-Body-Interaction (Body as Device)	Additional Camera(s)	Keyboard	Manipulation	Navigation	Selection	Application Control	Text Entry	Authentication	Interaction with Virtual Agent	Interaction with external Devices (Printer, other Comp., Smart Devices, etc.)	Interaction with Tags (QR-Code, etc.)	Singlemodal	Intermodal / Multimodal	General Interaction Design (Filtering Light, Resolution, Focus Distance, etc.)	UX Guidelines	Working/ Industrial Environment	Private/ Domestic Environment	Public/ Everyday Environment (Activities, Stores, etc.)	Security, Privacy & Safety (indirect Interaction)	General Interaction Challenges (Click/Touch Detection, Boundary Awareness, Interaction Overload, etc.)	Technical Ch allenges (Latency, Material etc.)	(Smart) Environment Challenges (Context-Awareness, etc.)
Greenfeld et al. (2018)																						X	X							
Lu et al. (2020)	X			X										X	X					X	(X)									
Lee and Hui (2018)	X	X	X				X	X			X	X	X	X	X					X	(X)							X		
Wang et al. (2021)	X		X	X							X		(X)	X						X	X									
Hoppe et al. (2017)	X	X			X						X									X										
Vi et al. (2019)																						X	X							
Özacar et al. (2017)	X			X	X								X							X	(X)									
Xu et al. (2019b)	X										(X)		X	X						X	(X)									
Lee and Chu (2018)		X				X					X			(X)						X										
Wang et al. (2019)	X	X	X											X			X													
Xu et al. (2020)	X										X											X						X		
Oh et al. (2020)	X	X					X				X	X	X							X										
Frutos-Pascual et al. (2019)	X		X								X									X		X								
Kymäläinen (2016)																						(X)		X	X	X				
Xiao et al. (2018)	X	X									X									X								X	X	
Sun et al. (2020)		X				X							X							X										
Ham et al. (2014)	X	X					X				X	X	X	X						X	X									
Lee et al. (2021a)																						(X)				X				X
Lin et al. (2016)	X												X	X				X	X	X	(X)	(X)								
Messaci et al. (2022)	X										X	X	X	X						X										
Sorgalla et al. (2018)	X													X				X	X	X					X					
Thomas et al. (2020)	X		X									X		X						X	X	X		X						
Meli et al. (2018)	X						X		X		X									X	(X)									
Knierim et al. (2019)																		X							X		X	X		
Teo et al. (2018)	X								X		X		X							X				(X)						
Jiang et al. (2018)	X	X								X				X	X					X										
Hadjidemetriou et al. (2019)	X													(X)		X				X										
Lebeck et al. (2018)																						(X)	(X)				X	X		
Bailly et al. (2019)	X													X						X										
Saidi et al. (2019)	X	X			X			X			X		X							X										
Lee and Kim (2017)		X			X									X	X					X										
Kim et al. (2018)																								X	X	X				
Bertarini (2014)	X	X					X																				X	X		
Berkemeier et al. (2019)																						X		X						

Part II

Text Entry in Augmented Reality

3 Introduction

This part of the thesis focuses on a foundational aspect of AR UX by investigating text entry in AR applications. Recent demonstrations of AR have highlighted engaging spatial features while avoiding text input mechanisms (Apple, 2024a; Meta, 2024b). This trend is not indicative of diminishing importance but stems from the absence of a satisfactory solution for text input within a comprehensive AR system. The introduction of any novel technological device necessitates a reevaluation of interaction modalities, including those for text input. AR devices, with their array of sensors, provide a plethora of opportunities for both uni- and multimodal interactions. Nonetheless, it is critical to thoroughly understand the problem space before proposing any solutions. The initial approach to this problem was to follow a Design Science Research (DSR) process (Kuechler & Vaishnavi, 2008) to generate design knowledge concerning the learnability and performance of AR keyboards. Drawing upon transfer of learning theory and HCI literature on virtual keyboards, meta requirements were derived and initial design principles were suggested in Chapter 4.

Hands-free gaze typing emerges as a promising method to maintain mobility while enabling both rapid and precise text input. However, previous studies have encountered difficulties due to participants' lack of experience with gaze typing. To complement the findings from Chapter 4, a study design that eliminates the learning curve and provides insights into future gaze typing performance is presented in Chapter 5. Specifically, it examines the impact of varying dwell times on typing performance over time. By addressing user-specific constraints, the study lays the groundwork for developing user-adaptive gaze typing systems that minimize fatigue. Due to time constraints and prioritization in favor of a comparison study, the suggested design was not implemented in this thesis.

Based on the design requirements, a multimodal AR keyboard prototype was developed that combines eye-gaze input with a pinch gesture for text entry. The pinch gestures are built upon the concept of touch-typing, that is, mapping each character to a specific finger. By reusing already known interaction patterns, the goal is to minimize the learning curve and maximize input speed. However, initial trials of the prototype were characterized by a slow learning curve, high cognitive load, and unergonomic hand postures, leading to the decision not to further develop this prototype.

The insights gained from the iterative cycles of design, implementation, and evaluation led to the hypothesis that the requirements for text entry depend on the application scenario and are not solely based on performance. However, since performance has been the focus of prior research, a subsequent study was developed that prioritizes user preferences. The field has already seen a variety of text entry solutions, but there is a lack of comparability between these solutions (Dube & Arif, 2019; Speicher et al., 2018). Therefore, three promising text entry solutions were selected, optimized based on previous development insights and requirements, implemented, and compared in a study. As with previous studies, the non-stationary context was chosen as the application area, where the use of speech and physical keyboards is often not possible or impractical.

Thus, the study compares three promising device-free text-entry solutions for AR on the Microsoft HoloLens 2: dwell-based eye-gaze input, eye-gaze with pinch-gesture-commit input, and mid-air tap typing on virtual QWERTY keyboards. A controlled within-subjects lab experiment with 27 subjects measures typing performance, task load, usability, and preference across the three keyboards. Users expressed distinct preferences for the respective keyboards and evaluated the advantages and disadvantages differently. Considering diverse usage scenarios, subjects would even prefer these input modes over speech or physical keyboard input. The results suggest that virtual keyboard design should be tailored to individual user preferences. Therefore, this study provides essential insights into designing AR keyboards for heterogeneous user groups. The study is discussed in detail in Chapter 7.

Additionally, to facilitate the remote supervision of the study and improve the scalability of the session, a software tool was created – Xperisight. Xperisight provides remote access to Unity-based XR applications to oversee multiple sessions in parallel via one unified dashboard. Without influencing subjects by their presence, experimenters can access relevant information, remotely control the devices, and be called for help if questions or errors arise. The application of Xperisight in the study described in Chapter 7 effectively halved the overall required experiment time.

Overall, this part of the thesis contributes to the understanding of text entry in AR applications, advancing research with new perspectives and providing insights for practitioners who design future AR systems.

4 Typing the Future: Designing Multimodal AR Keyboards

This chapter comprises the version of record of the following paper: Schenkluhn, M., Peukert, C., & Weinhardt, C. (2022b). Typing the Future: Designing Multimodal AR Keyboards. *Wirtschaftsinformatik 2022 Proceedings*, *11*. https://aisel.aisnet.org/wi2022/hci/hci/11 It was awarded as the best short paper at the conference. Changes include formatting, numbering of the research question and chapters, minor changes for consistency, and correction of spelling errors.

4.1 Introduction

In future, AR devices will be ubiquitous in our everyday life assisting users in various use-cases (Bardeen et al., 2018; Company, 2020; Gurman et al., 2017). While the industry is waiting for lightweight, powerful, and unobtrusive AR glasses to emerge, several aspects of the next-generation devices already ask for new ideas and improvements today (Masood & Egger, 2020). One important aspect is text input (Dudley et al., 2019). With every new device category, researchers were exploring adequate ways for users to enter text into the computer (e.g., smartphones (Balakrishnan & CNBC, 2017), smartwatches (Espósito, 2020; Oney et al., 2013), Smart-TVs (Choi & Li, 2016), or smart speakers (Peng & Sarazen, 2017)) as previous methods often did not perform sufficiently. This pattern holds true for AR and poses a major UX design challenge for the already complex transition from traditional systems to this next generation platform (Dube & Arif, 2019). Text input will prevail because speech recognition is not suitable in many use cases, like in noisy environments or when entering confidential information (Masood & Egger, 2020; Pyae & Joelsson, 2018; Turk, 2014). Especially for expert users, it is highly important that a system facilitates a performant, learnable, portable, non-fatiguing, and unobtrusive way of text input (Dube & Arif, 2019). Yet, the main representative of state-of-the-art AR headsets, Microsoft's HoloLens 2, does not provide a fast, reliable, and user-friendly keyboard. The gesture-based mid-air keyboard lacks haptic feedback, touch-typing capabilities, and visually blocks most of the field of view. However, equipped with various sensors, AR devices open up a plentitude of input modalities that application designers may leverage when developing user-centered AR systems, e.g., gaze-, gesture-, or contextual input.

Previous work has made attempts to create text entry techniques with the goal of finding a tailored solution for the AR and VR context (Dube & Arif, 2019; Jiang & Weng, 2020; Jiang et al., 2018; Kim & Kim, 2017; Kuester et al., 2005; Prätorius et al., 2015; Rosenberg & Slater, 1999; Yi et al., 2015; Yu et al., 2018). However, new approaches often struggle to both perform well and be learned quickly (Dube & Arif, 2019). Typing speed, accuracy, and learning rate are common metrics for measuring the successful application of new text input techniques and are the foundation of user acceptance (Dube & Arif, 2019). Moreover, several approaches rely on external hardware, e.g., trackers, controllers, or keyboards, limiting the mobility which is essential for ubiquitous AR (Grubert et al., 2018; Jiang et al., 2018; Mourouzis et al., 2014). Thus, we argue that there is a need for the IS and HCI community to address and research this issue in order for user-centered AR to succeed. As prior approaches often struggled, the underlying design issues and requirements must be analyzed before suggesting novel modes of text entry. Accordingly, this research endeavor pursues the overall objective of investigating how AR keyboards need to be designed. In particular, we examine the following RQs:

- **RQ1.1**: *How to design a mobile virtual keyboard for AR systems to increase text entry performance?*
- **RQ1.2**: *How to design a mobile virtual keyboard for AR systems to increase learnability?*

To address the RQs, we commenced a DSR project to thoroughly examine the theoretical knowledge base and practical challenge, instantiate and evaluate a design artifact, and, eventually, produce design knowledge (Hevner et al., 2004). Our research is grounded in Transfer of Learning (ToL) theory and informed by prior HCI research on virtual keyboards. In this paper, we focus on the first three steps of the first design cycle to derive meta requirements and design principles from relevant issues and present a first version of the artifact featuring touch-typing and multimodal input.

4.2 Theoretical Background

Virtual Keyboards for Augmented Reality

Although consumer-ready AR headsets that are lightweight, small, affordable, and have long-lasting battery life are not yet available, many companies experiment with AR devices such as intermediate smartphone-based solutions or more capable headsets like the Microsoft HoloLens 2 to develop future use cases (Apple, 2024b; Bohn, 2019). In their review, Dube and Arif (2019) provide a comprehensive overview of text entry techniques in VR. Their suggested input categories and most of the accompanying issues, such as haptic feedback, new layout acceptance, low performance frustration, and physical demand, also apply to AR. The review separates physical from virtual techniques and the regular qwerty keyboard layout (according to the first row of characters on the English keyboard) from other approaches outlining that non-qwerty layouts tend to perform worse and require longer training periods (Dube & Arif, 2019). This issue is attributed to a network effect, as most users are familiar with qwerty layouts (Farrell & Klemperer, 2007). Overall, they conclude that, next to speed and accuracy, a well-designed keyboard needs to pay attention to haptic feedback, comfort, physical and cognitive demand, and potential frustration due to low performance (Dube & Arif, 2019).

Transfer of Learning

Depending on the prior knowledge, the teaching method, and the learning target, existent knowledge can have a positive or negative impact on learning (Perkins & Salomon, 1992). Hajian (2019) summarizes four theories in the field of the transfer of learning. There are several aspects that increase the likelihood of successful learning transfer from one context to another. For instance, transfer is more likely to be successful if the learning target and context are similar to the knowledge origin (Hajian, 2019; Perkins & Salomon, 1992). The theory of low and high road transfer describes two related mechanisms of how transfer can occur (Perkins & Salomon, 1992): Comparable to the two systems of thinking, low road transfer triggers intuitive responses of a well-known concept in a slightly different context (Kahneman, 2011; Perkins & Salomon, 1992). In contrast, high road transfer requires "mindful abstraction from the context of learning or application and a deliberate search for connections" (Perkins & Salomon, 1992, p. 8). Low and high road transfer can be exploited by the concepts of hugging and bridging (Perkins & Salomon, 1992). By applying hugging, the prior skill should be well-trained and tightly linked to the learning target. Bridging encourages the learner to actively abstract knowledge from the first context to apply it in the latter. Overall, these insights impact design decisions for the development, teaching, and evaluation of the artifact as a leading theory.

Multimodal Interaction

Multimodal interaction is natural to humans (Turk, 2014). When we give directions to a foreigner for example, we use spoken language and articulate by using our hands. Research distinguishes parallel and sequential multimodality, depending on the simultaneous or successive application of at least two modes of interaction (Turk, 2014). In general, multimodality has several advantages e.g., regarding user preference, flexibility, and reliability (Turk, 2014). Furthermore, multimodal interaction was already applied in AR to improve user experience (Chen et al., 2017; Kaiser et al., 2003; Nizam et al., 2018). However, the area of combining multiple non-voice interaction modes is rather unexplored to date (Nizam et al., 2018; Turk, 2014).

4.3 Method & First Activities in Design Science Research Cycle 1

To tackle the proposed RQs, we initialized a DSR project following the framework of Kuechler and Vaishnavi (2008). By means of creating a virtual keyboard artifact specifically for AR systems, we aim for knowledge gain to inform both research and practice. DSR offers the adequate research paradigm by providing structured, comprehensive, and iterative frameworks for the construction and observation of a previously non-existent artifact. Within this article, we will present the results of the first three activities of the first design cycle. Based on reviewing relevant literature from the HCI domain (particularly research on virtual keyboards for VR and AR systems), we identified issues (Ix) (**awareness of problem**). Next, taking the issues, virtual keyboard design knowledge, and ToL theories into account, we have derived Meta Requirements (MRxs) and proposed initial Design Principles (DPxs) as depicted in Figure 4.1 (**suggestion**). The MRs and DPs are then used to implement a situated software artifact (**development**) for evaluation (Gregor & Hevner, 2013). Overall, we plan to employ two full design cycles. In the following, we describe the already conducted activities in more detail.

4.3.1 Problem Awareness & Suggestion

Across literature on virtual keyboards, several issues were already pointed out that need to be taken into consideration (Dube & Arif, 2019). Depending on the keyboard design, directly mapping more than the 26 letters of the English alphabet to 10 fingers or few buttons on a controller is a challenge (**I1**). Hence, previous research with direct mappings was limited to digits (Prätorius et al., 2015), finger combinations ("chords") (Bowman et al., 2002), or overloading fingers with multiple characters (Kuester et al., 2005). Yet, solving this issue by capturing multiple touch regions on each finger might lead to complex and ambiguous gesture recognition (**I2**) (Prätorius et al., 2015). Therefore, we suggest a multimodal approach (**MR1**). Using two adequate input modalities results in enough combinations to capture all letters without the necessity to assign multiple touch regions per finger or choosing low performing chorded keyboards (Dube & Arif, 2019). More specifically, parallel multimodal interaction can increase entry performance as input combinations can occur simultaneously (**MR2**). Interacting via two simple modalities may also require less cognitive effort than one complex mode (Turk, 2014). This motivates the first suggested DP: Provide the Augmented Reality Keyboard (ARKB) with parallel multimodal input in order to quickly access the full alphabet while ensuring mobility. (**DP1**) Establishing non-qwerty keyboard layouts comes with further issues. Complex and new techniques can lead to a higher mental load (**I3**) while the training poses an increased entry barrier (**I4**) (Dube & Arif, 2019). This decreased learnability can be ascribed to the dissimilarity between traditional text input and the new technique which complicates ToL (Hajian, 2019). While there might be layouts that could be easier to learn and master for beginners, most users have prior typing experience with the qwerty layout and alternatives show low performance (**I5**) (Dube & Arif, 2019). Therefore, it is imperative to reuse and build on prior knowledge as much as possible. On the one hand, the goal is to exploit low road transfer (i.e., hugging) with a similar design and by addressing internalized intuitive knowledge (**MR3**). On the other hand, high road transfer is exploited (i.e., bridging) by actively pointing out the differences and how to foster them to abstract knowledge (**MR4**). New non-qwerty layouts could even imply effects of negative transfer (Perkins & Salomon, 1992). Consequently, we suggest the following DP: Provide the ARKB with interactions based on transferable prior knowledge to increase learnability. (**DP2**) Some entry techniques like gaze-based interaction have an inherent performance cap resulting from the required dwelling time that separates intended fixation from unintentional triggers during the search for characters (so-called Midas Touch effect) (**I6**) (Vrzakova & Bednarik, 2013). Having to wait for the system can lead to user frustration (Dube & Arif, 2019). Therefore, the system's recognition rate should be faster than users' entry speeds (**MR5**). Moreover, the event triggers for each character should be time independent (**MR6**), i.e., not requiring two subsequent actions or waiting times. Non-haptic techniques inherit the same issue of a typically lower input performance compared to haptic techniques (**I7**), thus, the system should provide haptic feedback (**MR7**) (Bowman et al., 2002; Dube & Arif, 2019; Dudley et al., 2019). Especially for independent AR glasses, stationary tracking devices hinder mobile usability (**I8**) (Mourouzis et al., 2014). The same issue arises for hardware input devices such as controllers (Yu et al., 2018), wrist-cameras (Prätorius et al., 2015), or gloves (Kuester et al., 2005) that block or limit the users' hands, need to be picked up, and stored (**I9**). Hence, the AR device should also be independent from external trackers or input devices (**MR8**). Finally, the device should be unobtrusive to keep the hands free when no text entry is performed (**MR9**). Thus, we suggest the third DP: Provide the ARKB with fast, haptic, independent, and unobtrusive mechanics to reduce obstacles for learnability and performance. (**DP3**) In conclusion, DP1 ensures the feasibility, DP2 the learnability, and DP3 the (final) performance of the approach.

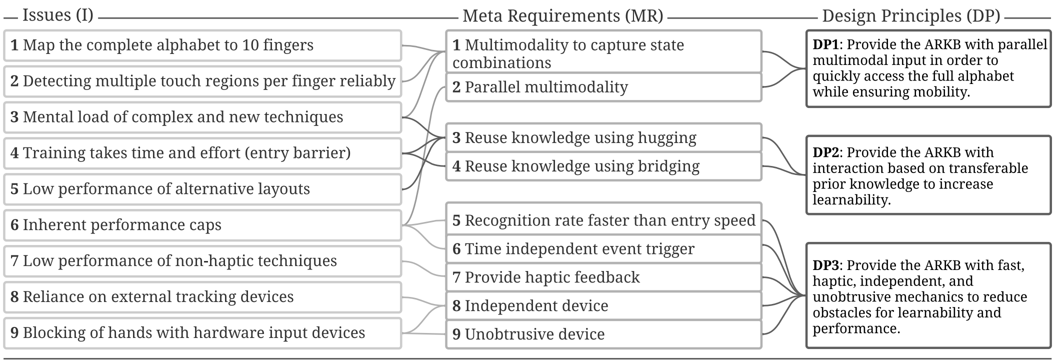

Figure 4.1.: Derived Issues, suggested MRs and DPs for the ARKB artifact development

4.3.2 Development

For the instantiation of the three DPs, we suggest a gaze- and gesture-based virtual ARKB artifact. The layout should be qwerty to be in line with DP2 and the similarity required by hugging. Moreover, each finger is responsible for the same character set like in regular touch typing. For instance, the left pinky is assigned to q, a, and z and the left middle finger to e, d, and c. The respective key is "pressed" by pinching thumb to finger. To account for the characters t or g, both index and middle fingers are pressed simultaneously. This movement is highly trained (Prätorius et al., 2015) and, thus, likely to transfer. In this case, the thumb provides a form of haptic sensation. Furthermore, the regular qwerty layout for the characters is divided into three layers (qwe, asd, yxc) (Kuester et al., 2005). The selection of the different layers is handled by gazing at one of three virtual areas projected by the AR device. The artifact is implemented in Unity for deployment on a Microsoft HoloLens 2[1]. The HoloLens has eye-tracking and hand-tracking capabilities without the need for an additional device to comply with DP3. Based on suggestions from Yi et al. (2015), we analyze the relative speed between thumb and each finger to detect a “key press”. Then, the area the user is currently gazing at is queried which selects the correct character.

4.4 Concluding Note & Future Research

In this research-in-progress, we contribute to the knowledge base by deriving MRs and initial DPs from prior research on virtual keyboards for AR and ToL theory to implement a multimodal AR keyboard artifact. Further, the current state of the artifact indicates that it is able to recognize both finger taps and gaze-selection solely based on the integrated sensors of a HoloLens 2 at a sufficient rate to provide fast text input. The final implementation of the artifact will then be evaluated in a lab experiment in which we will measure common features, e.g., typing speed over time (**evaluation**) (Dube & Arif, 2019). Based on the obtained findings, we will then be able to draw conclusions regarding the feasibility of the prototype system and applicability of the DPs including the implications of ToL for virtual AR keyboard designs (**conclusion**). In a subsequent design cycle, we want to instantiate the DPs in another artifact for generalization from an artefactual contribution towards a nascent design theory (Gregor & Hevner, 2013). Additionally, further investigations will be made by integrating predictive text and revision capabilities (Dube & Arif, 2019; Li et al., 2021). Hence, fast and enjoyable typing in AR might just be one gaze and tap away.

[1]A preview video of the artifact is available here: https://youtu.be/Aw93rxjk1iU

5 A Look Behind the Curtain: Exploring the Limits of Gaze Typing

This chapter comprises the version of record of the following paper: Schenkluhn, M., Peukert, C., & Weinhardt, C. (2022a). A Look Behind the Curtain: Exploring the Limits of Gaze Typing. In *NeuroIS Retreat 2022* (pp. 251–259). Springer International Publishing. Changes include formatting, numbering of the research question and chapters, minor changes for consistency, and correction of spelling errors. Reproduced with permission from Springer Nature.

5.1 Introduction

Eye tracking is becoming increasingly relevant in research[1] and business as sensors are becoming more powerful and cost-effective (Future, 2021). Primarily, eye tracking technology is used to study human attention allocation behavior to improve user experiences, marketing campaigns, or detect fatigue while driving (Duchowski, 2002). However, eye tracking can also be leveraged as active element of user interaction. Gaze typing allows users to type solely by using their eyes, i.e., by fixating the respective letters on a virtual keyboard for a certain time (Dube & Arif, 2019). This hands-free interaction mode is particularly useful for future AR and VR applications as traditional input devices are usually not available (Dube & Arif, 2019). With built-in eye tracking sensors, no additional controllers, keyboards, or gloves are required. Therefore, gaze typing enables mobility and concurrent task execution.

Various studies explore the parameters of gaze typing and its achievable speeds to improve usability and performance (Dube & Arif, 2019; Majaranta et al., 2004; Penkar et al., 2012). Moreover, longitudinal studies show the importance of the

[1]Number of yearly results on Scopus for “Eye Tracking” is steadily increasing since 2004 from 294 to 3,074 in 2021.

learning effect on typing speeds (Majaranta et al., 2009; Tuisku et al., 2008). However, the performance development while writing longer texts has not yet received much attention. Exploring this effect with users, who are unfamiliar with the system, is confounded as they need to simultaneously put effort into learning the system. Still, a hypothetical expert gaze typist will likely experience fatigue over extended periods of time at their peak typing speed thus limiting the long-term performance.

Hence, we will approach the issue from a different perspective. By excluding the learning effect and measuring time-dependent performance at varying dwell times during typing in a realistic AR context, we explore human limitations in gaze typing. The overall goal is to create gaze typing that is proactively adapting instead of retrospectively reacting to user fatigue. This would enable users to type short texts at their peak performance and economically utilizing cognitive resources for long texts. With the proposed study design in this article, we want to make a first step towards reaching this overarching research goal and seek to answer the following **RQ 2**: *What is the influence of dwell time and text length on gaze typing performance?*

In this research-in-progress paper, we propose a design for a laboratory study to answer the research question and demonstrate its implementation in an AR system.

5.2 Theoretical Background

Eye Tracking

Eye movement can be largely characterized by fixations and saccades (Pannasch et al., 2008). According to Pannasch et al. (2008), saccades are defined as "fast sequential movements, [that] are necessary to bring the fovea from one point to another" [p.1] and fixations as "periods in between of saccades, when the eyes are relatively stable" [p.1]. Saccades can be triggered by different events and are further categorized into visually guided and memory-guided saccades among others (Rommelse et al., 2008). Visually guided saccades are either reflexive to a sudden visual event or scanning unknown areas. Memory-guided saccades move the eyes to gaze towards a memorized location without external stimulus (Rommelse et al., 2008).

Today, eye tracking is usually performed by capturing the corneal reflection with video cameras (Duchowski, 2007). Eye trackers are used for interactive and diagnostic purposes in application domains such as neuroscience, aviation, automobile driving, print advertising, and user experience research (Duchowski, 2002). The interactive use of eye trackers, called gaze-based interaction, has also gained relevance in VR and AR applications when traditional means of input are not available or when user mobility would be limited (Dube & Arif, 2019). Gaze-based interaction can be the sole mode of input or used for target selection in combination with a controller or gesture to confirm a gaze-selected action.

Gaze Typing

New device factors often ask for new types of text entry, e.g., numpads on mobile phones, touch keyboards on smartphones, and voice recognition on smart speakers. There are several approaches to text entry in AR and VR ranging from physical keyboards to wrist- or gesture-based, and novel 3D techniques (Dube & Arif, 2019). Each technique has benefits and drawbacks in different contexts. New layouts require more training than layouts a user is already familiar with (Schenkluhn et al., 2022b). In general, the goal of new layouts is to maximize the text entry rate and minimizing the error rate at the same time (Dube & Arif, 2019).

Gaze typing relies solely on eye tracking to enter text. A common form is the display of a virtual keyboard on a screen. By gazing at a key for more than a given threshold, the eyes "press" the key and type the respective character (Dube & Arif, 2019). The main advantage of gaze typing is its hands-free character, which has been leveraged in several accessibility studies (Hansen et al., 2004; Zhang et al., 2017). Additionally, the eyes can move quickly and accurately. Studies therefore indicate a potential of high text entry speeds and low error rates (Kristensson & Vertanen, 2012; Kurauchi et al., 2016; Majaranta et al., 2009).

The time between the beginning of the initial fixation and the activation of the key press is called dwell time and is an important variable when designing gaze typing interaction. The dwell time is necessary to differentiate meaningful from unmeaningful gaze events, i.e., intended typing from a character search. As soon as the first fixation is registered within the Area of Interest (AoI) of one key, the system measures the visit time on this particular key. If one of the successive fixations targets an area outside the AoI before the dwell time threshold has been reached, the keypress is aborted. Otherwise, if all successive fixations stay within the AoI longer than the dwell time, the key is pressed once.

Short dwell times can cause the Midas touch effect, that is, the unintended selection of every key the user is passing over (Jacob, 1991; Penkar et al., 2012). Thus, longer dwell times reduce the number of errors. However, overly long dwell times lead to adverse effects as the user is not able to hold the gaze for long periods ("gaze-hold problem") (Penkar et al., 2012). Additionally, long dwell times slow down text entry speeds as the user must stay fixated to one character until the dwell time has passed. Longitudinal studies show that trained users are able to deal with shorter dwell times around 200 msec while still maintaining low error rates (Majaranta et al., 2009; Tuisku et al., 2008). Hence, this tradeoff between entry speed, error rate, and training determines the success of the application of eye gazing in VR and AR. Feasible dwell time depends on different factors. Following the human performance model of Kristensson and Vertanen (2012), the gaze typing interaction consists of overhead time and dwell time. The overhead time includes saccades to transition between keys and error correction. While the dwell time is internal to the application, the overhead time depends on the user.

5.3 Design Considerations for an Experimental Design to Study Time-Dependent Gaze Typing Performance

To explore the limitations of gaze typing of future experienced typists, we plan to conduct a laboratory experiment. In general, the typing performance is highly influenced by the dwell time (Kristensson & Vertanen, 2012; Rajanna & Hansen, 2018). Accordingly, it should be set as low as possible to support fast text entry. In contrast, the ability of users to concentrate is limited. Therefore, we presume that proficient gaze typists experience fatigue while typing a text section and cannot maintain the speed associated with a low dwell time.

Task

MacKenzie (2010) discusses different tasks for text entry evaluation. While a text creation task is closer to typical usage, it has several issues for performance measurement. Text creation includes aspects unrelated to the keyboard interaction such as thinking about content, phrasing, and grammar (MacKenzie, 2010). Additionally, error detection is more complex as the user intention cannot be inferred during text entry (MacKenzie, 2010). Finally, text creation complicates comparability between subjects because of differences in vocabulary and, therefore, the distribution of letters and words (MacKenzie, 2010).

For these reasons, scientists typically rely on text copying tasks that try to mimic text creation (Kuester et al., 2005; MacKenzie, 2010; Rajanna & Hansen, 2018; Yu et al., 2018). A typical task is as follows: The study participant is presented with one sentence for example from the phrase set of MacKenzie and Soukoreff (2003). After memorizing the sentence, the participant is asked to write the text as quickly as possible with the given keyboard (MacKenzie, 2010).

Gaze typing experiments usually study the performance of prototypes with respect to typing speed and accuracy (Dube & Arif, 2019). Due to typically novel keyboard designs and the unfamiliarity of participants with them, the performance of a potential expert typists cannot be trialed in one session. Longitudinal studies show that the average typing speed increases over multiple sessions, although participants cannot be considered experts afterwards (Majaranta et al., 2009; Tuisku et al., 2008). Novice typists have to search for characters on the keyboard during typing and are not accustomed to the dwell time. During this process, the fast eye movements can be considered as scanning saccades in combination with longer fixations required for pattern identification. In contrast, potential expert typists can use memory-guided saccades to quickly and accurately jump between keys after sufficient training. Thus, this study eliminates the training effect regarding interaction type, keyboard layout, and typing task, by approximating memory-guided saccades with reflexive saccades. Instead of displaying a phrase that the participant must type, the keyboard visually highlights the next key that shall be "pressed." This highlighting cues a reflexive saccade. In essence, participants follow the highlighted keys to gaze type sentences from the phrase sets of (MacKenzie & Soukoreff, 2003) at varying dwell times without relying on their gaze typing skill.

Text copying tasks introduce additional subtasks such as comparing the specified text with the typed text, which reduces text entry performance (MacKenzie, 2010). However, this study design does not require participants to memorize and compare texts when following the highlighted keys. Thus, we expect a higher external validity regarding real text creation.

Evaluation Procedure

The experiment begins with a calibration of the eye tracker. The participant is introduced to the task and performs a trial round. The task is to follow and focus the highlighted character as quickly and accurately as possible. After the dwell time threshold has been reached the next character of the sentence is highlighted on the keyboard. Afterwards, the task is performed at decreasing dwell time levels. The levels are 600 msec, 450 msec, 300 msec, 150 msec, and 0 msec (Rajanna & Hansen, 2018). This could be adjusted after our pretest. There are two consecutive trials at each dwell time level to account for inter-treatment fatigue. As we want to measure the time-dependent effects on intra-treatment performance, there is a relaxation phase of 3 minutes between each treatment to reduce effects on inter-treatment performance (Kurauchi et al., 2016).

Each treatment ends regularly after 5 minutes. Participants can abort the treatment if they are not able to maintain the increasingly fast speeds. If the error rate exceeds 15 % the participant is unlikely to keep up and the treatment ends as well. This threshold requires finetuning during the pretest, too.

Measurements

The performance is measured as a function of time to derive a relation between dwell time, text length, and performance. Measures include the overhead time between dwells, the number of lost focuses for one keypress, the minimal distance between the gaze-intersection with the keyboard and the key border, and the error rate. After each round, the participants answer the NASA-TLX questionnaire to supplement self-reported task loads (Hart & Staveland, 1988).

Participants and Compensation

Participants will be recruited from the participant panel of a large European university. For the 100-minute experiment, we plan to compensate each participant with 20 €. Further, the three participants with the highest performance will be rewarded with an additional 5 € to motivate concentration. We decided against an entirely performance-based compensation to avoid pushing participants over their limits in an unrealistic manner.

Application

We chose to implement the experiment in an AR context. Most research in the domain of virtual keyboards is currently conducted using Virtual Reality technology (Dube & Arif, 2019). However, AR systems are more dependent on mobile means of text entry as VR systems are limited in their mobility in order not to interfere or collide with real objects. In a professional VR application, a hardware keyboard can most likely be used with a mapped virtual representation within the VR environment. The application is implemented in Unity 2020.3 for an off-the-shelf Microsoft HoloLens 2 AR HMD using the Mixed Reality Toolkit 2.7.3. The HoloLens 2 is a state-of-the-art standalone AR device with built-in eye tracking capabilities.

The eye tracker updates at 30 Hz, is specified with a spatial accuracy of 1.5°, and has been evaluated previously for accuracy and precision (Aziz & Komogortsev, 2022). Although it is not as capable as external sensors, it is sufficiently accurate for this use case. The API only grants access to the combined fixation point. The video stream of the IR cameras, independent data per eye, and eye blinking data is not available. While this is a drawback, future eye trackers in consumer devices will likely comprise of the same limitations for privacy reasons. We implemented the application based on Microsoft's usability recommendations[2] for colors, contrast, object positioning and size, and audio. Figure 5.1 depicts the gaze typing keyboard in a simulated environment. A video[3] of the application shows different dwell time levels combined with an overhead time that simulates the human reaction time. The measurement display and the overhead time are included for demonstration purposes and will be deactivated for the experiment. Furthermore, Figure 5.2 lists all relevant measurements of the keyboard and its components in an isometric projection. The keyboard fits within the field of view of the HoloLens 2 without the need for users to turn their heads.

Figure 5.1.: Augmented Reality application with highlighted character

[2]https://docs.microsoft.com/en-us/windows/mixed-reality/design/ (Accessed: 12.04.2024)
[3]https://www.youtube.com/watch?v=GQayxnlKVqU

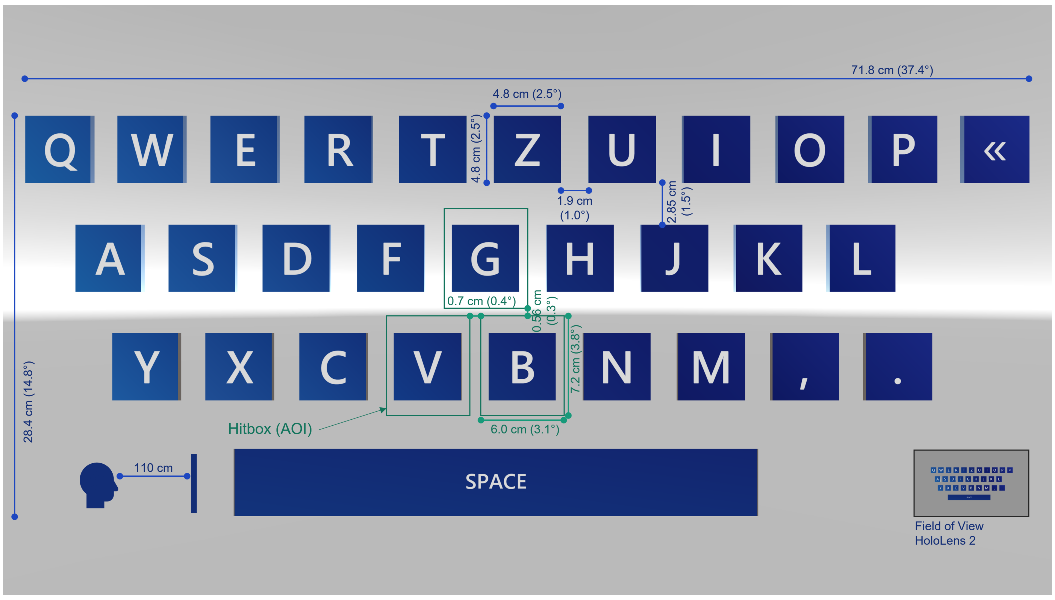

Figure 5.2.: Dimensions of the gaze typing keyboard

5.4 Discussion

This paper proposes a novel experiment design to explore human limits to inform future gaze typing implementations. Depending on the dwell time level, we expect that participants will start to struggle concentrating on the task after some time. The performance measures will likely represent this effect. For lower dwell time levels, the effect is expected to appear earlier during the task.

By conducting two consecutive trials at each dwell time level, we expect an approximated step function. The second round on the same level might show a slight decrease in performance. If this effect is too prominent in the pretest, the relaxation phase will be extended. As the overhead time can vary between participants, we do not expect that there will be one cutoff point where all participants experience concentration loss. However, understanding these differences between individuals will be the prerequisite for designing proactive user-adaptive gaze typing. A proactive system would be able to increase dwell time to decrease task load before typists experience fatigue.

There are some limitations to this experiment design. The limited spatial and temporal resolution of the eye tracker was already mentioned. Moreover, the overhead time in this experiment does not contain the tasks of character processing and finding on the keyboard layout and error correction. Thus, this factor must be added when comparing the results with other gaze typing studies even if expert typists are able to minimize it. Furthermore, there is a possible secondary training effect on concentration to type longer texts that this study cannot eliminate.

5.5 Concluding Notes and Future Research

The results of this study will enable the development of proactive user-adaptive eye gazing systems by complementing previous studies with a different perspective. Additionally, future systems do not have to rely on the exact fixation of singular characters (Kristensson & Vertanen, 2012; Kurauchi et al., 2016). Intelligent dwell-free gaze typing similar to swipe keyboards on smartphones could even improve gaze typing performance (Kurauchi et al., 2016). The particularization of the human performance model (Kristensson & Vertanen, 2012) similar to the keystroke-level model (Card et al., 1980) with a focus on the cognitive processes could also help to unveil cognitive limitations. By better understanding human limits, the usability and comfort of these systems can be improved with the results of this study leveraging gaze typing as attractive and competitive means of text entry in AR and VR environments.

6 Xperisight: Parallelizing Extended Reality Studies Without Losing Control

6.1 Introduction

XR laboratory experiments provide valuable insights into human-computer interactions with novel AR, VR, and MR technology. However, conducting such experiments efficiently and effectively poses significant challenges for researchers. Even though many labs have the space and equipment to run more than one experiment session in parallel, it is not feasible for one experimenter to supervise multiple sessions without losing control. Software bugs, user discomfort, and further inquiries regarding the experiment or device need the immediate attention of the experimenter. Hence, they often require them to stay with the participant throughout the experiment. At the same time, the effect of experimenters on participants should be as little as possible to avoid, e.g., social-desirability effects (Mullen et al., 1991; Rosenthal, 1976; Williamson & Williamson, 2017) and providing varying instructions and support between participants unintentionally. Additionally, staying with the participant during experiments is often time-consuming and inefficient with regard to the many other tasks researchers face.

Thus, researchers either need to rely on smaller sample sizes or invest substantial amounts of time and resources to conduct large experiments (Jicol et al., 2023; Peukert et al., 2019). Addressing these challenges is crucial for advancing the field of XR laboratory experiments within information systems and human-computer interaction research and ensuring the reliability and generalizability of research findings.

This chapter presents Xperisight, a novel tool for Unity-based XR applications designed to address some of these challenges by facilitating parallel sessions in XR laboratory experiments. Xperisight leverages Unity APIs to present application and device information, offers a help button for participants, and allows for scene management and eye-tracking calibration control remotely in and from a dashboard with

Figure 6.1.: The Xperisight Dashboard with information from two XR devices.

a zero-code setup. The tool was successfully tested in five different Unity XR projects for the Microsoft HoloLens 2, the Varjo XR-3, the HTC Vive Pro Eye, and Windows standalone and successfully applied the tool in an experiment with 29 participants that is presented in Chapter 7.

6.2 Related Work

Due to the high effort of XR studies, XR researchers in the human-computer interaction domain often choose to use only small samples of participants. Caine (2016) analyzed sample sizes across the studies published at CHI 2014. They find that the most common sample size is 12 while 70 % of the studies have less than 30 participants and discuss the implications of small N and underpowered quantitative studies (Caine, 2016). On top of the fact that small N studies can get successfully published, conducting large experiments takes a lot of effort. For example, a study by Peukert et al. (2019) required five weeks to collect 132 samples as each VR session had to run in succession. Another study reports that each of their successive 360 conducted VR sessions took between 20 to 25 minutes (Jicol et al., 2023). Still, XR research must insist on conducting rigorous studies with appropriate sample sizes for the scientific integrity and validity of the domain (Caine, 2016).

In general, there are many tools to facilitate the effort required for conducting experiments. For example, oTree (Chen et al., 2016) is an open-source software

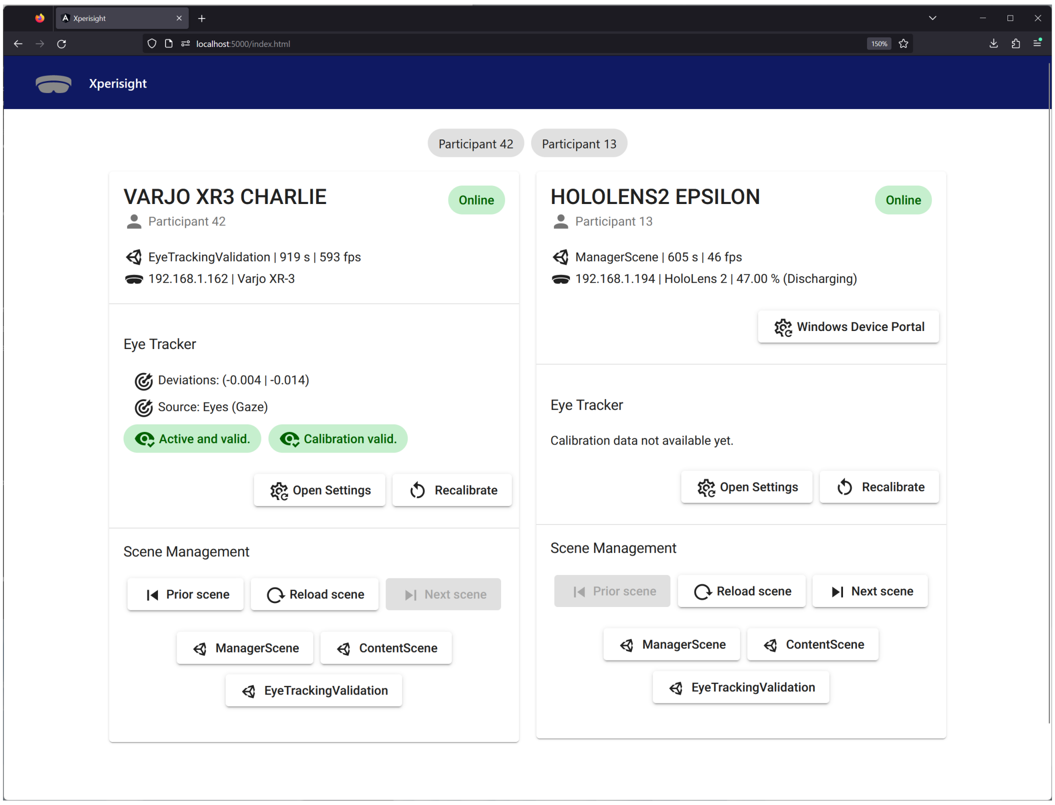

Figure 6.2.: Xperisight dashboard

framework for conducting lab, online, and field experiments. It allows researchers to easily design and implement large-scale experiments in a web-based environment. With oTree, the experimenter has full control over the experiment session, including the ability to monitor participant progress, intervene when necessary, and adjust experimental parameters in real-time (Chen et al., 2016).

Unfortunately, the toolset of oTree and similar tools is not compatible with XR applications without writing custom interfaces which would add to the high effort of creating XR applications for experiments. However, several tools exist that provide individual features required by researchers specifically for XR lab experiments. The ExpTrialMng (Kim et al., 2022) supports randomized trial orders and logging of experiment data. Ubiq-exp (Steed et al., 2022) extends Ubiq (Friston et al., 2021), a Unity networking library, with functionality specifically for remote or distributed experiments. The authors differentiate supervised and unsupervised and single-participant and multiple-participant sessions. Moreover, they describe requirements for a support tool and implement features such as distributed logging, questionnaires, and multiplayer functionality including avatars. At the same time, Steed et al. (2022) report often running experiments directly from the Unity Editor in order to remain control over the application. These tools focus on Unity[1], a game engine used by many researchers for XR applications as it supports and is supported by most XR device manufacturers such as Meta Quest, Microsoft, HTC, or Varjo (Radiah et al., 2021).

Overall, the availability of powerful tools for non-XR experiments does not translate into the XR space. There are tools that facilitate and there is research that investigates the *creation process* of XR experiments including questionnaires in XR (Bebko & Troje, 2020; Schwind et al., 2019), multiparticipant (Radu et al., 2021) and remote (Radiah et al., 2021) features.

However, tools to *oversee* XR experiments remain limited. While there are benefits in creating encapsulated, self-contained XR experiment applications that do not require the presence or even oversight of an experimenter, many researchers still argue in favor of XR lab experiments due to the feasibility, data collection and integrity, and control among others at least in some cases (Ratcliffe et al., 2021).

As I did not find a tool that could enable real-time oversight and control functionality throughout the literature and software review when it was required for the XR lab experiments described in Chapters 7 and 11 I decided to implement such a tool and contribute it to the XR community.

6.3 System Architecture and Design

The primary objective of the tool is to provide essential features to oversee and control existing Unity applications for XR experiments with as little setup effort as possible. Based on existing non-XR tools and prior XR experiment experience, the following key requirements were derived. Experimenters shall be able to leave the room and still see the experiment progress, application, and device health, and can be called if help is required. Additionally, if errors occur, experimenters shall be able to restart Unity scenes, i.e., specific sections of the application, and, if in use, recalibrate eye tracking cameras without the need to access a desktop running the application or even to put on a standalone XR device themselves. Access to these functions should work both from a potential control room desktop and on mobile devices in order to have mobile access to the dashboard, e.g., during the first setup of the XR device.

[1]https://unity.com/ (Accessed: 12.04.2024)

The Xperisight implementation consists of three components: A Unity library, a web dashboard using Angular, and a Python Flask server to connect them via HTTP requests. The Unity library provides a blueprint (Unity Prefab) that contains the functionality to collect all required information and send it to the server. Additionally, it queries the server for instructions such as scene changes, and performs them. The decision against the use of WebSockets was made to avoid additional dependencies and decrease the implementation complexity on the Unity side. This compromise is reasonable since the number of exchanged packets is relatively small. Similar to other tools mentioned in section 6.2 I decided to focus on Unity-based XR applications. However, the server application and dashboard can be used with any game engine via the HTTP API and a custom handler in the XR application.

The web dashboard retrieves the relevant information from the server for each client and displays them side-by-side as depicted in figure 6.1 and 6.2. Additionally, commands can be selected for each client in the dashboard and sent to the server which stores it in an instruction registry. The dashboard is optimized both for mobile and desktop screen sizes and adjusts content responsively. Moreover, it provides visual feedback when sending instructions to a device or if the application is not reachable anymore, e.g., due to a crash or connection issues. A high-level system architecture is depicted in figure 6.3.

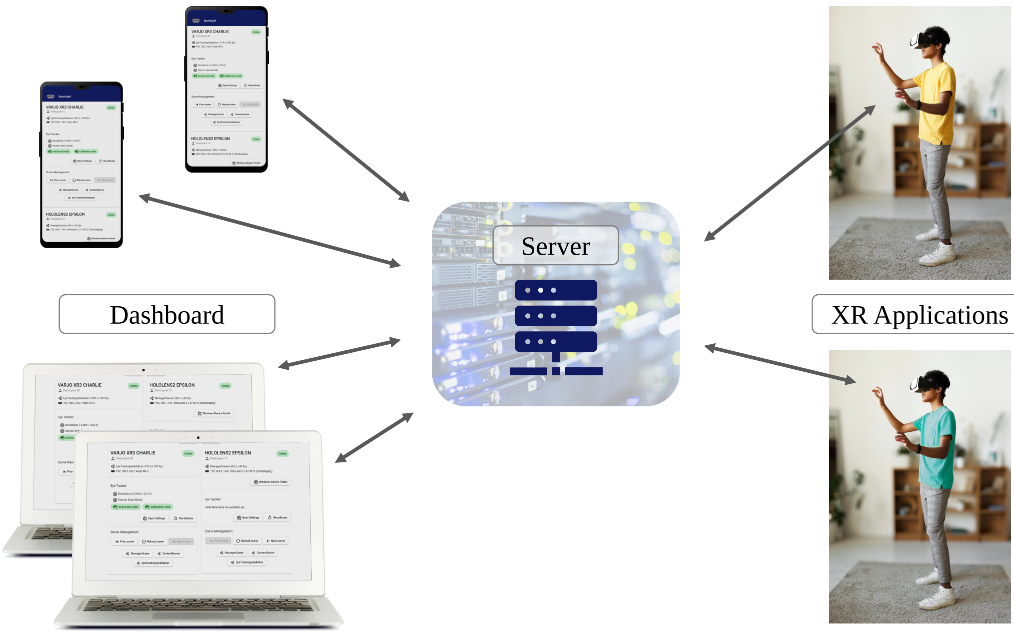

Figure 6.3.: Xperisight system architecture

Thanks to the usage of an intermediary server, a n-to-n relation between dashboards and Unity applications is possible with a synchronized state. Thus, one or multiple experimenters can open the dashboard on a computer in the lab's control room and a tablet or smartphone at the same time.

6.4 Usage and Implementation Details

Like different levels in a computer game, Unity scenes enable developers to segment their applications into distinct sections. In XR experiments, there could be for instance an introduction and a tutorial scene in addition to one scene for each treatment. After importing the package into their Unity application, experimenters can simply drag and drop the Xperisight prefab in their start scene and configure it in Unity's visual inspector. Once loaded, the prefab remains active even across scenes. The tool supports single and additive scene loading to support the use of a persistent manager scene. To communicate with the server, the IP address can be either set in Unity at build time if the server has a static address, by implementing a configuration UI that tells the Xperisight API which IP to use, or by writing the IP into a configuration file from which the application reads it. The author suggests that providing the IP via the configuration file is the easiest approach to avoid a custom UI that experimenters need to implement for the given XR device if the IP is not known at build time. Once set, the IP is stored on the device and persistent across sessions.

The dashboard is implemented as a Single Page Application (SPA) using Angular 13[2] and Google's Material Design[3] language. In the dashboard, each device is displayed as a card following the same layout. Experimenters can view information such as the device name, current Frames Per Second (FPS), device battery level, current Unity scene, and duration of the stay in the current scene. If the XR application generates a participant ID, it can be exposed to the Xperisight Unity script and will be automatically displayed in the dashboard. Otherwise, Xperisight generates a unique ID to identify each session in addition to the device name. Inspired by the flight attendant call button in airplanes, a feature for participants to call for help if questions or errors occur was included. This button can for example be included in every scene at a fixed position or in a hand menu for easy access. By toggling the button in the XR application, the respective dashboard card will flash red to raise the awareness of the experimenter as depicted in figure 6.4. When a user quits the XR application, it is displayed as offline and the card can be removed if desired. These

[2]https://angular.io/ (Accessed: 12.04.2024)
[3]https://m3.material.io/ (Accessed: 12.04.2024)

Figure 6.4.: Pressed help button (left) and Xperisight Dashboard card (right) reflecting the state of the help button.

interactions are synchronized between all devices that display the dashboard via the server application to maintain a common application state.

The Unity prefab queries all available scenes and informs the server about them. In the dashboard, the experimenter can load or reload specific scenes remotely. Depending on the number of available scenes, the scene selection is dynamically displayed as buttons (up to five scenes) or as a dropdown list (more than five scenes).

6.5 Eye Tracking Validation

Many state-of-the-art XR headsets have built-in eye trackers. Eye trackers are interesting for researchers both to passively observe the focus of participants during an experiment and enable an active mode to interact with objects in 3D space. The accuracy of the eye trackers and their calibration to each participant can influence the user experience and the collected data (Duchowski, 2017). Thus, Xperisight includes the validation of the eye tracker calibration by displaying a target for participants to focus on and reporting a possible offset in the dashboard as depicted in figure 6.5. This task only takes a couple of seconds and can ensure that the calibration is (still) valid and that there is no drift, especially for longer experiment sessions. If the eye tracker hardware reports a calibration status, this information is also displayed in the dashboard. In case of issues with the eye tracking calibration, the calibration process can be restarted from the dashboard as well. The calibration validation is available as a prefab and a Unity scene.

For studies that do not use eye tracking, the respective section in the dashboard is automatically minimized to save screen real estate.

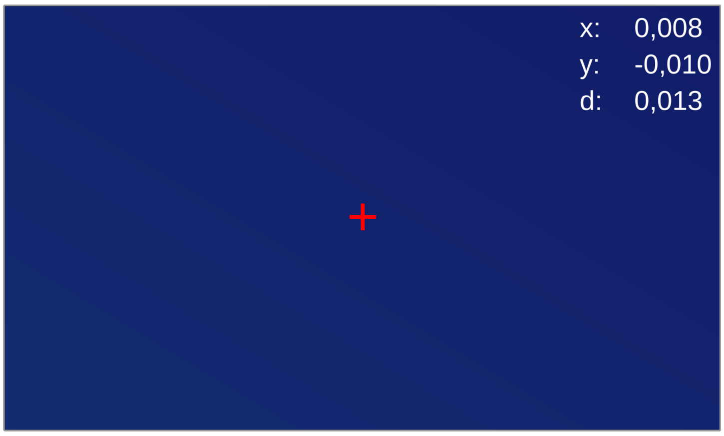

Figure 6.5.: Unity prefab for the validation of the eye tracking calibration with optional information for debugging in the top right-hand corner.

6.6 Mixed Reality Toolkit Addon

To minimize the required package dependencies in Unity, the Unity package was split into two separate subpackages. The first subpackage offers the core functionality and can be used with any Unity application using Unity 2019.3 and later.

However, there are different libraries for eye-tracking systems with distinct APIs. Therefore, a second, optional subpackage is provided as a reference implementation for Microsoft's MRTK that primarily targets the Microsoft HoloLens 2 but can be used with other XR hardware using OpenXR as well. It includes the eye tracking validation and a help button reference implementation using the MRTK.

6.7 Requirements

Xperisight has a few basic software requirements to run. The first Unity subpackage only requires the JavaScript Object Notation (JSON) library Newtonsoft JSON to (de)serialize objects when communicating with the server. Thus, no interference with other packages or Unity core functionality can occur in contrast to more comprehensive frameworks that attach to the main camera or interaction modalities. The second, optional Unity subpackage requires MRTK foundation 2.8.3. The server application requires Python to be installed and a separate Python environment for the dependencies is recommended. The web dashboard runs on any modern browser.

Xperisight is available as free open-source project on GitLab[4] including the Unity packages, the web dashboard, and the server application with a static build of the dashboard. The instructions to include Xperisight in XR projects are detailed in the repository as well.

6.8 Future Work

To assess the utility of Xperisight, a further evaluation of its functionality and reliability, as well as the user experience of the developers in including Xperisight within their Unity project and the user experience of the dashboard during the experiment would be useful.

The tool was already applied and tested internally in different scenarios. Based on these preliminary applications of Xperisight there are already first insights available. However, these intermediate results may be biased in favor of Xperisight as no independent entity was involved in testing, yet.

In a study on AR text entry, as described in Chapter 7, one experimenter supervised a total of 29 sessions with Xperisight, mostly running two sessions in parallel (Schenkluhn et al., 2023a). Each session took 75 minutes on average. After a brief setup and calibration of the Microsoft HoloLens 2, the experimenter left the room. Out of the 29 sessions, the help button was pressed four times to clarify questions and resolve the aforementioned issue and, hence, demonstrated its use even in a well-tested application with several iterations of pretests. Apart from these irregularities, the experimenter was able to work on other tasks while keeping the dashboard of Xperisight in view. As a result, the use of Xperisight not only nearly halved the total time required for the experiments, but also freed up much capacity otherwise blocked by supervision.

Additionally, Xperisight was successfully tested for its stability and easy integration into four other applications created by different researchers for different platforms. The devices encompass the Varjo XR-3, the HTC Vive Pro Eye, the Microsoft HoloLens

[4] https://gitlab.com/mschenkluhn-kit/xperisight

2 as in the experiment described above, and Windows standalone. The eye tracker calibration of the Varjo XR-3 could be validated without further setup as the application used the MRTK.

Overall, using multiple Unity scenes, especially for longer experiment sessions, enables granular movement between and oversight over sections of the experiment. As the dashboard displays the current scene of each device, it is easier to quickly grasp the current stage of the experiment and participant progress. Additionally, loading or reloading specific scenes in case of an error or when debugging the application is more granular. In the study mentioned above, one subject unintentionally clicked "Continue" without having fully completed the task as they intended. Thanks to the help button and a granular scene splitting, the experimenter was able to move the subject back to the last step of the previous task on their request without losing any data and without significant loss in time, effort, or validity through large interventions.

As it is advisable to leave the room during the experiment to reduce experimenter bias, the instructions must be clearly communicated. For this multi-stage experiment, audio instructions were recorded and played in combination with displaying the outline as text to ensure internal validity, i.e., a consistent experiment experience between participants and reduced potential eye fatigue compared to providing text instructions only. This approach encapsulates the experiment in the direction of fully unsupervised experiments while still maintaining the benefits of lab experiments (Ratcliffe et al., 2021; Steed et al., 2023).

Compared to the comprehensive feature set of, e.g., oTree, the Xperisight tool only provides features similar to the *Monitor* section in oTree. Unity applications can be more complex than oTree experiments, yet, it would be useful to have access to live experiment result data in the dashboard. While Xperisight does not provide an API to share data from the XR application with the dashboard, yet, the tool is extensible and the feature could easily be incorporated. However, this would require experimenters to programmatically expose this information to Xperisight in Unity scripts themselves.

The feature set of Xperisight does not provide the required tools to support experimenters with common issues of fully unsupervised remote experiments such as environmental factors, setup or hardware issues, or ambiguities in the operation of the device or application. Additionally, the communication between the participant and the experimenter apart from the help button needs to happen with different means and will potentially break the experiment flow and invalidate the session data.

In the future, logging of the collected information during the experiment could be added to trace potential issues during the data analysis. Furthermore, a live view and capture of the XR device could be added if appropriate for the respective experiment in terms of anonymity, data privacy, and ethics considerations.

6.9 Conclusion

In this paper, Xperisight, a novel tool for Unity-based XR applications that addresses the challenges of running efficient and replicable XR lab experiments was introduced. The limitations researchers face when conducting XR experiments and the need for monitoring tools in the XR domain were discussed. Xperisight is designed to provide essential functions for real-time monitoring and control of XR experiments, allowing experimenters to remotely monitor the state of devices and applications, manage scenes, and recalibrate eye-tracking cameras and, thus, avoid experimenter bias without losing control.

7

Does One Keyboard Fit All? Comparison and Evaluation of Device-Free Augmented Reality Keyboard Designs

This chapter comprises the version of record of the following paper: Schenkluhn, M., Peukert, C., Greif-Winzrieth, A., & Weinhardt, C. (2023a). Does One Keyboard Fit All? Comparison and Evaluation of Device-Free Augmented Reality Keyboard Designs. *Proceedings of the 29th ACM Symposium on Virtual Reality Software and Technology*, 1–11.

Changes include formatting, numbering of the research question and chapters, minor changes for consistency, and correction of spelling errors.

7.1 Introduction

Similar to the spread of smartphones, AR might become ubiquitous in the next couple of years when the technology matures. In addition to challenges in hardware and software development, foundational interactions such as text input must be properly designed to meet or even exceed heterogeneous expectations and requirements to satisfy a broad user base. None of the various proposed keyboard designs has yet been able to establish a standard for typing in AR or VR (Dube & Arif, 2019). Up to now, text entry research oftentimes especially focuses on entry rates and error rates (Arif & Stuerzlinger, 2009; Dube & Arif, 2019; Dudley et al., 2019). However, there might be more to the user experience of text input than performance, and the priorities may vary between different users and in different settings (Dube & Arif, 2019; Xu et al., 2019a).

Whereas physical keyboards with QWERTY layouts are the norm for text entry on laptops and desktop computers, touch-based QWERTY keyboards are primarily used on smartphones. Besides those, speech input can speed up text entry rates (Smith & Chaparro, 2015) but can be unavailable or inappropriate due to confidentiality, noisy settings, or in the public space. While a physical keyboard could be used for AR in a stationary setting, it does not prove useful in a mobile context, especially for short text messages or queries (Xu et al., 2019a). It is likely that there will be a standard input mode for AR typing in the future, however, the benefits and drawbacks of different keyboard designs for distinct user groups and in different settings still need to be evaluated. Despite the increasing importance of text input in AR and VR, the user experience of AR keyboard designs beyond performance measures has not been systematically investigated. Specifically, there is a lack of research that explores the trade-offs between different AR input modes and how they impact user experience and preferences, especially in mobile settings.

In our article, we present the results of a comparative laboratory study between three promising input modes for device-free AR typing, i.e., a dwell-based eye-gaze keyboard, eye-gaze with pinch-gesture-commit keyboard, and mid-air hand tap keyboard employing a virtual QWERTY layout. In this context, we contribute to a better foundation to balance and prioritize important characteristics such as text entry performance, comfort, control, and independence from user skill for future AR keyboard designs. We update established work through state-of-the-art technology and, thus, provide novel findings on user preferences.

In line with Dube and Arif (2019), this study thrives for reproducibility, comparability, and transparency in both keyboard implementation and study design to produce robust and generalizable results. Hence, we set out to explore the performance and user preference between the three keyboard designs as our research objectives with the following **RQ 3**: *Which text entry method is most suitable for mobile AR devices in terms of performance and user preference?*

7.2 Related Work

7.2.1 Text Entry Studies

Text entry design might appear trivial since users naturally rely on keyboards on a daily basis. However, new device form factors or special situations render the physical keyboard impractical and ask for innovative solutions. New approaches are typically evaluated by measuring text entry rate and error rate (Arif & Stuerzlinger, 2009; Wobbrock, 2010). The entry rate is measured as the time required to enter a phrase in relation to the number of characters of that sentence. It is expressed as Characters Per Second (CPS) or Words Per Minute (WPM) whereby a word is defined as consisting of five characters (Wobbrock, 2010). Error rates can be measured in several ways. We report errors as Corrected Error Rate (CER), Uncorrected Error Rate (UER), and Total Error Rate (TER) (Soukoreff & MacKenzie, 2003). This consideration differentiates errors that were corrected by the user (CER) from errors that were not corrected by the user (UER) and their sum (TER). All calculations are based on the formulas of Wobbrock (2010). The usual typing task for measuring text entry performance is *copy typing* (MacKenzie, 2010), i.e., a given phrase has to be identically copied as fast and accurate as possible by study participants. Although copy typing is rare for users in practice, the task has proven itself useful for evaluating keyboard performance isolated from text ideation and thinking processes (MacKenzie, 2010). Many studies rely on the phrase set of MacKenzie and Soukoreff (2003) for the reference texts that users need to copy type (Fashimpaur et al., 2020; Kimura et al., 2022; Kurauchi et al., 2016; Streli et al., 2022; Zhang et al., 2022).

Different entry task designs and metric calculations can lead to different results. Therefore, Dube and Arif (2019) recommend that all considerations should be communicated for transparency and comparability.

7.2.2 Text Entry in Augmented Reality

Text entry research has already investigated many aspects of AR and VR typing. Dube and Arif (2019) list various categories of text input modes for VR that in part translate to AR. Next to physical QWERTY keyboards that need a virtual representation to be visible in VR (Grubert et al., 2018; Jiang et al., 2018; Walker et al., 2016; Walker et al., 2017), researchers used game controllers (Yu et al., 2018), VR controllers (Jiang & Weng, 2020; Xu et al., 2019c), gloves (Mehring et al., 2004), or hand tracking (Yi et al., 2015) to create keyboards for VR. Unlike VR, AR does not have issues with proprioception – the localization of the hands – when typing on a physical keyboard due to occlusion. Moreover, AR can be fully mobile and less obtrusive as users are able to see their surroundings. Therefore, mobile text entry is an important factor for the future success of AR devices and their usability. One common hands-free approach is dwell-based eye-gaze typing (Lu et al., 2021; Majaranta et al., 2009; Rajanna & Hansen, 2018; Xu et al., 2019a). For this input mode, the AR device displays a virtual keyboard in front of the user. By dwelling, i.e., looking at a key for a given time period (dwell time), users can "press" a key solely by using their eyes.

Several studies propose and study specific prototypes as solutions for mobile AR text input with (Streli et al., 2022; Zhang et al., 2022) and without additional hardware (device-free) (Dudley et al., 2019; Dudley et al., 2018; Fashimpaur et al., 2020). Touch-typing keyboards assign characters to each finger to enable fast text entry for trained users (Yeo et al., 2017; Yi et al., 2015). In addition to character-level keyboards, where each character is entered individually, word-level keyboards allow the input of one word per interaction. For instance, word-level swiping keyboards are used by sequentially connecting each character with the finger on a touchscreen, in mid-air (Gupta et al., 2019; Markussen et al., 2014), or with the eyes on a virtual keyboard (Kurauchi et al., 2016). Usually, these approaches are improved with predictive algorithms to achieve faster entry rates, lower error rates, and compensate imprecise input. Recent studies investigated the issue of physical fatigue by proposing alignment concepts that leverage eye-hand coordination to improve text entry performance while reducing hand and eye fatigue (Lystbæk et al., 2022; Zhao et al., 2023).

Moreover, first comparison studies investigate different input modes. For instance, (Xu et al., 2019a) have compared four different input methods. The VR controller-based input (14.6 wpm) significantly outperforms the device-free approaches, i.e., head-gaze (5.62 wpm), hybrid (~8 wpm), and hand gesture (~7 wpm). However, the HTC Vive controller used in this experiment is already very accurate while eye tracking was not available and hand tracking technology has improved since. Lu et al. (2021) compare eye blinks (11.95 wpm), dwell (9.03 wpm), and swipe gestures (9.84 wpm) on an invisible keyboard with a HoloLens 2 whereby eye blinks outperformed the other two approaches. Yu et al. (2017) compare a gesture-commit, dwell, and word-level gesture swiping keyboard but rely on head-pointing instead of eye tracking. Head-pointing requires more effort as users have to move their head instead of the eyes for each character. Speicher et al. (2018) discuss the design space for text entry in VR and compare six different input modes based on head-pointing, mid-air finger tapping, and controller input. They find that there is a "lack of comparative performance evaluations in VR" (Speicher et al., 2018), that prior work is often not comparable due to differences in methodology, and emphasize the tradeoff between pure performance and user experience in different scenarios. Overall, a comparative study between hands-free and hand-based mobile text input modes for AR with similar sensor quality has not been performed, yet.

7.3 Augmented Reality Keyboard Design

There are several considerations and options for a comparative evaluation of different virtual keyboard designs. Physical keyboards and speech input already provide potent ways to enter text in AR in specific settings (Smith & Chaparro, 2015; Speicher et al., 2018). Thus, we expect that virtual keyboards will likely be used in mobile settings when carrying or wearing additional hardware such as physical keyboards (Walker et al., 2017), controllers (Yu et al., 2018), or gloves (Kuester et al., 2005) are impractical and the use of speech input might not be appropriate. Overall, we limit our design space to easy to learn and train, familiar, and non-predictive text entry and interaction modes that are already established with promising results in general or specifically for the VR domain, transferable to the AR space, usable in mobile settings, and underexplored in comparison with each other. Although there are plenty of novel designs (Dube & Arif, 2019), we rely on QWERTY-layouts due to their popularity to reduce the required effort for learning and training. Instead of head-pointing, we use eye-gaze input only as it requires less effort, is more natural, and faster. As we do not focus on a specific prototype implementation and want to allow reproducibility, we decided to use an off-the-shelf Microsoft HoloLens 2 – a state-of-the-art AR HMD. We looked into mid-air touch-typing which is possible and appropriate for touch-typists as shown by Fashimpaur et al.; Singhal et al.; Yi et al. (2020, 2022, 2015). However, based on our trials, the hand tracking system of the HoloLens 2 does not yet appear to be accurate enough for all fingers. More importantly, for touch-typing keyboards each character has a fixed assignment to one finger (e.g., "A" links to the left pinky). However, many touch-typists got used to a varied form of touch-typing by relying on less than 10 fingers or using other fingers for specific keys than the system suggests. For example, in our study only 22% of the participants applied true touch-typing while most typed with between four and eight fingers. Hence, non-touch-typists and some touch-typists would likely require thorough training and deviate from their habits before being able to use these keyboards (Dudley et al., 2019; Kuester et al., 2005; Streli et al., 2022).

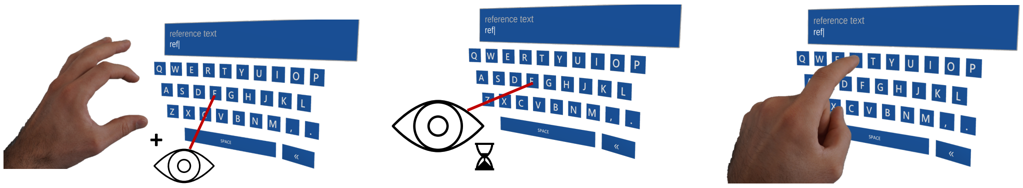

Figure 7.1.: Visualizations of the three AR keyboard prototypes. From left to right: Gaze and commit keyboard, Dwell keyboard, and Tap keyboard. The line representing the user's gaze is not visible in the application.

We therefore choose a dwell-based eye-gaze keyboard (Dwell), an eye-gaze with pinch-gesture-commit keyboard (Gaze and commit; in the following abbreviated as GnC), and a mid-air tap keyboard (Tap) for our comparison as depicted in Figure 7.1 (Mutasim et al., 2021).

Tap keyboard

The Tap keyboard is the default text input mode for the HoloLens 2 and is similar to VISAR for the HoloLens 1 (Dudley et al., 2018), PokeType (Fashimpaur et al., 2020), Freehand (FH) (Speicher et al., 2018), the keyboard presented in (Adhikary & Vertanen, 2021), AirTap (Lystbæk et al., 2022), and the index finger mid-air keyboard of Dudley et al. (2019). Users type by using both their index fingers to "press" the keys of a virtual keyboard that is displayed by the AR device. When a key is pressed, the key changes its color, moves back with the finger, and makes a sound in replacement of haptic feedback.

Dwell keyboard

The Dwell keyboard is an unobtrusive and hands-free input mode similar to Rajanna and Hansen (2018), Majaranta et al. (2009), Dwell-Typing (Lystbæk et al., 2022) and DwellType (Yu et al., 2017). During the development of the application, we noticed that the HoloLens 2 does not compensate the latency of the eye tracker. We created a tool similar to Stein et al. (2021) to get a rough approximation of the inherent latency. The approximate latency is 120 msec. We varied the dwell time in our pre-experiment sessions with multiple colleagues which resulted in a dwell time of 550 msec which is in line with Rajanna and Hansen (2018). Continuous fixations on one key enter another character every 550 msec. It is important to note that this setting applies to novice users. In a longitudinal study, participants could gaze type on average with a dwell time as low as 282 msec after ten sessions (Majaranta et al., 2009).

Gaze and commit (GnC) keyboard

The GnC keyboard is a combination of gaze-based selection by looking at the keys and pinching index finger and thumb together to "press" the key which the user currently looks at (Chatterjee et al., 2015; Pfeuffer et al., 2017). In comparison to the Dwell keyboard, the GnC keyboard is not susceptible to the Midas touch effect (Vrzakova & Bednarik, 2013) – the challenge to differentiate intentional from unintentional keypresses – thus, giving the user more control but requiring additional hand input. It is similar to the VISAR baseline (Dudley et al., 2018), TapType (Yu et al., 2017), to the controller variant of (Rajanna & Hansen, 2018) but with hand gesture-commit instead of the controller, or the hybrid scenario of Xu et al. (2019a). This interaction technique will also be the default interaction of the upcoming Apple Vision Pro (Apple, 2023).

General keyboard properties and design

To maximize comparability, all keyboards are developed as similar as possible in appearance and characteristics. The same input mode can be instantiated in several ways and with different properties, however, the "effects of various keyboard properties have not yet been fully studied" (Dube & Arif, 2019). Therefore, we consider existing knowledge, but must make assumptions about several attributes. We place the Dwell and GnC keyboards 2 m in front of the users to reduce discomfort from the vergence-accommodation conflict (Anthes, 2019). The Tap keyboard is positioned close to the user in order for them to reach the keys. All keyboards fit within the field of view as suggested by Rajanna and Hansen (2018). The keyboards are slightly curved to achieve similar distances between the user and each key. When gazed at and when "pressed," the keys respond with audio-visual feedback. For simplicity, we offer a reduced set of characters (a-z, space) like Vertanen et al. (2015) and text revision (backspace to delete characters). Furthermore, we decided against using any predictive system to study the input modes in isolation, including swiping keyboards. In line with He et al. (2022), we argue that features such as auto-correction can "confound accuracy." Several keyboards demonstrate the advantage of predictive systems especially for slower input modes (Dube & Arif, 2019; Fashimpaur et al., 2020; Zhang et al., 2022). However, this can be considered as a separate problem that can improve typing speed on top of the interaction itself.

We used the Mixed Reality Toolkit 2.7.3 in Unity 2020.3.34f1 LTS to develop the AR application for the HoloLens 2. We followed the MRTK design guidelines to not only achieve a functionally working prototype but also present an appealing appearance. A reliable implementation of each input mode is necessary to adequately compare the variants. Hence, we had multiple design and development iterations and user tests with colleagues and students before conducting the experiment. The gaze-based keyboards are placed 150 cm in front of the user. Each character key has a side length of 4 cm, 1.7 cm horizontal and 2.6 cm vertical spacing between characters, and the keyboard is slightly curved along a circle with a radius of 1.8 meters. The tap keyboard is scaled down to 70% and spawns 60 cm in front of the user.

7.4 Experiment Design

7.4.1 Design

The experiment follows a within-subjects design using a complete counterbalanced order of treatments as suggested by MacKenzie (2010) to reduce order and carryover effects. Hence, each participant learns and tests each of the three keyboard designs (i.e., GnC, Dwell, and Tap). Participants are seated at a desk in a spacious, neutral, and well-lit room, and on a movable and adjustable swivel chair. Questions are answered in a questionnaire on a regular laptop. Apart from the laptop, the desk is empty to leave enough space for AR holograms. Overall, we reduced any interaction between subjects and the experimenter to a minimum.

Experimental Procedure

The experimenter welcomes participants, asks them to read and sign the data privacy-related documents and consent form, helps with putting on and calibrating the AR glasses, and leaves the room after the first phrases have been entered successfully. Participants alternate between using an AR application specifically designed for this experiment and filling out a questionnaire. The AR application introduces participants to AR and the text entry task. As typing skill reference and to familiarize the participants with the task, they complete the first text entry test on a regular physical keyboard (Logitech MX Keys) connected to the AR device via Bluetooth. Then, for the first keyboard design, participants are taught the interaction mode with simple buttons, a game, and finally the keyboard where they can type freely as depicted in Figure 7.2. Afterwards, participants copy type five sentences as training and ten sentences as test set. Subsequently, they have to answer the treatment-questionnaire surveying different aspects about their experience with the respective keyboard design on a separate computer. Then, participants return to the AR device to repeat the same steps for the next keyboard design. After completing the procedure for all three keyboards, a post-treatment-questionnaire asks questions comparing the three keyboards and ends with demographic questions. The experimental procedure is depicted in Figure 7.3.

Figure 7.2.: Steps of the tutorial for each of the interaction modes. Participants learn to interact with simple buttons (left) and play a short game in which they have to pop objects by using the interaction (right) before typing with the keyboard.

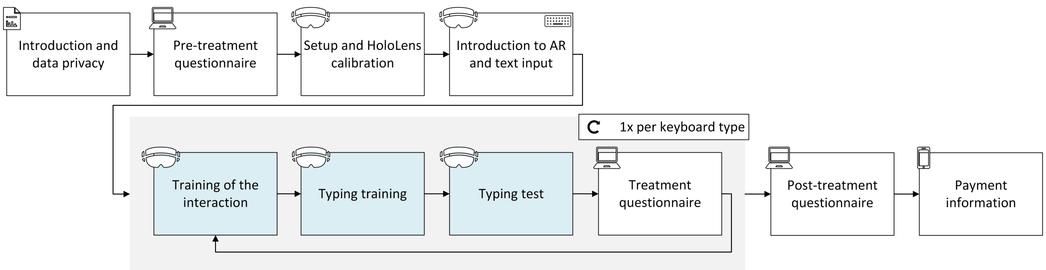

Figure 7.3.: Experiment process

The keyboard reference is useful as previous studies exhibited different levels of physical keyboard entry speeds and error rates indicating different experience in the respective samples (Dube & Arif, 2019; He et al., 2022). This can be viewed as a measure for participants' regular typing skill as typing skills can vary between age and culture. When answering the questionnaires, participants were asked to take off the AR device to reduce physical stress. We decided to implement the complete experiment procedure as an AR application instead of the keyboards only in order for participants to autonomously progress through the experiment without the experimenter being present. This encompasses tutorials, as recommended by Kuester et al. (2005), and voice instructions. The latter were recorded in a recording studio for each step and ensure a consistent interaction between subjects. At any time, participants can call the experimenter via a virtual help button displayed in AR if questions or issues occur.

Text entry task

For each keyboard, participants perform a copy text task with native phrases as described below. Participants are asked to type as quickly and accurately as possible. Error correction is allowed and required. Accepting uncorrected errors could lead to distorted results if participants vary in their intrinsic motivation to correct errors. Thus, participants must correctly enter each phrase to progress to the next phrase and complete a set. As a result, the UER equals zero for each sentence and the TER equals the CER. After conducting trial sessions, we choose five sentences for training and ten sentences for the typing test to capture the first impressions of each design. Additionally, several first time users of the AR device mentioned increasing eye strain and general discomfort over time due to the unfamiliar holographic display and weight of the device. This could lead to order effects in a within-subjects design if participants spent too much time with each keyboard design. The typing tests color correct and incorrect characters during typing as visual feedback to avoid additional eye strain by rigorously comparing the input and reference texts.

Phrase set

In our first pre-experiment trial, all four participants noted that they were slower and more erroneous in typing the English phrases from the standard phrase set of MacKenzie and Soukoreff (2003) than they would have with phrases in their native language. Franco-Salvador and Leiva (2018) argue that native phrase sets should be used for text entry tasks and present a sampling technique to derive phrases from a text corpus for different languages. Thus, we relied on a German phrase set for the language of the study from Franco-Salvador and Leiva (2018), which is based on movie captions. Additionally, the keyboard layout was changed to the more familiar QWERTZ layout, the German QWERTY adaption. For each keyboard, we sampled 15 unique sentences from the phrase set, filtered for similar word and character length distributions, and removed sentences with offensive language. Thus, every participant was presented with the same set and order of phrases.

Questionnaire

In the treatment-questionnaire, we ask questions regarding the subjective keyboard performance (Dudley et al., 2018), perceived enjoyment of the Technology Acceptance Model 3 (Venkatesh & Bala, 2008), the raw NASA task load index (Hart, 2006; Hart & Staveland, 1988), the System Usability Scale (Brooke, 1996), and for general feedback in open texts on aspects that participants liked, bothered, and what they would improve. Our self-composed questions are on a seven-point Likert scale unless otherwise noted. We added an additional question to the keyboard performance: "The keyboard could be easily used in a mobile context." Furthermore, we asked for the tutorials' helpfulness by asking "How effective was the tutorial in teaching the typing interaction?" and for transfer of learning based on Kuester et al. (2005) by asking "How effective mapped previous touch-typing skills to your skills with the keyboard you just used?"

The post-treatment-questionnaire asks for comparisons between characteristics of the three keyboards and in different scenarios. The usage scenarios, although not being exhaustive, address other external factors that may influence keyboard performance in mobile contexts (MacKenzie & Soukoreff, 2002; Speicher et al., 2018). We selected scenarios in which users are seated or standing, in a public or private environment, and tasks in which the hands are free or occupied. The questions are listed in Table 7.1.

After the questions regarding "Keyboards in context," participants are given the choice to either choose one of the previous keyboards (GnC, Dwell, Tap, or regular physical keyboard) or speech input instead for each of the contexts. Finally, participants optionally enter their gender and to which age range (18-24y, 25-29y, 30-34y, . . .) they belong.

Before conducting the experiment, the ethics commission and data privacy department of the university approved the experiment design.

7.4.2 Participants

Overall, 29 participants took part in our experiment. We recruited healthy participants from a participant pool of a large European technological university. Participants received a compensation of 17 € ($\approx$ $17.34) for participating in the 75-minute experiment with a chance to receive an additional reward of up to 3 € based on their relative average performance during the typing tests. Prior to the data analysis, we had to exclude one participant due to insufficient eye tracking accuracy caused by

Table 7.1.: Questions and response options of the post-treatment-questionnaire.

Construct name	Question text	Response options
Preference	"Which keyboard did you like the most?"	GnC, Dwell, Tap
Perceived performance	"Which keyboard was the fastest to enter text?"	GnC, Dwell, Tap
Perceived usability	"Which keyboard was easiest to use?"	GnC, Dwell, Tap
Perceived required attention	"Which keyboard required the most attention and focus?"	GnC, Dwell, Tap
Preference reason	"Why did you like the chosen keyboard the most?"	Open text
Keyboards in context	"Which keyboard would you most likely use for this task?"	
	Writing an email while sitting at your desk at work.	GnC, Dwell, Tap, Physical, Speech
	Replying to an Instant Message (e.g., WhatsApp) while standing in a subway.	GnC, Dwell, Tap, Physical, Speech
	Quickly commenting a post on social media while waiting for a friend at a restaurant.	GnC, Dwell, Tap, Physical, Speech
	Comparing product prices to Amazon.com while shopping at a local retailer.	GnC, Dwell, Tap, Physical, Speech
	Googling a recipe while cooking.	GnC, Dwell, Tap, Physical, Speech

interferences of the AR glasses with their corrective lenses. Further, one participant did not follow the instructions properly and failed to answer the attention checks correctly. Out of the remaining 27 participants, seven participants were female. 20 participants are in the age range of 18 to 24, six in the age range of 25 to 29, and one person is in the age range of 30 to 34. Regarding the typing style on a physical keyboard, 13 participants typed in a hunt-and-peck style with a maximum of four fingers, while the other participants applied hybrid- or touch-typing.

Throughout the study, two participants used the help button to ensure that they correctly understood the sequence of the experiment. One participant experienced a crash of the application after completing the physical keyboard test but could continue at the same stage without data loss. One participant asked to restart the survey after misreading instructions before starting with the second keyboard.

7.5 Results

We analyzed our data by calculating repeated measures Analysis of Variances (ANOVAs) between the three keyboards if the assumptions of normality and sphericity were met (denoted with F) and with a Friedman test (χ^2) otherwise. Only in case of significant differences between the keyboards, Bonferroni-corrected post-hoc tests [either paired t-tests (ANOVA) or Conover's test (Friedman) (Conover, 1999)] were performed and reported in the following. The significance level for all tests was set to $\alpha = 0.01$. We exclude the training phase of each keyboard for the analysis and only report measures for the actual test phase. An overview of all relevant measurements and questionnaire results is given in Table 7.2.

Text input measures

Participants typed on average with an entry rate of 11.67 wpm ($SD = 2.56$ wpm) and a Corrected Error Rate (CER) of 6.08% ($SD = 3.29\%$) on the GnC keyboard, 10.98 wpm ($SD = 1.92$ wpm) and 4.45% ($SD = 4.15\%$) on the Dwell keyboard, and 15.27 wpm ($SD = 2.07$ wpm) and 2.02% ($SD = 1.45\%$) on the Tap keyboard. The Tap keyboard is significantly faster than both Dwell ($p < .001$) and GnC ($p < .001$) keyboards and has a lower CER than the GnC keyboard ($p < .001$). The performance of the Tap keyboard is comparable to the results of Adhikary and Vertanen (2021). Figure 7.4 provides an overview on the entry rate and corrected error rate for the three keyboards. As reference, the average entry rate with the physical keyboard was 65.94 wpm ($SD = 13.19$ wpm) with a CER of 2.74% ($SD = 3.55\%$).

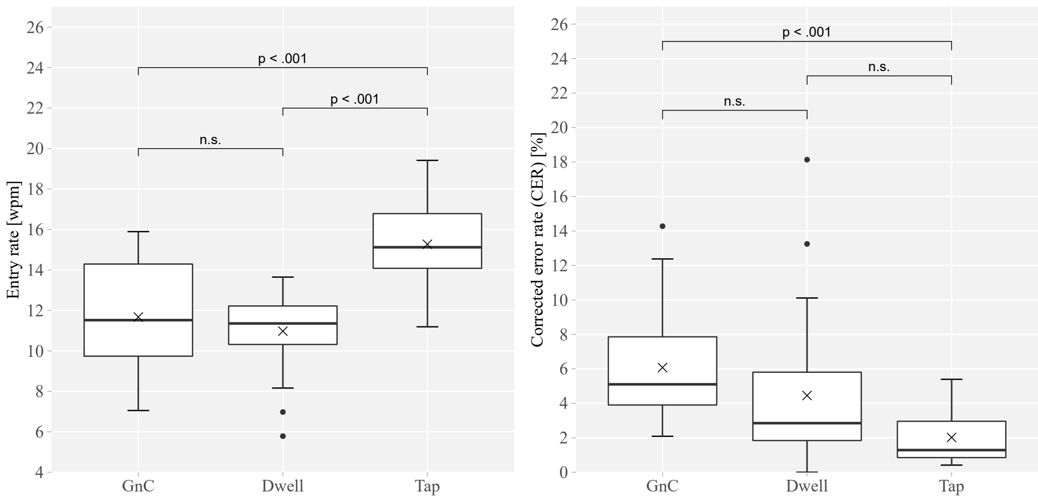

Figure 7.4.: Entry rate (left) and corrected error rate (right) of the three virtual keyboards (GnC, Dwell, Tap). "x" indicates the mean value and the bold horizontal line indicates the Median. Friedman test for entry rate ($\chi^2(2) = 31.63, p < .001$) and corrected error rate ($\chi^2(2) = 23.57, p < .001$). Post-hoc tests based on Conover's test with Bonferroni correction.

Table 7.2.: Statistics and questionnaire results. LS denotes scales measured on 7-point Likert scales ranging from $[-3; +3]$. Highlighted cells denote $p < .01$. Post-hoc tests based on Conover's test with Bonferroni correction.

Metric / scales	Gaze and commit Mean (SD)	Dwell Mean (SD)	Tap Mean (SD)	RM-ANOVA or Friedman test	Post-hoc tests
Entry rate [wpm]	11.67 (2.56)	10.98 (1.92)	15.27 (2.07)	$\chi^2(2) = 31.63, p < .001$	Tap-GnC, Tap-Dwell
Error rate (CER) [%]	6.08 (3.29)	4.45 (4.15)	2.02 (1.45)	$\chi^2(2) = 23.57, p < .001$	Tap-GnC
Raw NASA-TLX Score	50.52 (14.66)	44.35 (12.10)	45.68 (13.87)	$F_{2,52} = 2.22, p = .119$	
NASA-TLX: Mental load	53.70 (27.09)	56.30 (28.61)	32.04 (22.11)	$\chi^2(2) = 20.35, p < .001$	Tap-GnC, Tap-Dwell
NASA-TLX: Physical load	48.89 (26.21)	21.48 (24.29)	63.89 (19.48)	$\chi^2(2) = 26.08, p < .001$	GnC-Dwell, Tap-Dwell
NASA-TLX: Temporal load	50.74 (22.30)	54.63 (23.82)	55.56 (23.95)	$\chi^2(2) = 1.66, p = .435$	
NASA-TLX: Performance	43.70 (23.23)	37.59 (24.86)	37.59 (21.85)	$\chi^2(2) = 0.90, p = .639$	
NASA-TLX: Effort	56.67 (23.25)	54.81 (26.33)	52.41 (22.55)	$\chi^2(2) = 0.18, p = .913$	
NASA-TLX: Frustration	49.44 (27.68)	41.30 (24.20)	32.59 (26.54)	$\chi^2(2) = 3.10, p = .212$	
System usability scale score	66.91 (17.65)	75.62 (10.09)	73.83 (14.86)	$\chi^2(2) = 5.49, p = .064$	
Easy to type quickly. (LS)	0.26 (1.79)	0.07 (1.69)	0.00 (2.06)	$\chi^2(2) = 1.12, p = .572$	
Easy to type accurately. (LS)	0.07 (1.73)	0.89 (1.45)	1.15 (1.43)	$\chi^2(2) = 4.07, p = .131$	
Comfortable to use. (LS)	-0.22 (1.85)	0.81 (1.71)	-0.04 (1.65)	$\chi^2(2) = 4.07, p = .131$	
Usable in a mobile context (LS)	0.89 (1.58)	1.52 (1.65)	0.30 (1.61)	$\chi^2(2) = 10.21, p < 0.01$	Tap-Dwell
Perceived enjoyment (LS) (Cronbach's α)	0.31 (1.69), $\alpha = 0.942$	1.06 (1.22), $\alpha = 0.752$	0.91 (1.39), $\alpha = 0.874$	$F_{2,52} = 4.15, p = .021$	
Transfer of learning (LS)	0.0 (1.9)	0.19 (1.57)	0.52 (1.72)	$\chi^2(2) = 3.04, p = .219$	
Helpfulness of tutorials (LS)	1.67 (1.11)	2.48 (0.8)	2.07 (1.0)	$\chi^2(2) = 12.7, p < .01$	GnC-Dwell

Task load

The raw NASA-TLX scores were 50.52 ($SD = 14.66$) for GnC, 44.35 ($SD = 12.10$) for Dwell, and 45.68 ($SD = 13.87$) for Tap but not significantly different between the three keyboards ($F_2, 52 = 2.22, p = .119$). However, on the individual dimensions, participants reported a lower physical load for the Dwell keyboard ($M = 21.48, SD = 24.29$; $\chi^2(2) = 26.08, p < .001$) in comparison with the Tap ($M = 63.88, SD = 19.48; p < .001$) and GnC keyboards ($M = 48.89, SD = 26.21; p < .01$). Additionally, the mental load was lower for the Tap keyboard ($M = 32.04, SD = 22.11; \chi^2(2) = 20.35, p < .001$) compared to the Dwell ($M = 56.30, SD = 28.61; p < .001$) and the GnC keyboards ($M = 53.70, SD = 27.09; p < .01$). Figure 7.5 summarizes the results for the NASA-TLX sub-scales for the three virtual keyboards.

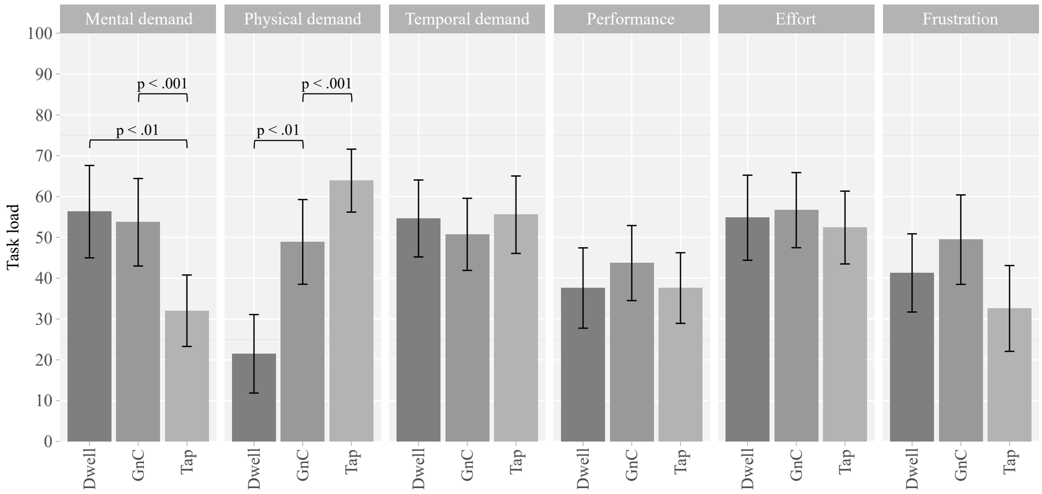

Figure 7.5.: NASA-TLX task load scores for the sub-scales (from left to right: mental demand, physical demand, temporal demand, performance, effort, frustration) of each of the three virtual keyboards. Whiskers indicate the 95% confidence interval.

Usability and preference

For the SUS score no significant differences between the GnC ($M = 66.91, SD = 17.65$), Dwell ($M = 75.62, SD = 10.09$), and Tap ($M = 73.83, SD = 14.86$) keyboards were found ($\chi^2(2) = 5.49, p = .064$). Following Sauro (2011), the usability of Dwell and Tap keyboards is rated above the average score of 68 and can thus be considered as “good,” while the GnC keyboard is slightly below average, i.e., considered as “OK.” Participants stated diverging preferences regarding the keyboards: eight participants liked the GnC, eight the Tap, and eleven the Dwell keyboard the most.

Similarly, the results for the keyboards' ease of use are also inconclusive (GnC: 5, Dwell: 12, Tap: 10). The gaze-based keyboards were rated to require the most attention and focus while typing (GnC: 10, Dwell: 16, Tap: 1). Regarding different usage scenarios, 24 participants (88.89%) would prefer to still use a physical keyboard over the virtual keyboards at their desk at work. However, particularly in public contexts, most participants prefer less obtrusive input modes, e.g., when writing an instant message in a subway (GnC: 2, Dwell: 19, Tap: 5, Physical: 1), when commenting a social media post while sitting in a restaurant (GnC: 7, Dwell: 12, Tap: 5, Physical: 3), or when searching online for a product while shopping at a local retailer (GnC: 8, Dwell: 11, Tap: 5, Physical: 3). In case speech recognition would be available in the previous scenarios, most participants would yet prefer one of the virtual or physical keyboards than using speech input ($66.67\% - 88.89\%$). Only for a scenario, in which participants would be at home cooking and want to search online for a recipe, the majority (88.89%) of the participants would prefer speech input over the other keyboards. Participants change their preference for one keyboard depending on the context. For instance, 12 participants would switch to the Dwell keyboard in the "messaging in a subway" scenario even if they in general preferred another keyboard. Cronbach's α for perceived enjoyment was above the commonly accepted threshold of 0.7 for all keyboards (GnC: 0.942, Dwell: 0.752, Tap: 0.874). The tutorials for the interaction modes were each rated as effective on average (GnC: $M = 1.67, SD = 1.11$; Dwell: $M = 2.48, SD = 0.80$; Tap: $M = 2.07, SD = 1.00$) on a 7-point Likert scale $[-3; +3]$.

Subjective typing performance and requirements

We determined the keyboard, which allowed the fastest typing for each participant ("actual") and compared it to the selection of their subjectively perceived fastest keyboard ("perceived"). 15 participants (55.56%) misjudged the keyboard on which they performed best. Three participants misjudged the fastest keyboard by more than 5 wpm ($> 40\%$). In all these cases, the participants actually typed the fastest on the Tap keyboard but chose either the GnC ($n = 10$) or the Dwell ($n = 5$) as perceived fastest keyboard. Regarding the reasons why participants chose one keyboard as their preference, comfort ($n = 14$), entry rate ($n = 11$), reliability ($n = 6$), and practicability in a mobile context (n = 5) were mentioned the most ($n \geq 5$) in an open text field. In the following paragraphs, we report aggregated answers to the open text fields regarding advantages, disadvantages, and improvement potential of each keyboard.

GnC keyboard

Five participants appreciate the control due to the commit gesture in comparison to the Dwell keyboard. Three participants experienced that they intuitively moved their eyes to the next key slightly before finishing the commit gesture, which resulted in an error and limited their entry rate. Additionally, ten participants mentioned that they want to be able to place their hands on their laps while typing, instead of holding the hands in front of their head. While the HoloLens 2 can detect the hands at an offset position, the hand tracking accuracy appears to be lower near the edges of the detection frame. Therefore, we did not communicate this option.

Dwell keyboard

On the one hand, participants like the simplicity (7x), learnability (3x), and hands-free character (10x) of the Dwell keyboard. On the other hand, participants found the input mode stressful (2x), eye-straining (7x), and complained about the dwell time which was either too long or too short for them (6x).

Tap keyboard

Participants mentioned the intuitive and familiar character (7x), and the audio-visual feedback when pressing the keys of the Tap keyboard (8x). Additionally, two participants explicitly mentioned lower eyestrain in comparison to the other keyboards. 12 participants highlighted the high physical strain as they had to hold up their arms continuously during typing. Six participants would prefer larger keys to increase the accuracy or smaller keys to keep the whole keyboard within the field of view. Thus, there were suggestions to add an option for adjusting the keyboard and key size. Moreover, the Tap keyboard placement should be lower than the gaze-based keyboards to reduce physical strain.

In general, 13 participants emphasized the maturity and reliability of the keyboard implementations. When comparing the responses between the keyboards, participants who struggled with one input mode still manage to satisfactorily type with one of the other keyboards.

7.6 Discussion

Participants achieved fast entry rates in comparison to similar studies given the limited amount of training they had with each keyboard (Dube & Arif, 2019; Lu et al., 2021; Xu et al., 2019a). Text entry speed will likely increase while the error rates decrease when participants would get the chance to use the keyboards for longer periods of time (Majaranta et al., 2009). Furthermore, it is plausible that the Dwell keyboard leads to a significantly lower physical load as it is hands-free and the Tap keyboard having a lower mental load as it is more familiar. Yet, it is interesting to mention that all other NASA-TLX sub-scales have similar medium loads with high variances between the participants. The SUS scores are better than we expected for novel keyboard designs, interaction modes, and working with an unfamiliar AR device for an extended period of time which can indicate a high maturity of the implementations. However, the high variance and average values for several of the subjective assessments such as perceived enjoyment are worth investigating further and lead to the conclusion that users have heterogeneous demands with respect to the keyboards in general, text lengths, and in different settings.

Users demand for appropriate keyboards

The participants of our study have displayed distinct preferences for the keyboards in general. This finding is in line with Fashimpaur et al. (2020) that outlined that participants had diverging opinions and preferences on a mid-air tap keyboard and a mid-air touch-typing keyboard. Many typing studies focus on entry and error rate, however, "speed and accuracy alone do not entirely reflect the effectiveness of a text entry method – usability, learnability, fatigue, and space requirement must also be taken into consideration" (Dube & Arif, 2019). The results of this study show that comfort, reliability, and practicability are important to users in mobile contexts next to the entry rate. Moreover, the requirements differ between individual users leading to different preferences for the respective input modes. While participants who prefer comfort would rather choose the Dwell keyboard, performance-oriented users choose the GnC or Tap keyboards. However, as the actual and perceived performances diverge, performance-oriented users do not only want to type quickly but also get the impression that they do. It is up to the designer of the keyboards to weight the actual performance against the subjective feeling of a keyboard being performant.

Users demand for appropriate keyboards for specific settings

While physical keyboards will likely remain relevant in future stationary setups due to the achievable typing speed, accuracy, and familiarity, reliable speech input can significantly speed up text input (Smith & Chaparro, 2015). Still, scenarios especially in public space remain where both input modalities are less practical than a virtual keyboard that does not have to compromise on mobility, confidentiality, or in subtlety (Speicher et al., 2018). Our implementations of Tap, Dwell, and GnC outperform Xu et al. (2019a) tapping keyboards who used a less advanced Meta 2 AR HMD and, thus, they conclude that device-free input types should only be used as a last resort. We argue that device-free input types are not only a last resort but rather a reasonable choice if the keyboards are tailored to those specific situations. Many participants would trade performance and comfort for subtlety when in a public space, preferring the hands-free Dwell keyboard over the other variants. Therefore, subtle but mobile approaches such as Zhang et al. (2022) might find interest among a specific group of future users.

Users demand for appropriate keyboards for different text lengths

Similar to the different contexts, participants would choose different keyboards for different typing tasks. All participants would perform a short query on Google or type a brief instant message with one of the three keyboards but refrain from using them for longer texts. This opens a gap for use cases such as working with longer texts on a train or similar where users would still need to carry physical keyboards with them or might use novel keyboards in the future.

Implications for specific interaction modes

Our findings suggest several implications for the different keyboard designs. Dwell keyboards can benefit from an adaptive dwell time as shown by Majaranta et al. (2009) or even dwell-free swiping approaches for more experienced users as suggested by Kristensson and Vertanen (2012) and implemented by Kurauchi et al. (2016), Xu et al. (2019a), and Kurauchi et al. (2020). Additionally, when users tap their fingers (or click a button on a controller) on a gaze and commit keyboard, the system should allow that the eyes could already move to the next character even if the commit-action has not been completed. Users intuitively try to reduce the fixation duration when optimizing input speed. When designing tap keyboards, the placement, size, and ability for users to resize keys and keyboard are important. For instance, Shen et al. (2022) demonstrate a keyboard that adapts the layout size to improve performance.

Limitations and future work

Although several participants mentioned the maturity of our keyboard implementations, production-ready versions would need further extensions. The character-set should include capitalization, numbers, and special keys. Also, well-developed error correction features as suggested by Li et al. (2021) are necessary. For example, the cursor could be placed by gazing at misspelled characters. The keyboard performance would also benefit from predictive features such as next-word prediction or auto-correction (Dube & Arif, 2019; Zhang et al., 2022). Additionally, the participants were seated during the study to increase comfort and control for movement and positioning. Asking participants to imagine arbitrary, non-seated scenarios and assess their usage of the virtual keyboards in these scenarios is only exploratory. The results indicate research potential in the tradeoff between such scenarios, however, further research must focus on the elements composing these scenarios. Even though the mobile aspects of the study refer to a non-stationary setup and environment in which it is safe to type instead of, e.g., typing while walking, further research can investigate the impact of standing or moving positions on AR text entry. As our participants are sampled from a western technological university pool of mainly students the participants do not represent the age or gender structure of the general population. We tested each measure for gender differences (Mann-Whitney U test) without significant results. We did not find reports of major gender differences in related studies, however, future work could investigate potential differences with smaller effect sizes in a larger study.

The implications for research are that the focus on keyboard performance must be broadened to the general user experience with text entry interfaces, especially, if users type in non-stationary, dynamic environments where a compromise on entry speed is tolerated. Additionally, keyboard design must account for different users with distinct requirements. This includes the accessibility domain, where users with specific needs could profit from personalized text entry (Creed et al., 2023). For instance, if one text entry mode cannot be used, users could switch to another mode without the need for external software.

We suggest further research in hybrid text input modes that addresses the needs of both novice and expert, as well as comfort- and speed-oriented typists. Derived from the implications of this study, our suggestion is researching a hybrid keyboard by combining Tap and GnC swipe keyboards like current smartphone keyboard implementations (touch-input + swiping). Most users from our sample can type fast and accurately on a Tap keyboard without previous knowledge especially if the text length is short to medium. Advanced users with a preference for a GnC keyboard could instead use eye-swiping by pinching index finger and thumb together for each word. Both interaction modes can be used with the same keyboard as they do not interfere with each other. For specific use-cases, separate buttons could activate a purely dwell-based and a speech-based approach for hands-free text input. We expect a solution like this to target different user groups and scenarios more comprehensively and inclusively, especially if a hands-free input mode is included.

Text entry studies are subject to several design decisions and tradeoffs. Depending on the study design, these tradeoffs put an emphasis on different characteristics of the results. However, the combination of several studies forms a comprehensive picture of the domain. An evaluation of a single prototype and its variations enables participants to spend more time with the keyboard and determine the effects of specific characteristics, but can struggle with comparability (Dube & Arif, 2019). In longitudinal studies, subjects take the time and effort to properly learn typing with a keyboard over multiple sessions and researchers can observe learning rates and upper limits of input modes. Yet, this is rare in reality, e.g., when looking at our sample of which 13 participants (48.15%) did not apply touch-typing when typing with a physical keyboard. We followed Speicher et al. (2018), Lu et al. (2021), and Xu et al. (2019a) in designing a comparison study. This design has confounding effects of novelty and learning as each participant only types a couple of phrases with each keyboard. However, this first impression is valuable and important for the long-term preference and user experience of a system. In addition, study participants were unfamiliar with all the three virtual keyboards. Thus, our study succeeds in providing controlled manipulation between the keyboards leading to comparable results. Still, testing the three keyboards in a between-subjects longitudinal study could gain interesting additional insights into the development of long-term entry speeds, error rates, and potential shifts in preference of experienced users that this study cannot provide.

Five users struggled with the hand tracking accuracy and reliability of the HoloLens 2, which resulted in an overall worse performance with and lower preference for the keyboard (especially the commit gesture of the GnC keyboard). In general, the hand-tracking capabilities in low lighting or low contrast environments and when fingers are occluded are a limitation of the HoloLens 2. Moreover, the eye tracker's latency, temporal resolution, and spatial accuracy is limited (Aziz & Komogortsev, 2022). Future studies could use better hand tracking sensors such as an Ultraleap

Leap Motion Controller in addition to the build-in capabilities of AR devices, as this is an essential aspect for gesture-commit keyboards. However, future lightweight and highly integrated devices might also compromise on sensor quality in favor for mobility and comfort. Thus, these challenges could arise in future consumer-grade hardware, too.

Practical implications

In light of our study, choosing only the Tap keyboard may appear as a reasonable choice for system designers. However, it is important to emphasize that such a decision could result in many users being unsatisfied and ultimately hinder the overall experience of future mixed reality systems. Due to the eye-gaze input, the Dwell and GnC keyboards were described as "futuristic," "fun," and "magical," opening the door for exciting rather than satisfactory experiences.

7.7 Conclusion

With the continuous advancement of AR technology, it is becoming increasingly important to explore potential solutions for the well-known problem of text input in AR. Particularly in mobile settings, users may demand keyboards that do not require to carry any additional hardware or external tracking equipment. Therefore, our study set out to compare three promising device-free text-entry solutions for AR in a controlled laboratory experiment while also capturing preferences for the different keyboards. Our results reveal that participants have different preferences regarding their choice of the keyboard in general, but also depending on factors such as usage scenario or text length. Therefore, we conclude that it will be interesting – especially for future research – not to focus only on a single universal keyboard, but to also think about adaptive keyboard designs that may offer, e.g., different input modalities depending on external factors. Furthermore, many participants had difficulties to judge with which keyboard they could achieve the highest performance.

The AR keyboard of the future will likely be defined by a company while building one of the first widely successful, formative AR devices. Until then, researchers have the chance – and in our humble opinion also the obligation – to gain enough knowledge so that future users of AR keyboards do not have to settle for an unsatisfactory but established keyboard.

Part III

Applying Augmented Reality to Smart Home Control

8 Introduction

AR has the potential to change the way we interact with digital content and the physical world. Particularly, the latter aspect fundamentally alters the philosophy of interaction, as it enables direct interactions with physical objects rather than through abstract user interfaces on digital screens. With AR, it is irrelevant whether these objects have physical buttons or screens, as the interfaces can be augmented through AR. These objects can encompass all IoT devices in any context, provided there is a connection to the AR HMD.

For the scientific investigation of cross-device interaction, the Smart Home serves as an ideal testbed. Smart Homes are often characterized by a multitude of interconnected IoT devices fulfilling various functions. This necessitates a wide range of interactions to accommodate different affordances. At the same time, Smart Homes can be structured to have a unified interface through which an AR HMD can communicate with the entire system. Furthermore, for human studies, a Smart Home provides a suitable testbed as it represents a private space where interactions are not influenced by social interactions (Knierim et al., 2019; Schenkluhn et al., 2023a). Moreover, it can be argued that for random participants, it is a more accessible environment than, for example, a production hall, an automotive workshop, or a hospital, where this form of cross-device interaction is also conceivable. Since aspects of Human-Computer Interaction are being investigated where the specific implementation of the IoT device is secondary, it can be assumed that findings from the Smart Home context can be transferred to other contexts.

One research area in Smart Home research is Ambient Assisted Living (AAL). AAL systems aim to support people with disabilities or older individuals in leading an independent life and to enable them to live alone in their own homes for longer. In designing systems targeted at people with disabilities or older individuals, improvements for the general populace that might otherwise be overlooked are also anticipated in addition to more inclusive design overall (Steinfeld & Maisel, 2012).

Therefore, the first step in this part is an investigation of the expected benefits and challenges of using AR in Smart Homes targeted at people with disabilities or older individuals. Based on the results of this investigation, the focus shifts to the automation of functions in Smart Homes.

Many systems are often not considered interoperable or user-friendly (Coppers et al., 2020; Jakobi et al., 2018; Setz et al., 2021; Solaimani et al., 2013). This issue becomes even more relevant as every interaction is more costly in terms of time and effort for people with disabilities or senior citizens. Thus, Chapter 10 explores the potential of automations that can replace typical daily interactions in AAL or Smart Home settings in general based on the users' location. Specifically, it proposes the innovative approach of utilizing the indoor positioning capabilities of AR HMDs to detect, track, and identify residents for the purpose of automatically controlling various IoT devices in Smart Homes. An implementation of this feature on an off-the-shelf Microsoft HoloLens 2 without additional external trackers is demonstrated, and the results of a feasibility study are presented.

A novel challenge when interacting with IoT devices at their spatial position is identifying the devices' 3D positions. Therefore, Chapter 11 introduces and evaluates three concepts for identifying IoT device positions with varying degrees of automation. This mixed-methods laboratory study with 28 participants revealed that despite being recognized as the most efficient option, the majority of participants opted against a fast, fully automated detection, favoring a balance between efficiency and perceived autonomy and control. This decision is linked to psychological needs grounded in self-determination theory and the strengths and weaknesses of each alternative, motivating a user-adaptive solution are discussed. Additionally, the study observed a "wow-effect" in response to AR interaction for Smart Homes, suggesting potential benefits of a human-centric approach to the Smart Home of the future.

These aspects are presented and discussed in detail in the following three chapters.

9 Leveraging Stakeholder Engagement in the Co-Creation of Augmented Reality Applications for Assistive Solutions

This chapter is based on a study conducted in collaboration with Jurek Muff as part of his seminar thesis at the Institute of Information Systems and Marketing at the Karlsruhe Institute of Technology (KIT). The study was planned, prepared, and carried out jointly based on my initial drafts. Jurek Muff took over the organizational part, the transcription and evaluation primarily and in regular coordination with each other. This chapter comprises a detailed revision of Jurek Muff's term paper.

9.1 Introduction

AR has witnessed a resurgence in recent years, propelled by innovative developments such as Microsoft's HoloLens 2, Apple Vision Pro, and promising startups like Magic Leap (Leap, 2024; Norouzi et al., 2019). This resurgence is further supported by the increased integration of smart devices within domestic environments, leading to the application of AR-based methods in private settings (Bitkom Research, 2022; Deloitte, 2018; Park et al., 2022). In these contexts, AR-based interfaces facilitate various functions, including the control of home devices, visualization of smart device data, and support for automation creation (Ariano et al., 2022; El-Moursy et al., 2022; Zheng et al., 2022). Such interfaces may offer distinct advantages over traditional interaction methods, such as mobile apps or Intelligent Virtual Assistants (IVAs), in terms of user experience, accessibility, and intuitiveness (Flick et al., 2021; Welch et al., 2019).

A notable application of AR technologies is within the realm of Ambient Assisted Living (AAL), which leverages technology and smart environments to enhance the quality of life for individuals, especially the elderly or those with disabilities, in their living spaces (Calvaresi et al., 2017; Ghorbani et al., 2019). Although AR holds significant promise for advancing AAL objectives, current research predominantly focuses on an expert-driven development approach, with limited emphasis on stakeholder engagement and the incorporation of Inclusive Design principles (Hayhurst, 2018; Kanno et al., 2018; Thakur & Han, 2021). To maximize the benefits of AR in AAL, adopting a user-centric co-creation process that emphasizes stakeholder engagement and Inclusive Design principles is imperative to circumvent a "One Size Fits All" approach (Calvaresi et al., 2014; Hayhurst, 2018; Meiland et al., 2014). The involvement of stakeholders, including AAL inhabitants, caregivers, and healthcare professionals, is important for identifying unique needs and challenges within the AAL environment, thereby optimizing the usability, effectiveness, and acceptance of AR applications (Calvaresi et al., 2017; Jones, 2018; Mansson et al., 2020). Stakeholders contribute valuable insights regarding daily routines and challenges (Jones, 2018). Incorporating their perspectives, feedback, and ideas ensures that the design and functionality of AR applications meet the specific needs and preferences of AAL inhabitants, facilitating independent living and enhancing well-being (Bhalla, 2014; Fuglerud & Sloan, 2013; Schnall et al., 2016).

AR technology has the potential to offer meaningful support in areas such as health monitoring, medication reminders, home automation, social connectivity, and personalized assistance (Alabood & Maurer, 2022; Al-Shaqi et al., 2016; De Belen et al., 2019; Ghorbani et al., 2019). However, without meaningful stakeholder involvement, general AR applications may inadvertently introduce barriers or challenges. Features, interfaces, or interactions that do not consider the abilities, preferences, or routines of AAL inhabitants could impede the successful adoption and utilization of AR technology. Moreover, including people with diverse abilities and disabilities in the development process can benefit all users by highlighting new requirements and opportunities for improvement (Persson et al., 2015).

Although initial steps have been taken to engage informal care partners in the design process for AAL solutions (Hwang et al., 2015), such an approach remains underexplored in the context of AR. Furthermore, previous efforts have often focused on specific stakeholder groups, such as the elderly or caregivers, without fostering collaborative co-creation processes among different parties (Calvaresi et al., 2017; Hwang et al., 2015; Sandoval & Favela, 2017). Nonetheless, there is an emerging trend towards involving more diversified stakeholder groups in the design process, often including both elderly individuals and their caregivers or family members (De Podestá Gaspar et al., 2018).

This chapter seeks to address the research gap regarding the co-creation process involving stakeholders for the ideation and design of AR-based applications in the living environment. To achieve this objective, the paper aims to address the following **RQ 4**: *What are possible use cases and requirements for AR-based applications that are accessible and inclusive for people with different abilities and disabilities?*

To do so, the main objective is to conduct focus groups with stakeholders like AAL inhabitants, caregivers, family members, or healthcare professionals. Thereby, the goal is to engage the stakeholders in the ideation process of further AR application developments relying on the Inclusive Design approach to develop use cases that consider the needs and capabilities of all users.

9.2 Background

9.2.1 Ambient Assisted Living (AAL)

AAL is an emerging field at the intersection of healthcare, technology, and environment design (Calvaresi et al., 2017). It aims to leverage advanced technologies, such as sensors, AI, IoT, and communication systems, to create intelligent living environments that enable individuals to maintain their independence, improve their quality of life, and reduce the burden on caregivers and healthcare systems (Calvaresi et al., 2017; Lloret et al., 2015; Marques, 2019). Advancements in healthcare have led to longer life expectancies, resulting in a larger proportion of elderly individuals in many countries (Nations, 2019). This demographic trend presents unique challenges for healthcare systems worldwide (Harper, 2014). Due to the growing preference among older adults to age in their own homes and communities, there is an increasing demand for innovative solutions that can support independent living, contributing to the future relevance of AAL (Mulliner et al., 2020). However, the concept of AAL is not limited to the elderly but is also aimed at people with different types of disabilities to assist them in living independently in their private homes or care facilities (Geman et al., 2015). Accordingly, the specific possibilities and assistance provided in the context of AAL vary, depending on the individual situation, target group, and the stakeholders involved (Blackman et al., 2016). Morita et al. (2018) developed an AAL system for activity recognition for the elderly based on which daily reports are generated automatically, thus facilitating the daily work of the caregivers in a nursing home. Similar systems to track activities and abnormal behaviors of disabled or elderly people and alert caregivers to abnormalities have been implemented in different contexts and variations (Bleda et al., 2018; Cebanov et al., 2019; Zdravevski et al., 2017). Accordingly, monitoring Activities of Daily Living (ADL) with the aim of tracking health status, identifying risks related to aging, physical impairments, and living independently as well as support for everyday tasks is one of the main goals of AAL environments (Calvaresi et al., 2017; Jaschinski & Allouch, 2014). The domain ADL further incorporates notifications and reminders to the user to perform tasks such as taking required medicine at the correct time, thus strengthening their autonomy (Giménez Manuel et al., 2022). Another area of application for AAL is indoor navigation, which is intended to enable people with Alzheimer's disease or impaired vision to navigate independently in their dwelling or care facility (Alabood & Maurer, 2022; De Belen et al., 2019; Tsirmpas et al., 2015). Considering that social connectedness was identified as a significant aspect for a high quality of life and successful aging, there are various AAL environments that target this aspect in particular (Blackman et al., 2016; Bouma et al., 2007; Gabriel & Bowling, 2004). The "Building Bridges" project, for example, explores how technology can support senior citizens in staying socially connected (Wherton & Prendergast, 2009), while "Digital Family Portrait Display" aims to promote peace of mind for other family members by raising awareness of the daily activities of older adults (Rowan & Mynatt, 2005).

Just as the applications in the AAL environment vary, so do the technologies used (Cicirelli et al., 2021). The increasing proliferation of IoT devices, for example, offers great potential for smart AAL environments, especially in the area of activity recognition and ADL (Maskeliūnas et al., 2019). Portable sensors or wearables enable fall detection or health monitoring and RFID technology is used, for example, to annotate objects or for 3D localization (Cicirelli et al., 2021; Rashidi & Mihailidis, 2012). The realm of AAL is a rapidly evolving domain, with emerging technologies such as blockchain technology or 5G being continuously incorporated to enable new avenues or enhance existing ones (Florea et al., 2022; Hermens, 2016). In summary, ongoing research and innovation in AAL is centered on developing advanced sensing technologies, improving data analysis capabilities, and enhancing user interfaces to create intuitive and user-friendly AAL solutions (Cicirelli et al., 2021). However, the successful implementation of AAL systems also requires interdisciplinary collaboration between researchers, clinicians, engineers, designers, policymakers, and end users (De Podestá Gaspar et al., 2018). Especially with respect to the use of AR methods in the AAL context, this represents an important prospective challenge, as current solutions mostly adopt an expert-based, technically oriented perspective (Alabood & Maurer, 2022; Ghorbani et al., 2019; Hwang et al., 2015).

9.2.2 Universal Design, Inclusive Design, and Accessibility

The term Universal Design (UD), or "design for all," encapsulates a deliberate effort to inclusively consider the widest range of end-user requirements throughout the lifecycle of product or service development, eliminating the need for subsequent design modifications (Mace, 1997; Steinfeld & Maisel, 2012; Stephanidis, 2001). UD is guided by principles including equitable use, flexibility in use, simple and intuitive use, perceptible information, tolerance for error, low physical effort, and size and space for approach and use (Burgstahler, 2009; Story & Mueller, 2001). Its aim is to cater to the needs of as broad a user base as possible through a one-size-fits-all strategy, acknowledging the diversity in abilities, ages, and characteristics within the general population (Keates et al., 2000; Steinfeld & Maisel, 2012).

In contrast, Inclusive Design represents a more recent and evolving concept that adopts a wider perspective (Center, 2017). It acknowledges the diverse needs, preferences, and backgrounds of individuals, striving to create products and environments that celebrate diversity and ensure equitable access and participation for everyone (Holmes, 2020). This design philosophy emphasizes engaging diverse user groups in the design process and prioritizes user-centeredness, collaboration, and co-creation. Unlike UD, which originated in architecture, Inclusive Design emerged in the mid-1990s from a blend of initiatives, experiments, and insights across various disciplines (Clarkson & Coleman, 2015; Goldsmith, 2000). Microsoft has notably adopted an Inclusive Design approach as a framework for designing digital products in recent years (Microsoft Design, 2016).

Both UD and Inclusive Design aim to create accessible and usable products, environments, and systems for individuals with diverse abilities. Accessibility, a core concept for both philosophies, refers to the extent to which a product, environment, or system is accessible and usable by individuals with disabilities or limitations (Holmes, 2020; Iwarsson & Ståhl, 2003; Persson et al., 2015). It involves eliminating barriers and providing equal opportunities for participation in various activities by addressing physical, sensory, cognitive, and technological aspects of design (Iwarsson & Ståhl, 2003; Persson et al., 2015; Sauer et al., 2020).

A key distinction between these concepts lies in their principles, scope, and design methodologies. Accessibility primarily focuses on adhering to specific standards and guidelines to ensure equal access for individuals with disabilities, often through a compliance-driven approach that adds specific features or accommodations to existing designs (Iwarsson & Ståhl, 2003; Regan, 2004). UD employs a proactive strategy by incorporating inclusive features from the beginning of the design process (Burgstahler, 2009). Inclusive Design advances this by actively involving diverse user groups and stakeholders in the design process, embracing their unique perspectives and requirements (Moon et al., 2019). It emphasizes co-design, collaboration, and user-centeredness, leading to more personalized and adaptable solutions (Swan et al., 2022).

In summary, while accessibility, UD, and Inclusive Design all strive towards creating inclusive and accessible environments, they differ in their foundational principles, scopes, and methodologies (Subasi et al., 2009). Accessibility aims to remove barriers for individuals with disabilities. UD seeks to meet the needs of a wide user base through proactive design strategies but typically results in a single solution intended to accommodate as many users as possible (Keates et al., 2000; Schulz et al., 2014). In contrast, Inclusive Design extends beyond accessibility and universal usability by considering social, cultural, and contextual factors and involving diverse user groups in the design process (Holmes, 2020; Keates & Clarkson, 2003). This approach may yield varied design solutions tailored to different users to ensure no one is excluded from using a product or system (Keates et al., 2000).

9.3 Method

This study adopted a qualitative research methodology, leveraging a focus group approach to elicit insights from AAL stakeholders regarding the daily challenges faced by individuals and potential applications where AR could offer support. Qualitative research facilitates a profound comprehension of participants' viewpoints, yielding rich, context-specific data. The focus group method was selected over expert interviews to foster interactive discussions among participants, allowing them to expand on each other's contributions and offer varied perspectives. The objective is to explore individuals' perceptions of technology within their personal and communal lives and their attitudes towards the challenges they encounter. While there are various methods to conduct focus groups, there are no universal guidelines, as the approach is highly dependent on the topic at hand. For this study, the guidelines and insights provided by Krueger (2014) and Schulz et al. (2012) were primarily consulted.

The execution of the focus group and subsequent data handling procedures received approval from the data protection unit at the Karlsruhe Institute of Technology (KIT). This included the anonymization of all personal data to prevent the identification of individuals or institutions.

9.3.1 Recruitment and Participants

Upon determining that focus groups were an apt method to explore the Research Question (RQ), we identified relevant stakeholders to contribute valuable insights, ideas, and perspectives on the topic. Key participants included individuals with temporary or permanent disabilities and older adults, who face daily challenges and could offer innovative ideas and suggestions for overcoming these obstacles. Advocacy services for people with disabilities were also considered for their broad interaction with a diverse group of affected individuals, potentially providing a more holistic view. Medical staff and caregivers, whether in individual or institutional care settings, were identified as relevant due to their daily encounters with various needs and care recipients. Architects specializing in accessible design were deemed important for their insights into the physical and architectural challenges and opportunities. To ensure a diverse range of perspectives, experiences, and expertise among the participants, purposive sampling was employed. Participants were selected based on inclusion criteria of having a minimum of 1 year of experience working with people with disabilities or elderly individuals and diversity in professional backgrounds to capture a comprehensive range of insights on potential use cases, requirements, and challenges for AR-based assistive solutions.

Potential participants were individually contacted via email or phone, receiving a clear and concise explanation of the study's purpose, procedures, and potential benefits. They were invited to participate voluntarily in the focus group discussions. Literature suggests that focus groups should consist of 6 to 12 participants to balance dynamics and manageability (Powell & Single, 1996; Wong, 2008). Groups smaller than this may lack dynamism, while larger groups could lead to management difficulties and potential fragmentation into smaller discussions (Krueger, 2014). Consequently, we organized the focus group sessions into two categories: one for directly affected individuals and another for different stakeholders.

Over 60 experts from various stakeholder groups were invited to the focus group discussions. One-third of these experts initially expressed interest in participating.

However, due to scheduling conflicts, only five participants were available for one of the proposed time slots. Two last-minute cancellations further reduced attendance to three participants. Despite the low turnout, the session proceeded as planned, yielding valuable insights and perspectives on the research topic.

A second attempt to conduct a focus group session with the same interested participants was made. However, this session received no responses, likely due to its scheduling close to the Christmas holiday season at the end of November 2023, and it was not conducted. Additionally, the focus group intended for directly affected individuals was postponed multiple times and ultimately canceled due to time constraints within the dissertation project.

9.3.2 Design

A semi-structured focus group guide was developed, consisting of open-ended questions that covered various aspects of the day-to-day challenges of elderly people and people with disabilities, assistive solutions, and prospects of AR methods for assistance (see Appendix B.1). Based on the research objective, the following general topics were identified:

- Assistive solutions
  - Challenges in everyday life
  - Opportunities for assistive solutions
- AR for assistive solutions
  - AR for challenge management
  - AR design and implementation

The focus group discussion was divided into two thematic blocks to gather comprehensive and unbiased insights from participants. This division aimed to facilitate a progressive exploration of the subject matter, minimizing potential bias associated with premature introduction of AR technology. In the first block, participants engaged in open dialogue about the daily challenges faced by individuals requiring assistive solutions, intentionally omitting AR technology to encourage sharing of experiences without influence from specific technologies.

Upon completing the initial block, the session introduced AR technology, its applications, and potential future developments, providing participants with a concise overview. This introduction prepared the ground for a focused discussion on utilizing AR to address the identified challenges and opportunities. The use of open-ended questions in this study serves to stimulate discussion among participants when necessary. These questions are designed to elicit thoughtful and unrestricted responses, allowing natural conversation flow without constraining participants' thoughts or perspectives and minimizing direct prompts from the moderator.

9.3.3 Data Collection

To ensure the reliability and validity of the data, focus group discussions were audio-recorded with the consent of the participants. High-quality recordings for post-processing and redundancy were achieved using a Logitech Rally conference system microphone and a Tascam DR100 MK III voice recorder. The focus group session lasted for approximately 120 minutes, providing sufficient time for comprehensive discussions. Two researchers attended the session, fulfilling the roles of moderator and note-taker. The moderator was responsible for facilitating the session and promoting discussion among the participants, while the note-taker documented non-verbal cues, group dynamics, and any additional contextual information pertinent to the analysis.

The discussions took place in a neutral, well-lit room. The session commenced with an introduction to the research team, an overview of the study's objectives, and the establishment of discussion ground rules (see the interview guide in Appendix B.1). Participants were then invited to introduce themselves, their institutions, and briefly describe their job responsibilities. Following introductions, the moderator navigated the session through thematic blocks, employing open-ended questions to foster discussion and encourage participants to share their experiences, insights, and ideas. At the session's conclusion, the note-taker summarized the key themes identified during the discussions. This summary aimed to accurately reflect the participants' ideas, perspectives, and experiences, ensuring their contributions were faithfully represented. Participants were given the opportunity to provide feedback on this summary, allowing for validation of the findings and correction of any misunderstandings that may have occurred during data collection. Subsequently, participants were asked to offer broader feedback on the study, including their impressions of its design, methodology, and execution. They were also encouraged to discuss the study's potential implications and practical applications.

Table 9.1.: Participant Information

ID	Gender	Institution	Institution Task	Years of Experience	Stakeholder Group
1	Male	Residential counseling for people with disabilities	Development of residential solutions for affected individuals according to their interests	2,5	Advocacy Services
2	Female	Autism accompaniment	Curative, individual day care for people with autism	10	Medical Staff & Caregivers
3	Male	Regulatory advocacy for people with disabilities	Advisory services for policy makers on the implementation of inclusion	7 (43 in related fields)	Government Representatives

9.3.4 Data Analysis

Qualitative data from the recordings were transcribed verbatim using the AI-based open source Automatic Speech Recognition (ASR) system "Whisper" developed by OpenAI (2022, September 16/2023), with the large language model option selected. The transcripts from the focus group were then analyzed qualitatively using established inductive coding techniques (Elo & Kyngäs, 2008; Onwuegbuzie et al., 2009). This analysis facilitated the identification of recurring themes, patterns, and unique insights from the discussion.

The analysis began with open coding, which involved a detailed examination of the transcript, annotation, and categorization. These initial categories were documented on coding sheets. During axial coding, these categories were organized under higher-order headings to construct a comprehensive framework (Williams & Moser, 2019). The process culminated in the abstraction of categories to articulate a detailed description of the research topics. Similar subcategories were merged to form broader categories, which were then further organized into main categories. This hierarchical structure enabled a structured and coherent representation of the data, enhancing the understanding of the emergent patterns and insights. All coding procedures were conducted using the software "f4analyse" provided by dr. dresing & pehl GmbH (n.d.).

9.4 Results

The focus group consisted of three participants: one female and two males. Table 9.1 presents the details provided by the participants regarding their personal information, affiliated institution, and tasks they are involved in. Furthermore, the transcript was translated from German to English to enable quotations within this document.

Table 9.2 presents the topics identified from the discussion, along with their corresponding codes. The subsequent section delves into these key topics, concluding with an analysis of the overarching challenges, requirements, and recommendations that surfaced during the participants' dialogue.

9.4.1 Routines and Exceptions

Participants highlighted the significance of routines in enhancing structure, efficiency, and independence in the lives of people with disabilities. However, these individuals often encounter difficulties when managing multiple tasks. For example, planning a grocery shopping trip involves several sequential steps that require careful consideration, such as determining the needed items based on dietary preferences and household necessities, compiling a shopping list to ensure no essential item is missed, and managing finances to ensure sufficient funds are available for purchases. These tasks, while seemingly straightforward for healthy adults, present multifaceted challenges for people with disabilities.

Assistive systems are viewed as valuable tools in helping individuals maintain their routines by providing timely reminders, prompts, and step-by-step guidance. The use of assistive technologies can enhance a person's independence and autonomy by enabling them to perform tasks independently, reducing the need for constant external assistance. This not only leads to increased security and long-term behavioral changes but also facilitates learning processes. A well-learned routine can be remembered and executed even after many years. Moreover, assistive systems can alleviate the workload of caregiving staff by taking over some tasks, allowing caregivers to focus more on personal care and individualized attention.

Participants also discussed the challenges posed by exceptions or unexpected events in routines. For instance, minor changes in train timetables or cashless payment terminals can significantly disrupt the lives of people with disabilities. Assistive technologies need to be capable of effectively managing these exceptional circumstances by providing decision support, suggesting alternative approaches, or recommending actions to mitigate disruptions.

The discussion further revealed that beyond handling exceptions, individuals with disabilities may face challenges such as a lack of motivation, which impedes their productivity. Simple reminders are insufficient in addressing this issue, indicating the need for assistive solutions that can effectively motivate users.

Table 9.2.: Coding of the Transcript

Code	Subcode	Description
T1_Routines_Exceptions		Support for routines and exception handling
	T1a_Routines_Except	Issues with routines and dealing with exceptions
	T1b_Flexibility_Individ	Need for highly flexible individualized systems
	T1c_Need_Valid	Continous assesment of abilities needs
	T1d_Action_Recomm	Recommendations for actions
	T1e_Adap_Learn	Need for adaptive systems for routines/exceptions
	T1f_Smart_Home	Incorporation of Smart Home
T2_Read_Writ_Arith_Mon		Support for reading, writing, arithmetic handling money
	T2a_Calculation_Reading	Issues with calculations, monetary values, reading
	T2b_Adap_Learn	Need for adaptive systems for reading/calculations
	T2c_Flexibility_Individ	Need for highly flexible individualized systems
T3_Lang_Comm_Social		Support for communication and social interactions
	T3a_Comm_Barriers	Issue of communication barriers
	T3b_Social_Interac	Issues in social interactions
T4_GEN_Requirements		General requirements
	T4_GEN_Need_Valid	Continous assesment of abilities needs
	T4_GEN_Complex_Instruc	Reduced complexity of instructions
	T4_GEN_Flexibility_Individ	Need for highly flexible individualized systems
T4_GEN_Challenges		General challenges
	T4_GEN_Financial	Financial challenges
	T4_GEN_Complex_Instruc	Issues related to complex instructions
	T4_GEN_Maint_Quality	Issues related to product quality maintainance
	T4_GEN_Data_Priv	Issues related to data privacy
T4_GEN_Recommendations		General recommendations
	T4_GEN_Prof	Learning from professional contexts
	T4_GEN_Dev_All	Developing for broad population
	T4_GEN_Focus_Topics	Identifying key issues specific target groups
	T4_GEN_Inc_Affected	Including people with disabilities in focus groups

Developing assistive systems for managing routines and exceptions involves addressing the diverse needs and capabilities of individuals with disabilities. Customization and flexibility are important, as what works for one person may not be suitable for another. An overly rigid or generic system could undermine independence and self-reliance by taking over tasks that individuals are capable of performing themselves. Finding the right balance between adapting the person to the assistive system and designing the system to accommodate the individual's unique needs and capabilities is essential.

9.4.2 Reading, Writing, Arithmetic & Handling Money

During discussions, it became clear that a significant number of individuals with disabilities encounter obstacles in performing basic daily activities. These include tasks such as reading, writing, arithmetic, and financial management.

The nature of these difficulties varies widely, complicating these everyday tasks for affected individuals. For instance, reading may be challenging due to visual impairments or learning disabilities, while writing difficulties may arise from motor skill limitations or cognitive impairments. Similarly, arithmetic and financial management tasks can be daunting for those with numerical challenges or difficulties in processing abstract concepts. Participants identified a significant potential for assistive solutions in addressing these challenges. One such solution is the use of automatic text recognition technology, which can transform written text into auditory formats, thus facilitating access to information for those with reading difficulties. Moreover, systems that represent monetary values visually in an intuitive manner were discussed as beneficial for helping individuals manage their finances independently. Additionally, the potential of these assistive technologies to support people with disabilities during shopping was explored. For example, in a supermarket, such systems could assist in verifying change by visually comparing the prices of items purchased, the amount tendered, and the change received, alerting the user to any discrepancies to ensure they receive the correct amount.

Participants expressed a desire for assistive systems that are adaptive and capable of learning, providing personalized support based on the unique usage patterns of individuals with disabilities. The importance of creating assistive technologies that can evolve with the user's changing needs and preferences was emphasized. Moreover, the significance of incorporating a learning component within these systems was highlighted. A learning-capable assistive system can enhance its performance over time by analyzing feedback and outcomes from previous interactions. This feedback loop allows the system to adapt and refine its responses, delivering more precise and effective support.

9.4.3 Language, Communication & Social Interactions

The discussion illuminated an aspect of social interactions, particularly the challenges arising from diverse languages and communication styles. It was noted that individuals often have their unique way of communicating, which can lead to tensions during interactions. People with autism were identified as facing particular challenges in language comprehension.

The primary issue is not the languages per se but the multitude of communication channels and the nuanced use of words, as described by the four-sides model by Schulz von Thun et al. (1981). For those with autism, navigating these complex communication aspects can lead to irritations and misunderstandings during social interactions. The use of idioms, figurative language, and sarcasm, in particular, can be confusing for individuals with autism, resulting in potential misunderstandings or feelings of being overwhelmed. Participants discussed initial solutions, such as Metacom symbols (Kitzinger, 2022), which offer a stylistic means to convey meanings more effectively than words alone. These symbols are especially useful in representing abstract concepts like love or peace, which are difficult to articulate through verbal communication alone. The importance of simple language was also emphasized, as it enhances text comprehension for many individuals with disabilities. While some websites provide an option to display content in simple language, the manual implementation of this feature is limited in scope. To improve accessibility further, there was a call for a technical solution capable of automatically generating content in an accessible format. Such a solution could revolutionize information accessibility for people with various abilities, enabling them to independently access a wider array of content. The ultimate goal, as expressed by participants, is to use these tools to streamline and improve social interactions for individuals with disabilities. Furthermore, it was acknowledged that engaging with new assistive systems and learning their use offers substantial benefits to many people with disabilities.

9.4.4 General Challenges, Requirements & Recommendations

This section delves into prevalent challenges and requirements associated with assistive systems. Challenges are defined as obstacles or difficulties that need to be addressed to achieve a certain goal, representing barriers that could hinder the development of assistive solutions. Requirements, conversely, denote the essential needs, criteria, or conditions necessary for realizing the objectives of a project or solution, encapsulating the critical features, functionalities, or characteristics vital for the successful deployment of assistive solutions. It is important to note that categorizing aspects strictly as either challenges or requirements is challenging and inherently subjective, indicating that some elements may be interpreted as both. Additionally, this section summarizes general recommendations and insights from participants, which transcend specific domains or use cases.

Challenges & Requirements

During the discussion, participants identified several challenges and requirements for the development of assistive systems, which are applicable across various application areas. A key requirement is the development of personalized solutions that accurately meet the unique needs and abilities of individuals with disabilities. This necessitates a deep understanding of each person's specific requirements, which may change over time, thus adding complexity to the design process due to the need for continuous reassessment and adaptation. Another significant challenge concerns the introduction and instruction for new assistive systems. It was emphasized that the learning process must be user-friendly, as navigating new, technologically advanced systems can be overwhelming for many users, particularly for those with disabilities who may struggle with online manuals. Ensuring the usability and accessibility of instructional materials is essential in overcoming this barrier. Furthermore, individuals with disabilities may face difficulties in learning and adapting to new situations or systems, which can hinder the adoption of assistive technologies due to the increased effort and time needed for familiarization. The physical characteristics of assistive devices were also discussed. Participants highlighted the need for stability and durability in these devices, noting that they are more prone to damage and wear when used by people with disabilities. This underscores the importance of considering product quality and resilience in the development of assistive systems to ensure their longevity. Despite the need for increased durability, proper maintenance and support are also important, often relying on staff or caregivers to keep the systems operational. Simplicity and ease of use were identified as critical factors in making assistive systems user-friendly for people with disabilities, as complex interfaces or functionalities may deter users.

The cost of assistive devices presents a significant challenge. Specialized devices designed for people with disabilities are often much more expensive than generic products. This cost disparity creates financial barriers to accessing these technologies. Additionally, even commercially available AR devices can be costly, posing an obstacle for individuals and organizations looking to integrate these tools into their practices. This is compounded by the costs associated with introducing and maintaining the devices. Data protection and privacy emerged as pressing challenges in the context of assistive solutions. Supporting individuals effectively often requires access to extensive personal data, including sensitive information. Ensuring robust data protection measures is paramount to protecting user privacy and building trust in assistive technologies.

Recommendations

Focus group participants discussed several points that were not easily categorized. One significant observation was that systems or technologies not initially designed as assistive tools are often highly accepted by people with disabilities. Specifically, technologies commonly used by a large part of society, without being explicitly labeled as assistive technologies, were found to provide valuable support for individuals with disabilities. Voice messages were cited as an example of such a technology.

Although voice messages were not developed with the primary intention of assisting individuals with disabilities, they have been immensely beneficial in facilitating communication for this group. The convenience of voice messages has made them accessible and advantageous for people facing barriers to traditional written communication. Participants noted that when a technology is widely adopted by the general population, it is more likely to be embraced and utilized by individuals with disabilities, allowing them to participate in mainstream communication practices and routines. Thus, mainstream technologies can inadvertently become assistive tools, enhancing accessibility and inclusivity for a wider range of users. Furthermore, the discussion highlighted the potential of assistive solutions for people with disabilities in professional contexts. Participants recognized that the structured and organized nature of the professional environment is conducive to seamlessly incorporating assistive technologies. In the workplace, tasks often adhere to a predetermined structure and guidelines, which can facilitate the effective implementation of assistive solutions. While individual accommodations are necessary to address diverse abilities, the primary objective remains to complete work in a specific manner to maintain productivity and efficiency.

The participants also acknowledged the challenges in creating assistive systems that cater to a broad and diverse group of people with disabilities due to the high level of individualization required. As a practical approach, they suggested focusing on key issues or specific groups with shared challenges, such as individuals with learning disabilities, and developing targeted assistive solutions tailored to their needs. By concentrating on particular segments of the disability community, developers can gain a deeper understanding of the unique obstacles and requirements faced by those individuals. This focused approach enables more effective problem-solving and customization, resulting in solutions that more comprehensively address the specific needs of the identified group.

9.5 Discussion

9.5.1 Principal Findings

Participants in our focus group were receptive to adopting new technological assistive solutions, recognizing their potential to enhance the lives of individuals with disabilities. They emphasized the significance of continuous technological and research advancements for developing more effective and inclusive assistive solutions. Specific use cases discussed highlighted the profound impact assistive technologies could have on enhancing daily experiences and promoting independence for individuals with disabilities. Understanding the complexities of everyday tasks is essential for designing effective assistive solutions that empower individuals and enhance their independence. The potential of future AR devices, equipped with sensors and cameras, lies in their ability to understand the context of tasks and provide tailored support. Combined with AI, these devices can adapt to users' needs, offering personalized assistance in real-time through visual cues and step-by-step guidance.

In home environments, AR devices integrated into Smart Home systems could offer comprehensive solutions for managing daily routines and tasks. Smart Home sensors, in conjunction with AR device sensors, can accurately understand the user's environment and provide adequate support, such as reminders about appliances left on or instructions for replacing smoke detector batteries. Assistive systems should inspire and incentivize users to stay engaged and accomplish their tasks, going beyond merely addressing exceptions and disruptions. Adaptive methods that learn routines and respond to exceptions could introduce a multidimensional paradigm in adaptive routine management with AR, facilitating the understanding of routines and analyzing deviations. This dynamic approach ensures individuals are equipped to navigate unexpected disruptions effectively.

The incorporation of AR has the potential to impact the motivational landscape for people with disabilities significantly. Visual cues embedded within the immediate environment can serve as a more potent motivational mechanism than standard audio prompts, fostering a stronger commitment to routines or tasks. Support in reading, arithmetic, and money management is important, requiring visual processes due to the nature of these tasks. In communication and social interaction, support focuses on automatic translations into simplified language or symbols to enhance accessibility. Overlaying translated text onto real-world objects or scenarios enables individuals to engage more effectively with their surroundings. However, there are fundamental challenges and essential requirements across all application areas for the successful use of AR-based assistive solutions. These include financing, maintenance, device quality, data protection, and addressing individual requirements and diverse target groups. Ensuring robust data protection measures is crucial, with a recommendation for Smart Home solutions that process sensitive data locally.

The participants' statements underscore the importance of a human-centered design approach in developing AR-based assistive solutions for enhancing autonomy and independence in the field of AAL. While catering to all individual needs may be unrealistic, leveraging widely used technologies offers an opportunity to prioritize accessibility, data protection, and straightforward explanations. This approach aligns with UD principles but may not comprehensively address disability-specific issues.

Alternatively, designing AR-based assistive solutions tailored to specific target groups considers their unique requirements and abilities, aligning with Inclusive Design principles. These two approaches are complementary, suggesting the development of universally accessible AR systems while also collaborating with people with disabilities on specific application areas. It is important not to group people with disabilities and older people together due to distinct requirements. Further dividing target groups according to specific impairments can lead to more effective support solutions tailored to each group following Inclusive Design principles.

Despite Smart Home technologies being relatively uncommon among individuals with disabilities due to financial and maintenance barriers, combining these technologies with AR-based assistive solutions holds potential. Accessibility should be considered from the outset in developing AR methods for the Smart Home context, ensuring usability by people with disabilities as Smart Home devices become more prevalent in care facilities.

9.5.2 Limitations

Focus groups, as a qualitative research method, provide insights into participants' perspectives and experiences, yet their findings are constrained by several limitations. The results and their prioritization are specific to the participants involved and cannot be generalized to a larger population due to the small sample size and non-random selection of participants, limiting the broader applicability of the findings. Additionally, the dynamics within the focus group could influence participants' responses, potentially introducing bias. In this particular focus group, the absence of interdependencies among participants helped to minimize this risk.

The most significant limitation of this study is the small number of focus group participants, affected by spontaneous cancellations. To enhance the validity and reliability of the research, conducting two or three focus group sessions is recommended rather than relying solely on one. This approach would provide a more comprehensive understanding and deeper insights into assistive solutions for people with disabilities. Furthermore, the subjective nature of participants' statements and the subjectivity involved in coding and processing discussion results mean that the analysis's topic hierarchies and groupings should be viewed as suggestive rather than definitive. Different perspectives or approaches could lead to alternative interpretations and categorizations of the data, highlighting the inherent interpretative and subjective nature of qualitative research. Researchers should therefore be cautious in drawing conclusions based solely on one perspective.

It is also important to recognize that the focus groups may not have included all relevant stakeholders, potentially omitting valuable insights. Notably, a significant stakeholder group—people with disabilities or older individuals—was not included in this focus group, representing a significant limitation also acknowledged by the participants. This omission underscores the importance of interpreting the findings with an awareness of this limitation.

9.5.3 Future Research

The focus group highlights several avenues for future research in the development of AR-based assistive solutions for individuals with disabilities or the elderly, as well as in the creation of general AR applications that incorporate UD principles.

Conducting further focus group studies and engaging a wider array of stakeholders would significantly contribute to improving both the volume and the quality of the insights gathered. Incorporating perspectives not represented in the initial focus group would lead to a more comprehensive understanding of the topic. Organizing distinct, homogeneous focus groups with directly affected individuals, such as people with disabilities or the elderly, could prove particularly beneficial. This method would facilitate a more detailed examination of the specific needs and capabilities of each target group, in line with Inclusive Design principles. Given the growing ubiquity of smart devices, engaging with these stakeholders to discuss their unique requirements presents a valuable opportunity. Additionally, future research could leverage the initial findings to investigate the technical viability of various proposed ideas and application domains. Such efforts would advance the development of AR-based assistive solutions.

9.6 Conclusion

In conclusion, the findings from the focus group underscore the potential advantages of AR-based assistive solutions for individuals with disabilities. The discussions identified specific areas where AR technologies offer promising use cases not addressed by existing technologies, illustrating the significant impact AR could have on enhancing the daily lives of people with disabilities. However, it was also noted that stakeholders showed reservations in certain domains, indicating the necessity for further research and refinement of assistive solutions in these areas. A critical insight from the focus group was the recognition of the need for assistive technologies to be highly personalized. Participants emphasized the importance of selecting specific disabilities and target groups to develop tailored solutions that meet their unique needs without providing excessive or insufficient support. This personalized approach is especially vital in everyday assistance, as different disabilities require distinct functionalities and features to offer meaningful support and promote independence.

Furthermore, the discussions revealed an interesting interplay between UD and Inclusive Design principles. Depending on the goals and contexts of assistive solutions, participants showed a preference for either a UD approach, which aims to create solutions accessible to a wide range of users, or an Inclusive Design strategy, which focuses on customizing systems for specific target groups. This highlights the need for flexibility and adaptability in the development of assistive technologies to meet the varied and changing needs of individuals with disabilities. In summary, the insights from the focus group provide valuable knowledge for advancing AR-based assistive solutions. The study emphasizes the importance of continued research, involving a diverse range of stakeholders and conducting more focused and homogeneous focus groups with actual beneficiaries, such as people with disabilities and older adults. Such efforts are essential for further refinement and exploration of technical possibilities, moving closer to the successful implementation of AR-based assistive technologies that empower individuals with disabilities and improve their quality of life.

10 Augmented Reality-based Indoor Positioning for Smart Home Automations

This chapter comprises the version of record of the following paper: Schenkluhn, M., Peukert, C., & Weinhardt, C. (2023b). Augmented Reality-based Indoor Positioning for Smart Home Automations. *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–6.

Changes include formatting, numbering of the research question and chapters, minor changes for consistency, and correction of spelling errors.

10.1 Introduction

Smart Home systems are gaining popularity and it becomes easier to integrate solutions from various vendors thanks to new open standards such as Matter (Purdy, 2022; Statista, 2023). However, Smart Home systems equipped with smart lights, smart thermostats, or smart speakers require meaningful data input to enable smart automations. Without linking actors to sensor data, these systems remain anything but "smart," and users are forced to detour through opening vendor-specific smartphone apps to perform even simple tasks. One promising source to trigger automations is data about a user's location. The room temperature, for instance, can be automatically lowered to save energy when all residents leave the house. A system that is aware of its users' indoor positions can provide even more automated features regarding, e.g., user comfort, safety, or security. Thereby, these automations may not only increase convenience for healthy residents at home, but also allow senior citizens and people with disabilities to regain autonomy and improve their quality of life.

Current Smart Home systems usually solve the task of detecting users with the help of Passive Infrared (PIR) motion sensors. Although being affordable, PIR sensors suffer from several drawbacks (Kemper & Linde, 2008). Besides PIR sensors, more advanced so-called Indoor Positioning Systems (IPSs) are available, but they are mostly intended for professional applications. Still, most IPS need – depending on room size and required resolution – one or more sensors and/or receivers per room, which often leads to expensive solutions with the need for professional installation and setup.

In this paper, we shed light on a future scenario by looking at the positioning information obtained through an AR HMDs with tracking capabilities across Six Degrees of Freedom (6DoF). While such devices are already available and can be used for research today, they are still bulky, expensive, and not suitable for everyday use. Several companies such as Apple (Huddleston Jr., 2022), Meta (Labs, 2021), and Bosch (Sensortec, 2023) follow and invest into the vision of "Ubiquitous Augmented Reality" (Newman et al., 2007). Therefore, we make the assumption that unobtrusive AR HMDs will be available in the near future for this scenario and that users will wear these HMDs all-day as a potential replacement for smartphones and smartwatches. 6DoF AR HMDs are equipped with the necessary sensors to accurately position themselves within a house or an apartment to reliably anchor virtual objects in a room. Prior research has used this ability for instance for indoor navigation in public buildings (Liu et al., 2016). Hence, in this paper, we want to answer the following **RQ 5**: *How can sensor technology of AR glasses be leveraged for precise indoor user positioning, and what implications does this have for Smart Home automations?*

We suggest applying AR self-positioning to the Smart Home domain and its automations. Thus, reliable user tracking in a Smart Home environment without the need to install and setup dedicated external devices is available as soon as user-friendly, i.e., lightweight, comfortable, and affordable AR HMDs become feasible.

10.2 Related Work

Smart Home systems have extensively been researched in the past years (Marikyan et al., 2019). In general, there are a wide variety of application areas within the Smart Home context, such as safety, sustainability, security, comfort, or entertainment

(Marikyan et al., 2019). Especially their application for senior citizens or people with disabilities as AAL systems has been of high interest due to the aging population and rising costs of healthcare (Rashidi & Mihailidis, 2012). AAL provides assistive tools such as health and activity monitoring or fall and wandering prevention (Rashidi & Mihailidis, 2012).

10.2.1 Indoor Positioning Systems in Smart Homes

Many Smart Home applications require or can be substantially improved by referring to the residents' position. For example, Smart Home systems detecting room occupancy can control lighting or Heating, Ventilation, and Air Conditioning (HVAC) more energy-efficiently (Alam et al., 2020; Zou et al., 2018). Additionally, alarm systems can differentiate between residents and intruders (Jose & Malekian, 2017). In the context of AAL, activities of daily living can be monitored (Labonnote & Høyland, 2017; Stavropoulos et al., 2020; Zou et al., 2018), abnormalities detected in combination with vital sensors (Mshali et al., 2018; Singh et al., 2020), and care givers can be relieved for example by notifying them only for tasks where help is required (Aloulou et al., 2013). In general, this means that more affected people can remain living at home, which can increase the quality of life (Rialle et al., 2002). Depending on the technology, IPS can detect, track, or even identify residents at home (Denis et al., 2019). Active IPS use tags or devices that are carried by the user while passive, device-free systems do not require the user to carry anything (Alam et al., 2020). Solutions are based on different technologies such as infrared radiation, radio frequency, visible light, physical excitation, computer vision, or Light Detection and Ranging (LiDAR) (Alam et al., 2020; Zafari et al., 2019). Existing infrastructure based on, e.g., Wi-Fi or Bluetooth can be used as IPS to reduce costs, however, this may have adverse effects on the primary purpose of the technology, i.e., data transmission (Zafari et al., 2019).

While passive and unobtrusive positioning is more user-friendly, these systems come with many challenges especially when tracking or identifying multiple subjects (Alam et al., 2020; Denis et al., 2019). Complex systems typically require professionals for setup and maintenance (Aloulou et al., 2013). Hence, smart floors with a dense network of sensors underneath the floor, cameras that require proper lighting and can suffer from occlusion and blind spots, or systems sensible to environmental changes (Vlasenko et al., 2014) can become problematic (Alam et al., 2020). From a practical standpoint, Zafari et al. (2019) call for system requirements such as minimal calibration, resilience against environmental changes (e.g., moving furniture), real-time positioning, and low computational complexity and energy-consumption with low costs at the same time. Especially for senior citizens, usability, acceptance, and safety are named as the most important factors (Mshali et al., 2018). However, general acceptance and interoperability of AAL systems still appears to be an issue (Stavropoulos et al., 2020).

10.2.2 Augmented Reality-based Indoor Positioning Systems

AR devices need to acquire their pose in 3D space to display and anchor virtual objects aligned with the physical environment (Billinghurst et al., 2015). This self-positioning is achieved by applying Simultaneous Localization and Mapping (SLAM) algorithms often based on different sensors to create a 3D representation of the environment and derive the users' location within this environment at the same time (Durrant-Whyte & Bailey, 2006; Morar et al., 2020). As satellite-based positioning systems (e.g., GPS) are not available indoors, numerous research projects have considered the self-positioning feature of AR combined with the ability to display spatial information to create AR indoor navigation systems to find a destination in an unknown environment such as a hospital (Drewlow et al., 2022; Huang et al., 2020), a university (Huey et al., 2011; Subakti & Jiang, 2016), provide help for people with low vision (Chi et al., 2022), or an escape route in an emergency situation (Yoo & Choi, 2022). Additionally, positional information can be used to display context-aware information in museums (Lin et al., 2019) or to support workers in factories (Flatt et al., 2015) or facility management (Baek et al., 2019).

AR as an interface for IoT devices in Smart Homes has been studied by several authors lately (El-Moursy et al., 2022; Heun et al., 2013b; Mayer et al., 2014; Park et al., 2020; Ullah et al., 2012; van der Vlist et al., 2013). Most of these studies focus on novel 3D-based interaction modes to control these devices, display information at their physical location, or create automations between them. However, prior research has not yet considered the potential of using the users' position for Smart Home automations, which can substitute certain user interactions altogether and actually integrate the user in their Smart Home environment.

10.3 Methods and Approach

To explore the technical and practical feasibility of AR HMDs for indoor location-based Smart Home automations, we derived requirements from prior research and own considerations to build an early prototype.

10.3.1 Design Considerations

The user tracking requirements for Smart Home automations vary depending on the use case. A higher resolution enables Smart Home systems to recognize user intents more accurately and ensure that automations are not accidentally triggered. Moreover, the AR HMDs should be able to also track themselves in dark environments for example to turn on lights when entering a room. Additionally, automations need to be aware of every resident to avoid turning off lights in a room that is still occupied when one person leaves. In general, the assumption of everyone wearing an AR HMD at all times for accurate tracking depends on the attributes of the headset, such as weight, comfort, or runtime. Still, a fallback mechanism is required when guests are visiting or small children without AR HMDs are living there. Thus, this solution should be implemented as part of a multi-modal approach and regular switches are still useful.

10.3.2 Prototype Description and Introduction of Use Cases

For our prototype implementation, we chose to use a 6DoF AR HMD with a high resolution and pose the assumption that residents wear the HMD most of the time to enable and test various novel use cases. Therefore, we selected the Microsoft HoloLens 2 as AR HMD for this prototype as it represents the state of the art in AR technology and supports these requirements. In the first step, we use the HoloLens 2 to create a 3D scan of the test apartment. The scan is imported and authored in Unity 2020.3.34f. To align the virtual scan and the physical environment, we use multiple spatial anchors from Microsoft's World Locking Tools that enable cross-device and cross-session consistency. Each room and more specific areas of interest are marked with boxes (colliders). Entering or exiting a collider publishes an event to the Smart Home system to notify it about the user's position. The Smart Home system for

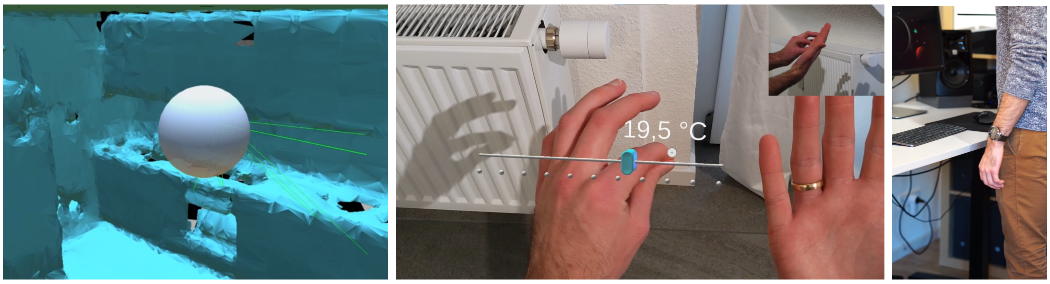

Figure 10.1.: Three potential use cases for location-based Smart Home automations. Left to right: Tracking resident position within the kitchen to turn on countertop lighting when approached (left); automatically selecting the IoT devices in the current room for user interactions (middle); moving height-adjustable desk to user's standing or sitting position (right).

this prototype is a Home Assistant[1] v2023.1.2 instance that controls off-the-shelf IKEA smart lights and Tado thermostats. While the AR HMD only sends location events similar to a motion sensor (i.e., inputs), Home Assistant handles the resulting actions in automation scripts (i.e., processing). In the following, we introduce three different potential future use cases that rely on room-level and centimeter-level detection, tracking, and identification.

Overall, the user can interact with the system as follows: For the first use case, users call a slider to control the room temperature by looking at their open hand as depicted in Figure 10.1. The selected temperature setting is automatically matched to the room in which the user is currently located. Thus, users do not have to browse a list of all thermostats and rooms to control the temperature. Additionally, when entering or leaving a room, the system turns the lights on and off. We use a second HoloLens 2 to simulate two residents and their interaction to, e.g., keep the lights on when one person remains in the room. In our prototype, both AR HMDs only communicate through Home Assistant and are not aware of each other. The second use case uses the high spatial resolution to control the countertop lighting in a kitchen. When a resident approaches the countertop, the lights turn on. Moving outside the collider area turns the lights off again. The third use case relies on height and the additional rotation information provided by the AR HMD to control a height-adjustable desk. By standing in front of the desk, it can move up to an ergonomic standing position. In turn, if the user sits down in front of the desk, it moves down to a seated position. To avoid false positives, the mechanism is only activated if the user is oriented towards the desk[2].

[1] https://www.home-assistant.io/ (Accessed: 12.04.2024)
[2] A demonstration of the prototype can is available here: https://youtu.be/z0OErqb09J8

10.3.3 Latency comparison with regular PIR motion sensors

The positioning performance potential of camera-based SLAM systems has already been demonstrated (Mur-Artal & Tardós, 2017). As PIR sensors are already often used in Smart Homes but can only detect and not track or identify residents[3], we wanted to compare the latency of a PIR sensor with our prototype for the use case of turning on lights in a realistic setting. To avoid residents entering a dark room and having to wait until the lights turn on, this operation should take below one second. For the study, a test subject enters a room at around 1 m/s as measured by markers on the floor and a metronome.

We tried an off-the-shelf IKEA Trådfri motion sensor, however, this device has a "recharge time" of around three minutes during which it is blind and does not report motion to save energy. Therefore, we built a non-battery-powered solution based on a standard HC-SR501 motion sensor[4] connected to an ESP32 microcontroller that runs ESPHome v2022.10.0 and publishes sensor data to the Home Assistant instance via Wi-Fi. Home Assistant is among the largest open-source projects on GitHub in general and in the Smart Home system domain (Escobar, 2022). The HC-SR501 and the ESP32 were chosen as they are widely available development hardware that support ESPHome and, thus, can be easily integrated into Home Assistant. We measure the end-to-end latency between entering the room and the timestamp of the reported event in Home Assistant when it becomes available for further automation tasks and count false negatives if one system misses an entry event. The motion sensor is placed just behind a corner close to the path of the test subject. For the AR HMD, we defined the boundary of the room at the intersection of the motion sensor's field of view and the middle of the path as depicted in Figure 10.2.

Since the timestamps of the motion sensor and the AR HMD only provide a relative latency between both systems, we filmed the entering event with a slow-motion camera at 480 fps to obtain more precise data. An LED light was added to the microcontroller to indicate that the motion sensor registered a movement. Thus, we measured the latency as number of frames between the point of entering and the LED turning on in the video. The first part of the study was conducted in a well-lit room from daylight and the artificial room lighting. While the PIR sensor filters visible light, the HoloLens 2 requires a well-lit room (Microsoft, 2022). Hence, we controlled for the light in a second part measuring only the latency difference

[3]Tracking is possible with a dense array of PIR sensors, e.g., Kim et al. (2009).
[4]Datasheet: https://www.epitran.it/ebayDrive/datasheet/44.pdf (Accessed: 12.04.2024)

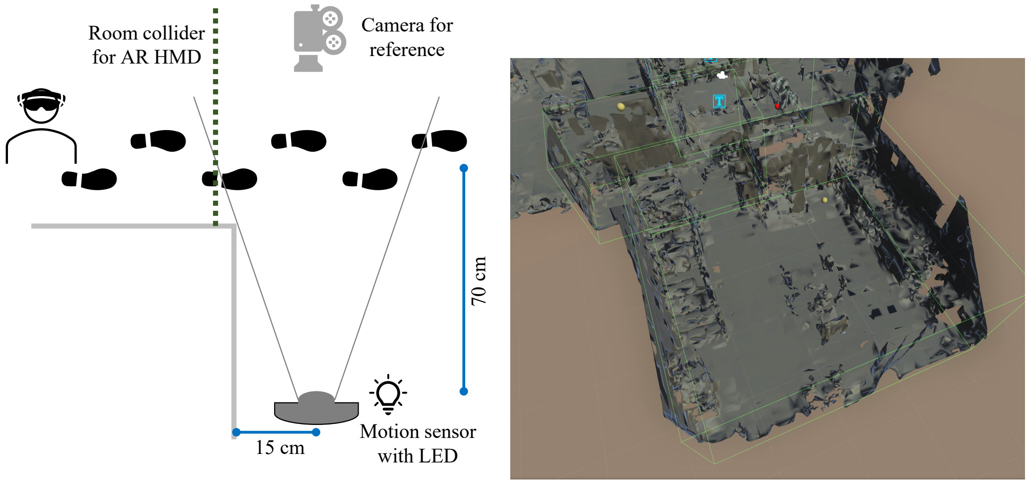

Figure 10.2.: Study setup (left) and the 3D scan of the test apartment with box colliders in green (right).

between PIR sensor and AR HMD to test if dark environments hinder the practical application of AR as IPS. The room was solely artificially lit with dimmable lights.

In general, we expect the PIR sensor to react quicker than the AR HMD due to the lower computational complexity of the mechanism. For the second part of the study, we expect the detection latency to increase and the HoloLens 2 to fail at tracking itself below a certain lighting threshold.

10.4 Results

For the first study part and for each lighting condition in the second part, the test subject entered the room ten times. The variance of both systems would require more measurements for an adequate statistical evaluation which was beyond the scope of an initial proof-of concept study aiming at testing the general feasibility of the approach. Thus, the results should be viewed as exploratory and preliminary. The results are depicted in Figure 10.3. In the first part of the study with the slow-motion camera reference, the PIR sensor (M = 729.09 msec, SD = 183.90 msec) was significantly slower than the AR HMD (M = 424.26 msec, SD = 154.84 msec; $t(10) = 9.09, p < .001$) with the specific positioning depicted in Figure 10.2. In the second part with artificial lighting only, the latency differences between both systems are distinctly smaller (ΔM = 22.71 msec, SD = 74.02 msec) and could also be attributed to variances in data transmission to or data processing in Home

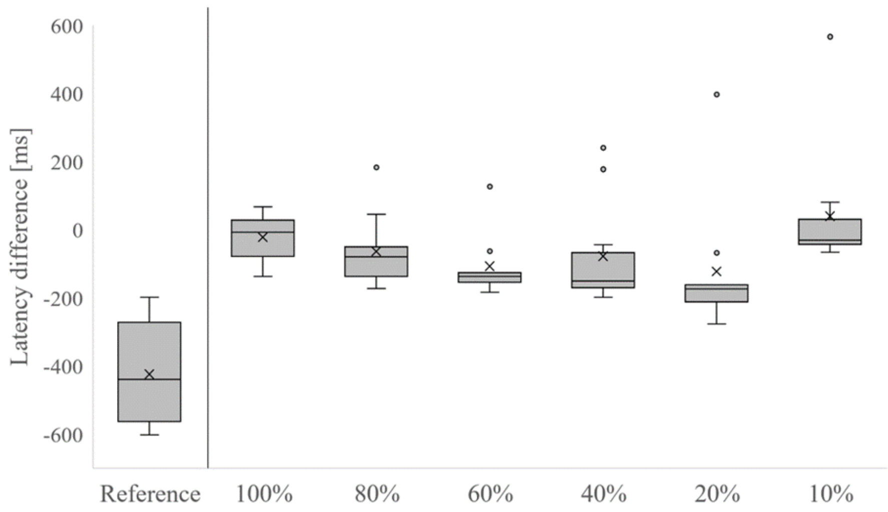

Figure 10.3.: Results of the conceptual study.

Assistant. We used a BH1750 light sensor[5] connected to a second microcontroller to measure the illuminance. During the first study part, the sensor measured 78 lx. In the second part, we stepwise dimmed the lights from 100% power (measured at 46 lx) to 0% ($0lx$) in steps of 20%. At 60% ($17lx$) and 20% ($4lx$), the HoloLens 2 missed one out of ten entry events. Even at 5% ($1.5lx$), the HoloLens 2 could detect entry events in all ten runs and the latency is still comparable to the PIR sensor ($\Delta M = 14.61$ msec, $SD = 35.00$ msec). However, the primary use case of displaying holograms does not reliably work anymore below 40% ($8lx$) resulting in a jumping and unstable depiction of the holograms. Finally, when entering a dark room ($0lx$) from a dimly lit room, three out of five entry events were missed. Hence, the HoloLens 2 becomes unreliable whereas the PIR sensor detects every entry motion.

10.5 Discussion and Future Work

Benefits

The localization based on AR HMDs brings several advantages for Smart Home users. Compared to PIR sensors and many IPS alternatives, no external hardware is required to determine the user's exact position. Thus, the solution scales only with the number of users and is independent from living area and number of rooms. As

[5]Datasheet: https://www.mouser.com/datasheet/2/348/bh1750fvi-e-186247.pdf (Accessed: 12.04.2024)

the Smart Home system can differentiate between users based on authentication on the AR HMD or alternatively their height, the system can personalize rooms and user interfaces (Marques et al., 2019). Moreover, PIR sensors struggle with detecting non-moving human subjects (e.g., on a couch or in the bathroom) which complicates room occupancy estimation (Andrews et al., 2020). As IPS know where each resident is, determining the occupancy of each room is trivial and the reliability of related automations is higher.

Use Cases

The positioning resolution of AR HMDs enables a variety of automation use cases in Smart Homes. Accurate room occupancy data does not only allow for exact light controlling but also occupancy prediction to pre-heat or cool rooms to save energy without compromising on user comfort. Music can follow the user around the home by only playing on speakers near them. The aforementioned room personalization can encompass light color and temperature, ambient sound, room temperature, or the art selection of digital picture frames adapted to the user's preferences or the compromise of several simultaneous residents. User locations could also be used for security aspects. Doors could automatically unlock when authenticated users are nearby. Additionally, alarm systems in unoccupied rooms of larger homes could automatically be armed (Jose & Malekian, 2017).

AAL literature suggests more use cases based on accurate indoor positioning (Stavropoulos et al., 2020). Automations that adjust the route of autonomous vacuum cleaners to avoid blocking residents with motor impairments or automatically lock and unlock the bathroom could improve life quality and avoid typical day-to-day inconveniences. In cases of emergency, an AR HMD could display a safe exit route, e.g., based on smoke sensors. Moreover, emergency personnel could access the location of residents if such systems are connected. This support in emergency situations would be even more critical for people with motor impairments.

AR HMDs in the Smart Home open a potential for many more applications. With the principles of Universal Design, simple improvements cannot only help improve the lives of people with disabilities but also improve comfort and safety for everyone at the same time. An overview of possible location-based automations depending on required resolution and localization type is given in Table 10.1.

Table 10.1.: Examples of possible location-based automations depending on required resolution and localization type. (1 inch = 2.54 cm)

Required resolution	Localization type	Example
Room-level	Detection	Turn on lights HVAC control
Room-level	Detection and identification	Personalization (lights, media, etc.) Activate alarm system in unoccupied rooms
Centimeter-level	Detection	Open doors Remove potential obstacles such as a vacuum robot (AAL)
Centimeter-level	Tracking	Control specific lights in a room Predict movements to trigger automations before an actual event
Centimeter-level	Detection and identification	Lock and unlock front door
Centimeter-level	Tracking and identification	Set adjustable objects to user preference (e.g., adjustable desk) Recognize Activities of Daily Living (ADL)

Limitations

The suggested use cases and potentials rely on the availability of a user-friendly AR HMD, which is not yet available today. It might appear far-fetched, yet light-weight, unobtrusive AR devices will likely become available in the near future. Thus, using sensor data from already worn AR HMDs does not pose additional effort or cost and is, therefore, less expensive than using a sensor network, even if the individual sensors are inexpensive.

The HoloLens 2 used for our prototype does not have sensors that sustain its features in dark environments. Thus, when walking into a dark room, the device often loses track of its position and could incorrectly trigger automations for other rooms. Additionally, the rather slow update rate of several tracking sensors and their processing struggles with fast movements of the user. Our exploratory study demonstrated that the camera-based self-tracking performs consistently fast when compared to the popular PIR sensor. Particularly in low-lighting conditions, the performance turned out to be better than expected. The results are also limited to the experimental setting and would be different if the collider would be larger or the PIR sensor would be moved. Nevertheless, additional tracking sensors that do not rely on room illumination should be integrated for AR HMDs to be applicable to realistic home environments. Still, the HoloLens 2 is capable of demonstrating future opportunities and use cases of user-friendly AR HMDs.

Today, Indoor Position Systems are frequently criticized for potential privacy issues (Alam et al., 2020). Indeed, privacy can be also a challenge for our concept, especially with high-resolution systems that rely on cameras. However, AR-based systems could request user consensus before positional data leaves the device at all. Settings could be applied depending on location (e.g., no tracking within the bathroom). This aspect is also applicable for public environments where a user might, e.g., allow a shopping mall to access the current location to provide indoor navigation, thereby setting permissions for different places. Therefore, developers of AR HMDs should consider data privacy as a fundamental principle always having the users' interests in mind.

Future Work

In a broader context, we want to not only study holistic AR interfaces for Smart Home and AAL systems that include indoor positioning for automations, but also display user interfaces for appliances and IoT devices that only have limited screen space (washing machines, dish washers, etc.) and allow the user to naturally interact with these devices via gestures, eye movement, or speech. Multi-modal and natural interactions, adaptive and personalized user interfaces, and the interoperability of these systems open exciting future research areas.

10.6 Conclusion

Assuming its availability, the application of AR HMDs for indoor positioning in Smart Homes appears both simpler and superior to current solutions and potential alternatives. Reliable positioning of residents allows for several use cases that may improve life quality and autonomy in AAL contexts, and user comfort and security in Smart Home settings in general. The prototype presented in this paper demonstrates the feasibility and performance of the novel idea of an AR HMD for indoor positioning in the context of Smart Home automations on an off-the-shelf device. Still, a lot of effort is required to miniaturize the technology and make this idea a reality.

11 Connecting Home: Human-Centric Setup Automation in the Augmented Smart Home

This chapter comprises the version of record of the following paper: Schenkluhn, M., Knierim, M. T., Kiss, F., & Weinhardt, C. (2024). Connecting Home: Human-Centric Setup Automation in the Augmented Smart Home. *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*, 16. Changes include formatting, numbering of the research question and chapters, minor changes for consistency, and correction of spelling errors.

11.1 Introduction

Smart Home interactions currently come with multiple hurtful UX challenges as users rely mostly on smartphones for interacting with connected household devices (Bitkom Research, 2022). This forces users into long journeys: find the phone, unlock it, find the vendor-specific application, locate the target device using the vendor-specific Graphical User Interface (GUI), find the desired functionality within those supported by the device, and finally trigger the action. *Smart light switches* that often have several buttons to control light temperature, color, intensity, or custom actions simplify this lengthy process at the expense of usability: users must memorize (sometimes quite complex) button combinations to control their home.

One particularly cumbersome setup aspect is locating individual devices and accessing their controls. A possible way to overcome this problem is through the display of user interfaces in visual proximity to the target device via AR[1]. With current

[1] e.g., Smart AR Home: https://smartarhome.com/, Reality Editor: https://realityeditor.org/ (Accessed: 12.04.2024)

Figure 11.1.: Augmented Reality allows a direct interaction with Smart Home devices if their 3D position is known to the system. Our approach supports users with the spatial setup process of their Smart Home devices.

developments in the consumer market (e.g., Apple's launch of the Vision Pro) hinting at the long-promised consumer-market-grade maturity of AR technologies, AR applications are approaching large-scale deployment, particularly in domestic spaces. Consequentially, investigating AR solutions to Smart Home problems seems a rather promising approach, since AR could provide several UX benefits, such as on-the-fly interactions and more natural and intuitive interaction designs. Further, AR not only removes the spatial dissociation between the target device and its user interface, but it also simplifies the user journey immensely, by offering larger areas for displaying visual contents and interactive elements. And importantly, an AR interface would strongly reduce reliance on smartphones, which have been increasingly considered a negative presence in households (Kiss & Schmidt, 2019; Park, 2005; Richards et al., 2015).

Despite the advantages and comfort offered by AR-based Smart Home interaction, this field of application is still at an early stage of development. Assembling a Smart Home requires setting up many products, often from different manufacturers and with diverse characteristics. Many vendors alleviate the installation process through "plug and play" products, which are configured into the Smart Home system with varying degrees of automation (e.g., Amazon's Frustration-Free Setup[2]). However, the configuration of AR elements in Smart Home applications is typically done by hand. This is particularly inconvenient in terms of matching the positions of physical

[2]https://developer.amazon.com/frustration-free-setup (Accessed: 12.04.2024)

devices to coordinates in the spatial models of AR frameworks. Multiple technical solutions have been proposed to simplify this task, such as the usage of visual markers and QR codes, or indoor location mechanisms. These solutions present drawbacks in terms of product design and production costs, as well as a compromise in practicality (e.g., QR codes must be scanned individually).

In this paper, we propose a technique to solve the problem of device localization. We take advantage of the sensors available on AR devices, and the actuators present on connected household appliances. By making appliances blink, buzz, or call for attention the best way they can, we enable AR devices to identify them individually and calculate their physical coordinates in the real world. Our solution is manufacturer-independent, allows for a high level of automation, and requires no additional hardware. It can be retroactively applied to many legacy devices and requires no significant costs of implementation for future designs.

In designing a feasible setup method, it is paramount to consider the experience that the users have during the configuration of Smart Homes. This initial contact with Smart Home technology can have a conditioning effect on long-term subjective perception of interactions and, in extreme cases, can result in discouraging levels of frustration. To that end, recent research on *positive computing* (Peters et al., 2018) emphasizes the importance of looking beyond classic usability factors like *ease of use*, especially for the interaction with pervasive technologies that accompany people in their lives. Thereby, the innate psychological needs for autonomy, competence, and relatedness ought to be recognized in the design of technology interactions like AR Smart Home setups. This recognition could ensure that users experience a lasting, positive connection to their Smart Home, established right from the start (Adams et al., 2017). However, as this idea has not yet been pursued with Smart Home interactions, it is paramount to explore different interaction designs and gain a better understanding of their impacts on users' needs and preferences.

As a starting point for the design spectrum, we herein mainly considered the degree of setup automation, as we expected substantial differences in how this dimension could affect psychological needs. Initially, we considered that a manual setup could be the one that maximally fulfills these needs, as it provides full control over the setup process (providing a high degree of autonomy), could instill a sense of mastery by completing the setup actively (providing a high degree of competence), and could create a sense of connection to the system through this engagement (providing a high degree of relatedness) – a sensation also known as the IKEA effect (Norton et al., 2012). However, we also expect that a fully manual setup could easily become demanding and frustrating, especially when the number of smart devices that need to be set up increases, making us aware of likely trade-offs between the recognition of psychological needs and classical usability dimensions (Saket et al., 2016). In contrast, a fully automated setup might be easier and more convenient, yet might move the user “out of the loop”. Thereby, a fully automated setup could also be experienced as alienating and disconnecting. Therefore, we considered how to possibly mitigate these trade-offs and achieve an effective balance between psychological needs recognition and usability through a semi-automatic setup process, leading us to a final set of three interaction design variants. Altogether, these design considerations gave rise to these Research Questions (RQs), aiming to explore emerging experience trade-offs:

- **RQ6.1**: *Does a manual Smart Home spatial setup design maximize psychological need recognition, and do classical usability dimensions undermine the benefit of this characteristic?*
- **RQ6.2**: *Does a fully automated spatial setup maximize classical UX dimensions like ease of use, mental workload, and frustration, but reduce the attractiveness of the interaction design by thwarting psychological needs?*
- **RQ6.3**: *Does a combination of manual and automated features strike an effective balance between psychological needs recognition and classical UX dimensions, effectively enhancing technology acceptance?*

We pursue answers to these questions through a mixed-method study that includes a prototypical implementation of the device localization system for the manual setup scenario, and a Wizard-of-Oz study for the (semi-) automatic scenarios. By recruiting a diverse sample (age 19-64, 54% female, mixed residence types and experience with Smart Homes and AR) and conducting the study in a state-of-the-art Smart Home lab, we enable an experience of how this AR Smart Home setup would be experienced in a real-world setting in the future.

Our work contributes to the HCI community in the following ways:

- We provide a system to support and facilitate device localization during AR configuration of Smart Homes. The code for the project is provided with the article.
- Overall, we find that users report high levels of engagement with this AR-based Smart Home interaction, highlighting the approach as a promising design option for future work in the HCI domain.

- Furthermore, through the combination of quantitative and qualitative evidence, we provide a comprehensive and in-depth account of users' experiences, highlighting that indeed, the combination of control and automation provided a good mixture of need recognition and usability, indicating high levels of technology adoption.
- At the same time, we also find substantial experience contrasts, for example that some participants do report a strong sense of connection to the Smart Home environment whereas others remain indifferent about it. Paired with the observation of different preferences for the degree of active interaction, we outline implications and design recommendations for following work.

11.2 Related Work

Smart Homes are characterized by connecting several devices, automation features, and remote control (Jiang et al., 2004) with goals like helping users or increasing hedonic value, e.g., through aesthetic home improvements (Jensen et al., 2018). The Smart Home extends to numerous device categories such as lights, speakers, thermostats, blinds, household appliances, sensors, and more (Robles & Kim, 2010). Users can choose between individual devices for selected functions, ecosystems from specific vendors or consortiums such as Home Connect[3] usually using a common Smart Home hub, or integrator solutions such as Home Assistant[4] that combine fragmented ecosystems (Jakobi et al., 2017). In recent years, Smart Home research has evolved from engineering disciplines to several other fields, such as HCI, and inspires interdisciplinary research (Yao et al., 2023).

11.2.1 Smart Home Research in HCI

Yao et al. (2023) identify five trends in Smart Home research within the HCI community: *interaction design*, *user behavior*, *smart devices*, *design exploration*, and *data, privacy, and security*.

Data privacy and security is currently the most prominent research stream (Yao et al., 2023). While Smart Home devices connected to the internet pose various privacy and security risks (Acar et al., 2020), users generally trust IoT devices and manufacturers (Zheng et al., 2018).

[3] https://www.home-connect.com/ (Accessed: 12.04.2024)
[4] https://www.home-assistant.io/ (Accessed: 12.04.2024)

Regarding user behavior, Woźniak et al. (2023) observed distinct roles from passive users to active users, and administrators. Administrators face a trade-off between professionally installed, well-integrated, and pre-configured systems and more flexible, adaptive, cheaper retro-fit systems that usually require more effort for device selection, setup, and configuration (Jakobi et al., 2017). While tasks like selection, setup, and configuration are usually carried out by interested users from the administrator role, they still face significant challenges and have to build up knowledge for their Smart Home (Jakobi et al., 2017). Household members who only use the system typically rely on several vendor-specific Smart Home apps on smartphones or wall-mounted tablets (Sikder et al., 2022), voice-based assistants such as Amazon Alexa (Edu et al., 2020), buttons on the devices themselves, or remote controls to interact with Smart Home devices. Still, they often require training and need to remember button combinations, voice commands, app layouts, and the affordances of smart devices in general. Thus, researchers are demanding more natural interactions (Yao et al., 2023).

11.2.2 AR, Indoor Positioning & The Smart Home

Integrating AR technologies into Smart Homes is a promising area of application, which has seen many efforts in diverse areas like elderly care (Arcelus et al., 2007), energy management (Zhou et al., 2016), and nutrition support (Luo et al., 2008).

A subset of this work aims to provide insights and recommendations for AR integration with Smart Homes in general terms. Mahroo, Greci and Sacco propose a framework for AR-based interaction with Smart Homes and their components (Mahroo et al., 2019). Their work focuses on the defining features of this application, namely the spatial aspects, such as the alignment of mixed elements, and the interconnection of the components. Jo and Kim delve further into the technical aspects, identifying the main components to achieve synergetic integration (Jo & Kim, 2019).

Devices are usually assigned an area (e.g., a room within the house), and can be grouped for it (e.g., turning on all lights in a room at once). Thus, the exact location of each device is not known to the system, but also not required in a traditional setup. However, the three-dimensional position of a device is necessary for advanced use cases. Especially for applications that connect AR glasses to the Smart Home, the precise location of the devices is required for an unmediated, natural interaction. There are numerous technologies for indoor localization ranging from radio-frequency-based approaches to inertial sensors, ultrasound, and visible light communication (Basri & Elkhadimi, 2020; Kim et al., 2021b). Ultra-wide-band systems can precisely track beacons placed on IoT devices (Minoli & Occhiogrosso, 2018) and visible light communication can track devices without congesting radio-frequency bands (Tiwari et al., 2015).

Yet, AR devices can locate themselves within a 3D coordinate system without the need for additional external devices (Durrant-Whyte & Bailey, 2006), thus, enabling automations in the Smart Home that further reduce the number of required interactions and determine the relative localization of other devices (Schenkluhn et al., 2023b). For instance, "Smart ARbnb" (Gecevicius et al., 2021) provides transparency of device capabilities and automations for guest users by detecting light patterns of small LEDs next to each smart device with their smartphone camera. Similarly, several papers discuss effective locators for use cases such as spatial automation creation (Heun et al., 2013a; Seiger et al., 2019), privacy awareness (Prange et al., 2021; Song et al., 2020), or providing context-sensitive, relevant information (Ghorbani et al., 2019; Sharma et al., 2022). Leveraging the spatial aspect of AR, Wu et al. (2020) developed Megereality, a model for gestural interaction using multiple devices in AR. Their work attempts to break the barrier between the physical and digital realms by using metaphors and embodying abstract processes.

Presently, this existing work focuses on running systems. Thereby, installation, configuration, matching, and integration of AR components with their physical counterparts is performed by an administrator and rarely discussed. However, the setup is a critical aspect of Smart Home popularity and, although it is likely done just once, it can have a significant detrimental UX effect (Jakobi et al., 2017). The challenge of configuring spatial Smart Home settings in AR was considered by van der Vlist et al. (2013) in their work on *semantic connections*. This concept attempts to facilitate a better user understanding of their Smart Home configuration using visible lines and symbols displayed with a small projector. Another approach allows users to set individual privacy settings by pointing an AR device towards any IoT device during setup (Bermejo Fernandez et al., 2021). Lyu et al. (2022) created HomeView to automatically derive a digital twin of Smart Homes based on AR captures, reducing the need for continuous manual reconfiguration of device positions.

From the literature, it is clear that breaking the division between the real world and the spatial model is critical, yet challenging. This duality becomes particularly relevant for AR applications in Smart Homes since it is key to enabling the kind of interaction that can truly benefit the user. Thus, solving the problem of matching spatial coordinates with Smart Home devices presents an opportunity for a valuable contribution to both the AR and Smart Home communities. Furthermore, as the device setup is the entry point for many Smart Home experiences, anticipating the UX impacts of interaction designs is vital for an effective innovation at this intersection of AR and Smart Homes.

11.2.3 Self-Determination Theory

We hypothesize, that a Smart Home setup process must satisfy the homeowners' psychological needs to enable a lasting positive UX and adoption. Self-Determination Theory (SDT), initially proposed by Ryan and Deci (2000), posits that a positive life experience is fundamentally rooted in the fulfillment of psychological needs. Central to SDT is the idea that individuals have innate psychological needs, and the satisfaction of these needs can foster optimal growth and well-being (Ryan & Deci, 2000). These needs are:

1. **Autonomy**: The sense of volition and being the origin of one's behavior.
2. **Competence**: The feeling of effectiveness in one's actions.
3. **Relatedness**: The feeling of connection and belonging with others.

While the theory has been extensively applied and confirmed in the education (Xia et al., 2022) and work domains (Gagné et al., 2022; Olafsen et al., 2017), HCI scholars too have found it to be a useful vehicle for the design and evaluation of positive user experiences, especially in games (Ballou et al., 2022; Ryan et al., 2006; Tyack & Mekler, 2020), but also in general as an extension to classical UX considerations (Peters et al., 2018). Understanding and incorporating these psychological needs can significantly influence user experience. For instance, a system or interface that supports a user's sense of competence can enhance engagement, satisfaction, and persistence in interaction. Likewise, providing users with choices (supporting autonomy) and fostering a sense of community or connection (supporting relatedness) can further enhance user engagement and satisfaction (Peters et al., 2018).

In some more specific instances, previous HCI work has explicitly investigated how psychological needs recognition can improve the design of interactions with intelligent technologies like chatbots (Yang & Aurisicchio, 2021), robots (Lu et al., 2023), and recommendation agents (Vreede et al., 2021), showing that the recognition of psychological needs creates higher engagement, deeper interaction, and longer-lasting acceptance of such intelligent systems.

While the approaches to need fulfillment in interaction design differ somewhat from application to application, there appears to be a certain consensus, that autonomy can be fostered by providing control, for example by allowing customization and meaningful choices whenever possible so that users feel they have a say in how they interact with the technology (Peters et al., 2018; Yang & Aurisicchio, 2021). For competence support, it is recommended that interactions enable gradual skill development and provide positive feedback and reinforcement for completing tasks successfully to enhance users' feelings of mastering a particular task (Peters et al., 2018; Tyack & Mekler, 2020). Relatedness is, on the one hand, primarily fostered by incorporating social elements into the interaction design that enable interaction with others, such as social media integration, collaboration features, or community forums, to create a sense of connection with other users (Peters et al., 2018; Tyack & Mekler, 2020). On the other hand, relatedness is also considered as a connection to the technology, which can be enhanced by tailoring the system to the individuals' preferences. This personal touch supposedly enhances the sense of connection between the user and the technology (Peters et al., 2018; Yang & Aurisicchio, 2021).

Besides these previous works, psychological needs have not yet been considered in the context of Smart Home technologies. However, we argue, that this is a vital application domain as it is known that thwarting psychological needs reduces general well-being (Gagné et al., 2022; Olafsen et al., 2017), we argue that the interaction that individuals have with the technologies in their own homes must be designed to support these needs due to the pervasiveness of the interaction in everyday life. Furthermore, we argue that the recognition of these needs will have an important influence at the very early stages of a Smart Home interaction. In a sense, first interactions with a Smart Home should leave a pleasant impression to elicit positive spillover effects for following everyday interactions.

11.3 Application and Experimental Setup

To develop an effective solution for AR-based Smart Home setups, we created the Prototypical Augmented Reality Configuration System (PARCS), a system capable of determining smart device positions. The PARCS is manufacturer-independent and works under the assumption of a working Smart Home setup without any initial knowledge about the position of any device. PARCS combines the actuators present in Smart Home appliances with the sensing capabilities commonly provided by HMDs. Each Smart Home device provides a distinctive signal by e.g., switching

LED power indicators on and off, emitting specific sounds, or visually distinctive movements, thus allowing cameras and microphones integrated into an HMD being able to pick up those cues and calculate their position.

11.3.1 The Prototype

For this experiment, we implemented the PARCS based on a Microsoft HoloLens 2 (v2020.3.34f). We used Unity as the main development environment, with Microsoft's Mixed Reality Toolkit (MRTK v2.8.3[5]) as the supporting framework. As a proof-of-concept, we implemented the functionality to support the detection of smart light-bulbs (Philips Hue E27) using computer vision (OpenCV v4.7.0[6]). The Smart Home hub itself consists of a Raspberry Pi 3 running Home Assistant (v2023.5.3). The popular open-source project Home Assistant offers several thousand integrations, including 141 smart light ecosystems (Assistant, 2024).

As a use case, we implemented the positioning of smart lights within an already configured Smart Home environment without knowledge about specific device positions. Smart lights were our primary choice as they usually occur several times in a Smart Home, give immediate visual feedback to users, and were the most natural device category to build a camera-based position estimation prototype for due to distinct visual characteristics (blinking) and simple, unified APIs. The HMD connects to the Smart Home hub and sends commands to the individual Smart Home devices via the Home Assistant REST API. To detect an individual device, the smart light is turned on and off repeatedly. This approach is manufacturer-independent, as the Smart Home hub abstracts and exposes each smart light as a light entity with a fixed feature set. The "turn on" and "turn off" commands are available for all smart lights by definition.

The AR application queries the most recently triggered motion sensor, if available, to determine and suggest the area that the user is currently in. Otherwise, the user can select the respective area or room manually. Then, a list of all smart lights in the area is retrieved and turned off. Using the front-facing RGB camera of the HMD, the contour of bright surfaces or reflections is detected using a technique adapted from the work of Suzuki (1985). Once the planar coordinates of the camera's image are calculated, these are projected on the 3D mesh generated by the HoloLens' depth camera, determining the coordinates of the bright spot. The application marks these spots to ignore and avoid false positives later on. Next, the first smart light is

[5] https://github.com/microsoft/MixedRealityToolkit-Unity (Accessed: 12.04.2024)
[6] https://opencv.org/ (Accessed: 12.04.2024)

turned on and off repeatedly for detection and the user is asked to look towards the device. After each "turn on" command, the application considers the 3D position of each new bright spot as a potential candidate for the device and removes bright spots that remain after turning the device off again. Hence, if only one candidate remains consistently, the process terminates, stores the position of the device, and continues with the next one. This approach is executed locally and in real-time on the HoloLens without any perceivable detriment to the HMD's frame rate. Images are captured at 15 frames per second, and each image is analyzed within 4 frames of the application's update loop ($\leq 67msec$). Depending on the time the user requires to look towards the flashing device, the process can take less than 5 seconds per device.

The general design of the interaction was created following the HoloLens 2 guidelines from the official MRTK documentation [7]. By these recommendations, interaction with near elements and hand menus was controlled using finger-pointing. The positioning of the spheres to mark the spatial coordinates of the Smart Home devices was based on the go-go interaction technique, to reach distant locations and minimize the required movements (Poupyrev et al., 1996). We abstained from further embellishments to minimize external factors in the behavior observed during the study.

The source code[8] of the implementation and a depiction of the process at the end of the accompanying video are made available with the article.

11.3.2 Three Interaction Design Variants

Beyond the light detection feature, we adopted the Wizard-of-Oz technique to both focus our research on the user interaction experiences and also to extend the PARCS' feature set. Specifically, we simulated a perfectly functional application that could allow the user to control 50 smart lights and 4 smart speakers. To gain insights into the potential trade-offs between psychological needs and classical UX dimensions, we developed an experiment comparing degrees of system automation, as we expect this dimension to substantially impact psychological need fulfillment (see RQ1-3 in the Introduction). We designed three variants (see Figure 11.2):

[7] https://learn.microsoft.com/en-us/windows/mixed-reality/mrtk-unity/mrtk2/ (Accessed: 12.04.2024)
[8] https://gitlab.com/mschenkluhn-kit/parcs (Accessed: 12.04.2024)

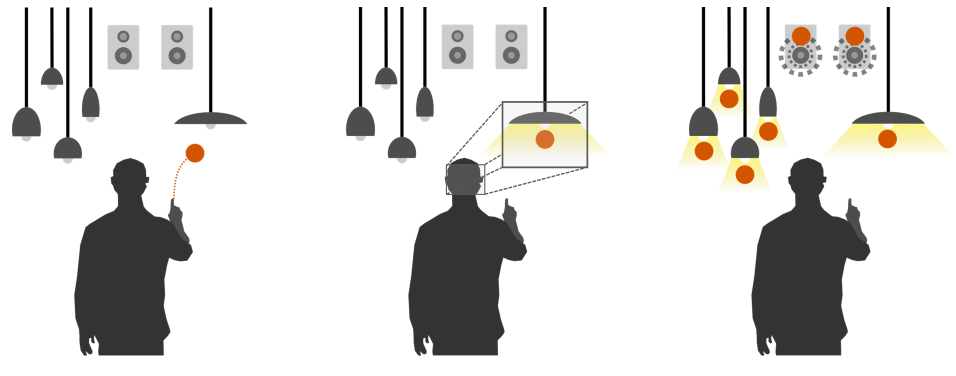

Figure 11.2.: Schematics of the three interaction design variants: manual (left) – showing that a user places a sphere on a lamp for a manual spatial configuration of the respective device, semi-automatic (center) – the AR cameras detect a flashing smart light automatically if the user briefly focuses on the device to set the spatial position one device at a time, and automatic (right) – showing that all devices emit signals for a simultaneous spatial setup of each device.

Manual

The manual setup is proposed as the more need-recognizing condition for the interaction and lacks intelligent support. The user interacts with one device at a time (e.g., lights and audio devices). We used audio devices in addition to smart lights to stimulate another sense as contrast. Devices attract the attention of the user through their feature sets (e.g., lights turning on/off, audio devices playing sounds). The user then positions a virtual sphere on the device, which functions both as an anchor for the Smart Home system and a visual interface for the user. Spheres are initially positioned in abundance on the floor and can be chosen indistinctly, to avoid the spawning and search of new spheres. Once the users are satisfied with the position of the sphere, they open a hand menu by making a gesture to confirm the positioning and cue the system to move on to the next device. This process is repeated for each of the available devices. It is important to note that the coordinates of the device are obtained from the user's manual positioning of the sphere.

We expected this interaction variant to best fulfill the psychological needs by offering complete control over the process (ensuring autonomy) and fostering a sense of mastery and engagement (related to competence and relatedness through active engagement with the system). However, we also expect potential challenges with a fully manual approach, particularly as the number of devices increases, prompting us to consider trade-offs between psychological needs and traditional usability dimensions.

Automatic

In contrast to the manual design, the fully automatic variant reduces the users' involvement to the minimum. In this condition, the recognition of all devices is parallelized, and all devices emit their signals simultaneously. While the sensors available on the HoloLens 2 make this variant technically feasible, the effort to develop such a system surpasses the scope of our work. Thus, to provide this functionality, we resorted to the Wizard-of-Oz technique and simulated the automated location of devices. This is achieved by actuating all the Smart Home devices simultaneously for 25 seconds. After that time, all devices are turned back to their idle states and all interaction spheres are shown in their correct (pre-recorded) positions. We expect that a fully automated setup might be easier and more convenient (higher classical UX), yet might move the user "out of the loop". Thereby, a fully automated setup could also be experienced as off-putting and disconnecting, thwarting psychological needs.

Semi-Automatic

Finally, the semi-automatic interaction can be seen as an *assisted* approach that could bridge the UX/needs trade-off discussed for the previous two design variants. To achieve this effect, we designed the semi-automatic interaction to feature control and automation on demand.

Similarly, as in the manual condition, devices connected to the Smart Home are configured sequentially, one at a time. Each device is actuated individually until the users fix their head gaze towards the device for at least 2 seconds. The successful spatial setup of the device is indicated by the appearance of a control sphere on the device and a short sound signal.

After configuring the device, users are prompted to choose between continuing the configuration for each single device, or setting up all devices from the same category (e.g., lights or audio devices) simultaneously. After the user confirms their position, the device is automatically recognized and its position is calculated and recorded. If the user chooses the second option, all devices of the category are actuated simultaneously (e.g., all lights blink), and the user configures each of them by fixing their gaze in the direction of the devices. Once it is configured, each device stops immediately emitting signals, thus allowing the user to choose a different device from the remaining ones.

Independently of this choice, the semi-automation of the PARCS is limited to calculating the position of the device, while the rest of the process is still controlled by the users.

For our experiment, the position of devices is already known to the Wizard-of-Oz system. This significantly simplifies the recognition process by limiting users' gaze tracking and reaction when it hovers over the invisible target for the goal device for more than two seconds.

11.3.3 The Smart Home Environment

The experiment was conducted at our lab (to ensure anonymity, we exclude distinctive details from this manuscript. A thorough description of the infrastructure would be added in a camera-ready version).

The used space is a dedicated room with a surface of 74 m^2 (around 800 $sqft$), fully dedicated to the purpose of replicating a real Smart Home environment. The interior design resembles a modern open apartment with a fully functional kitchen, a living room with comfortable sitting options, a dining area with a large table, and multiple props to reproduce the appearance of an inhabited home. The Smart Home devices are managed using Home Assistant and include:

- 60 distinct lights (spots, panels, ambiance luminaries, all controlled via DALI)
- 3 Philips Hue lights
- 4 smart TVs
- 9 smart speakers
- Smart blinds
- Smart oven
- Smart vent hood

Other devices, such as door locks, atmospheric sensors, smart appliances, or cameras, were not used in the study and thus not listed.

11.4 Experiment Design

We designed the experiment to reproduce a realistic use case scenario while attempting to consistently collect reliable data to reach our research goals. To achieve this, we devised a scripted procedure consisting of three tasks, one per condition, and used standardized questionnaires to collect quantitative data. Additionally, we collected qualitative data over individual semi-structured interviews with all participants of the experiment.

11.4.1 Procedure

The participation had a total duration of approximately 60 minutes for each participant. Participation and travel time to the remote location of the lab were compensated for a fixed total of 70€. This amount was suggested by the recruiting agency in consideration of the increased logistics and travel time required for participation during working hours.

Preparation

Participants were welcomed, briefed, and prompted to provide written informed consent for their participation. Details regarding data privacy were collected, processed, and stored following European GDPR and approved by our data protection office. The participants then received a short introduction to the concepts of Smart Homes and AR. This was followed by an explanation of the problem of assigning real-world positions to the devices connected to a Smart Home system and how this can be achieved using AR.

Before starting the tasks, participants were asked to fill out a questionnaire collecting information about prior experience with Smart Homes and AR, and categories of Smart Home devices in possession and planned to be purchased.

Next, participants were asked to wear the HoloLens 2 and follow the calibration procedure. This was followed by two interactive tutorials. The first one was based on the default MRTK Hand Interaction Sample Scene[9], including the use of the *hand menu*[10] gesture. This tutorial acquaints the user with the general interaction concept

[9]https://learn.microsoft.com/en-us/windows/mixed-reality/mrtk-unity/mrtk2/features/example-scenes/hand-interaction-examples (Accessed: 12.04.2024)
[10]https://learn.microsoft.com/en-us/windows/mixed-reality/design/hand-menu (Accessed: 12.04.2024)

and, in particular, with the elements relevant to this user study. The second tutorial teaches the participants how to turn lights on and off using the interactions learned during the previous tutorial.

Task

The order in which each interaction design variant was administered was counterbalanced across participants to compensate for learning effects. For the manual variant (see Section 11.3.2), participants were asked to position the spheres manually for 50 lights and 4 speakers. This condition of the task was limited to 12 minutes for the sake of brevity, and to keep the participation within a reasonable time frame. We included all available lights in the lab for consistency to avoid participants completing the task before the time limit has passed. After the time passes, the task is interrupted independently of the achieved progress.

For the semi-automatic variant, participants were asked to use the interaction described in Section 11.3.2. The task consisted of assigning the same 50 lights and 4 speakers used in the manual condition. This task was also limited to 12 minutes. The automatic variant followed the methodology described in Section 11.3.2. Thus, the duration was limited to less than a minute.

After concluding the task for each application variant, participants were asked to fill out multiple questionnaires: Technology-based Experience of Need Satisfaction (TENS) (Peters et al., 2018), the short version of the User Experience Questionnaire (UEQ) (Schrepp et al., 2017), the Technology Acceptance Model (TAM) (Park et al., 2018; Venkatesh & Bala, 2008), and the NASA Task Load Index (NASA-TLX) (Hart & Staveland, 1988).

The UEQ-S and NASA-TLX are well-established tools in HCI to measure subjective user experience and subjective workload, respectively. We used TAM to assess *perceived values*, *perceived enjoyment*, *perceived usefulness*, and *intention to use* (Venkatesh & Bala, 2008). Following the literature and to keep the questionnaires short, we used only one item with the highest factor load for each of the target topics.

We used a subset of the TENS questionnaire, namely the TENS-Interface and the TENS-Life. The TENS-Interface questionnaire assesses *autonomy* and *competence*. In the TENS-Interface questionnaire, the third self-determination theory construct of relatedness is optional. Yet, we wanted to explore if a direct interaction model and the setup process would have effects on the *relatedness* not to other people, but rather the Smart Home environment itself. Therefore, the TENS-Life subscale was adapted and used to assess perceived *relatedness*.

After the completion of the task for the three variants, we collected data about each participant's gender, age interval, and type of home. Additionally, they filled out the Affinity for Technology Interaction scale questionnaire (ATI) (Franke et al., 2019). During a short semi-structured interview, participants provided insights regarding general observations, preferences, and efficiency ranking of the alternatives as well as overall user experience feedback. Interviews were recorded, transcribed with Whisper AI [11], manually checked for errors, formatted, and coded. The interview guide is available in Appendix A.1.

11.4.2 Participants

We recruited 28 participants from a specialized agency. We targeted the general adult population within a radius of 50km of the lab. 13 participants identified as male, while the remaining 15 identified as female. The age range was 19 to 64 years, with an average of 36. Regarding their living accommodations, 18 participants reported living in an apartment, 8 lived in a house, and 2 occupied a room in a shared flat. 17 participants reported having at least one Smart Home device, and 14 of them have been using Smart Home technology for longer than 2 years. Participants that use Smart Home technologies have devices of an average of 5 smart devices categories (range is 2 to 12) out of an open list of 16 categories based on Home Assistant's physical entity types[12]. 11 participants indicated they would buy more Smart Home devices in the future, 9 were undecided, and 8 would need to inform themselves before deciding to buy more.

Regarding experience with AR technologies, 16 participants claimed to have no prior experience with HMDs. 10 participants had used AR HMDs once or twice, and 2 participants had used AR HMDs more than two times.

The ATI (Affinity for Technology Interaction) score resulted in a mean value of 1.579 (Range: $(-1.000, 2.889)$ on Scale $(-3, 3)$, with Cronbach's alpha: $.851$).

[11] https://github.com/openai/whisper (Accessed: 12.04.2024)
[12] https://developers.home-assistant.io/docs/core/entity/ (Accessed: 12.04.2024)

Table 11.1.: Result analysis: for each scale and condition, the calculated average and standard deviation, along the results of the Friedman and Bonferroni-corrected post-hoc tests. Highlighted cells denote $p < .01$.

Metric / scales	Manual (SD)	Semi (SD)	Auto (SD)	Friedman test	Posthoc tests
TENS: Competence	0.771 (0.794)	1.343 (0.461)	1.371 (0.454)	$\chi^2(2) = 10.358, p < .05$	Man-Semi, Man-Auto
TENS: Autonomy	0.993 (0.786)	1.264 (0.561)	0.879 (0.759)	$\chi^2(2) = 8.804, p < .05$	Semi-Auto
TENS: Relatedness	0.179 (1.215)	0.405 (1.101)	0.036 (1.225)	$\chi^2(2) = 6.0, p = 0.05$	
UEQ-S: Hedonic	2.545 (0.601)	2.795 (0.385)	2.562 (0.912)	$\chi^2(2) = 10.945, p < .05$	Man-Semi, Man-Auto
UEQ-S: Pragmatic	1.08 (1.247)	2.33 (0.532)	2.295 (0.704)	$\chi^2(2) = 29.22, p < .05$	Man-Semi, Man-Auto
UEQ-S: Overall	1.812 (0.846)	2.562 (0.351)	2.429 (0.712)	$\chi^2(2) = 20.058, p < .05$	Man-Semi, Man-Auto
Perceived Enjoyment	1.821 (1.679)	2.357 (1.224)	2.214 (1.397)	$\chi^2(2) = 4.351, p = 0.114$	
Perceived Performance	0.179 (1.765)	2.107 (1.37)	2.321 (0.983)	$\chi^2(2) = 36.026, p < .05$	Man-Semi, Man-Auto
Intention to use	0.679 (1.611)	1.607 (1.286)	1.464 (1.374)	$\chi^2(2) = 18.123, p < .05$	Man-Semi, Man-Auto
NASA-TLX: Mental	33.214 (29.193)	12.679 (16.244)	12.5 (19.65)	$\chi^2(2) = 15.918, p < .05$	Man-Semi, Man-Auto
NASA-TLX: Physical	29.821 (26.994)	10.0 (14.207)	4.464 (6.85)	$\chi^2(2) = 27.798, p < .05$	Man-Semi, Man-Auto
NASA-TLX: Time	25.179 (22.256)	10.179 (16.693)	14.643 (26.768)	$\chi^2(2) = 16.247, p < .05$	Man-Semi, Man-Auto
NASA-TLX: Performance	26.786 (24.578)	15.0 (20.0)	14.464 (26.223)	$\chi^2(2) = 6.977, p < .05$	Man-Auto
NASA-TLX: Load	30.357 (24.905)	11.429 (13.666)	7.143 (12.128)	$\chi^2(2) = 24.0, p < .05$	Man-Semi, Man-Auto
NASA-TLX: Frustration	30.536 (25.724)	13.214 (16.844)	16.071 (22.375)	$\chi^2(2) = 15.364, p < .05$	Man-Semi, Man-Auto
NASA-TLX: Overall	29.315 (21.998)	12.083 (13.812)	11.548 (12.903)	$\chi^2(2) = 20.434, p < .05$	Man-Semi, Man-Auto

11.5 Results

We analyzed the collected data using non-parametric Friedman tests since the assumptions of normality and sphericity for ANOVA were not met for all tests. In the cases where significant differences between conditions were found, we applied Conover's test with Bonferroni-correction for post-hoc analysis (Conover, 1999). The significance level was considered at the usual value of 0.05 for all tests. An overview of the results can be seen in Table 11.1.

11.5.1 Psychological Needs: TENS

The TENS-Interface (Competence and Autonomy) and TENS-Life (Relatedness) scales are measured using a 5-point Likert scale from -2 to 2. On average, participants reported medium to high values for perceived competence and autonomy for all conditions (see Figure 11.3). The Cronbach's alpha levels are 0.666 for competence, 0.687 for autonomy, and 0.938 for relatedness. The participants felt significantly more competent ($\chi^2(2) = 10.358, p < .01$) while using the semi-automatic ($M = 1.343, SD = 0.461; p < 0.05$) or automatic ($M = 1.371, SD = 0.454; p < 0.05$) variants compared to the manual alternative ($M = 0.771, SD = 0.794$). The participants felt significantly more autonomy ($\chi^2(2) = 8.805, p < .05$) when using the semi-automatic variant ($M = 1.264, SD = 0.561; p < 0.05$) over the automatic alternative ($M = 0.879, SD = 0.759$). Results for perceived relatedness to the home environment show even distributions without significant differences between the conditions.

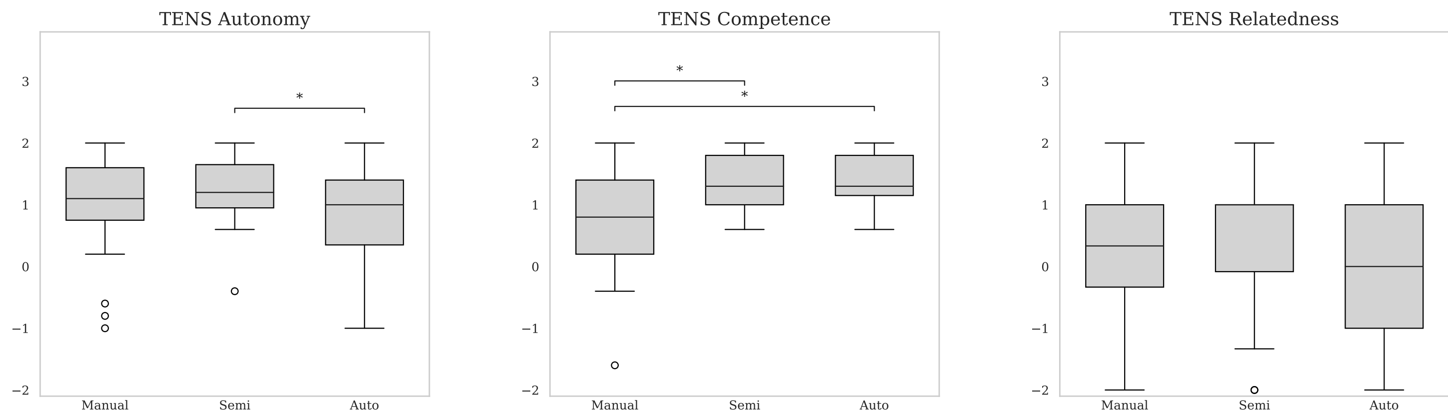

Figure 11.3.: Psychological need self-reports (TENS): Autonomy (left), competence (right), and relatedness scores (bottom) for the three interaction design variants. Boxes represent 25th percentile, median, and 75th percentile. Whiskers indicate $1.5 \cdot IQR$ (interquartile range) from the box borders. Outliers are denoted as circles. The p-values indicate the significance levels of the differences: $p < 0.05$ (*), $p < 0.01$ (**), and $p < 0.001$ (***).

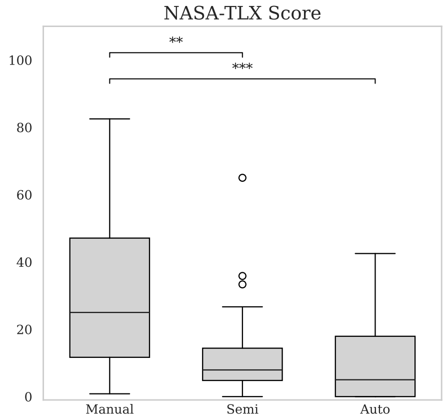

Figure 11.4.: Task load self-reports (NASA-TLX): Aggregated scores for the three interaction design variants. Boxes represent 25th percentile, median, and 75th percentile. Whiskers indicate $1.5 \cdot IQR$ (interquartile range) from the box borders. Outliers are denoted as circles. The p-values indicate the significance levels of the differences: $p < 0.05$ (*), $p < 0.01$ (**), and $p < 0.001$ (***).

11.5.2 User Experience and Task Load

The scores for the Task Load Index were significantly different between the manual condition and the semi-automatic condition (see Figure 11.4), except for the *performance* subscale. Cronbach's alpha for the task load was 0.901. On a scale from 0 (no load) to 100 (high load), the overall task load scores for the semi-automatic condition ($M = 12.083, SD = 13.812; p < 0.01$) and automatic condition ($M = 11.548, SD = 12.903; p < 0.001$) were significantly lower ($\chi^2(2) = 20.434, p < .001$) than in the manual variant ($M = 29.315, SD = 21.998$).

The UEQ-S is measured with a 7-point Likert scale, with values between -3 and 3 (see Figure 11.5). Cronbach's alpha is 0.813 for the *hedonic* and 0.716 for *pragmatic* subscales. The collected UEQ values are consistently high for all items across all conditions. The overall UEQ-S score is significantly higher ($\chi^2(2) = 20.058, p < .001$) for the semi-automatic ($M = 2.562, SD = 0.351; p < 0.001$) and automatic variant ($M = 2.429, SD = 0.712; p < 0.01$) compared to the manual alternative ($M = 1.812, SD = 0.846$). While both semi-automatic and automatic options have significantly higher scores on the *pragmatic* and *hedonic* UEQ-S subscales, only the *pragmatic* scores show a relevant difference. The *hedonic* user experience is rated very high for all three conditions ($M > 2.5$). Notably, all 28 participants rated the semi-automatic experience with the highest score for the decision between "usual" and "leading edge".

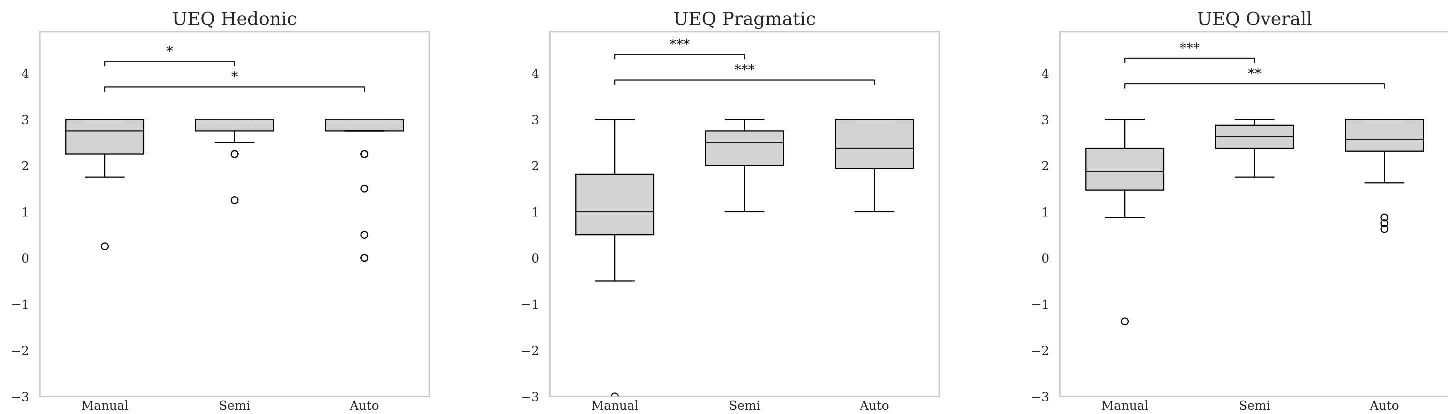

Figure 11.5.: User experience self-reports (UEQ): Hedonic (left), pragmatic (right), and overall (bottom) scores for the three interaction design variants. Boxes represent 25th percentile, median, and 75th percentile. Whiskers indicate $1.5 \cdot IQR$ (interquartile range) from the box borders. Outliers are denoted as circles. The p-values indicate the significance levels of the differences: $p < 0.05$ (*), $p < 0.01$ (**), and $p < 0.001$ (***).

11.5.3 Perceived Enjoyment, Performance, and Intention to Use

The collected values for *perceived enjoyment* were high overall, with $M > 1.8$ on a scale with range $[-3, 3]$. We found no significant differences between the scores for each condition. In terms of *perceived performance*, the collected data shows a significantly higher value ($\chi^2(2) = 36.026, p < .001$) for the semi-automatic variant ($M = 2.107, SD = 1.370; p < 0.001$) and automatic variant ($M = 2.321, SD = 0.983; p < 0.001$) when compared to the manual alternative ($M = 0.179, SD = 1.765$). Regarding the *intention to use* of the presented technology the collected data shows a significantly higher value ($\chi^2(2) = 18.123, p < .001$) for the semi-automatic variant ($M = 1.607, SD = 1.286; p < 0.001$) and automatic variant ($M = 1.464, SD = 1.374; p < 0.01$) compared to the manual alternative ($M = 0.679, SD = 1.611$) but no significant difference between automatic and semi-automatic were found.

11.5.4 Interviews

A *thematic analysis* was conducted on the data collected in the interviews, using an *inductive coding* approach (Braun & Clarke, 2006). In total, two and a half hours of audio-recorded interviews were transcribed (total duration: 02:27:26, average duration: 00:05:16, SD: 00:02:34). Two researchers coded 6 of the interviews independently (ca. 20% of the total), sampling interviews randomly. Duplicates were expelled, and a final coding tree was jointly developed and refined through an in-depth discussion of results. Subsequently, one researcher coded the rest of the interviews. Based on the coding tree, the following six overarching categories were identified, comprising a total of 18 themes. Figure 11.6 shows the distribution of the occurrences of each category and theme. In the following paragraphs, we summarize the categorization and provide exemplary quotes for each of the themes.

First Impressions

A subset of the material is related to the initial impressions of participants when interacting with the prototype. Many participants ($n = 18$) emphatically expressed a strong enjoyment of the interaction of using AR HMDs to set up a Smart Home. Within this group were present both experienced and novice users of AR. We called this theme **Wow Factor**.

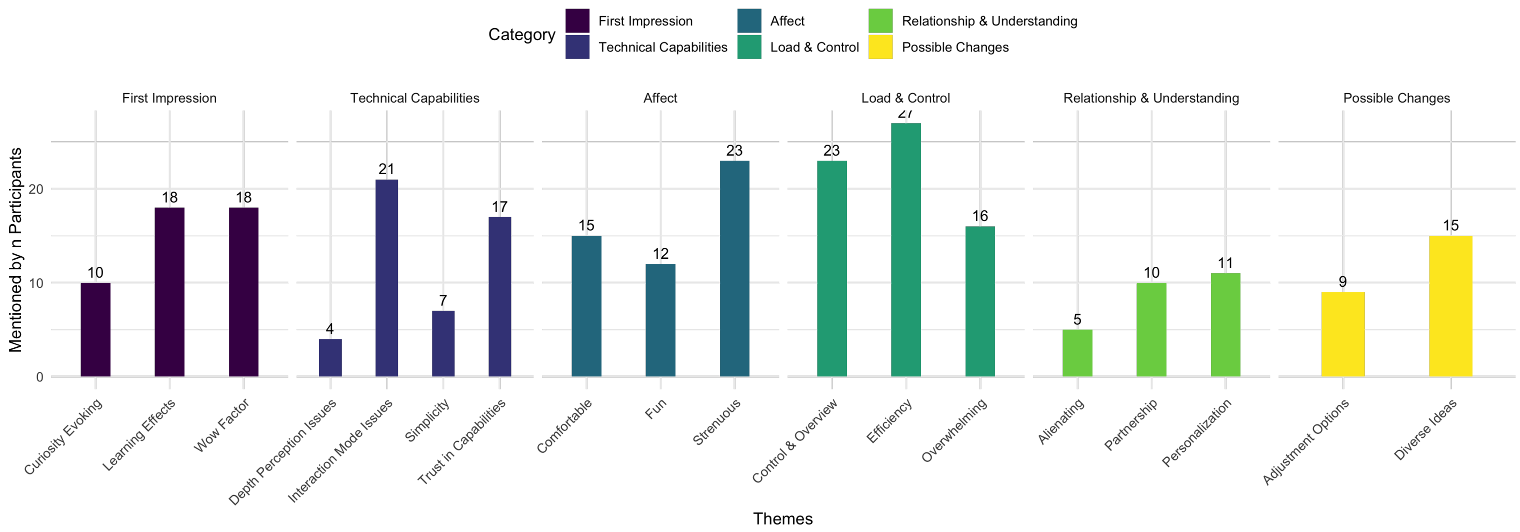

Figure 11.6.: Occurrences of themes in the interviews. Bars and corresponding numbers on top refer to the number of participants mentioning these themes. Themes are clustered in categories by color. The codebook can be found in Appendix A.2.

"So, in general first. It was definitely a very interesting experience, to be honest. And it's truly impressive what's possible and how it might actually look in the future." – P28

Extending these thoughts, we defined a theme as **Curiosity Evoking**, for statements about how the interaction mode evoked curiosity and exploration, in order to get to know the system and the Smart Home environment ($n = 10$).

"I wanted to try it out. I just looked to see what would happen. And then, after, I don't know, what did I click, I had seven or eight lights, so I clicked on it quite late and thought I'd give it a try." – P19

At the same time, a repeatedly occurring theme was the need to learn how to properly use the HoloLens in the setup process ($n = 18$). We classified this as **Learning Effects**. Importantly, participants stated that initial challenges with the interaction could be overcome quickly within the time of these first interactions – or that they believed additional practice would surely enable them to use the system well.

"I had to first get used to what the device wanted from me. And practice that. It's a matter of practice for me." – P21

To that end, participants repeatedly remarked about some initial difficulties with the interaction mode ("the pinching" motion for positioning the bulbs in the room was sometimes mentioned as error-prone; $n = 21$) and errors in the manual positioning due to depth perception conflicts ($n = 4$) where they thought they had placed a bulb at a further location that was later revealed to be incorrect but not visible from the initial vantage point. However, no participant considered these challenges as a major issue, but rather an annoying nuisance emerging during the first moves with the manual configuration of the Smart Home.

Technical Capabilities

The technical capabilities of the prototype were another recurring topic. As described above, some users declared experiencing **Depth Perception Issues** ($n = 4$), indicating that the UX suffers detrimental effects caused by technical limitations.

"The problem was in the depth, but also somehow the position in the room in general. So the perspective didn't always quite fit." – P7

We did not measure the offset between the actual placement in the manual task and a potential *correct* position. The correct placement is partly of subjective character as users have to choose where they want to interact with the device. However, all participants placed the spheres near the correct device without exception.

Many participants expressed having experienced **Interaction Mode Issues** ($n = 21$). In particular, "the pinching" motion for positioning the bulbs in the room was often mentioned as error-prone. General detection of gestures by the HoloLens seemed to be a recurring issue:

> *"So sometimes it didn't work right away to bring up the menu, or bringing up the menu worked, but then tapping on it didn't." – P20*

In contrast, the automatic detection features were generally considered to function well and smoothly. Many participants reported having **Trust in the Capabilities** of the system ($n = 17$). As we used a Wizard-of-Oz study method, we should point out that this trust in the system's capabilities is likely underlying other impressions about the usability and preference for interaction modes.

> *"In hindsight, I did think, okay, what if something goes wrong. But I felt, or I got the impression, that it then found things well. Yes, so I would trust the system." – P11*

The **Simplicity** of the interaction (for the fully automatic variant) was also highlighted by participants as a positive feature ($n = 7$).

> *"I found it quite exciting to see how fast some things can happen, how everything is captured automatically." – P22*

Affect, Load & Control, Relationship & Understanding, and Diverse Ideas

Beyond these more general observations, the remaining emerging themes are best discussed in connection to the different conditions. Here, especially the manual and automatic characteristics of the setup were contrasted by the study participants.

The manual aspect of the home configuration was often appraised as playful and **Fun** ($n = 12$), often mentioned together with the curiosity about the system's functioning (see above), and an interest in feeling an achievement through the setup process (that is not given by automatic configuration) or a sense of **Personalization** ($n = 11$) connected to the setup of the Smart Home.

"I also liked the manual version because it has this certain playful aspect to it, and honestly, you don't set up new devices that often." – P22

"Well, I believe the version where I can set it up myself is just more individualized." – P4

Similarly, the advantage of staying in **Control** and keeping an **Overview** of the process was mentioned ($n = 23$).

"I did a bit, walked around the apartment a bit. I felt responsible for the setup, but didn't have to do everything myself." – P18

However, another major theme for the manual setup was its **Strenuous** and demanding nature ($n = 23$).

"It was just frustrating with the whole setup of the individual devices." – P17

This was mentioned as the major downside of the manual setup experience, together with its low level of **Efficiency** ($n = 27$). For example, several participants raised doubts about the utility of a manual setup if it were employed for many Smart Home devices or repeated setups. In contrast, the automatized setup features were mostly appraised as delivering high **Efficiency**.

"Of course, the most efficient is the automatic version. I walk through the room, and the thing is done. I don't really have to choose anything; I don't have to make any decisions." – P7

As participants reported high trust in the system's technical capabilities, this setup mode appeared to many as the quickest and easiest way to process the task. However, in the fully automatic condition especially, participants described the experience as **Overwhelming** ($n = 16$), losing their overview of the configuration or experiencing an **Alienating** ($n = 5$) sensation as the system takes over the task completely.

"I think, for example, I would not recommend this to my mother; she would probably freak out if something like this happened in her apartment." – P18

"It was also a bit strange, especially when all the things started to light up or draw attention to themselves." – P28

Between these two extremes, the majority of the participants appraised the semi-automatic condition as the best of both worlds. This is reflected in the identified preferences for either condition (see Section 11.5.5). However, we believe that this preference is not merely emerging as a consensus between the two approaches but rather as a productive integration of nuanced aspects of them. We observed that participants often mentioned a preference for combining both manual and automatic processes, and also benefiting from reduced levels of both aspects, resulting in a more **Comfortable** interaction ($n = 15$).

> *"The second one [the semi-automatic condition] was the most relaxed, I could pick a few devices that I want and the rest is done automatically." – P5*

For example, the group-wise setup process was often appraised as providing a necessary overview that brings users "on board" with the partially automatic configuration, through which a sense of cooperation and **Partnership** emerged ($n = 10$).

> *"There, I just have the feeling of having accomplished something and having contributed, and the device doesn't do everything on its own." – P7*

In this spirit, we also want to highlight that participants discussed related **Diverse Ideas** ($n = 15$) for integrating the features from the three conditions further and did not just declare a preference for one over the other. For example, participants remarked that further gamification of the manual approach would be interesting or that the **Adjustment Options** ($n = 9$) of choosing the setup approach based on mood, time pressure, or user in the household would be beneficial over employing just one of the modes. Furthermore, it was a recurring theme that extending the system to allow the opposite order (automatic first, manual adjustment second) would be a vital feature.

> *"So, if there were, let's say, a game module included, where I could participate in some AR gaming situation with the glasses, okay, that would surely be great." – P21*

> *"Yes, then it would be good if you could adjust it a bit." – P13*

11.5.5 Actions & Preferences

The behavior of the participants during the experiment was recorded. Within the 12 minutes of the manual condition, we observed that participants placed 20.93

entities on average ($SD = 5.74$). In the semi-automatic condition, participants were given the option to choose to parallelize the detection of the rest of the device category (i.e., lights and speakers) after each device detection. 15 participants chose to automate all remaining lights after one detection, 6 participants tried up to 5 individual detections, and 7 participants performed between 6 and up to 20 individual detections.

Overall, 21 participants (75%) stated a preference for the semi-automatic alternative over the two other variants, followed by 6 participants (21.43%) favoring the automatic option, and 1 participant (3.57%) preferring the manual option. For most participants, the automatic alternative made second place (60.71%) and the manual alternative last place (82.14%). 27 participants (96.43%) rated the automatic version as the most efficient option. One person rated the semi-automatic version as the most efficient one with the comment that they would individually check and correct each position after using the automatic variant and, thus, require more time than with the semi-automatic alternative.

11.6 Discussion

Our mixed-method results provided rich insights into the anticipated trade-offs between designing a Smart Home AR setup and classical UX dimensions (RQ1-3). Importantly, beyond our research questions, we identified valuable findings through the design exploration. To provide structure to the discussion of our findings, we group the themes as follows:

11.6.1 Psychological Needs & UX Trade-Offs

The data collected during the interviews combined with the answers to the TENS questionnaire and the *intention to use* suggests that *perceived competence* and *autonomy* may have a role in the preference rating between the three interaction variants.

Here, it is not possible to exclude technical limitations being an additional factor in this equation. The HoloLens 2 offers a limited Field of View (FoV): 43° horizontal FoV and 29° vertical, roughly a third of human typical vision (Howard & Rogers, 1995). This constraint becomes particularly challenging for hand interaction, since gestures must be consistently performed within the HoloLens cameras' FoV. Especially in this room-scale application, this can lead to significantly higher levels of frustration, lower pragmatic user experience ratings, and also have an impact on perceived competence.

Both the semi-automatic and automatic variants were rated with overall low load and high user experience scores, confirming our expectations for RQ2 and RQ3, that these more automated variants would lead to better classical UX experiences (whereas the manual condition showed poorer UX perceptions as outlined in RQ1). Further, both the semi-automatic and automatic variants scored high levels of *perceived competence*. This is aligned with the preferences stated explicitly by the study participants, who largely prefer these two variants over the manual option. Overall, we were a bit surprised about the lower levels of competence in the manual design variant as we expected higher psychological need satisfaction in the manual condition overall (RQ1). It appears, that our participants did not experience the manual setup as competence building, possibly because of some initial challenges with learning the controls, and also because the setup progress was fairly slow. While anticipated differently, this does potentially highlight the trade-off that high control can undermine competence needs if it slows the user in achieving their tasks.

Furthermore, the manual and semi-automatic variants showed similarly high levels of *perceived autonomy*, showing that an effective balance between automation and manual control can be achieved that still acknowledges autonomy. This observation further supports our expectation that a more manual variant would increase psychological need satisfaction (RQ1), at least for the autonomy dimension. Also, comparing the automatic against the semi-automatic variant, the *perceived autonomy* metric suggests that users value being involved in the interaction. This is attested by the preference for the semi-automatic variant over the more efficient automatic alternative, where users are passive observers. Thus, it is possible to argue that in this particular case, the fulfillment of psychological needs has precedence over pure functional effectiveness or efficiency. Most importantly, this result epitomizes the expected trade-offs for a fully automated setup variant (RQ2) and affirms our consideration in RQ3 that a combination of manual and automatic features could strike a more effective balance of psychological need fulfillment and classical UX design considerations.

The exploratory use of the *relatedness* subscale with the alternated subject of the Smart Home instead of other people did not show significant differences between the conditions due to a large variance in the ratings. Interestingly, the interviews provided context to this variance, since different lines of thought between participants can be reconstructed. On the one hand, some participants reported a strong feeling of *connectedness* to the Smart Home environment through the immediate and direct interaction with it (evidence that would support the expectation of RQ1 that a manual interaction could create stronger need fulfillment). This is even more remarkable considering the setting of the experiment being a lab inside a remote corporate complex. On the other hand, some participants felt disconnected from reality by using the AR HMD:

> *"The screen creates a distance. At the same time, you are in the middle of it, but like in another world. So, to me it is a different reality." – P24*

Of course, it remains to be explored if this effect is temporary and may fade away once the user gets used to AR. This prompts a further, more intriguing question about the nature of the relationship between users and Smart Homes mediated by AR. Combined with an increasing level of agency in Smart Homes and artificial intelligence applications, a high level of *connectedness* can result in dramatic changes in how people conceptualize homes.

On a more general note, the discussed results highlight the importance of psychological needs when considering factors for AR and Smart Home interactions. Application designers must be sensitive to the potential diverse emotional and social effects of AR (Ariso, 2017; Slater et al., 2019), especially in home environments.

11.6.2 Wow-Effect: Novelty and Ceiling Effect

In the accompanying handbook for the UEQ Scale [13], the authors warn that it is unlikely to observe any average score above 2 due to different opinions and people's tendency to avoid extremes. Yet, the semi-automatic ($M = 2.562$) and the automatic version ($M = 2.429$) are well beyond this threshold. Additionally, participants characterized the experience as fun, futuristic, exciting, or fascinating, and 18 participants explicitly described the interaction as a great experience overall. We relate this to both novelty and *ceiling* effects, as no participant reported having experienced a similar AR application before. Although AR applications have been used and studied for decades, the particular application of configuring a Smart Home seemed to be particularly attractive to the study's participants. This perhaps underlines the potential for AR to establish a close connection between the user and their surrounding.

However, we cannot eliminate the possibility of positive bias caused by the experience of participating in the study at a modern research facility, or by the relatively high compensation.

[13] https://www.ueq-online.org/Material/Handbook.pdf (Accessed: 12.04.2024)

11.6.3 All Alternatives Have Their Benefits

Different characteristics of the design variants make them interesting for users, even if the overall variant is not their first choice. This is supported by participants' statements during the interviews. The automatic variant is attractive due to its efficiency, with many participants being torn between this option or meeting their psychological need for autonomy and competence through the semi-automatic alternative. The manual variant's potential for *gamification* was mentioned by 12 participants during the interviews. While some stated that the *gamification* character is not important to them, all participants who mentioned this characteristic stated that it is either important to them or to another family member. Further, it was suggested that in the case of the setup process taking longer, the fun character should be emphasized for an overall better experience.

The choice of the optimal solution will depend on the circumstances of the interaction while performing a given task in a given situation. These circumstances may pose different time constraints, different expectations towards duration and playfulness, and different expectations towards the accuracy of positioning or *tidiness*, thus shifting the weight from one factor to another. This is supported by the statements recorded during the interviews. For example, 9 participants stated their interest in personalizing device positions after the automatic placement, and one participant even took the time to meticulously check the position of each entity after the automatic configuration.

Regardless, based on the gained insights, we can formulate some recommendations for future iterations of this application. The most important is to *keep the user in the loop*. It is paramount to give the user options about the degree and type of automation, and include options to adjust positions after placement. When providing fully automated placement, the process needs to be made visually transparent, make the user feel in control, and eventually offer the user to control or monitor the first few devices to understand the process.

11.6.4 Use Cases of the Setup Process

The proposed system presents clear benefits for the initial setup of multiple static devices, since the automated solutions can save significant amounts of effort and frustration. In the future, the system could allow to easily update the position of movable devices and notify the user if a device changed its location (e.g., based on wireless signal intensity). Additionally, during an initial setup, the AR HMD could visually record the position of devices and automatically detect them at the new position via image detection mechanisms. Dynamic devices capable of self-tracking, such as vacuum-cleaning robots, can be synchronized with the HMD aligning their coordinate systems and then providing live position updates.

11.6.5 Future Use Cases of AR in the Smart Home

The information about the location of connected devices within a Smart Home can enable further applications well beyond the scope of our proposed design. We envision AR applications controlling not only individual entities but complete groups of entities in direct interactions. Further, interaction can simplify lengthy or complex tasks through automatic grouping of entities using different criteria (e.g., type of device, location in a given area, user preference, etc.). This can be further extended using artificial intelligence to create dynamic filters or the automatic creation of routines. This allows to, for example, toggle lights when entering a room or run a specific service when in the proximity of a device (Schenkluhn et al., 2023b). Finally, this can enrich the user experience in households with multiple members, empowering individual users to create both personalized and collective experiences.

11.7 Limitations and Future Work

As stated before, the HoloLens 2 hand-tracking FoV and quality present a clear constraint for the proposed interaction. This problem can be addressed using downwards-facing cameras, as in the Apple Vision Pro. This device will likely improve the issues faced by the participants of our study. Furthermore, we did not measure the actual performance of the light detection implementation. As the HoloLens 2 does not have state-of-the-art sensors and cameras, performance metrics would not be representative of this approach. Still, detecting IoT device positions based on tags or even a precise ultra-wide-band indoor positioning solution is likely faster than the approach presented in this paper. However, when including the time required for setting up and calibrating such a system, we argue that our approach is faster, less error-prone, and more user friendly.

Another important limitation is that our approach only works for devices capable of attracting attention. Lights, blinds, audio devices, fans, or anything with a display can be instrumentalized to emit an identifiable signal. Many large home appliances, such as ovens, washing machines, or hood vents can become detectable. However, some devices can only remain silent and still, making their identification by our system more difficult. We see this challenge as hard to overcome but also of relative criticality: our system captures a large range of Smart Home devices, and especially those that come in large quantities (e.g., lights).

A further limitation to consider is the context of the study. Despite the high score of *connectedness* that some participants reported, the study was conducted in a lab setting. This aims to replicate a modern flat with many Smart Home devices, but it remains a foreign place for the study participants. A field study in actual home environments could offer a higher validity and deeper insights that could become visible only in such an environment.

Here, it is important to highlight the exploratory nature of the study. Future studies should look into long-term usage, as well as the incorporation and assessment of further functionality (e.g., adjusting placement of automatically positioned devices, automatic grouping, and incorporation of artificial intelligence elements).

Finally, this study was conducted using a Wizard-of-Oz technique to present the participants with a credible interaction. Although our prototype is capable of detecting lights on a per-device basis (similar to the semi-automated option), we plan to implement and test a fully parallelized automated version in the future.

11.8 Conclusion

In this paper we investigated two main topics: firstly, we proposed a solution for the spatial configuration of Smart Homes using AR, developed a prototype with a basic functionality, and evaluated the concept through a controlled experiment. Secondly, we investigated the effect of psychological needs, specifically *autonomy*, *competence*, and *relatedness*, as a factor of user preference for interaction design.

In the conducted user study, participants performed the task of setting up Smart Home devices spatially using an AR HMD. The task was performed under three different conditions: manual positioning, semi-automatic positioning, and automatic positioning, which we compared towards their support of psychological needs and classical UX dimensions.

The collected data indicates a general preference for the semi-automatic positioning method, despite the automatic alternative being faster and more efficient. The participants' statements recorded during post-participation interviews suggest that this preference stems from their psychological needs being best addressed by the semi-automatic variant. This is aligned with the reported TENS scores for *autonomy*, *competence*, and *relatedness*.

Additionally, the interaction design proposed for the configuration of Smart Homes was received positively by the participants. Supported by the collected data, this suggests that our technique for locating Smart Home devices is a viable alternative to typically manual approaches.

Based on the feedback collected through interviews and further insights obtained through the analysis of the quantitative data, we derived some recommendations for future applications in similar contexts.

Part IV

Finale

12 A Metaverse Future of Human-Computer Interaction

Communication has been key in human relationships for centuries. The way we communicate has changed over time, from stories that have been passed on over generations, to the first written letters, the first phone call, and from the first email to the first video call to interactions through avatars in virtual worlds. However, communication with machines is fundamentally different from human communication. At first, machines appear idiosyncratic, arbitrary, and stubborn. Yet, they follow strict and predictable rules. As soon as a user deviates only marginally from the rules or is imprecise in one's choice of words, the user is misunderstood and the results deviate from what was expected and what most humans would have understood correctly instead. One historical example is the mix-up between metric and imperial units by the programmers of NASA's Mars Climate Orbiter, resulting in its disintegration upon entering the Martian atmosphere (NASA, 1999).

Unfortunately, there is no universal language for humans to learn to understand machines. Behind every machine, there are humans who have programmed it. Thus, every machine and every application potentially speak a different language and follow a distinct set of rules. This diversity poses a challenge, especially for novice users who are not familiar with the specifics of software that is new to them.

Past efforts have led the communication interface between humans and machines to evolve from machine code to assembly, to high-level languages, and graphical user interfaces (Myers, 1998; Roller, 2022). No-code platforms enable users to program without writing a single line of code (Sufi, 2023). With the advent of large language models, machines accommodate humans even further by learning and thoroughly understanding their language. Taken even further, future Brain-Computer Interfaces (BCIs) aim to establish an immediate link between machines and the human brain rendering the need for explicit language obsolete (Jarosiewicz et al., 2015). Breaking this language barrier is one of HCI's visions of the future (Quigley et al., 2013).

Text entry, a fundamental problem of HCI and one form of human-computer communication, would not be necessary anymore. Text could be extracted directly from the user's thoughts and intentions instead of from the strenuous process of typing on keyboards, touch screens, or even in mid-air. Search queries would not need to go through voice assistants that might give the right answer, or misunderstand the user's request leading to a back-and-forth of clarifications and corrections to iteratively elicit the correct information from some corner of the internet.

Until we get to this point – which will introduce its own sets of ethical concerns (Burwell et al., 2017) – system designers and developers need to make approximations and compromises to bridge the gap between human and machine language. Norman (2013) describes seven stages of action that a user goes through when interacting with a system. It outlines the discrepancy between the user's intentions, their subsequent action, and the actual system response. Deriving the user's original intention from their interaction with the system is a key challenge (Sadikov et al., 2010; Xia et al., 2018; Zha et al., 2010). As users act differently, the system's response needs to be adaptive and account for the context.

Another vision of HCI is calm, ubiquitous, and invisible computing (Quigley et al., 2013; Weiser, 1991). Instead of technology being the center of attention and seeking excitement, it should be unobtrusive and blend into the background (Quigley et al., 2013). To attain this goal, Weiser (1991) describes the need for "cheap, low-power computers that include equally convenient displays, software for ubiquitous applications, and a network that ties them all together". It could be argued that smartphones, smartwatches, and tablets already fulfill these requirements. With processing power that outperforms the Saturn V's Apollo Guidance Computer that steered humankind to the moon by several magnitudes (Kendall, 2019), bright, sharp, and colorful displays, and network technology such as 5G that connects billions of devices via the internet in our pockets, the vision appears to have become reality in recent years. However, the availability of information and communication is still tied to handheld devices and physical screens, and the software is neither made for ubiquitous applications nor is it unobtrusive.

This discrepancy is addressed by AR. Ubiquitous AR describes a vision of the future where information is available everywhere and at any time, and where the digital and physical worlds are seamlessly intertwined (Azuma, 2017; Milgram & Kishino, 1994). AR is not a new concept (Sutherland, 1965). Yet after decades of research (Billinghurst et al., 2015), product launches (Apple, 2024a; Leap, 2024; Microsoft, 2020; Weidner, 2023), and unmet expectations and setbacks (Metz, 2022; Nunes &

Arruda Filho, 2018) along the hype cycle (Gartner, 2018; Herdina, 2020), AR is on the verge of becoming feasible in everyday life.

AR could represent the next evolutionary step in HCI. It addresses the issues of today's technology by offering continuous access to information, everywhere and non-stationary, hands-free, context-aware, and through natural interaction.

However, AR will likely be embedded as one technology in a comprehensive set of technologies that form the future digital landscape. AR, while beneficial as a standalone platform offering isolated applications, reaches its full potential when integrated with other technologies, serving as a comprehensive interface to a ubiquitous digital world. This future form can also be described along the Virtuality Continuum (Milgram & Kishino, 1994). On one hand, fully virtual worlds that exist independent from the physical world can be accessed and experienced through VR. On the other hand, the integration of AR with digital services and physical IoT devices adds additional layers to the physical world, which are adaptively available to users. This concept is known as the metaverse (**Rosedale2008**; **Zuckerberg2021**).

While numerous technological challenges remain to be surmounted in order to realize a vision of the Metaverse, the architecture of the Metaverse is particularly dependent on governmental and corporate policy decisions. Therefore, the following section addresses four different potential architectures of the Metaverse and the interoperability among different parties within a future Metaverse.

12.1 The Metaverse is not an Island

This section comprises the version of record of the following article: Peukert, C., & Schenkluhn, M. (2023). Das Metaverse ist keine Insel. *Wirtschaftsinformatik & Management*, 1–10. The initial idea for the metaphor of an archipelago was conceived by Christian Peukert and further developed and written as a joint effort with equal contributions. Changes include formatting, numbering of the research question and chapters, minor changes for consistency, and correction of spelling errors. Additionally, the document was translated from German to English. Reproduced with permission from Springer Nature.

Since Facebook's rebranding to Meta, the concept of the metaverse has attracted considerable attention both in academia and industry (Meta, 2021). Fuel to the debate about the metaverse is not only added by analysts attributing remarkable economic potential to it but also by statements from various technology companies, describing the metaverse as the successor to the internet as we know it today. Despite this attention, there is currently no consensus on what exactly the concept of the metaverse entails, and speculation abounds as to what form it will ultimately take (Ravenscraft, 2023).

To characterize the metaverse, reference is often made to various features. Ball (2021), regarded as one of the most notable pioneers of the metaverse, attributes characteristics such as persistence and synchrony, interoperability among multiple 3D worlds, a massively scalable network of worlds, a virtually unlimited user number, real-time rendering, and a deeply immersive experience to the metaverse. It has long been emphasized that the metaverse cannot be a single, albeit highly immersive virtual world but must instead be a nearly limitless network of interconnected virtual worlds (Dionisio et al., 2013). Additionally, the merging of physical with Virtual Reality is sometimes cited as another characteristic (Lee et al., 2021b).

Strictly speaking, to this date – at least based on these characteristics – there is no metaverse in the true sense, as no application can meet these requirements. Especially the key characteristic of *interoperability*, the general ability to link individual components of the metaverse, is largely unmet on several levels.

12.2 Understanding the Metaverse as an Archipelago

To make the concept of interoperability in the context of the metaverse more tangible, we will illustrate its development using the metaphor of an archipelago and show potential scenarios for the interaction of different islands. Interoperability can be considered across various dimensions, including *platform, inventory, identity, game engine, and end devices*. Table 12.1 provides an overview of these dimensions and their respective properties. The compatibility of islands across these dimensions then defines the interoperability of the archipelago.

Before examining the interaction between various islands and thus the archipelago, it is crucial to develop a basic understanding of the properties of individual islands.

Table 12.1.: Dimensions of Interoperability, Characteristics, and their Metaphorical Representation in the Archipelago

Dimension	**Characteristics** (Example)	**Archipelago** (Example)
Platform (World, Laws, Multiplayer, etc.)	Self-hosted/Hosting by Provider, Singleplayer/Multiplayer, PvE/PvP, Physics, etc.	Island State (Island and Laws)
Inventory	Avatar, virtual Items, Land register entry, etc.	Property and Goods
Identity	Account, Self-Sovereign Identity, etc.	Citizenship
Game-Engine	Unity Engine, Unreal Engine, Minecraft, etc.	Language
End device	VR, AR, Smartphone, Desktop, etc.	Means of Transport

12.2.1 Describing one Island among the Archipelago

An island embodies various properties along the aforementioned dimensions. For example, an **island state** comprises the island itself – the landscape – and the underlying legal system. The size of the island determines the number of residents and the general shape of the island, i.e., what the world on the island looks like. Through the legal system, laws can be established, which determine the range of actions available to residents. Projected back onto the metaverse, an individual island state represents a platform provided by a provider. Examples of platforms include Facebook's Horizon Worlds, The Sandbox, any computer game, or an immersive e-commerce shop, which may differ by the rules applicable therein. By registering an account with the platform provider, one can acquire **citizenship** for the island. Moreover, each island has *one* **language**, representing different game engines needed to display 3D worlds. Additionally, each island has a transport network allowing various **means of transport** to explore different parts of an island. In relation to the metaverse, these transport modes can be analogized to the end devices through which users access the metaverse. Immersive technologies like VR and AR, along with traditional device types such as mobile devices and desktops, could serve as gateways to metaverse experiences, and like in the metaphor, all or only selected devices might support different experiences. Ultimately, life on an island is made possible by its citizens, who can rely on their **inventory**, reflecting the individual property of users. This may include assets represented in the island's currency and various objects, from virtual clothing to land property. However, land properties typically play a special role as they are usually bound to the island and cannot be moved arbitrarily.

Until now, the focus has only been on describing a single island. As long as an island remains autonomous, life can thrive in this closed ecosystem. However, if islands seek connections to others, either voluntarily or involuntarily, through the construction of bridges, for example, numerous interoperability questions for these previously closed ecosystems arise.

12.2.2 The Current State of the Archipelago

Numerous closed island states or "Walled Gardens" exist today. Even if it is possible to speak the same language on different islands or navigate with the same means of transport on more than one island, other citizenships are rarely recognized, and the transfer or even trading of property across island borders is simply not possible. The metaverse's potential only unfolds by opening these borders, meaning current applications cannot be considered metaverses or *the* metaverse: *The metaverse is not an island*. Overcoming these borders is primarily not a technical hurdle but a political decision of the respective islands. Blueprints for building bridges between islands exist or could be developed by mutual agreement. The various languages could be standardized and translated into one another, and the transport network could be designed to allow access to all means of transport. The acceptance of other citizenships is likewise conceivable and desirable from the citizens' perspective to make it easier to travel to new islands. Therefore, overcoming these borders represents the potential currently attributed to the metaverse's hype.

However, the islands pursue their strategic intentions individually, which partially counteracts the general and comprehensive opening of borders. For this reason, various developments are conceivable, which will be illustrated in the following. In particular, we will focus on development scenarios that can be considered potential cornerstones.

12.2.3 Future Scenarios for Interlinking the Archipelago

The transfer and trading of property represent significant logistical challenges for islands, hence the focus of the next section will particularly be on inventory management. It is essential for goods to be transported between islands easily, without damage, and exempt from duties, while also preventing their duplication or theft. Thus, the organization of property becomes a question of trust in the system and its participating institutions. Figure 12.1 presents four potential development scenarios for the future, which will be discussed in more detail below.

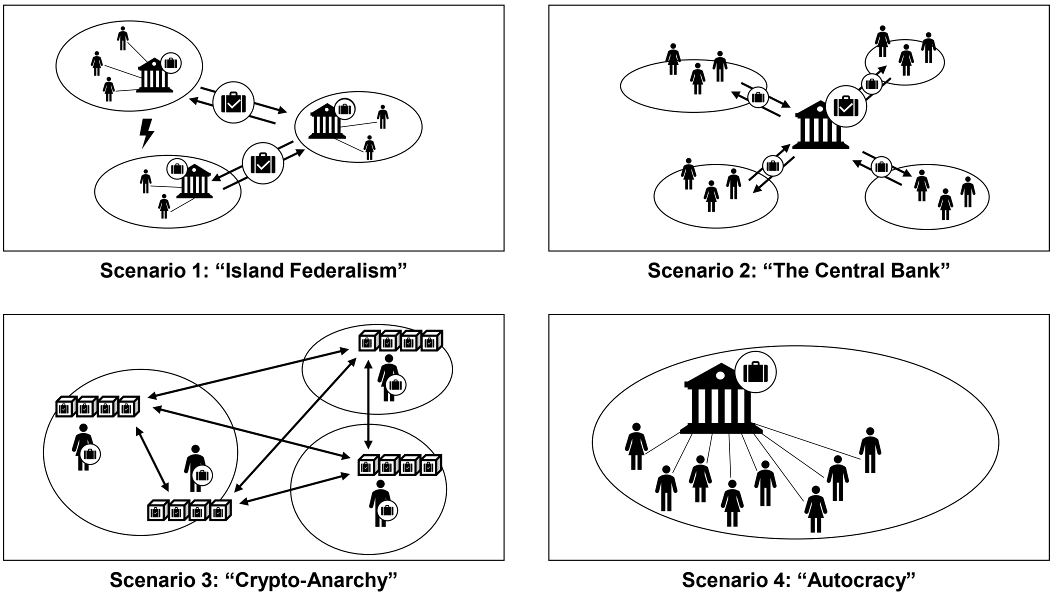

Figure 12.1.: Development Scenarios for the Interoperability of Property in the Metaverse

Scenario 1: “Island Federalism”

In this first scenario, various islands network with each other to facilitate the provision and movement of property. This could be achieved bilaterally for a smaller number of islands, but standardization can be an option as the number of islands increases.

Implications for Interoperability: Each island can regulate the production of goods differently. The islands determine themselves with whom they want to cooperate. However, they must either trust each other or use mechanisms that create trust, so malicious actors cannot secretly duplicate limited goods. Goods produced on one island could be taken to others, while the management of property remains the original island’s responsibility. Such restrictions can vary between islands. Since bilateral agreements are required, this scenario is only suitable for linking a few islands. The storage of goods occurs centrally on the individual islands, requiring appropriate administrative authorities. Furthermore, newly founded islands must establish diplomatic relations with all existing islands or groups of islands.

Implications for Users: Citizens depend on the islands’ decisions regarding the management and provision of property. Moreover, if an island is abandoned, access to the property could be lost. To visit non-cooperating providers’ islands, multiple citizenships with the respective property management may be necessary, and objects might need to be purchased multiple times for use on each island.

Scenario 2: "The Central Bank"

In this scenario, a regulated entity, the "central bank," manages the property of all citizens of the archipelago and thus all transactions between the islands. Thus, a standardized system, decoupled from individual island states, is created, linking all participating islands.

Implications for Interoperability: All goods and their property information are stored at this central institution and kept available at all times. Goods that are newly introduced into the archipelago or created and intended to be transferable must be registered with the central bank. It retains an overview of all goods and their owners, thus acting as the controlling entity. All islands cooperating with this central bank are closely interconnected and can effortlessly exchange goods. New islands wanting to use the central bank must accept and implement predefined cooperation rules. Rejection or sanctioning can have relevant consequences for islands. Moreover, a failure of the central bank would make the bridges temporarily impassable.

Implications for Users: Citizens must link their identity to the central bank for property management. This could include identifying with an island citizenship or a direct identification form of the central bank. Even if an island is abandoned, citizens retain access to their property (except for objects "permanently" tied to that island like real estate). Through its role, the central bank holds a powerful position, requiring the trust of the islands and their citizens. Thus, citizens' movements between islands could also be monitored at a single point. Therefore, this scenario suggests a transparent reporting system and consistent regulation.

Scenario 3: "Crypto-Anarchy"

In a decentralized Crypto-Anarchy[1] scenario, each citizen of the archipelago carries a briefcase marking their property. The interaction of numerous, distributed, and as independent as possible individuals ensures through a secure voting mechanism that the transfer and trading of goods works. Owners only need to trust the consensus procedure, not each other or other institutions.

Implications for Interoperability: For each creation and all transactions of goods, citizens vote on their legitimacy. Unlike the current use of these decentralized voting mechanisms, which occur only within the bounds of individual islands, this vote

[1]The term anarchy should be understood here in its original sense, in contrast to the widespread, colloquial concept of chaos and unrest. In contrast to "regular" anarchy, however, crypto-anarchy is still oriented towards capitalism.

spans all participating islands and their citizens. The vote is automated, so while citizens have a say, they do not need to invest time. This decentralized architecture secures the bridges against the failure of some participants. On the other hand, limiting the duplication of individual goods as counterfeit can be difficult, as their authenticity cannot be centrally and automatically verified. Counterfeits could represent a copy slightly deviating from the original, making similarity detection through machine processes infeasible. At the same time, for citizens, counterfeits would be indistinguishable from the original. A manual, democratic vote for every new good is unrealistic, so the decision must then be delegated to some trustworthy authorities. Moreover, at least trust in the organization of the decentralized system is necessary (e.g., in the developers), and in certain cases, manipulations of property relations are possible with great effort. Additionally, although this system regulates the allocation of goods to their owners, it does not ensure the actual storage of goods. Therefore, the provision of goods must also be ensured decentrally, which represents a significant effort. This is the only scenario where Non-Fungible Tokens (NFTs) traded on a blockchain serve as a foundational technology. Hence, NFTs are deployed not only within island boundaries, as is mostly the case today, but also beyond. Other technologies from the Web3 context are conceivable as well.

Implications for Users: Apart from a briefcase proving identity and property, no further citizenship of an island or institution is generally necessary. Depending on where the good is stored, owners have unrestricted disposal of the good, as no centralized entity is required for transactions and the proof of ownership. While owners of a valuable good are reassured that it belongs to them, others can plagiarize this good. Even if the corresponding NFT clearly shows who owns the original, these counterfeits could be identical in appearance and function, meaning the value remains purely ideational upon verification of originality. Moreover, transactions of each citizen across islands are visible and traceable to all others through blockchain. Overall, it remains unclear whether the technical complexity associated with the use of blockchain technology can be simplified for individual users as easily as a central institution could, without having to delegate rights and trust to another party.

Scenario 4: "Autocracy"

The final scenario describes the development towards a single platform that prevails over all others. This involves a drastic growth of this one platform and the continuous disappearance or displacement of all others. The island thus evolves into a single large continent, leaving no room for other associated islands.

Implications for Interoperability: In this scenario, the challenge of interoperability is negligible since all decision-making power is concentrated with the one platform operator, and there are no interfaces to other islands, as the continent has no special incentive for cooperation. For new islands, it would thus also be more beneficial to integrate into the ecosystem of the continent due to the network effect rather than founding a new island. A classic database suffices to ensure the property relationships of goods, and NFTs find no meaningful application concerning interoperability. The citizens' property management is fully realized by a central institution.

Implications for Users: For citizens, this development would be a mixed blessing. They benefit from the classic advantages of a closed ecosystem, where compatibility and uniformity are ensured by the provider. Among the negative aspects, foremost is the complete dependency due to the lack of alternatives. The monopolistic position of this provider can then lead to high transaction costs when trading goods and the creation of new goods on the continent being strictly regulated. Moreover, the lack of competitive pressure could lead to less innovation and lower user-friendliness. The provider can also comprehensively monitor all movements of citizens on the continent. Altogether, this scenario resembles the dystopian worlds from science fiction novels *Snow Crash* (Stephenson, 1994) and *Ready Player One* (Cline, 2011).

Comparing all scenarios, it can be summarized that only in the third scenario do NFTs fulfill a functionally sensible purpose for interoperability. In all other scenarios, the inventory of goods is regulated by one or more (central) institutions.

12.3 Looking Towards an Uncertain Future

In this chapter, we presented four potential development scenarios for the realization of interoperability of goods in the metaverse, intended primarily as conceptual cornerstones. It should be emphasized that future combinations or variations of these scenarios could occur, which are changeable. Furthermore, external efforts for regulations were not considered in the scenarios, which could influence reality. Current efforts to develop standards for various aspects of the metaverse include collaborations within the Metaverse Standards Forum[2], where various stakeholders from academia and practice have come together to develop open standards for future interoperability of the metaverse. The future role of NFT technology remains open, even if current efforts integrate it into many concepts. However, our outlined scenarios demonstrate that its deployment, depending on the architecture of the

[2] https://metaverse-standards.org/, Accessed: 12.04.2024

metaverse, is neither a necessary choice nor an uncompromising solution for cross-platform management of digital property. Trust in each other – as in reality in the cooperation of states, companies, individuals, or in today's internet – will also play a crucial role in the metaverse. One of the central questions on the path to an interoperable metaverse will remain how digital goods can be created, owned, and traded in the metaverse – especially across platform boundaries. Crucial is who answers these questions. Since the interests and influence of platform operators, individual users, and society diverge in different areas depending on the scenario, framework conditions should be created to ensure fair interaction among actors. If the metaverse gains societal influence and significance as currently predicted, and companies generally hold significant power of design, a timely examination of the metaverse by governments is necessary. It is particularly desirable from users' perspectives before decisions are primarily made based on market economic interests. The course of action may involve limiting sub-optimal scenarios or regulating and supervising key elements of the system within the democratization of the metaverse. It should also be ensured that the creation of new islands and their networking with other islands is low-threshold and cost-effective, maintaining the pressure for innovation. Companies should intrinsically commit to interoperability and openness at all levels, ensuring the metaverse remains an attractive destination in the long term.

Conclusion 13

> *Det mesta är ännu ogjort. Underbara framtid! (Most things still remain to be done. A glorious future!)*
>
> — **Ingvar Kamprad,** Founder of IKEA

13.1 Summary and Implications

This thesis has addressed various aspects of AR UX, focusing on innovative text input methods and the interaction with Smart Home technologies. In this Chapter, the results and implications of this research are summarized, and the answers to the Research Questions (RQs) are presented. Nine RQs were investigated in the context of this dissertation. These questions have been answered in detail in the previous chapters, so this Chapter will primarily derive overarching implications.

Recognizing the inadequacy of text input methods in AR, Chapter 4 proposed meta requirements and design principles aimed at optimizing learnability and performance and grounded in transfer of learning theory and existing HCI literature on virtual keyboards to answer RQ1.1 and RQ1.2 (Schenkluhn et al., 2022b).

> **Research Question 1.1 (RQ1.1)**
> How to design a mobile virtual keyboard for AR systems to increase text entry performance?

The results highlight various factors that should be considered in the development of virtual text input methods to achieve high performance. One factor is the use of parallel, multimodal input, allowing for increased input speed through the combination of multiple modes such as gaze and gestures. Additionally, leveraging users' prior knowledge and experience with existing text input methods can help achieve high input speed more quickly. Regarding the technical performance of hardware and software, requirements for response time and form factor are also established.

Moreover, the use of haptic feedback is recommended to facilitate input and reduce error rates (Schenkluhn et al., 2022b).

> **Research Question 1.2 (RQ1.2)**
> How to design a mobile virtual keyboard for AR systems to increase learnability?

To improve learnability, two approaches are recommended based on transfer of learning theory (Perkins & Salomon, 1992). First, the use of familiar interactions and layouts should be encouraged to facilitate the concept of hugging. Second, explicit attention to differences and how to manage them should be provided to users to utilize the concept of bridging (Perkins & Salomon, 1992; Schenkluhn et al., 2022b).

RQ2 aimed to investigate the impact of dwell time and text length on gaze typing performance.

> **Research Question 2 (RQ2)**
> What is the influence of dwell time and text length on gaze typing performance?

An experiment was designed to measure fatigue effects and performance in gaze typing, independent of learned and trained prior experiences. Since there is no pool of participants with experience in such text input methods, it is important to exclude learnability effects as much as possible. An AR prototype was developed to conduct the study. Due to time constraints and the prioritization of other studies, only the experimental design was published (Schenkluhn et al., 2022a), but the study was not conducted, leaving this research question unanswered in this work.

To answer RQ3, a comparative analysis of three distinct text-entry solutions – dwell-based eye-gaze input, eye-gaze with pinch-gesture-commit input, and mid-air tap typing on the Microsoft HoloLens 2 – was conducted and presented in Chapter 7 (Schenkluhn et al., 2023a). This controlled within-subjects lab experiment with 27 participants assessed typing performance, task load, usability, and preference, highlighting the importance of personalization in keyboard design for AR environments.

Research Question 3 (RQ3)
Which text entry method is most suitable for mobile AR devices in terms of performance and user preference?

The investigation of RQ3 revealed that none of the three keyboards significantly outperformed the others. While the Tap keyboard had better input speed and sometimes lower error rates, each of the three keyboards was chosen as the first choice by approximately one-third of the participants. Significant differences were found in ratings of mental and physical load and whether the keyboard could be well used in a mobile context. Overall, the results show that user preferences vary greatly and that there is no one-size-fits-all solution. Therefore, it is advisable to offer various text input options, especially for specific application contexts and text lengths. However, if a system developer had to choose one keyboard, the Tap keyboard would be a logical option due to its better initial performance. Yet, it is argued that a combination of Tap and swipe-based gaze-and-commit input would be a more promising choice, as they can be implemented together in a virtual keyboard, combining the advantages of both input methods and thus leaving the choice to users (Schenkluhn et al., 2023a).

In the second part of the work, the fourth RQ turns to the requirements and needs of people with disabilities and older adults for the potential use of AR in Smart Home applications.

Research Question 4 (RQ4)
What are possible use cases and requirements for AR-based applications that are accessible and inclusive for people with different abilities and disabilities?

The results from the focus group in Chapter 9 show that participating experts place particular emphasis on simple and intuitive operation, personalizability and adaptivity, a wide range of supportive functions, and a good understanding of the current application context. Furthermore, affordability, administration, stability, and data security of systems play an important role. Especially regarding adaptivity, a good balance between automation and manual activities must be found to not underchallenge users according to their abilities. Although the focus group was limited in size and not randomly selected, the results indicate that automations in the Smart Home are a promising solution to reduce the required number of interactions at home. However, the degree of automation must be individually adjustable.

Furthermore, the potential of AR HMDs in automating Smart Home interactions through indoor positioning was explored:

> **Research Question 5 (RQ5)**
> How can sensor technology of AR glasses be leveraged for precise indoor user positioning, and what implications does this have for Smart Home automations?

To answer RQ5, an AR prototype was developed to investigate the feasibility of indoor positioning using the Microsoft HoloLens 2 in Chapter 10 (Schenkluhn et al., 2023b). Using the device's spatial mapping capabilities, a flat was 3D-scanned and annotated. Compared to a regular PIR-based motion sensor, the AR prototype exhibited similarly short latency to detect user presence. Additionally, the AR prototype enables the system to identify the user and track their position and orientation on a centimeter-level resolution in real-time. This set of features without the need for additional base stations spread around the house enables the implementation of several automation scenarios, such as precisely turning on and off lights or adjusting room temperature based on user location. Another possible usage scenario is predicting users' locations and intentions to preemptively adjust the environment to their needs (Schenkluhn et al., 2023b).

Finally, Chapter 11 investigates various degrees of automation in spatial Smart Home setup. In RQ6.1 – RQ6.3, classical User Experience dimensions such as task load were compared against psychological needs from Self-Determination Theory.

> **Research Question 6.1 (RQ6.1)**
> Does a manual Smart Home spatial setup design maximize psychological need recognition, and do classical usability dimensions undermine the benefit of this characteristic?

> **Research Question 6.2 (RQ6.2)**
> Does a fully automated spatial setup maximize classical UX dimensions like ease of use, mental workload, and frustration, but reduce the attractiveness of the interaction design by thwarting psychological needs?

Research Question 6.3 (RQ6.3)
Does a combination of manual and automated features strike an effective balance between psychological needs recognition and classical UX dimensions, effectively enhancing technology acceptance?

To investigate these research questions, an AR prototype was again developed, enabling the examination of three different degrees of automation in spatial setup. The mixed-methods laboratory study with 28 participants discovered a preference for solutions that balanced efficiency with user autonomy and control for 75% of participants, underscoring psychological needs as part of UX evaluations, thus answering RQ6.3 (Schenkluhn et al., 2024). RQ6.2 was answered by showing that a fully automated spatial setup design can enhance classical UX dimensions like ease of use, mental workload, and frustration over a manual setup process. However, while the semi-automated option performed significantly better in terms of perceived autonomy, the fully automated setup outperformed the manual setup in terms of perceived competence. The inherent hypothesis of RQ6.1 was not confirmed, as the manual setup process did not perform significantly better in terms of psychological need recognition than the other two alternatives. Overall, the results suggest that users should be able to choose among different levels of automation to address differences in preference (Schenkluhn et al., 2024).

Similar to Heisenberg's uncertainty principle in quantum mechanics (Heisenberg, 1927), no single study can paint a sharp picture of truth alone; only through the combination and repetition of multiple studies, perspectives, and subfields does a clearer overall picture gradually emerge. Controlled experiments' ceteris paribus assumption allows for closer examination of certain dimensions but only illuminates a small part of truth at a time. Although both main themes of this work examine UX in AR from different areas, the results point to recurring themes and implications.

A clear overlap is found in the statement "No size fits all." In neither laboratory study could one solution emerge as a clear favorite due to participants' heterogeneous evaluations of individual attributes leading to different preferences. This underscores the need for adaptive and personalized solutions that meet individual users' needs and preferences. As discussed in differences between Universal Design and Inclusive Design, there are limits to optimization across all user groups.

Particularly regarding performance and efficiency of systems studied, it was noted that depending on the context, a better User Experience was often preferred over the highest text entry speed or automation performance. Developers and designers should consider this when focusing primarily on performance optimization. As evident in comparing text entry methods in Chapter 7, perceived performance can also significantly differ from actual performance (Schenkluhn et al., 2023a).

Another recurring theme is the "Wow effect," observed in both laboratory studies. A significant portion of participants in both studies were impressed by the possibilities of AR technology and innovative interaction possibilities it offers (Schenkluhn et al., 2023a, 2024). This initial trust in technology can help increase the group of early adopters and acceptance of AR applications, especially if compromises in UX must be initially accepted. Since the studies dealt with users' first impressions, it remains open how lasting this effect is. Especially as technology becomes more prevalent, it should be expected that the Wow effect diminishes, and users pay more attention to the usability and efficiency of applications (Rauschnabel & Ro, 2016).

13.2 Research Limitations

This dissertation encompasses a series of studies aimed at advancing the understanding and application of AR. While these studies collectively contribute significant insights, they are subject to a range of limitations that must be acknowledged and considered in the interpretation of their findings.

A common limitation for both conducted lab studies is the geographical limitation to Germany, restricting the cultural diversity of the participant pool, which may affect the generalisability of the findings. Additionally, the laboratory setting of the studies may not fully replicate the complexities and variabilities of real-world environments, thus constraining the applicability of the results to everyday AR use. The absence of longitudinal and field studies within the dissertation restricts insights into long-term user engagement, adaptation, learnability, and effectiveness of AR applications. Overall, the decision for a broader examination of the field with the trade-off of a smaller sample size for each study limited the depth and detail of statistical analyses. However, the sample sizes in the HCI domain and AR research, in particular, are often relatively small due to the limited scalability of the experiment sessions (Caine, 2016).

From a technical perspective, the reliance on a single AR hardware platform, the Microsoft HoloLens 2, may limit the generalizability of the findings to other AR devices. While the HoloLens 2 is a popular choice in comparable research and is yet to be superseded by newer optical-see-through models, especially the hand- and eye-tracking capabilities were identified as limiting factors by participants in both lab studies.

Specific limitations of the individual studies are discussed in the respective chapters. In summary, conducting the study proposed in Chapter 5 would have provided valuable insights into the human limitations of dwell-based text entry, but the study was not carried out due to time constraints and a focus on the study presented in Chapter 7. The study's design, which involved participants being seated, limits insights into the usability of AR text entry interfaces in more dynamic, non-seated environments. The study's reliance on a relatively homogeneous student participant pool further restricts the generalizability of its findings.

The focus group study in Chapter 9 was limited by the small, non-randomly selected sample and the potential for bias introduced by group dynamics. The insights gained from this study are primarily qualitative and possibly subjective as the limited availability of key stakeholder groups and time constraints prevented a more comprehensive exploration of the topic as part of this dissertation.

In conclusion, addressing these limitations through future research is essential for broadening our understanding and enhancing the design, usability, and adoption of AR technologies across diverse user groups and real-world contexts.

13.3 Research Outlook

Based on the findings, limitations, and implications of this dissertation, numerous further research perspectives emerge that can advance the understanding and design of the UX of AR. In particular, the "Wow-Factor" observed in both laboratory studies promises significant potential for the development of AR interfaces that excite users.

Longitudinal and field studies remain a relevant gap, hindered by technical and organizational challenges and the lack of suitable lightweight and unobtrusive AR glasses. While the focus on first impressions in this dissertation is relevant for the acceptance and adoption of AR, early consideration of learning effects and long-term applications is essential for users to benefit sustainably from AR technologies.

Similar to the limitations, specific further research directions are discussed in the respective chapters. Overall, the insights from this dissertation offer numerous starting points for future research that can advance the understanding and design of the UX of AR. In particular, findings on the adaptability and personalization of AR interfaces, as shown in the laboratory studies, highlight the need to focus on individual user preferences and needs instead of purely maximizing efficiency and performance in application development. An investigation into the differences between perceived and actual task performance in various application fields is an interesting direction that became apparent in both laboratory studies. With the advancement of AI technologies, the interaction between AI and autonomous manual activities should also be further explored to better address users' psychological needs, promising greater intrinsic motivation. Another interesting direction for future research is the investigation of predictors of the usage characteristics in order to classify users correctly at the beginning of use and to offer them the appropriate option adaptively and automatically.

Regarding specific application fields, future research should more comprehensively investigate long-term usage and acceptance in mobile text entry, where users must balance trade-offs between efficiency, social acceptance, privacy, and comfort. Contextualized experiments in real environments could provide valuable insights that complement and extend the results of laboratory studies.

Although the focus group in Chapter 9 could only be conducted in a small setting, it nevertheless provided interesting insights into users' requirements and needs that should be further explored in future studies. In addition to the promising potentials identified in the focus group that require further research, the challenges mentioned must also be tackled to unlock the potential for people of all ages and abilities.

The exploration of holistic AR interfaces for Smart Homes presents another exciting research direction. Future studies should aim to improve interaction methods with smart devices through multi-modal and natural user interfaces, enhance the adaptability and personalization of these interfaces, and ensure their interoperability. The potential for using the advanced sensing capabilities of AR HMDs for external devices is particularly promising and will likely inspire many application scenarios and use cases in the future.

All covered topics not only highlight the significant tasks that lie ahead on the path to ubiquitous AR but also inspire enthusiasm for the future of HCI.

13.4 The End

This dissertation emphasizes the significant potential of AR in transforming HCI landscapes. By focusing on user-centric design principles and recognizing the diversity in user preferences and needs, this work contributes to the development of more intuitive, accessible, and efficient interaction paradigms within AR environments. As we approach a new era in digital interaction, the integration of AR with intelligent systems signals a future where technology is seamlessly woven into the fabric of daily life, enhancing human capabilities and experiences.

The timeline for ubiquitous AR becoming a reality remains uncertain. Many forecasts and promises from recent years have not been fulfilled, leading to more cautious predictions (Boland, 2021). Palmer Luckey (2015), the founder of Oculus, may also be correct in the context of AR when he stated that "VR will become something everyone wants before it becomes something everyone can afford" – similar to the "Wow-Effects" described in this work.

AR, in its own right as well as in the context of a potential metaverse, will need to balance the fine line between supporting and empowering humans on one side, and dictating and creating dependency on the other, to be beneficial to society and humanity. This work reinforces the necessity of developing software tailored to users to promote their acceptance and adoption of new technology. Moreover, similar to the smartphone, there is significant power when the potentials of technology unfold, and users become increasingly tied to the technology (Harari, 2018). With the even closer integration of AR into the user's field of view, as well as through understanding context via cameras and gaze behavior through eye-tracking, the technology will increasingly intervene in users' lives. Combined with additional sensors from wearables, the user can become as transparent as never before. Not only the digital traces that accumulate from device use or physical actions will be traceable, but also the thoughts and emotions of the user derived from comprehensive sensor information – at any given moment.

Yet, dystopian considerations need not reach so far. Already, constant interruptions and distractions through notifications and perpetual availability take their toll (Pang, 2013). The increasing acceleration and upsurge lead to overwhelm and lack of resonance, reflected in various areas of society (Rosa, 2016). Although the world is objectively becoming a better place (Rosling et al., 2018), paradoxically, many people feel subjectively dissatisfied and unhappy (Easterbrook, 2003). The call for Calm Technology (Weiser & Brown, 1996) is, therefore, more relevant than ever in light of the even tighter integration of technology and user. Technology should, instead of demanding attention, offer the tranquility we urgently need (Hartl, 2021).

In a market economy, the responsibility primarily lies with legislative bodies to protect citizens' interests while simultaneously creating a level playing field for companies by eliminating incentives for exploiting users' data and attention. However, as technology advances faster than legislation can keep up in many cases, and established practices and business models are harder to regulate, developers and designers are also called upon to be aware of their responsibility and to design technology that serves primarily the users and society.

Despite the challenges and risks, the outlook for the future can be optimistic. AR holds potentials on various levels that are not yet fully estimable but have already shown promising results in current studies. We should leverage them to make the digital ubiquity a better place.

Supplementary Material for Chapter 11

A

A.1 Interview Guide

Instructions

- Set aside the questionnaire tablet after the participant finished filling out the post-treatment questionnaire and the HoloLens to ensure a clean table and an unobstructed line of sight between the interviewer and the interviewee.
- Provide an explanation of the interview process: "We are near the end of the study. We will now have a short interview in which you have the opportunity to share your experiences during the study."
- Remind the participant about the commencement of the audio recording and its processing.
- Start the audio recording.

Interview

1. Please share your overall experience with the task and the system.
2. Please rank the three application variants in order of preference, and explain the reasons for your choice.
3. Please rank the three task variants based on their efficiency, and explain the reasons for your choice.
4. For the semi-automatic variant, you chose to use the simultaneous detection after x number of individual lamp detections. Please explain your decision-making process.
5. Please share any observations or thoughts on the usability of the system. Did anything stand out?

A.2 Interview Codebook

Category	Themes	N Instances	N Participants	Exemplary Quote
First Impression	Wow Factor	22	18	"So, in general first. It was definitely a very interesting experience, to be honest. And it's truly impressive what's possible and how it might actually look in the future." – P28
	Curiosity Evoking	13	10	"I wanted to try it out. I just looked to see what would happen. And then, after, I don't know, what did I click, I had seven or eight lights, so I clicked on it quite late and thought I'd give it a try." – P19
	Learning Effects	25	18	"I had to first get used to what the device wanted from me. And practice that. It's a matter of practice for me." – P21
Technical Capabilities	Depth Perception Issues	5	4	"The problem was in the depth, but also somehow the position in the room in general. So the perspective didn't always quite fit." – P7
	Interaction Mode Issues	31	21	"So sometimes it didn't work right away to bring up the menu, or bringing up the menu worked, but then tapping on it didn't." – P20
	Trust in Capabilities	30	17	"In hindsight, I did think, okay, what if something goes wrong. But I felt, or I got the impression, that it then found things well. Yes, so I would trust the system." – P11
	Simplicity	12	7	"I found it quite exciting to see how fast some things can happen, how everything is captured automatically." – P22
Affect	Strenuous	43	23	"It was just frustrating with the whole setup of the individual devices." – P17
	Comfortable	23	15	"The second one [the semi-automatic condition] was the most relaxed, I could pick a few devices that I want and the rest is done automatically." – P5
	Fun	22	12	"I also liked the manual version because it has this certain playful aspect to it, and honestly, you don't set up new devices that often." – P22
Load & Control	Control & Overview	56	23	"I did a bit, walked around the apartment a bit. I felt responsible for the setup, but didn't have to do everything myself." – P18
	Overwhelming	31	16	"I think, for example, I would not recommend this to my mother; she would probably freak out if something like this happened in her apartment." – P18
	Efficiency	56	27	"Of course, the most efficient is the automatic version. I walk through the room, and the thing is done. I don't really have to choose anything; I don't have to make any decisions." – P7
Relationship & Understanding	Personalization	24	11	"Well, I believe the version where I can set it up myself is just more individualized." – P4
	Partnership	20	10	"There, I just have the feeling of having accomplished something and having contributed, and the device doesn't do everything on its own." – P7
	Alienating	2	5	"It was also a bit strange, especially when all the things started to light up or draw attention to themselves." – P28
Possible Changes	Diverse Ideas	38	15	"So, if there were, let's say, a game module included, where I could participate in some AR gaming situation with the glasses, okay, that would surely be great." – P21
	Adjustment Options	17	9	"Yes, then it would be good if you could adjust it a bit." – P13

Figure A.1.: Interview Analysis Codebook with Exemplary Quotes

Supplementary Material for Chapter 9

B

B.1 Interview Guide

Introduction (20 Min)

This section provides an introduction and further explanations regarding the purpose of the focus group. It also addresses organizational details (e.g., recording, time constraints) and the structure of the focus group.

Objective and Structure of the Focus Group

a) Further explanations on the purpose of the focus group: Gathering perspectives from stakeholders on AR applications for everyday assistance solutions.

b) Clarification of organizational details:

1. Duration of the Focus Group
    - Planned timeframe: 90 minutes (scheduled for 120 minutes, including buffer time).
2. Structure of the Focus Group
    - Introduction (background and experiences of participants, conceptual foundation for the research topic, example of AR application).
    - Brainstorming phase on participants' experiences with challenges faced by individuals with disabilities in daily life, as well as the use of AR methods for everyday assistance solutions.
    - If necessary, delve into emerging topics, or deepen discussions on topics from the Deep Dive list depending on group dynamics.
    - Conclusion/Next Steps.

3. Recording of the Focus Group
    - A scientific paper will be developed based on the discussion.
    - The discussion will be recorded (audio only, no video).
    - All data will be anonymized and will not allow for identification of individuals or institutions.
    - If individual statements are to be quoted verbatim, consent will be obtained again individually.
4. The results of the study will be made available upon request.

Personal Introduction/Who is Present?

Initially, the background and experience of the participants are captured to understand their general stance towards everyday assistance solutions.

- Moderator
- Note-taker
- Participants: Representative for concerns of people with disabilities, individual support for people with disabilities, a representative for students with disabilities and chronic illnesses, residential group leader.
- Background and experience of participants: Organization/Institution, number of employees/care recipients, tasks and main responsibilities, experiences with AR and/or everyday assistance solutions, experiences with Smart Home technologies, perception of everyday assistance solutions/AR (e.g., How do you assess the significance of AR/assistance solutions?).

Brainstorming Phase (60 Min)

Brainstorming (Consider your daily experiences and challenges...)

Block 1: Challenges in Daily Life

1. What are the biggest challenges you observe in people with disabilities or older adults in their daily living environment? Please describe specific situations or scenarios.

2. What are the major challenges for accessibility in apartments/spaces/facilities for people with disabilities or older adults, and how can they be overcome?

Block 2: Possibilities of Assistance Solutions

3. From your perspective, what are the key points or outcomes that should be aimed for with assistance solutions for an independent life? How would you define success in this context?
4. How can assistance solutions enable individuals to maintain their independence and improve their quality of life, in your view? Are there specific areas or activities where you see the greatest potential impact?
5. What role could caregivers, healthcare professionals, or family members play? How can assistance solutions effectively support their needs and concerns as well?
6. Can you think of a specific use case or example where assistance solutions could make a significant difference?
7. What are your expectations regarding the integration of technology and Ambient Assisted Living (AAL)? What potential benefits and risks do you see? (Question as a transition)

Block 3: AR in Addressing Challenges

8. How do you think AR can be used to address the challenges you mentioned earlier, considering factors such as architectural barriers, inaccessible areas, or limited mobility?
9. How can AR support individuals with disabilities or older adults in navigating and interacting with their living environment, taking into account factors like physical barriers and communication issues?
10. How do specific disabilities affect daily activities and tasks in a living environment, and how can AR technologies be adapted to address these specific challenges?
11. How can AR promote independence and self-reliance in people with disabilities and older adults in performing activities of daily living, such as personal care, cooking, cleaning, or medication management? (e.g., through step-by-step guidance, reminders, or interactive instructions)

12. In what ways can AR applications facilitate communication and social interaction for people with disabilities/older adults in their living environment, considering features like real-time subtitles, sign language recognition, or virtual communication platforms?

13. How can AR assist individuals with cognitive disabilities in organizing their living spaces, managing routines, and remembering important tasks or appointments - e.g., through visual cues, reminders, or interactive schedules?

Block 4: AR – Design and Implementation

14. Have you had experiences with user interfaces and user interactions that people with disabilities or older adults interact with? Were there problems or interfaces that were particularly easy and intuitive to use?

15. How can AR support the engagement and involvement of AAL stakeholders, including caregivers, medical professionals, and family members, in caring for and supporting AAL residents by promoting collaboration, information exchange, and remote monitoring opportunities?

16. What potential challenges or considerations exist regarding user acceptance and training when implementing AR technologies in living environments, and how can these challenges be addressed to ensure broad usability and accessibility?

17. What potential ethical and privacy considerations exist when implementing AR in AAL environments, and how can these concerns be addressed to safeguard the dignity, autonomy, and privacy of AAL residents?

Deep Dive/In-depth Exploration (optional)

18. What are the potential benefits of using AR-based navigation and orientation systems in AAL environments that help AAL residents with cognitive impairments or memory difficulties move more independently and with less anxiety in their living spaces?

19. In what ways can AR applications improve the safety of AAL residents in their living environment, e.g., through real-time monitoring, fall detection, emergency alerts, or hazard detection?

Conclusion and Next Steps (10 Min)

Conclusion

- Wrap-up and repetition of key insights.
- Participants' statements.
- Are there any further questions or suggestions/recommendations you would like to make regarding the focus group itself or the study as a whole?
- The results of the study will be made available upon request. Please provide your contact details, e.g., email address, on the provided list.

Thank you for your time and valuable insights!

Bibliography

Acar, A., Fereidooni, H., Abera, T., Sikder, A. K., Miettinen, M., Aksu, H., Conti, M., Sadeghi, A.-R., & Uluagac, S. (2020). Peek-a-boo: I see your smart home activities, even encrypted! In *Proceedings of the 13th ACM Conference on Security and Privacy in Wireless and Mobile Networks* (pp. 207–218). Association for Computing Machinery.

Adams, N., Little, T. D., & Ryan, R. M. (2017). Self-Determination Theory. In M. L. Wehmeyer, K. A. Shogren, T. D. Little, & S. J. Lopez (Eds.), *Development of Self-Determination Through the Life-Course* (p. 4754). Springer Netherlands.

Adhikary, J., & Vertanen, K. (2021). Typing on midair virtual keyboards: Exploring visual designs and interaction styles. In *Human-Computer Interaction–INTERACT 2021: 18th IFIP TC 13 International Conference, Bari, Italy, August 30–September 3, 2021, Proceedings, Part IV 18* (pp. 132–151). Springer.

Alabood, L., & Maurer, F. (2022). An IoT-based Immersive Smart Home System for Seniors with Neurocognitive Disorders. *EMPATHY@ AVI*, 15–20. Retrieved April 12, 2024, from https://ceur-ws.org/Vol-3172/short3.pdf

Alam, F., Faulkner, N., & Parr, B. (2020). Device-free localization: A review of non-RF techniques for unobtrusive indoor positioning. *IEEE Internet of Things Journal*, *8*(6), 4228–4249.

Aloulou, H., Mokhtari, M., Tiberghien, T., Biswas, J., Phua, C., Kenneth Lin, J. H., & Yap, P. (2013). Deployment of assistive living technology in a nursing home environment: Methods and lessons learned. *BMC medical informatics and decision making*, *13*(1), 1–17.

Al-Shaqi, R., Mourshed, M., & Rezgui, Y. (2016). Progress in ambient assisted systems for independent living by the elderly. *SpringerPlus*, *5*(1), 624.

Andrews, J., Kowsika, M., Vakil, A., & Li, J. (2020). A motion induced passive infrared (PIR) sensor for stationary human occupancy detection. In *2020 IEEE/ION Position, Location and Navigation Symposium (PLANS)* (pp. 1295–1304). IEEE.

Anthes, G. (2019). Augmented reality gets real. *Communications of the ACM*, *62*(9), 16–18.

Apple. (2023). *WWDC 2023 Apple Keynote*. Retrieved April 12, 2024, from https://www.youtube.com/watch?v=GYkq9Rgoj8E

Apple. (2024a). *Apple Vision Pro*. Apple. Retrieved April 12, 2024, from https://www.apple.com/apple-vision-pro/

Apple. (2024b). *Augmented Reality*. Retrieved April 12, 2024, from https://www.apple.com/augmented-reality/

Apple. (2024c). *Paid Applications Agreement (Schedules 2 and 3 of the Apple Developer Program License Agreement)*. Retrieved April 12, 2024, from https://developer.apple.com/support/downloads/terms/schedules/Schedule-2-and-3-20230828-English.pdf

Apple. (2024d). *Submit your apps to the App Store for Apple Vision Pro*. Apple Developer. Retrieved April 12, 2024, from https://developer.apple.com/visionos/submit/

Arcelus, A., Jones, M. H., Goubran, R., & Knoefel, F. (2007). Integration of Smart Home Technologies in a Health Monitoring System for the Elderly. In *21st International Conference on Advanced Information Networking and Applications Workshops (AINAW'07)* (pp. 820–825).

Ariano, R., Manca, M., Paternò, F., & Santoro, C. (2022). Smartphone-based augmented reality for end-user creation of home automations. *Behaviour & Information Technology*, *42*(1), 124–140.

Arif, A. S., & Stuerzlinger, W. (2009). Analysis of text entry performance metrics. In *2009 IEEE Toronto International Conference Science and Technology for Humanity (TIC-STH)* (pp. 100–105). IEEE. https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=5444533

Ariso, J. M. (2017). Is Critical Thinking Particularly Necessary when Using Augmented Reality in Knowledge Society? An Introductory Paradox. In J. M. Ariso (Ed.), *Augmented Reality: Reflections on its contribution to knowledge formation* (p. 322). De Gruyter.

Assistant, H. (2024). *Integrations*. Retrieved April 12, 2024, from https://www.home-assistant.io/integrations/#light

Aziz, S., & Komogortsev, O. (2022). An Assessment of the Eye Tracking Signal Quality Captured in the HoloLens 2. In *2022 Symposium on Eye Tracking Research and Applications* (pp. 1–6).

Azuma, R. T. (1997). A survey of augmented reality. *Presence: Teleoperators & Virtual Environments*, *6*(4), 355–385.

Azuma, R. T. (2017). Making augmented reality a reality. *Propagation Through and Characterization of Atmospheric and Oceanic Phenomena*, JTu1F–1. Retrieved April 12, 2024, from https://opg.optica.org/abstract.cfm?uri=pcaop-2017-JTu1F.1

Azuma, R. T., Baillot, Y., Behringer, R., Feiner, S., Julier, S., & MacIntyre, B. (2001). Recent advances in augmented reality. *IEEE computer graphics and applications*, *21*(6), 34–47.

Babich, N. (2020). *UX Design Principles for Augmented Reality | Adobe XD Ideas*. Retrieved April 12, 2024, from https://xd.adobe.com/ideas/principles/emerging-technology/ux-design-principles-for-augmented-reality/

Baek, F., Ha, I., & Kim, H. (2019). Augmented reality system for facility management using image-based indoor localization. *Automation in construction*, *99*, 18–26.

Bailly, C., Leitner, F., & Nigay, L. (2019). Head-controlled menu in mixed reality with a HMD. In *Human-Computer Interaction–INTERACT 2019: 17th IFIP TC 13 International Conference, Paphos, Cyprus, September 2–6, 2019, Proceedings, Part IV 17* (pp. 395–415). Springer.

Balakrishnan, A., & CNBC. (2017). *How the first iPhone blew the BlackBerry away with an amazingly accurate on-screen keyboard*. Retrieved April 12, 2024, from https://www.cnbc.com/2017/06/26/creation-of-the-first-iphone-touchscreen-interview-with-scott-forstall-tony-fadell-and-greg-christie.html

Ball, M. (2021). *Framework for the Metaverse*. MatthewBall.co. Retrieved April 12, 2024, from https://www.matthewball.co/all/forwardtothemetaverseprimer

Ballou, N., Deterding, S., Tyack, A., Mekler, E. D., Calvo, R. A., Peters, D., Villalobos-Zúñiga, G., & Turkay, S. (2022). Self-Determination Theory in HCI: Shaping a Research Agenda. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery.

Bardeen, L., Services, H. A., & Review, H. B. (2018). *Mixed Reality: A new dimension of work*. Retrieved April 12, 2024, from https://hbr.org/resources/pdfs/comm/microsoft/MixedRealityNewDimensionOfWork.pdf

Basri, C., & Elkhadimi, A. (2020). A review on indoor localization with internet of things. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *44*, 121128.

Bastian, M. (2017). *Apple übernimmt das deutsche Eye-Tracking-Unternehmen SMI*. Retrieved April 12, 2024, from https://mixed.de/apple-uebernimmt-das-deutsche-eye-tracking-unternehmen-smi/

Bastian, M. (2020a). *Apple stellt Magic Leaps entlassene Angestellte ein*. Retrieved April 12, 2024, from https://mixed.de/apple-stellt-magic-leaps-entlassene-angestellten-ein/

Bastian, M. (2020b). *Bericht: Apple kauft VR-Streaming-Spezialisten NextVR*. Retrieved April 12, 2024, from https://mixed.de/bericht-apple-kauft-vr-streaming-spezialisten-nextvr/

Bastian, M. (2020c). *Google kauft sich Glass 3 in Kanada ein - Bericht*. Retrieved April 12, 2024, from https://mixed.de/bericht-google-uebernimmt-techbrillen-hersteller-north/

Bebko, A. O., & Troje, N. F. (2020). bmlTUX: Design and control of experiments in virtual reality and beyond. *i-Perception*, *11*(4), 2041669520938400.

Berkemeier, L., Zobel, B., Werning, S., Vogel, J., Remark, F., Ickerott, I., & Thomas, O. (2019). Heuristic Theorizing in Software Development: Deriving Design Principles for Smart Glasses-based Systems.

Bermejo Fernandez, C., Lee, L. H., Nurmi, P., & Hui, P. (2021). PARA: Privacy Management and Control in Emerging IoT Ecosystems using Augmented Reality. In *Proceedings of the 2021 International Conference on Multimodal Interaction* (pp. 478–486). Association for Computing Machinery.

Bertarini, M. (Ed.). (2014). *Smart glasses: Interaction, privacy and social implications*. Citeseer.

Betker, J., Goh, G., Jing, L., Brooks, T., Wang, J., Li, L., Ouyang, L., Zhuang, J., Lee, J., Guo, Y., & Ramesh, A. (2023). Improving image generation with better captions. *Computer Science. https://cdn. openai. com/papers/dall-e-3. pdf*, *2*(3), 8.

Bhalla, G. (2014). How to plan and manage a project to co-create value with stakeholders. *Strategy & leadership*, *42*(2), 19–25. Retrieved March 28, 2024, from https://www.emerald.com/insight/content/doi/10.1108/SL-01-2014-0006/full/html?fullSc=1&fullSc=1&fullSc=1&mbSc=1&utm_campaign=Emerald_Strategy_PPV_November22_RoN

Billinghurst, M. (2021). Grand Challenges for Augmented Reality. *Frontiers in Virtual Reality*, *2*.

Billinghurst, M., Clark, A., & Lee, G. (2015). A survey of augmented reality. *Foundations and Trends in Human-Computer Interaction*, *8*, 73–272.

Binary Banana. (2018). *Smart AR Home*. Retrieved April 11, 2024, from http://smartarhome.com

Bitkom Research. (2022). *Das intelligente Zuhause: Smart Home 2022*. Retrieved April 12, 2024, from https://www.bitkom.org/sites/main/files/2022-09/220912_Bitkom_Smart_Home_Chartbericht_2022_final.pdf

Blackman, S., Matlo, C., Bobrovitskiy, C., Waldoch, A., Fang, M. L., Jackson, P., Mihailidis, A., Nygård, L., Astell, A., & Sixsmith, A. (2016). Ambient Assisted Living Technologies for Aging Well: A Scoping Review. *Journal of Intelligent Systems*, *25*(1), 55–69.

Bleda, A. L., Maestre, R., Beteta, M. A., & Vidal, J. A. (2018). AmICare: Ambient intelligent and assistive system for caregivers support. *2018 IEEE 16th International Conference on Embedded and Ubiquitous Computing (EUC)*, 66–73. Retrieved April 12, 2024, from https://ieeexplore.ieee.org/abstract/document/8588850/

Bohn, D. (2019). *Microsoft's HoloLens 2: A $3500 mixed reality headset for the factory*. Retrieved April 12, 2024, from https://www.theverge.com/2019/2/24/18235460/microsoft-hololens-2-price-specs-mixed-reality-ar-vr-business-work-features-mwc-2019

Boland, M. (2021). *Is Ubiquitous AR a Decade Away?* AR Insider. Retrieved April 12, 2024, from https://arinsider.co/2021/06/16/is-ubiquitous-ar-a-decade-away/

Bouma, H., Fozard, J. L., Bouwhuis, D. G., & Taipale, V. (2007). Gerontechnology in perspective. *Gerontechnology*, *6*(4), 190–216. Retrieved March 28, 2024, from https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=ac798e136b2ad14897c841ac6a260a4b24c401a6

Bowman, D. A., Rhoton, C. J., & Pinho, M. S. (2002). Text input techniques for immersive virtual environments: An empirical comparison. In *Proceedings of the human factors and ergonomics society annual meeting* (pp. 2154–2158, Vol. 46). SAGE Publications Sage CA: Los Angeles, CA.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, *3*(2), 77101.

Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability evaluation in industry*, *189*(194), 4–7. https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html

Burgstahler, S. (2009). Universal Design: Process, Principles, and Applications. *DO-IT*. Retrieved March 28, 2024, from https://eric.ed.gov/?id=ED506550

Burwell, S., Sample, M., & Racine, E. (2017). Ethical aspects of brain computer interfaces: A scoping review. *BMC Medical Ethics*, *18*(1), 60.

Caine, K. (2016). Local standards for sample size at CHI. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 981–992).

Cakmakci, O., Qin, Y., Bosel, P., & Wetzstein, G. (2021). Holographic pancake optics for thin and lightweight optical see-through augmented reality. *Optics Express*, *29*(22), 35206–35215.

Calvaresi, D., Cesarini, D., Sernani, P., Marinoni, M., Dragoni, A. F., & Sturm, A. (2017). Exploring the ambient assisted living domain: A systematic review. *Journal of Ambient Intelligence and Humanized Computing*, *8*(2), 239–257.

Calvaresi, D., Claudi, A., Dragoni, A. F., Yu, E., Accattoli, D., & Sernani, P. (2014). A goal-oriented requirements engineering approach for the ambient assisted living domain. *Proceedings of the 7th International Conference on PErvasive Technologies Related to Assistive Environments*, 1–4.

Card, S. K., Moran, T. P., & Newell, A. (1980). The keystroke-level model for user performance time with interactive systems. *Communications of the ACM*, *23*(7), 396–410.

Carmigniani, J., Furht, B., Anisetti, M., Ceravolo, P., Damiani, E., & Ivkovic, M. (2011). Augmented reality technologies, systems and applications. *Multimedia tools and applications*, *51*(1), 341–377.

Carter, B. T., & Luke, S. G. (2020). Best practices in eye tracking research. *International Journal of Psychophysiology*, *155*, 49–62. Retrieved April 12, 2024, from https://www.sciencedirect.com/science/article/pii/S0167876020301458

Cebanov, E., Dobre, C., Gradinaru, A., Ciobanu, R.-I., & Stanciu, V.-D. (2019). Activity recognition for ambient assisted living using off-the-shelf motion sensing input devices. *2019 Global IoT Summit (GIoTS)*, 1–6. Retrieved April 12, 2024, from https://ieeexplore.ieee.org/abstract/document/8766379/

Center, I. D. R. (2017). *The Inclusive Design Guide*. Retrieved April 12, 2024, from https://guide.inclusivedesign.ca/

Chadwick, D., & Wesson, C. (2016). Digital Inclusion and Disability. In A. Attrill & C. Fullwood (Eds.), *Applied Cyberpsychology* (pp. 1–23). Palgrave Macmillan UK.

Chatterjee, I., Xiao, R., & Harrison, C. (2015). Gaze+ gesture: Expressive, precise and targeted free-space interactions. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction* (pp. 131–138).

Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88–97.

Chen, Z., Li, J., Hua, Y., Shen, R., & Basu, A. (2017). Multimodal interaction in augmented reality. In *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (pp. 206–209). IEEE.

Chi, Z., Niu, Z., & Sawabe, T. (2022). Enabling Augmented Reality Incorporate with Audio on Indoor Navigation for People with Low Vision. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* (pp. 882–883). IEEE.

Choi, Y. M., & Li, J. (2016). Usability Evaluation of a New Text Input Method for Smart TVs. *Journal of Usability Studies*, *11*(3).

Chowdhary, K. R. (2020). Natural Language Processing. In *Fundamentals of Artificial Intelligence* (pp. 603–649). Springer India.

Cicirelli, G., Marani, R., Petitti, A., Milella, A., & D'Orazio, T. (2021). Ambient assisted living: A review of technologies, methodologies and future perspectives for healthy aging of population. *Sensors*, *21*(10), 3549. Retrieved March 28, 2024, from https://www.mdpi.com/1424-8220/21/10/3549

Clarkson, P. J., & Coleman, R. (2015). History of inclusive design in the UK. *Applied ergonomics*, *46*, 235–247. Retrieved March 28, 2024, from https://www.sciencedirect.com/science/article/pii/S0003687013000410

Cline, E. (2011). *Ready Player One*. Ballantine Books.

Company, B.
bibinitperiod. (2020). *Augmented or Virtual Reality*. Retrieved April 12, 2024, from https://www.bain.com/insights/customer-experience-tools-augmented-or-virtual-reality/

Conover, W. J. (1999). *Practical nonparametric statistics* (Vol. 350). john wiley & sons.

Coppers, S., Vanacken, D., & Luyten, K. (2020). FORTNIoT: Intelligible Predictions to Improve User Understanding of Smart Home Behavior. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *4*(4), 1–24.

Creed, C., Al-Kalbani, M., Theil, A., Sarcar, S., & Williams, I. (2023). Inclusive AR/VR: Accessibility barriers for immersive technologies. *Universal Access in the Information Society*, 1–15.

Dador, D. (2023). *Torrance hospital revolutionizes hip surgery with augmented reality in the operating room - ABC7 Los Angeles*. Retrieved April 12, 2024, from https://abc7.com/torrance-hospital-augmented-reality-hip-surgery/13960575/

De Belen, R. A. J., Bednarz, T., & Favero, D. D. (2019). Integrating Mixed Reality and Internet of Things as an Assistive Technology for Elderly People Living in a Smart Home. *Proceedings of the 17th International Conference on Virtual-Reality Continuum and Its Applications in Industry*, 1–2.

De Podestá Gaspar, R., Bonacin, R., & Gonçalves, V. P. (2018). Designing IoT Solutions for Elderly Home Care: A Systematic Study of Participatory Design, Personas and Semiotics. In M. Antona & C. Stephanidis (Eds.), *Universal Access in Human-Computer Interaction. Virtual, Augmented, and Intelligent Environments* (pp. 226–245, Vol. 10908). Springer International Publishing.

Deloitte. (2018). *Smart Home Consumer Survey 2018*. Retrieved April 12, 2024, from https://www2.deloitte.com/content/dam/Deloitte/de/Documents/technology-media-telecommunications/Deloitte_TMT_Smart_Home_Studie_18.pdf

Denis, S., Berkvens, R., & Weyn, M. (2019). A survey on detection, tracking and identification in radio frequency-based device-free localization. *Sensors*, *19*(23), 5329.

Dey, A., Billinghurst, M., Lindeman, R. W., & Swan, J. E. I. (2018). A Systematic Review of 10 Years of Augmented Reality Usability Studies: 2005 to 2014. *Frontiers in Robotics and AI*, *5*.

Dionisio, J. D. N., Iii, W. G. B., & Gilbert, R. (2013). 3D Virtual worlds and the metaverse: Current status and future possibilities. *ACM Computing Surveys*, *45*(3), 1–38.

dr. dresing & pehl GmbH. (n.d.). *F4analyse*. Retrieved March 28, 2024, from https://www.audiotranskription.de/f4analyse/

Drewlow, J., Däppen, M., & Lehmann, M. (2022). Navigation with Augmented Reality in a Hospital. *Studies in Health Technology and Informatics*, *292*, 111–114.

Dube, T. J., & Arif, A. S. (2019). Text entry in virtual reality: A comprehensive review of the literature. In *International Conference on Human-Computer Interaction* (pp. 419–437). Springer.

Duchowski, A. (2002). A breadth-first survey of eye-tracking applications. *Behavior Research Methods, Instruments, & Computers*, *34*(4), 455–470.

Duchowski, A. (2007). Eye tracking techniques. In *Eye tracking methodology* (pp. 51–59). Springer. https://link.springer.com/chapter/10.1007/978-1-84628-609-4_5?noAccess=true

Duchowski, A. T. (2017). *Eye tracking methodology: Theory and practice*. Springer.

Dudley, J., Benko, H., Wigdor, D., & Kristensson, P. O. (2019). Performance envelopes of virtual keyboard text input strategies in virtual reality. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (pp. 289–300). IEEE.

Dudley, J. J., Vertanen, K., & Kristensson, P. O. (2018). Fast and precise touch-based text entry for head-mounted augmented reality with variable occlusion. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *25*(6), 1–40.

Duenser, A., Rapha, Seichter, H., & Billinghurst, M. (2007). Applying HCI principles to AR systems design.

Durrant-Whyte, H., & Bailey, T. (2006). Simultaneous localization and mapping: Part I. *IEEE robotics & automation magazine*, *13*(2), 99–110.

Easterbrook, G. (2003). *The progress paradox: How life gets better while people feel worse* (1st ed). Random House.

Edu, J. S., Such, J. M., & Suarez-Tangil, G. (2020). Smart home personal assistants: A security and privacy review. *ACM Computing Surveys (CSUR)*, *53*(6), 136.

El-Moursy, A. A., Sibai, F. N., Rehman, J., Gouda, O. M., Gaber, A. T., & Khedr, A. M. (2022). Home Automation Using Augmented Reality (HAAR). *Wireless Personal Communications*, *124*(2), 1525–1555.

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing*, *62*(1), 107–115.

Escobar, E. (2022). *The state of open source on GitHub*. Retrieved April 12, 2024, from https://octoverse.github.com/2022/state-of-open-source

Espósito, F. (2020). *Shift Keyboard introduces a new way to write messages on Apple Watch*. Retrieved April 12, 2024, from https://9to5mac.com/2020/01/07/shift-keyboard-introduces-a-new-way-to-write-messages-on-apple-watch/

European Commission. (2024). *The Digital Services Act package | Shaping Europe's digital future*. Retrieved April 12, 2024, from https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package

Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on Contestable and Fair Markets in the Digital Sector and Amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act) (Text with EEA Relevance) (2022). Retrieved April 12, 2024, from http://data.europa.eu/eli/reg/2022/1925/oj/eng

Evans, B. (2020). *The VR winter*. Retrieved April 12, 2024, from https://www.ben-evans.com/benedictevans/2020/5/8/the-vr-winter

Evans, G., Miller, J., Pena, M. I., MacAllister, A., & Winer, E. (2017). Evaluating the Microsoft HoloLens through an augmented reality assembly application. In I. S. for Optics and Photonics (Ed.), *Degraded Environments: Sensing, Processing, and Display 2017* (p. 101970, Vol. 10197). International Society for Optics and Photonics.

Farrell, J., & Klemperer, P. (2007). Coordination and lock-in: Competition with switching costs and network effects. *Handbook of industrial organization*, *3*, 1967–2072.

Farshid, M., Paschen, J., Eriksson, T., & Kietzmann, J. (2018). Go boldly!: Explore augmented reality (AR), virtual reality (VR), and mixed reality (MR) for business. *Business Horizons*, *61*(5), 657–663.

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).

Flatt, H., Koch, N., Röcker, C., Günter, A., & Jasperneite, J. (2015). A context-aware assistance system for maintenance applications in smart factories based on augmented reality and indoor localization. In *2015 IEEE 20th conference on emerging technologies & factory automation (ETFA)* (pp. 1–4). IEEE.

Flick, C. D., Harris, C. J., Yonkers, N. T., Norouzi, N., Erickson, A., Choudhary, Z., Gottsacker, M., Bruder, G., & Welch, G. (2021). Trade-offs in augmented reality user interfaces for controlling a smart environment. *Symposium on Spatial User Interaction*, 1–11.

Florea, A.-I., Anghel, I., & Cioara, T. (2022). A review of blockchain technology applications in ambient assisted living. *Future Internet*, *14*(5), 150. Retrieved March 28, 2024, from https://www.mdpi.com/1999-5903/14/5/150

Fortune Business Insights. (2024). *Augmented Reality [AR] Market Size, Share, Trends | Forecast*. Retrieved April 12, 2024, from https://www.fortunebusinessinsights.com/augmented-reality-ar-market-102553

Franco-Salvador, M., & Leiva, L. A. (2018). Multilingual phrase sampling for text entry evaluations. *International Journal of Human-Computer Studies*, *113*, 15–31.

Franke, T., Attig, C., & Wessel, D. (2019). A personal resource for technology interaction: Development and validation of the affinity for technology interaction (ATI) scale. *International Journal of Human-Computer Interaction*, *35*(6), 456467.

Friston, S. J., Congdon, B. J., Swapp, D., Izzouzi, L., Brandstätter, K., Archer, D., Olkkonen, O., Thiel, F. J., & Steed, A. (2021). Ubiq: A system to build flexible social virtual reality experiences. In *Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology* (pp. 1–11).

Frutos-Pascual, M., Creed, C., & Williams, I. (2019). Head mounted display interaction evaluation: Manipulating virtual objects in augmented reality. In *Human-Computer Interaction–INTERACT 2019: 17th IFIP TC 13 International Conference, Paphos, Cyprus, September 2–6, 2019, Proceedings, Part IV 17* (pp. 287–308). Springer.

Fuglerud, K. S., & Sloan, D. (2013). The Link between Inclusive Design and Innovation: Some Key Elements. In M. Kurosu (Ed.). D. Hutchison, T. Kanade, J. Kittler, J. M. Kleinberg, F. Mattern, J. C. Mitchell, M. Naor, O. Nierstrasz, C. Pandu Rangan, B. Steffen, M. Sudan, D. Terzopoulos, D. Tygar, M. Y. Vardi, & G. Weikum (**typeredactors**), *Human-Computer Interaction. Human-Centred Design Approaches, Methods, Tools, and Environments* (pp. 41–50, Vol. 8004). Springer Berlin Heidelberg.

Future, M. R. (2021). *Eye Tracking Market, By Type (Mobile, Remote), By Applications (Human-Computer Interaction, Virtual Reality), By Verticals (Aerospace, Retail, Automotive, Government & Defense) - Forecast 2027*. Retrieved April 12, 2024, from https://www.marketresearchfuture.com/reports/eye-tracking-market-3958

Gabriel, Z., & Bowling, A. N. N. (2004). Quality of life from the perspectives of older people. *Ageing & Society*, *24*(5), 675–691. Retrieved March 28, 2024, from https://www.cambridge.org/core/journals/ageing-and-society/article/quality-of-life-from-the-perspectives-of-older-people/F5BB3981F0791BF694AA06206BF9A297

Gagné, M., Parker, S. K., Griffin, M. A., Dunlop, P. D., Knight, C., Klonek, F. E., & Parent-Rocheleau, X. (2022). Understanding and shaping the future of work with self-determination theory. *Nature Reviews Psychology*, *1*(7), 378392.

Gallagher, W. (2023). *Apple Vision Pro draws investors back to struggling VR/AR market*. AppleInsider. Retrieved April 12, 2024, from https://appleinsider.com/articles/23/07/26/apple-vision-pro-draws-investors-back-to-struggling-vrar-market

Gartner. (2018). *5 Trends Emerge in the Gartner Hype Cycle for Emerging Technologies, 2018*. Retrieved April 12, 2024, from https://www.gartner.com/smarterwithgartner/5-trends-emerge-in-gartner-hype-cycle-for-emerging-technologies-2018/

Gecevicius, T., Chuang, Y., & An, J. (2021). Smart ARbnb: Smarthome Interface for Airbnb with Augmented Reality and Visible Light Communication. In *Proceedings of the 2021 Workshops on Computer Human Interaction in IoT Applications co-located with the International Conference on Embedded Wireless Systems and Networks (EWSN 2021) and the 13th ACM SIGCHI Symposium on Engineering Interactive Computing Systems (EICS 2021)*. CEUR-WS.org.

Geman, O., Sanei, S., Costin, H.-N., Eftaxias, K., Vyšata, O., Procházka, A., & Lhotská, L. (2015). Challenges and trends in Ambient Assisted Living and intelligent tools for disabled and elderly people. *2015 International Workshop on Computational Intelligence for Multimedia Understanding (IWCIM)*, 1–5. Retrieved April 12, 2024, from https://ieeexplore.ieee.org/abstract/document/7347088/

Geronazzo, M., & Serafin, S. (Eds.). (2023). *Sonic Interactions in Virtual Environments*. Springer International Publishing.

Ghorbani, F., Kia, M., Delrobaei, M., & Rahman, Q. (2019). Evaluating the possibility of integrating augmented reality and Internet of Things technologies to help patients with Alzheimer's disease. *2019 26th National and 4th International Iranian Conference on Biomedical Engineering (ICBME)*, 139–144. Retrieved April 12, 2024, from https://ieeexplore.ieee.org/abstract/document/9030404/

Giménez Manuel, J. G., Augusto, J. C., & Stewart, J. (2022). AnAbEL: Towards empowering people living with dementia in ambient assisted living. *Universal Access in the Information Society*, *21*(2), 457–476.

Goldsmith, S. (2000). *Universal design: A Manual of Practical Guidance for Architects*. Routledge. Retrieved March 28, 2024, from http://archive.org/details/UNIVERSAL_DESIGN

Google. (2019). *Google AR & VR | Home*. Google AR & VR. Retrieved April 12, 2024, from https://arvr.google.com/

Grand View Research. (2024). *Augmented Reality Market Size And Share Report, 2030*. Retrieved April 12, 2024, from https://www.grandviewresearch.com/industry-analysis/augmented-reality-market

Greenfeld, A., Lugmayr, A., & Lamont, W. (2018). Comparative reality: Measuring user experience and emotion in immersive virtual environments. In *2018 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)* (pp. 204–209). IEEE.

Gregor, S., & Hevner, A. R. (2013). Positioning and Presenting Design Science Research for Maximum Impact. *MIS Quarterly*, *37*, 337–355.

Grubert, J., Witzani, L., Ofek, E., Pahud, M., Kranz, M., & Kristensson, P. O. (2018). Text Entry in Immersive Head-Mounted Display-based Virtual Reality using Standard Keyboards. In IEEE (Ed.), *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*.

Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): A vision, architectural elements, and future directions. *Future generation computer systems*, *29*(7), 1645–1660. Retrieved April 11, 2024, from https://www.sciencedirect.com/science/article/pii/S0167739X13000241

Guo, M.-H., Xu, T.-X., Liu, J.-J., Liu, Z.-N., Jiang, P.-T., Mu, T.-J., Zhang, S.-H., Martin, R. R., Cheng, M.-M., & Hu, S.-M. (2022). Attention mechanisms in computer vision: A survey. *Computational Visual Media*, *8*(3), 331–368.

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).

Gurman, M., Webb, A., & Bergen, M. (2017). *Apple's Next Big Thing: Augmented Reality*. Retrieved April 12, 2024, from https://www.bloomberg.com/news/articles/2017-03-20/apple-s-next-big-thing

Hadjidemetriou, G., Belk, M., Fidas, C., & Pitsillides, A. (2019). Picture passwords in mixed reality: Implementation and evaluation. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–6).

Häfner, V. (2019). *PolyVR - A Virtual Reality Authoring Framework for Engineering Applications*.

Hajian, S. (2019). Transfer of Learning and Teaching: A Review of Transfer Theories and Effective Instructional Practices. *IAFOR Journal of education*, *7*(1), 93–111.

Ham, J., Hong, J., Jang, Y., Ko, S. H., & Woo, W. (2014). Smart wristband: Touch-and-motion–tracking wearable 3D input device for smart glasses. In *Distributed, Ambient, and Pervasive Interactions: Second International Conference, DAPI 2014, Held as Part of HCI Interational 2014, Heraklion, Crete, Greece, June 22-27, 2014. Proceedings 2* (pp. 109–118). Springer.

Hansen, J. P., Tørning, K., Johansen, A. S., Itoh, K., & Aoki, H. (2004). Gaze typing compared with input by head and hand. In *Proceedings of the 2004 symposium on Eye tracking research & applications* (pp. 131–138).

Harari, Y. N. (2018). *21 lessons for the 21st century* (First edition). Spiegel & Grau.

Hargittai, E. (2010). Digital Na(t)ives? Variation in Internet Skills and Uses among Members of the "Net Generation". *Sociological Inquiry*, *80*(1), 92–113.

Harper, S. (2014). Economic and social implications of aging societies. *Science*, *346*(6209), 587–591.

Hart, S. G. (2006). NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting* (pp. 904–908, Vol. 50). SAGE Publications Sage CA: Los Angeles, CA.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology* (pp. 139–183). Elsevier.

Hartl, J. (2021). *Eden Culture: Ökologie des Herzens für ein neues Morgen*. Herder.

Hassenzahl, M., & Tractinsky, N. (2006). User experience - a research agenda. *Behaviour & Information Technology*, *25*(2), 91–97.

Hatmaker, T. (2023). *Meta's Reality Labs lost $13.7 billion on VR and AR last year*. TechCrunch. Retrieved April 12, 2024, from https://techcrunch.com/2023/02/03/metas-reality-labs-lost-13-7-billion-on-vr-and-ar-last-year/

Hayhurst, J. (2018). How Augmented Reality and Virtual Reality is Being Used to Support People Living with Dementia—Design Challenges and Future Directions. In T. Jung & M. C. Tom Dieck (Eds.), *Augmented Reality and Virtual Reality* (pp. 295–305). Springer International Publishing.

He, Z., Lutteroth, C., & Perlin, K. (2022). TapGazer: Text Entry with Finger Tapping and Gaze-directed Word Selection. In *CHI Conference on Human Factors in Computing Systems* (pp. 1–16).

Heath, A. (2023). *Apple has bought an AR headset startup called Mira*. The Verge. Retrieved April 12, 2024, from https://www.theverge.com/2023/6/6/23751350/apple-mira-ar-headset-startup

Heisenberg, W. (1927). Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Zeitschrift für Physik*, *43*(3-4), 172–198.

Herdina, M. (2020). *Augmented Reality Disappeared From Gartner's Hype Cycle - What's Next?* Retrieved April 12, 2024, from https://arpost.co/2020/09/25/augmented-reality-gartners-hype-cycle/

Hermens, H. J. (2016). 5G for Personalized Health and Ambient Assisted Living. *5G Outlook–Innovations and Applications*, 195. Retrieved March 28, 2024, from https://library.oapen.org/bitstream/handle/20.500.12657/59729/9781000795240.pdf?sequence=1#page=228

Heun, V., Hobin, J., & Maes, P. (2013a). Reality editor: Programming smarter objects. In *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication* (p. 307310).

Heun, V., Kasahara, S., & Maes, P. (2013b). Smarter objects: Using AR technology to program physical objects and their interactions. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems* (pp. 961–966).

Hevner, A., March, S. T., Park, J., & Ram, S. (2004). Design science research in information systems. *MIS Quarterly*, *28*(1), 75–105.

Holmes, K. (2020). *Mismatch: How inclusion shapes design*. Mit Press.

Hoppe, A. H., van de Camp, F., & Stiefelhagen, R. (2017). Interaction with three dimensional objects on diverse input and output devices: A survey. In *HCI International 2017–Posters' Extended Abstracts: 19th International Conference, HCI International 2017, Vancouver, BC, Canada, July 9–14, 2017, Proceedings, Part I 19* (pp. 130–139). Springer.

Howard, I. P., & Rogers, B. J. (1995). *Binocular vision and stereopsis*. Oxford University Press, USA.

Huang, B.-C., Hsu, J., Chu, E. T.-H., & Wu, H.-M. (2020). Arbin: Augmented reality based indoor navigation system. *Sensors*, *20*(20), 5890.

Huddleston Jr., T. (2022). *Apple CEO Tim Cook doesn't like the metaverse—he predicts a different technology will shape the future*. Retrieved April 12, 2024, from https://www.cnbc.com/2022/10/03/apple-ceo-tim-cook-doesnt-like-metaverse-prefers-augmented-reality.html

Huey, L. C., Sebastian, P., & Drieberg, M. (2011). Augmented reality based indoor positioning navigation tool. In *2011 IEEE Conference on Open Systems* (pp. 256–260). IEEE.

Hwang, A. S., Truong, K. N., Cameron, J. I., Lindqvist, E., Nygård, L., & Mihailidis, A. (2015). Co-designing ambient assisted living (AAL) environments: Unravelling the situated context of informal dementia care. *BioMed research international*, *2015*. Retrieved March 28, 2024, from https://www.hindawi.com/journals/bmri/2015/720483/abs/

Intel. (2024). *Virtual Reality vs. Augmented Reality vs. Mixed Reality*. Intel. Retrieved April 12, 2024, from https://www.intel.com/content/www/us/en/tech-tips-and-tricks/virtual-reality-vs-augmented-reality.html

Iwarsson, S., & Ståhl, A. (2003). Accessibility, usability and universal design—positioning and definition of concepts describing person-environment relationships. *Disability and Rehabilitation*, *25*(2), 57–66.

Jacob, R. J. K. (1991). The use of eye movements in human-computer interaction techniques: What you look at is what you get. *ACM Transactions on Information Systems (TOIS)*, *9*(2), 152–169.

Jakobi, T., Ogonowski, C., Castelli, N., Stevens, G., & Wulf, V. (2017). The catch (es) with smart home: Experiences of a living lab field study. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (p. 16201633).

Jakobi, T., Stevens, G., Castelli, N., Ogonowski, C., Schaub, F., Vindice, N., Randall, D., Tolmie, P., & Wulf, V. (2018). Evolving Needs in IoT Control and Accountability: A Longitudinal Study on Smart Home Intelligibility. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *2*(4), 1–28.

Janssen, J.-K. (2017). *Warum Microsoft seine VR-Brillen "Mixed Reality" nennt*. c't Magazin. Retrieved April 12, 2024, from https://www.heise.de/hintergrund/Windows-Mixed-Reality-Warum-Microsoft-seine-VR-Brillen-Mixed-Reality-nennt-3820657.html

Jarosiewicz, B., Sarma, A. A., Bacher, D., Masse, N. Y., Simeral, J. D., Sorice, B., Oakley, E. M., Blabe, C., Pandarinath, C., Gilja, V., Cash, S. S., Eskandar, E. N., Friehs, G., Henderson, J. M., Shenoy, K. V., Donoghue, J. P., & Hochberg, L. R. (2015). Virtual typing by people with tetraplegia using a self-calibrating intracortical brain-computer interface. *Science translational medicine*, *7*(313), 313ra179.

Jaschinski, C., & Allouch, S. B. (2014). Ambient assisted living: Benefits and barriers from a user-centered perspective. Retrieved March 28, 2024, from https://research.utwente.nl/en/publications/ambient-assisted-living-benefits-and-barriers-from-a-user-centere

Jensen, R. H., Strengers, Y., Kjeldskov, J., Nicholls, L., & Skov, M. B. (2018). Designing the desirable smart home: A study of household experiences and energy consumption impacts. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (p. 114).

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)* (pp. 692–703). IEEE. https://www.researchgate.net/profile/Haiyan-Jiang-20/publication/342063698_HiPad_Text_entry_for_Head-Mounted_Displays_Using_Circular_Touchpad/links/5efc076c458515505080f943/HiPad-Text-entry-for-Head-Mounted-Displays-Using-Circular-Touchpad.pdf

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)* (pp. 132–137). IEEE.

Jiang, L., Liu, D.-Y., & Yang, B. (2004). Smart home research. In IEEE (**typecollaborator**), *Proceedings of 2004 international conference on machine learning and cybernetics (IEEE Cat. No. 04EX826)* (p. 659663).

Jicol, C., Clarke, C., Tor, E., Dakin, R. M., Lancaster, T. C., Chang, S. T., Petrini, K., O'Neill, E., Proulx, M. J., & Lutteroth, C. (2023). Realism and Field of View Affect Presence in VR but Not the Way You Think. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (pp. 1–17).

Jo, D., & Kim, G. J. (2019). AR Enabled IoT for a Smart and Interactive Environment: A Survey and Future Directions. *Sensors*, *19*(19).

Jones, P. (2018). Contexts of Co-creation: Designing with System Stakeholders. In P. Jones & K. Kijima (Eds.), *Systemic Design* (pp. 3–52, Vol. 8). Springer Japan.

Jose, A. C., & Malekian, R. (2017). Improving smart home security: Integrating logical sensing into smart home. *IEEE Sensors Journal*, *17*(13), 4269–4286.

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Kaiser, E., Olwal, A., McGee, D., Benko, H., Corradini, A., Li, X., Cohen, P., & Feiner, S. (2003). Mutual disambiguation of 3D multimodal interaction in augmented and virtual reality. In *Proceedings of the 5th international conference on Multimodal interfaces* (pp. 12–19).

Kammler, F., Brinker, J., Vogel, J., Hmaid, T., & Thomas, O. (2019). How Do We Support Technical Tasks in the Age of Augmented Reality? Some Evidence from Prototyping in Mechanical Engineering. *International Conference on Information Systems (ICIS)*.

Kanno, K. M., Lamounier, E. A., Cardoso, A., Lopes, E. J., & de Lima, G. F. M. (2018). Augmented reality system for aiding mild Alzheimer patients and caregivers. *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 593–594. Retrieved April 12, 2024, from https://ieeexplore.ieee.org/abstract/document/8446143/

Keates, S., & Clarkson, P. (2003). Countering design exclusion: Bridging the gap between usability and accessibility. *Universal Access in the Information Society*, *2*(3), 215–225.

Keates, S., Clarkson, P. J., Harrison, L.-A., & Robinson, P. (2000). Towards a practical inclusive design approach. *Proceedings on the 2000 Conference on Universal Usability - CUU '00*, 45–52.

Kemper, J., & Linde, H. (2008). Challenges of passive infrared indoor localization. In *2008 5th Workshop on Positioning, Navigation and Communication* (pp. 63–70). IEEE.

Kendall, G. (2019). *Your Mobile Phone s. Apollo 11's Guidance Computer*. Retrieved April 12, 2024, from https://www.realclearscience.com/articles/2019/07/02/your_mobile_phone_vs_apollo_11s_guidance_computer_111026.html

Kim, H., Kwon, Y.-T., Lim, H.-R., Kim, J.-H., Kim, Y.-S., & Yeo, W.-H. (2021a). Recent Advances in Wearable Sensors and Integrated Functional Devices for Virtual and Augmented Reality Applications. *Advanced Functional Materials*, *31*(39), 2005692.

Kim, H. H., Ha, K. N., Lee, S., & Lee, K. C. (2009). Resident location-recognition algorithm using a Bayesian classifier in the PIR sensor-based indoor location-aware system. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *39*(2), 240–245.

Kim, J., Kim, Y. J., & Lee, J. (2022). *ExpTrialMng: A Universal Experiment Trial Manager for AR/VR/MR Experiments based on Unity*.

Kim, K., Billinghurst, M., Bruder, G., Duh, H. B.-L., & Welch, G. F. (2018). 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.

Kim, K., Li, S., Heydariaan, M., Smaoui, N., Gnawali, O., Suh, W., Suh, M. J., & Kim, J. (2021b). Feasibility of LoRa for smart home indoor localization. *Applied Sciences*, *11*(1), 415.

Kim, Y. R., & Kim, G. J. (2017). Hovr-type: Smartphone as a typing interface in vr using hovering. In *2017 IEEE International Conference on Consumer Electronics (ICCE)* (pp. 200–203). IEEE.

Kimura, N., Gemicioglu, T., Womack, J., Li, R., Zhao, Y., Bedri, A., Su, Z., Olwal, A., Rekimoto, J., & Starner, T. (2022). SilentSpeller: Towards mobile, hands-free, silent speech text entry using electropalatography. In *CHI Conference on Human Factors in Computing Systems* (pp. 1–19).

Kiss, F., & Schmidt, A. (2019). Stressed by Design? The Problems of Transferring Interaction Design from Workstations to Mobile Interfaces. In *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare* (pp. 377–382). Association for Computing Machinery.

Kitzinger, A. (2022). *METACOM 9*. Retrieved April 12, 2024, from https://www.metacom-symbole.de/metacom_en.html

Klopfer, E., & Squire, K. (2008). Environmental Detectives—the development of an augmented reality platform for environmental simulations. *Educational Technology Research and Development*, *56*(2), 203–228.

Knierim, P., Woźniak, P. W., Abdelrahman, Y., & Schmidt, A. (2019). Exploring the potential of augmented reality in domestic environments. In *Proceedings of the 21st International Conference on human-computer interaction with mobile devices and services* (pp. 1–12).

Kohn, V., & Harborth, D. (2018). Augmented Reality - A Game Changing Technology For Manufacturing Processes? *European Conference on Information Systems (ECIS)*.

Kortekamp, S.-S., Werning, S., Thomas, O., & Ickerott, I. (2019). The Future of Digital Work - Use Cases For Augmented Reality Glasses. *European Conference on Information Systems (ECIS)*.

Kristensson, P. O., & Vertanen, K. (2012). The potential of dwell-free eye-typing for fast assistive gaze communication. In *Proceedings of the symposium on eye tracking research and applications* (pp. 241–244).

Krueger, R. A. (2014). *Focus groups: A practical guide for applied research*. Sage publications.

Kuechler, B., & Vaishnavi, V. (2008). On theory development in design science research: Anatomy of a research project. *European Journal of Information Systems*, *17*(5), 489–504.

Kuester, F., Chen, M., Phair, M. E., & Mehring, C. (2005). Towards keyboard independent touch typing in VR. In *Proceedings of the ACM symposium on Virtual reality software and technology* (pp. 86–95).

Kurauchi, A., Feng, W., Joshi, A., Morimoto, C., & Betke, M. (2016). EyeSwipe: Dwell-free text entry using gaze paths. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 1952–1956).

Kurauchi, A., Feng, W., Joshi, A., Morimoto, C. H., & Betke, M. (2020). Swipe&Switch: Text Entry Using Gaze Paths and Context Switching. In *Adjunct Publication of the 33rd Annual ACM Symposium on User Interface Software and Technology* (pp. 84–86).

Kymäläinen, T. (2016). Introduction to the special session: Design and research for advanced human augmentation. In *2016 12th International Conference on Intelligent Environments (IE)* (pp. 144–146). IEEE.

Labonnote, N., & Høyland, K. (2017). Smart home technologies that support independent living: Challenges and opportunities for the building industry–a systematic mapping study. *Intelligent Buildings International*, *9*(1), 40–63.

Labs, F. R. (2021). *Inside Facebook Reality Labs: Wrist-based interaction for the next computing platform*. Retrieved April 12, 2024, from https://tech.facebook.com/reality-labs/2021/3/inside-facebook-reality-labs-wrist-based-interaction-for-the-next-computing-platform/

Leap, M. (2024). *AR for Business | Augmented Reality*. Magic Leap. Retrieved April 12, 2024, from https://www.magicleap.com

Lebeck, K., Ruth, K., Kohno, T., & Roesner, F. (2018). Towards security and privacy for multi-user augmented reality: Foundations with end users. In *2018 IEEE Symposium on Security and Privacy (SP)* (pp. 392–408). IEEE.

Lee, C.-J., & Chu, H.-K. (2018). Dual-mr: Interaction with mixed reality using smartphones. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology* (pp. 1–2).

Lee, L.-H., Braud, T., Hosio, S., & Hui, P. (2021a). Towards augmented reality driven human-city interaction: Current research on mobile headsets and future challenges. *ACM Computing Surveys (CSUR)*, *54*(8), 1–38.

Lee, L.-H., Braud, T., Zhou, P., Wang, L., Xu, D., Lin, Z., Kumar, A., Bermejo, C., & Hui, P. (2021b). *All One Needs to Know about Metaverse: A Complete Survey on Technological Singularity, Virtual Ecosystem, and Research Agenda*. arXiv: 2110.05352 [cs]. Retrieved April 5, 2024, from http://arxiv.org/abs/2110.05352

Lee, L.-H., & Hui, P. (2018). Interaction methods for smart glasses: A survey. *IEEE access*, *6*, 28712–28732.

Lee, Y., & Kim, G. J. (2017). Vitty: Virtual touch typing interface with added finger buttons. In *Virtual, Augmented and Mixed Reality: 9th International Conference, VAMR 2017, Held as Part of HCI International 2017, Vancouver, BC, Canada, July 9-14, 2017, Proceedings 9* (pp. 111–119). Springer.

Leswing, K. (2020). *Apple buys virtual reality company NextVR*. CNBC. Retrieved April 12, 2024, from https://www.cnbc.com/2020/05/14/apple-buys-virtual-reality-company-nextvr.html

Li, Y., Sarcar, S., Zheng, Y., & Ren, X. (2021). Exploring Text Revision with Backspace and Caret in Virtual Reality. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–12).

Lin, S., Cheng, H. F., Li, W., Huang, Z., Hui, P., & Peylo, C. (2016). Ubii: Physical world interaction through augmented reality. *IEEE Transactions on Mobile Computing*, *16*(3), 872–885.

Lin, Y. H., Chang, S. H., Huang, T. C., Lin, Y. T., & Chen, Y. J. (2019). Augmented reality enhanced ubiquitous-learning in museum. *International Journal of Information and Education Technology*, *9*(8), 543–547.

Liu, K., Motta, G., & Ma, T. (2016). Xyz indoor navigation through augmented reality: A research in progress. In *2016 IEEE International Conference on Services Computing (SCC)* (pp. 299–306). IEEE.

Lloret, J., Canovas, A., Sendra, S., & Parra, L. (2015). A smart communication architecture for ambient assisted living. *IEEE Communications Magazine*, *53*(1), 26–33. Retrieved March 28, 2024, from https://ieeexplore.ieee.org/abstract/document/7010512/

Longo, U. G., De Salvatore, S., Candela, V., Zollo, G., Calabrese, G., Fioravanti, S., Giannone, L., Marchetti, A., De Marinis, M. G., & Denaro, V. (2021). Augmented reality, virtual reality and artificial intelligence in orthopedic surgery: A systematic review. *Applied Sciences*, *11*(7), 3253. Retrieved April 11, 2024, from https://www.mdpi.com/2076-3417/11/7/3253

Lu, X., Yu, D., Liang, H.-N., & Goncalves, J. (2021). iText: Hands-free text entry on an imaginary keyboard for augmented reality systems. In *The 34th Annual ACM Symposium on User Interface Software and Technology* (pp. 815–825).

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)* (pp. 344–349). IEEE.

Lu, Y., Chen, C., Chen, P., & Yu, S. (2023). Designing Social Robot for Adults Using Self-Determination Theory and AI Technologies. *IEEE Transactions on Learning Technologies*, *16*(2), 206218.

Luckey, P. (2015). *Reminder of something I have talked about before: VR will become something everyone wants before it becomes something everyone can afford.* Twitter. Retrieved April 12, 2024, from https://twitter.com/PalmerLuckey/status/679907802962214912

Luo, S., Xia, H., Gao, Y., Jin, J. S., & Athauda, R. (2008). Smart fridges with multimedia capability for better nutrition and health. In IEEE (**typecollaborator**), *2008 International Symposium on Ubiquitous Multimedia Computing* (p. 3944).

Lystbæk, M. N., Pfeuffer, K., Grønbæk, J. E. S., & Gellersen, H. (2022). Exploring gaze for assisting freehand selection-based text entry in ar. *Proceedings of the ACM on Human-Computer Interaction*, *6*, 1–16.

Lyu, Z., Yang, J., Lam, M. S., & Landay, J. A. (2022). HomeView: Automatically Building Smart Home Digital Twins With Augmented Reality Headsets. In *Adjunct Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology* (p. 16).

Mace, R. (1997). What is universal design. *The Center for Universal Design at North Carolina State University*, *19*, 2004. Retrieved March 28, 2024, from https://www.uwyo.edu/wind/_files/docs/resources/ud_review.pdf

MacKenzie, I. S. (2010). Evaluation of Text Entry Techniques. In I. S. MacKenzie & K. Tanaka-Ishii (Eds.), *Text entry systems: Mobility, accessibility, universality* (pp. 75–101). Elsevier. https://www.yorku.ca/mack/chapter4.html

MacKenzie, I. S., & Soukoreff, R. W. (2002). Text entry for mobile computing: Models and methods, theory and practice. *Human–Computer Interaction*, *17*(2-3), 147–198.

MacKenzie, I. S., & Soukoreff, R. W. (2003). Phrase sets for evaluating text entry techniques. In *CHI'03 extended abstracts on Human factors in computing systems* (pp. 754–755).

Mahroo, A., Greci, L., & Sacco, M. (2019). HoloHome: An augmented reality framework to manage the smart home. In *International Conference on Augmented Reality, Virtual Reality and Computer Graphics* (pp. 137–145). Springer.

Majaranta, P., Ahola, U.-K., & Špakov, O. (2009). Fast gaze typing with an adjustable dwell time. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 357–360).

Majaranta, P., Aula, A., & Räihä, K.-J. (2004). Effects of feedback on eye typing with a short dwell time. In *Proceedings of the 2004 symposium on Eye tracking research & applications* (pp. 139–146).

Mann, S., Furness, T., Yuan, Y., Iorio, J., & Wang, Z. (2018). *All reality: Virtual, augmented, mixed (x), mediated (x, y), and multimediated reality*.

Mansson, L., Wiklund, M., Öhberg, F., Danielsson, K., & Sandlund, M. (2020). Co-creation with older adults to improve user-experience of a smartphone self-test application to assess balance function. *International journal of environmental research and public health*, *17*(11), 3768. Retrieved March 28, 2024, from https://www.mdpi.com/1660-4601/17/11/3768

Marikyan, D., Papagiannidis, S., & Alamanos, E. (2019). A systematic review of the smart home literature: A user perspective. *Technological Forecasting and Social Change*, *138*, 139–154.

Markussen, A., Jakobsen, M. R., & Hornbæk, K. (2014). Vulture: A mid-air word-gesture keyboard. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1073–1082).

Marques, B., Dias, P., Alves, J., & Santos, B. S. (2019). Adaptive augmented reality user interfaces using face recognition for smart home control. In *International Conference on Human Systems Engineering and Design: Future Trends and Applications* (pp. 15–19). Springer.

Marques, G. (2019). Ambient assisted living and internet of things. *Harnessing the Internet of everything (IoE) for accelerated innovation opportunities*, 100–115. Retrieved March 28, 2024, from https://www.igi-global.com/chapter/ambient-assisted-living-and-internet-of-things/221284

Maskeliūnas, R., Damaševičius, R., & Segal, S. (2019). A review of internet of things technologies for ambient assisted living environments. *Future Internet*, *11*(12), 259. Retrieved March 28, 2024, from https://www.mdpi.com/1999-5903/11/12/259

Masood, T., & Egger, J. (2020). Adopting augmented reality in the age of industrial digitalisation. *Computers in Industry*, *115*, 103–112.

Mayer, S., Hassan, Y. N., & Sörös, G. (2014). A magic lens for revealing device interactions in smart environments. In *SIGGRAPH Asia 2014 Mobile Graphics and Interactive Applications* (pp. 1–6).

McKinsey. (2022). McKinsey Technology Trends Outlook 2022. Retrieved April 12, 2024, from https://www.mckinsey.com/spContent/bespoke/tech-trends/pdfs/mckinsey-tech-trends-outlook-2022-immersive-reality.pdf

Mehring, C., Kuester, F., Singh, K. D., & Chen, M. (2004). KITTY: Keyboard independent touch typing in VR. In *IEEE Virtual Reality 2004* (pp. 243–244). IEEE.

Meiland, F. J. M., Hattink, B. J. J., Overmars-Marx, T., De Boer, M. E., Jedlitschka, A., Ebben, P. W. G., Stalpers-Croeze, I. I., Flick, S., Van der Leeuw, J., & Karkowski, I. P. (2014). Participation of end users in the design of assistive technology for people with mild to severe cognitive problems; the European Rosetta project. *International psychogeriatrics*, *26*(5), 769–779. Retrieved March 28, 2024, from https://www.cambridge.org/core/journals/international-psychogeriatrics/article/participation-of-end-users-in-the-design-of-assistive-technology-for-people-with-mild-to-severe-cognitive-problems-the-european-rosetta-project/12D31FD90430B5AD9FC5D3B99E5E37A3

Meli, L., Pacchierotti, C., Salvietti, G., Chinello, F., Maisto, M., Luca, A., & Prattichizzo, D. (2018). Combining wearable finger haptics and augmented reality: User evaluation using an external camera and the microsoft hololens. *IEEE Robotics and Automation Letters*, *3*(4), 4297–4304.

Messaci, A., Zenati, N., Belhocine, M., & Otmane, S. (2022). Zoom-fwd: Efficient technique for 3D gestual interaction with distant and occluded objects in virtual reality. *Computer Animation and Virtual Worlds*, *33*(1), 2033.

Meta. (2021, October 28). *The Facebook Company Is Now Meta*. Meta. Retrieved April 12, 2024, from https://about.fb.com/news/2021/10/facebook-company-is-now-meta/

Meta. (2024a, February 14). *AR, VR and MR*. Meta for Work. Retrieved April 12, 2024, from https://forwork.meta.com/de/en/blog/difference-between-vr-ar-and-mr/

Meta. (2024b). *Meta Quest 3: Neues Mixed-Reality-VR-Headset – Jetzt kaufen | Meta Store*. Retrieved April 12, 2024, from https://www.meta.com/us/quest/quest-3/

Metz, R. (2022). *Magic Leap raised billions but its headset flopped. Now it's trying again | CNN Business*. CNN. Retrieved April 12, 2024, from https://www.cnn.com/2022/03/10/tech/magic-leap-new-headset/index.html

Microsoft. (2020). *What can HoloLens 2 do?* Retrieved April 12, 2024, from https://www.youtube.com/watch?v=XwOnHqiNAeU

Microsoft. (2022). *Holograms and interactions troubleshooting*. Retrieved April 12, 2024, from https://learn.microsoft.com/en-us/hololens/hololens-faq

Microsoft. (2024). *What is Augmented Reality (AR) | Microsoft Dynamics 365*. Retrieved April 12, 2024, from https://dynamics.microsoft.com/en-us/mixed-reality/guides/what-is-augmented-reality-ar/

Microsoft Design. (2016). *Inclusive*. Retrieved April 12, 2024, from https://inclusive.microsoft.design/tools-and-activities/Inclusive101Guidebook.pdf

Milgram, P., & Kishino, F. (1994). A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems*, *77*(12), 1321–1329.

Milgram, P., Takemura, H., Utsumi, A., & Kishino, F. (1995). Augmented reality: A class of displays on the reality-virtuality continuum. *Telemanipulator and Telepresence Technologies*, *2351*, 282–292. Retrieved April 12, 2024, from https://www.spiedigitallibrary.org/conference-proceedings-of-spie/2351/1/Augmented-reality--a-class-of-displays-on-the-reality/10.1117/12.197321.short

Minoli, D., & Occhiogrosso, B. (2018). Ultrawideband (UWB) technology for smart cities IoT applications. In IEEE (**typecollaborator**), *2018 IEEE international smart cities conference (ISC2)* (p. 18).

Moon, N. W., Baker, P. M., & Goughnour, K. (2019). Designing wearable technologies for users with disabilities: Accessibility, usability, and connectivity factors. *Journal of Rehabilitation and Assistive Technologies Engineering*, *6*, 205566831986213.

Moradi Dakhel, A., Majdinasab, V., Nikanjam, A., Khomh, F., Desmarais, M. C., & Jiang, Z. M. ((2023). GitHub Copilot AI pair programmer: Asset or Liability? *Journal of Systems and Software*, *203*, 111734.

Morar, A., Moldoveanu, A., Mocanu, I., Moldoveanu, F., Radoi, I. E., Asavei, V., Gradinaru, A., & Butean, A. (2020). A comprehensive survey of indoor localization methods based on computer vision. *Sensors*, *20*(9), 2641.

Morita, T., Taki, K., Fujimoto, M., Suwa, H., Arakawa, Y., & Yasumoto, K. (2018). BLE beacon-based activity monitoring system toward automatic generation of daily report. *2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, 788–793. Retrieved April 12, 2024, from https://ieeexplore.ieee.org/abstract/document/8480348/

Mourouzis, A., Kilintzis, V., Chouvarda, I., & Maglaveras, N. (2014). Using the LEAP 3D motion controller for implementing text entry for all. In *Proceedings of the 2014 European, Mediterranean and Middle Eastern Conference on Information Systems*.

Mshali, H., Lemlouma, T., Moloney, M., & Magoni, D. (2018). A survey on health monitoring systems for health smart homes. *International Journal of Industrial Ergonomics*, *66*, 26–56.

Mullen, B., Johnson, C., & Salas, E. (1991). Productivity loss in brainstorming groups: A meta-analytic integration. *Basic and applied social psychology*, *12*(1), 3–23.

Mulliner, E., Riley, M., & Maliene, V. (2020). Older people's preferences for housing and environment characteristics. *Sustainability*, *12*(14), 5723. Retrieved March 28, 2024, from https://www.mdpi.com/2071-1050/12/14/5723

Mur-Artal, R., & Tardós, J. D. (2017). Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras. *IEEE transactions on robotics*, *33*(5), 1255–1262.

Mutasim, A. K., Batmaz, A. U., & Stuerzlinger, W. (2021). Pinch, click, or dwell: Comparing different selection techniques for eye-gaze-based pointing in virtual reality. In *Acm symposium on eye tracking research and applications* (pp. 1–7).

Myers, B. A. (1998). A brief history of human-computer interaction technology. *Interactions*, *5*(2), 44–54.

NASA. (1999). *Mars Climate Orbiter*. Retrieved April 12, 2024, from https://science.nasa.gov/mission/mars-climate-orbiter/

NASA. (2021). *Nine Ways We Use AR and VR on the International Space Station - NASA*. Retrieved April 12, 2024, from https://www.nasa.gov/missions/station/nine-ways-we-use-ar-and-vr-on-the-international-space-station/

Nations, U. (2019). *World Population Ageing 2019*. UN. https://digitallibrary.un.org/record/3846855/files/WorldPopulationAgeing2019-Highlights.pdf

Newman, J., Bornik, A., Pustka, D., Echtler, F., Huber, M., Schmalstieg, D., & Klinker, G. (Eds.). (2007). *Tracking for distributed mixed reality environments*.

Nizam, S. S. M., Abidin, R. Z., Hashim, N. C., Lam, M. C., Arshad, H., & Majid, N. A. (2018). A review of multimodal interaction technique in augmented reality environment. *International Journal on Advanced Science, Engineering and Information Technology*, *8*(4-2), 1460–1469.

Norman, D. (2013). *The design of everyday things: Revised and expanded edition*. Basic books.

Norouzi, N., Bruder, G., Belna, B., Mutter, S., Turgut, D., & Welch, G. (2019). A Systematic Review of the Convergence of Augmented Reality, Intelligent Virtual Agents, and the Internet of Things. In F. Al-Turjman (Ed.), *Artificial Intelligence in IoT* (pp. 1–24). Springer International Publishing.

Norton, M. I., Mochon, D., & Ariely, D. (2012). The IKEA effect: When labor leads to love. *Journal of consumer psychology*, *22*(3), 453460.

Nott, G. (2022). *M&S takes AR wayfinding app trial public*. The Grocer. Retrieved April 12, 2024, from https://www.thegrocer.co.uk/marks-and-spencer/mands-takes-ar-wayfinding-app-trial-public/663880.article

Noyes, J. (1983). The QWERTY keyboard: A review. *International Journal of Man-Machine Studies*, *18*(3), 265–281.

Nunes, G. S., & Arruda Filho, E. J. M. (2018). Consumer behavior regarding wearable technologies: Google Glass. *Innovation & Management Review*, *15*(3), 230–246.

Oesterreich, T., & Teuteberg, F. (2017). Evaluating augmented reality applications in construction: A cost-benefit assessment framework based on VoFI. *European Conference on Information Systems (ECIS)*.

Oh, J. Y., Park, J.-H., & Park, J.-M. (2020). FingerTouch: Touch interaction using a fingernail-mounted sensor on a head-mounted display for augmented reality. *IEEE access*, *8*, 101192–101208.

Olafsen, A. H., Niemiec, C. P., Halvari, H., Deci, E. L., & Williams, G. C. (2017). On the dark side of work: A longitudinal analysis using self-determination theory. *European Journal of Work and Organizational Psychology*, *26*(2), 275285.

Oney, S., Harrison, C., Ogan, A., & Wiese, J. (2013). ZoomBoard: A diminutive qwerty soft keyboard using iterative zooming for ultra-small devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2799–2802).

Onwuegbuzie, A. J., Dickinson, W. B., Leech, N. L., & Zoran, A. G. (2009). A Qualitative Framework for Collecting and Analyzing Data in Focus Group Research. *International Journal of Qualitative Methods*, *8*(3), 1–21.

OpenAI. (2023). *Whisper*. Retrieved March 28, 2024, from https://github.com/openai/whisper

Özacar, K., Hincapié-Ramos, J. D., Takashima, K., & Kitamura, Y. (2017). 3D selection techniques for mobile augmented reality head-mounted displays. *Interacting with Computers*, *29*(4), 579–591.

Pang, A. S.-K. (2013). *The distraction addiction: Getting the information you need and the communication you want without enraging your family, annoying your colleagues, and destroying your soul* (1st ed). Little, Brown.
OCLC: ocn816563530.

Pannasch, S., Helmert, J. R., Roth, K., Herbold, A.-K., & Walter, H. (2008). Visual fixation durations and saccade amplitudes: Shifting relationship in a variety of conditions. *Journal of Eye Movement Research*, *2*(2).

Park, E., Kim, S., Kim, Y., & Kwon, S. J. (2018). Smart home services as the next mainstream of the ICT industry: Determinants of the adoption of smart home services. *Universal Access in the Information Society*, *17*, 175190.

Park, S., Cha, H.-S., Kwon, J., Kim, H., & Im, C.-H. (2020). Development of an online home appliance control system using augmented reality and an ssvep-based brain-computer interface. In *2020 8th International Winter Conference on Brain-Computer Interface (BCI)* (pp. 1–2). IEEE.

Park, S., Ha, J., Park, J., Lee, K., & Im, C.-H. (2022). Brain-controlled, AR-based Home automation system using SSVEP-based brain-computer interface and EOG-based eye tracker: A feasibility study for the elderly end User. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *31*, 544–553. Retrieved March 28, 2024, from https://ieeexplore.ieee.org/abstract/document/9978934/

Park, W. K. (2005). Mobile phone addiction. In *Mobile communications: Re-negotiation of the social sphere* (p. 253272). Springer.

Peng, T., & Sarazen, M. (2017). *Inside the Voice Inside Your Smart Speaker*. Retrieved April 12, 2024, from https://medium.com/syncedreview/inside-the-voice-inside-your-smart-speaker-e8eabb679740

Penkar, A. M., Lutteroth, C., & Weber, G. (2012). Designing for the eye: Design parameters for dwell in gaze interaction. In *Proceedings of the 24th Australian Computer-Human Interaction Conference* (pp. 479–488).

Perkins, D. N., & Salomon, G. (1992). Transfer of learning. *International encyclopedia of education*, *2*, 6452–6457.

Persson, H., Åhman, H., Yngling, A. A., & Gulliksen, J. (2015). Universal design, inclusive design, accessible design, design for all: Different concepts—one goal? On the concept of accessibility—historical, methodological and philosophical aspects. *Universal Access in the Information Society*, *14*(4), 505–526.

Peters, D., Calvo, R. A., & Ryan, R. M. (2018). Designing for motivation, engagement and wellbeing in digital experience. *Frontiers in Psychology*, *9*, 115.

Peukert, C., Pfeiffer, J., Meißner, M., Pfeiffer, T., & Weinhardt, C. (2019). Shopping in Virtual Reality Stores: The Influence of Immersion on System Adoption. *Journal of Management Information Systems*, *36*(3), 755–788.

Peukert, C., & Schenkluhn, M. (2023). Das Metaverse ist keine Insel. *Wirtschaftsinformatik & Management*, 1–10.

Peukert, C. D. (2020). *Next Generation Shopping: Experimental Studies on the Adoption and Design of Virtual Reality Shopping Environments*.

Pfeuffer, K., Mayer, B., Mardanbegi, D., & Gellersen, H. (2017). Gaze+ pinch interaction in virtual reality. In *Proceedings of the 5th symposium on spatial user interaction* (pp. 99–108).

Porter, M. E., & Heppelmann, J. E. (2017). Why every organization needs an augmented reality strategy. *Harvard Business Review's 10 Must Reads*, *85*, 85–108. Retrieved April 11, 2024, from https://bawar.net/data0/books/5cc4319927bb4/pdf/Harvard_Business_Review,_Joan_C__Williams,_Thomas.pdf#page=100

Poupyrev, I., Billinghurst, M., Weghorst, S., & Ichikawa, T. (1996). The go-go interaction technique: Non-linear mapping for direct manipulation in VR. In *Proceedings of the 9th annual ACM symposium on User interface software and technology* (p. 7980).

Powell, R. A., & Single, H. M. (1996). Focus groups. *International journal for quality in health care*, *8*(5), 499–504. Retrieved March 28, 2024, from https://academic.oup.com/intqhc/article-abstract/8/5/499/1843013

Prange, S., Shams, A., Piening, R., Abdelrahman, Y., & Alt, F. (2021). Priview-exploring visualisations to support users' privacy awareness. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (p. 118).

Prätorius, M., Burgbacher, U., Valkov, D., & Hinrichs, K. (2015). Sensing thumb-to-finger taps for symbolic input in vr/ar environments. *IEEE computer graphics and applications*, *35*(5), 42–54.

Prilla, M., Janßen, M., & Kunzendorff, T. (2019). How to Interact with Augmented Reality Head Mounted Devices in Care Work? A Study Comparing Handheld Touch (Hands-on) and Gesture (Hands-free) Interaction. *AIS Transactions on Human-Computer Interaction*, *11*(3), 157–178.

Purdy, K. (2022). *IoT harmony? What Matter and Thread really mean for your smart home*. Retrieved April 12, 2024, from https://arstechnica.com/gadgets/2022/10/matter-and-thread-could-fix-smart-home-compatibility-but-dont-get-excited-yet/

Pyae, A., & Joelsson, T. N. (2018). Investigating the usability and user experiences of voice user interface: A case of Google home smart speaker. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct* (pp. 127–131).

Quigley, A., Dix, A., Mackay, W. E., Ishii, H., & Steimle, J. (2013). Visions and visioning in chi: Chi 2013 special interest group meeting. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems* (pp. 2545–2548).

Radiah, R., Mäkelä, V., Prange, S., Rodriguez, S. D., Piening, R., Zhou, Y., Köhle, K., Pfeuffer, K., Abdelrahman, Y., & Hoppe, M. (2021). Remote VR studies: A framework for running virtual reality studies remotely via participant-owned HMDs. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *28*(6), 1–36.

Radu, I., Joy, T., & Schneider, B. (2021). Virtual makerspaces: Merging AR/VR/MR to enable remote collaborations in physical maker activities. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–5).

Rajanna, V., & Hansen, J. P. (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). https://dl.acm.org/doi/10.1145/3204493.3204541

Rashidi, P., & Mihailidis, A. (2012). A survey on ambient-assisted living tools for older adults. *IEEE journal of biomedical and health informatics*, *17*(3), 579–590.

Ratcliffe, J., Soave, F., Bryan-Kinns, N., Tokarchuk, L., & Farkhatdinov, I. (2021). Extended Reality (XR) remote research: A survey of drawbacks and opportunities. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–13).

Rauschnabel, P. A. (2021). Augmented reality is eating the real-world! The substitution of physical products by holograms. *International Journal of Information Management*, *57*, 102279.

Rauschnabel, P. A., Felix, R., Hinsch, C., Shahab, H., & Alt, F. (2022). What is XR? Towards a Framework for Augmented and Virtual Reality. *Computers in Human Behavior*, *133*, 107289.

Rauschnabel, P. A., & Ro, Y. K. (2016). Augmented reality smart glasses: An investigation of technology acceptance drivers. *International Journal of Technology Marketing*, *11*(2), 123.

Ravenscraft, E. (2023). What Is the Metaverse, Exactly? [magazine]. *Wired*. Retrieved April 5, 2024, from https://www.wired.com/story/what-is-the-metaverse/

Ray-Ban. (2024). *Entdecken Sie Ray-Ban | Meta Smart Glasses: Spezifikationen und Funktionen | Ray-Ban® DE*. Retrieved April 12, 2024, from https://www.ray-ban.com/germany/discover-ray-ban-meta-smart-glasses/clp

Regan, B. (2004). Accessibility and design: A failure of the imagination. *Proceedings of the International Cross-Disciplinary Workshop on Web Accessibility - W4A*, 29.

Rialle, V., Duchene, F., Noury, N., Bajolle, L., & Demongeot, J. (2002). Health" smart" home: Information technology for patients at home. *Telemedicine Journal and E-Health*, *8*(4), 395–409.

Richards, D., Caldwell, P. H. Y., & Go, H. (2015). Impact of social media on the health of children and young people. *Journal of paediatrics and child health*, *51*(12), 11521157.

Robles, R. J., & Kim, T.-h. (2010). Applications, systems and methods in smart home technology: A. *Int. Journal of Advanced Science And Technology*, *15*, 3748.

Rolland, J. P., Holloway, R. L., & Fuchs, H. (1995). Comparison of optical and video see-through, head-mounted displays. *Telemanipulator and Telepresence Technologies*, *2351*, 293–307.

Roller, J. (2022). *Coding From 1849 to 2022: A Guide to The Timeline of Programming Languages*. IEEE Computer Society. Retrieved April 12, 2024, from https://www.computer.org/publications/tech-news/insider-membership-news/timeline-of-programming-languages/

Rommelse, N. N. J., van der Stigchel, S., & Sergeant, J. A. (2008). A review on eye movement studies in childhood and adolescent psychiatry. *Brain and cognition*, *68*(3), 391–414.

Rosa, H. (2016). *Resonanz: eine Soziologie der Weltbeziehung* (3. Aufl). Suhrkamp.

Rosenberg, R., & Slater, M. (1999). The chording glove: A glove-based text input device. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *29*(2), 186–191. http://www.cs.columbia.edu/~feiner/courses/mobwear/resources/The%20Chording%20Glove.pdf

Rosenthal, R. (1976). *Experimenter Effects in Behavioral Research*. Irvington Publishers.

Rosling, H., Rosling, O., & Rosling Rönnlund, A. (2018). *Factfulness: Ten Reasons We're Wrong About the World - and Why Things are Better Than You Think* (Paperback edition). SCEPTRE.

Rowan, J., & Mynatt, E. D. (2005). Digital Family Portrait Field Trial: Support for Aging in Place. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 521–530.

Ryan, R., & Deci, E. (2000). Self-determination theory and the facilitation of intrinsic motivation. *American Psychologist*, *55*(1), 6878.

Ryan, R. M., Rigby, C. S., & Przybylski, A. (2006). The motivational pull of video games: A self-determination theory approach. *Motivation and Emotion*, *30*(4), 347363.

Sadikov, E., Madhavan, J., Wang, L., & Halevy, A. (2010). Clustering query refinements by user intent. In *Proceedings of the 19th International Conference on World Wide Web, WWW '10* (pp. 841–850).

Sahu, C. K., Young, C., & Rai, R. (2021). Artificial intelligence (AI) in augmented reality (AR)-assisted manufacturing applications: A review. *International Journal of Production Research*, *59*(16), 4903–4959.

Saidi, H., Serrano, M., Irani, P., Hurter, C., & Dubois, E. (2019). On-body tangible interaction: Using the body to support tangible manipulations for immersive environments. In *Human-Computer Interaction–INTERACT 2019: 17th IFIP TC 13 International Conference, Paphos, Cyprus, September 2–6, 2019, Proceedings, Part IV 17* (pp. 471–492). Springer.

Saket, B., Endert, A., & Stasko, J. (2016). Beyond usability and performance: A review of user experience-focused evaluations in visualization. In *Proceedings of the Sixth Workshop on Beyond Time and Errors on Novel Evaluation Methods for Visualization* (p. 133142).

Sandoval, D. C., & Favela, J. (2017). Co-designing ambient-assisted interventions using digital interlocutors for people with dementia. *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*, 813–821.

Sauer, J., Sonderegger, A., & Schmutz, S. (2020). Usability, user experience and accessibility: Towards an integrative model. *Ergonomics*, *63*(10), 1207–1220.

Sauro, J. (2011). *A practical guide to the system usability scale: Background, benchmarks & best practices*. Measuring Usability LLC. https://measuringu.com/interpret-sus-score/

Schenkluhn, M., Knierim, M. T., Kiss, F., & Weinhardt, C. (2024). Connecting Home: Human-Centric Setup Automation in the Augmented Smart Home. *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*, 16.

Schenkluhn, M., Peukert, C., Greif-Winzrieth, A., & Weinhardt, C. (2023a). Does One Keyboard Fit All? Comparison and Evaluation of Device-Free Augmented Reality Keyboard Designs. *Proceedings of the 29th ACM Symposium on Virtual Reality Software and Technology*, 1–11.

Schenkluhn, M., Peukert, C., & Weinhardt, C. (2022a). A Look Behind the Curtain: Exploring the Limits of Gaze Typing. In *NeuroIS Retreat 2022* (pp. 251–259). Springer International Publishing.

Schenkluhn, M., Peukert, C., & Weinhardt, C. (2022b). Typing the Future: Designing Multimodal AR Keyboards. *Wirtschaftsinformatik 2022 Proceedings*, *11*. https://aisel.aisnet.org/wi2022/hci/hci/11

Schenkluhn, M., Peukert, C., & Weinhardt, C. (2023b). Augmented Reality-based Indoor Positioning for Smart Home Automations. *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–6.

Schnall, R., Rojas, M., Bakken, S., Brown, W., Carballo-Dieguez, A., Carry, M., Gelaude, D., Mosley, J. P., & Travers, J. (2016). A user-centered model for designing consumer mobile health (mHealth) applications (apps). *Journal of biomedical informatics*, *60*, 243–251. Retrieved March 28, 2024, from https://www.sciencedirect.com/science/article/pii/S1532046416000241

Schrepp, M., Hinderks, A., & Thomaschewski, J. (2017). Design and evaluation of a short version of the user experience questionnaire (UEQ-S). *International Journal of Interactive Multimedia and Artificial Intelligence, 4 (6), 103-108*.

Schulz, M., Mack, B., & Renn, O. (2012). *Fokusgruppen in der empirischen Sozialwissenschaft: Von der Konzeption bis zur Auswertung*. Springer-Verlag.

Schulz, T., Fuglerud, K. S., Arfwedson, H., & Busch, M. (2014). A case study for universal design in the internet of things. In *Universal design 2014: Three days of creativity and diversity* (pp. 45–54). IOS Press. Retrieved March 28, 2024, from https://ebooks.iospress.nl/doi/10.3233/978-1-61499-403-9-45

Schulz von Thun, F., Ruppel, J., & Stratmann, R. (1981). *Miteinander reden* (60th ed., Vol. 1). Retrieved March 30, 2024, from https://sfbs.tu-dortmund.de/handle/sfbs/524

Schwind, V., Knierim, P., Haas, N., & Henze, N. (2019). Using presence questionnaires in virtual reality. In *Proceedings of the 2019 CHI conference on human factors in computing systems* (pp. 1–12).

Seiger, R., Gohlke, M., & Aßmann, U. (2019). Augmented reality-based process modelling for the internet of things with holoflows. In Springer (**typecollaborator**), *Enterprise, Business-Process and Information Systems Modeling: 20th International Conference, BPMDS 2019, 24th International Conference, EMMSAD 2019, Held at CAiSE 2019, Rome, Italy, June 3-4, 2019, Proceedings 20* (p. 115129).

Sensortec, B. (2023). Smartglasses Light Drive. https://www.bosch-sensortec.com/products/optical-microsystems/smartglasses-light-drive/#applications

Setz, B., Graef, S., Ivanova, D., Tiessen, A., & Aiello, M. (2021). A Comparison of Open-Source Home Automation Systems. *IEEE Access*, *9*, 167332–167352.

Shakir, U. (2023). *US Army orders more Microsoft AR headsets now that they no longer make soldiers want to barf*. The Verge. Retrieved April 12, 2024, from https://www.theverge.com/2023/9/13/23871859/us-army-microsoft-ivas-ar-goggles-success-new-contract-hololens

Sharma, R. A., Soltanaghaei, E., Rowe, A., & Sekar, V. (2022). Lumos: Identifying and Localizing Diverse Hidden 5IoT6 Devices in an Unfamiliar Environment. In *31st USENIX Security Symposium (USENIX Security 22)* (p. 10951112).

Shen, J., Hu, J., Dudley, J. J., & Kristensson, P. O. (2022). Personalization of a Mid-Air Gesture Keyboard using Multi-Objective Bayesian Optimization. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (pp. 702–710). IEEE.

Sikder, A. K., Babun, L., Celik, Z. B., Aksu, H., McDaniel, P., Kirda, E., & Uluagac, A. S. (2022). Who's controlling my device? Multi-user multi-device-aware access control system for shared smart home environment. *ACM Transactions on Internet of Things*, *3*(4), 139.

Singh, A., Rehman, S. U., Yongchareon, S., & Chong, P. H. J. (2020). Sensor technologies for fall detection systems: A review. *IEEE Sensors Journal*, *20*(13), 6889–6919.

Singhal, Y., Noeske, R. H., Bhardwaj, A., & Kim, J. R. (2022). Improving Finger Stroke Recognition Rate for Eyes-Free Mid-Air Typing in VR. In *CHI Conference on Human Factors in Computing Systems* (pp. 1–9).

Skarbez, R., Smith, M., & Whitton, M. C. (2021). Revisiting Milgram and Kishino's reality-virtuality continuum. *Frontiers in Virtual Reality*, *2*, 647997.

Skyquest. (2024). *Augmented Reality Market Size - Industry Forecast 2031*. Retrieved April 12, 2024, from https://www.skyquestt.com/report/augmented-reality-market

Slater, M., Sanchez-Vives, M. V., Rizzo, A., & Bergamasco, M. (2019). *The Impact of Virtual and Augmented Reality on Individuals and Society*. Frontiers Media SA.

Smith, A. L., & Chaparro, B. S. (2015). Smartphone text input method performance, usability, and preference with younger and older adults. *Human factors*, *57*(6), 1015–1028.

Solaimani, S., Bouwman, H., & Secomandi, F. (2013). Critical design issues for the development of Smart Home technologies. *Journal of Design Research*, *11*(1), 72.

Sommerauer, P., & Müller, O. (2018). Augmented reality for teaching and learning–a literature review on theoretical and empirical foundations. *European Conference on Information Systems (ECIS)*.

Song, Y., Huang, Y., Cai, Z., & Hong, J. I. (2020). I'm all eyes and ears: Exploring effective locators for privacy awareness in IoT scenarios. In *Proceedings of the 2020 CHI conference on human factors in computing systems* (p. 113).

Sorgalla, J., Fleck, J., & Sachweh, S. (2018). ARGI: Augmented Reality for Gesture-based Interaction in Variable Smart Environments. In *VISIGRAPP (2: HUCAPP)* (pp. 102–107).

Soukoreff, R. W., & MacKenzie, I. S. (2003). Metrics for text entry research: An evaluation of MSD and KSPC, and a new unified error metric. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 113–120).

Speicher, M., Feit, A. M., 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).

Speicher, M., Hall, B. D., & Nebeling, M. (2019). What is Mixed Reality? *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–15.

Statista. (2023). *Smart Home -Worldwide: Digital Markets Insights*. Retrieved April 12, 2024, from https://www.statista.com/outlook/dmo/smart-home/worldwide

Stavropoulos, T. G., Papastergiou, A., Mpaltadoros, L., Nikolopoulos, S., & Kompatsiaris, I. (2020). IoT wearable sensors and devices in elderly care: A literature review. *Sensors*, *20*(10), 2826.

Steed, A., Archer, D., Brandstätter, K., Congdon, B. J., Friston, S., Ganapathi, P., Giunchi, D., Izzouzi, L., Park, G. W. W., & Swapp, D. (2023). Lessons learnt running distributed and remote mixed reality experiments. *Frontiers in Computer Science*, *4*, 966319. Retrieved March 26, 2024, from https://www.frontiersin.org/articles/10.3389/fcomp.2022.966319/full

Steed, A., Izzouzi, L., Brandstätter, K., Friston, S., Congdon, B., Olkkonen, O., Giunchi, D., Numan, N., & Swapp, D. (2022). Ubiq-exp: A toolkit to build and run remote and distributed mixed reality experiments. *Frontiers in Virtual Reality*, *3*.

Steffen, J. H., Gaskin, J. E., Meservy, T. O., Jenkins, J. L., & Wolman, I. (2019). Framework of Affordances for Virtual Reality and Augmented Reality. *Journal of Management Information Systems*, *36*(3), 683–729.

Stein, N., Niehorster, D. C., Watson, T., Steinicke, F., Rifai, K., Wahl, S., & Lappe, M. (2021). A comparison of eye tracking latencies among several commercial head-mounted displays. *i-Perception*, *12*(1), 2041669520983338.

Steinfeld, E., & Maisel, J. (2012). *Universal Design: Creating Inclusive Environments*. John Wiley & Sons, Incorporated.

Stephanidis, C. (2001). User interfaces for all: New perspectives into human-computer interaction. *User interfaces for all-concepts, methods, and tools*, *1*(1), 3–17. Retrieved March 28, 2024, from https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=d08abee541705a0bcaedeb71344e410308a64b1d

Stephenson, N. (1994). *Snow Crash*. Penguin UK.

Story, M. F., & Mueller, J. L. (2001). Universal design of products. *New York, McGraw-Hill*. Retrieved April 12, 2024, from https://www.academia.edu/download/40536722/0071629238Universal.pdf#page=334

Streli, P., Jiang, J., Fender, A. R., Meier, M., Romat, H., & Holz, C. (2022). TapType: Ten-finger text entry on everyday surfaces via Bayesian inference. In *CHI Conference on Human Factors in Computing Systems* (pp. 1–16).

Subakti, H., & Jiang, J.-R. (2016). A marker-based cyber-physical augmented-reality indoor guidance system for smart campuses. In *2016 IEEE 18th International Conference on High Performance Computing and Communications; IEEE 14th International Conference on Smart City; IEEE 2nd International Conference on Data Science and Systems (HPCC/SmartCity/DSS)* (pp. 1373–1379). IEEE.

Subasi, Ö., Leitner, M., & Tscheligi, M. (2009). A Usability and Accessibility Design and Evaluation Framework for ICT Services. In A. Holzinger & K. Miesenberger (Eds.). D. Hutchison, T. Kanade, J. Kittler, J. M. Kleinberg, F. Mattern, J. C. Mitchell, M. Naor, O. Nierstrasz, C. Pandu Rangan, B. Steffen, M. Sudan, D. Terzopoulos, D. Tygar, M. Y. Vardi, & G. Weikum (**typeredactors**), *HCI and Usability for e-Inclusion* (pp. 102–110, Vol. 5889). Springer Berlin Heidelberg.

Sufi, F. (2023). Algorithms in Low-Code-No-Code for Research Applications: A Practical Review. *Algorithms*, *16*(2), 108.

Sun, M., Cao, M., Wang, L., & Qian, Q. (2020). Phonecursor: Improving 3d selection performance with mobile device in ar. *IEEE access*, *8*, 70616–70626.

Sutherland, I. E. (1965). The ultimate display. In *Proceedings of the IFIP Congress* (pp. 506–508, Vol. 2). New York.

Suzuki, S. (1985). Topological structural analysis of digitized binary images by border following. *Computer vision, graphics, and image processing*, *30*(1), 3246.

Swan, H., Pouncey, I., Pickering, H., & Watson, L. (2022). Inclusive design principles. *https://inclusivedesignprinciples. org/# principles*.

Tarafdar, P., Leung, A., Yue, W. T., & Bose, I. (2019). Impact of Immersive Interface Design on Consumer Perceptions during Online Product Presentation. *International Conference on Information Systems (ICIS)*.

Teo, T., Lee, G. A., Billinghurst, M., & Adcock, M. (2018). Hand gestures and visual annotation in live 360 panorama-based mixed reality remote collaboration. In *Proceedings of the 30th Australian conference on computer-human interaction* (pp. 406–410).

Tepper, O. M., Rudy, H. L., Lefkowitz, A., Weimer, K. A., Marks, S. M., Stern, C. S., & Garfein, E. S. (2017). Mixed reality with HoloLens: Where virtual reality meets augmented reality in the operating room. *Plastic and reconstructive surgery*, *140*(5), 1066–1070.

Thakur, N., & Han, C. Y. (2021). Indoor localization for personalized ambient assisted living of multiple users in multi-floor smart environments. *Big Data and cognitive computing*, *5*(3), 42. Retrieved March 28, 2024, from https://www.mdpi.com/2504-2289/5/3/42

Thies, I. M. (2015). User interface design for low-literate and novice users: Past, present and future. *Foundations and Trends® in Human–Computer Interaction*, *8*(1), 1–72. Retrieved April 11, 2024, from https://www.nowpublishers.com/article/Details/HCI-047

Thomas, O., Lambert, D., & Dayrit, B. (2020). Augmented reality space informatics system. In *HCI International 2020–Late Breaking Posters: 22nd International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings, Part II 22* (pp. 221–228). Springer.

Tian, Y., Guan, T., & Wang, C. (2010). Real-Time Occlusion Handling in Augmented Reality Based on an Object Tracking Approach. *Sensors*, *10*(4), 2885–2900.

Tiwari, S. V., Sewaiwar, A., & Chung, Y.-H. (2015). Smart home technologies using visible light communication. In IEEE (**typecollaborator**), *2015 IEEE International Conference on Consumer Electronics (ICCE)* (p. 379380).

Tsirmpas, C., Rompas, A., Fokou, O., & Koutsouris, D. (2015). An indoor navigation system for visually impaired and elderly people based on Radio Frequency Identification (RFID). *Information Sciences*, *320*, 288–305. Retrieved March 28, 2024, from https://www.sciencedirect.com/science/article/pii/S0020025514008007

Tuisku, O., Majaranta, P., Isokoski, P., & Räihä, K.-J. (2008). Now Dasher! Dash away! Longitudinal study of fast text entry by eye gaze. In *Proceedings of the 2008 symposium on Eye tracking research & applications* (pp. 19–26).

Turk, M. (2014). Multimodal interaction: A review. *Pattern recognition letters*, *36*, 189–195.

Tyack, A., & Mekler, E. D. (2020). Self-determination theory in HCI games research: Current uses and open questions. In *Proceedings of the 2020 CHI conference on human factors in computing systems* (p. 122).

Ullah, A. M., Islam, M. R., Aktar, S. F., & Hossain, S. A. (2012). Remote-touch: Augmented reality based marker tracking for smart home control. In *2012 15th International conference on computer and information technology (ICCIT)* (pp. 473–477). IEEE.

UXDA. (2020). *Case Study: World's First Mixed Reality VR / AR Banking UX/UI Design Concept*. Retrieved April 12, 2024, from https://medium.muz.li/case-study-worlds-first-mixed-reality-vr-ar-banking-ux-ui-design-concept-2f45d42e6a55

van der Vlist, B., Niezen, G., Rapp, S., Hu, J., & Feijs, L. (2013). Configuring and controlling ubiquitous computing infrastructure with semantic connections: A tangible and an AR approach. *Personal and ubiquitous computing, 17*(4), 783–799.

Varjo. (2023). *Mixed Reality Headset for Professionals – Varjo XR-4 Series*. Retrieved April 12, 2024, from https://varjo.com/products/xr-4/

Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision sciences*, *39*(2), 273–315.

Vertanen, K., Memmi, H., Emge, J., Reyal, S., & Kristensson, P. O. (2015). VelociTap: Investigating fast mobile text entry using sentence-based decoding of touchscreen keyboard input. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 659–668).

Vi, S., Da Silva, T. S., & Maurer, F. (2019). User experience guidelines for designing HMD extended reality applications. In *Human-Computer Interaction–INTERACT 2019: 17th IFIP TC 13 International Conference, Paphos, Cyprus, September 2–6, 2019, Proceedings, Part IV 17* (pp. 319–341). Springer.

Vlasenko, I., Nikolaidis, I., & Stroulia, E. (2014). The smart-condo: Optimizing sensor placement for indoor localization. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *45*(3), 436–453.

Vreede, T., Raghavan, M., & Vreede, G.-J. (2021). Design foundations for AI assisted decision making: A self determination theory approach. *Proceedings of the 54th Hawaii International Conference on System Sciences*, 166176.

Vrzakova, H., & Bednarik, R. (2013). That's not norma (n/l) a detailed analysis of midas touch in gaze-based problem-solving. In *Proceedings of the 2013 CHI Conference on Human Factors in Computing Systems* (pp. 85–90).

Wagner, D., Noui, L., & Stannard, A. (2019). *Why is making good AR displays so hard?* Retrieved April 12, 2024, from https://www.linkedin.com/pulse/why-making-good-ar-displays-so-hard-daniel-wagner

Walker, J., Kuhl, S., & Vertanen, K. (2016). Decoder-assisted typing using an HMD and a physical keyboard. *Extended Abstracts of the the ACM Conference on Human Factors in Computing Systems, CHI (Vol. 16)*, *16*.

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).

Wang, I., Smith, J., & Ruiz, J. (2019). Exploring Virtual Agents for Augmented Reality. In S. Brewster, G. Fitzpatrick, A. Cox, & V. Kostakos (Eds.), *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19* (pp. 1–12). ACM Press.

Wang, Z., Wang, H., Yu, H., & Lu, F. (2021). Interaction with gaze, gesture, and speech in a flexibly configurable augmented reality system. *IEEE transactions on human-machine systems*, *51*(5), 524–534.

Webster, A. (2024). *Apple opens App Store to game streaming services*. The Verge. Retrieved April 12, 2024, from https://www.theverge.com/2024/1/25/24050430/apple-app-store-game-streaming

Webster, J., & Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly*, *26*(2), -. www.jstor.org/stable/4132319

Weidner, J. B. (2023). *Why Google Glass Failed*. Investopedia. Retrieved April 12, 2024, from https://www.investopedia.com/articles/investing/052115/how-why-google-glass-failed.asp

Weiser, M. (1991). The computer for the 21st century. *Scientific American*, 94–104.

Weiser, M., & Brown, J. S. (1996). Designing Calm Technology. *PowerGrid Journal*, *1*(1), 75–85. Retrieved April 7, 2024, from https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=fdc2e87fcb4575bdf5154840ebd19c2dd490165c

Welch, G. F., Bruder, G., Squire, P., & Schubert, R. (2019). Anticipating widespread augmented reality: Insights from the 2018 ar visioning workshop. Retrieved March 28, 2024, from https://stars.library.ucf.edu/ucfscholar/786/?ref=blairmacintyre.me

Wherton, J., & Prendergast, D. (2009). The Building Bridges Project: Involving Older Adults in the Design of a Communication Technology to Support Peer-to-Peer Social Engagement. In D. Hutchison, T. Kanade, J. Kittler, J. M. Kleinberg, F. Mattern, J. C. Mitchell, M. Naor, O. Nierstrasz, C. Pandu Rangan, B. Steffen, M. Sudan, D. Terzopoulos, D. Tygar, M. Y. Vardi, G. Weikum, A. Holzinger, & K. Miesenberger (Eds.), *HCI and Usability for e-Inclusion* (pp. 111–134, Vol. 5889). Springer Berlin Heidelberg.

Williams, M., & Moser, T. (2019). The art of coding and thematic exploration in qualitative research. *International management review*, *15*(1), 45–55. Retrieved March 28, 2024, from https://www.academia.edu/download/82465402/imr-v15n1art4.pdf

Williamson, J. R., & Williamson, J. (2017). Understanding public evaluation: Quantifying experimenter intervention. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 3414–3425).

Wobbrock, J. O. (2010). Measures of Text Entry Performance. In I. S. MacKenzie & K. Tanaka-Ishii (Eds.), *Text entry systems: Mobility, accessibility, universality*. Elsevier.

Wong, L. P. (2008). Focus group discussion: A tool for health and medical research. *Singapore Med J*, *49*(3), 256–60. Retrieved March 28, 2024, from http://www.smj.org.sg/sites/default/files/4903/4903me1.pdf

Woźniak, M. P., Vöge, S., Krüger, R., Müller, H., Koelle, M., & Boll, S. (2023). Inhabiting Interconnected Spaces: How Users Shape and Appropriate Their Smart Home Ecosystems. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (p. 118).

Wu, H.-K., Lee, S. W.-Y., Chang, H.-Y., & Liang, J.-C. (2013). Current status, opportunities and challenges of augmented reality in education. *Computers & Education*, *62*, 41–49.

Wu, S., Byrne, D., & Steenson, M. W. (2020). Megereality": Leveraging Physical Affordances for Multi-Device Gestural Interaction in Augmented Reality. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (p. 14).

Xia, C., Zhang, C., Yan, X., Chang, Y., & Yu, P. S. (2018). Zero-shot user intent detection via capsule neural networks. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, EMNLP 2018* (pp. 3090–3099).

Xia, Q., Chiu, T. K. F., Lee, M., Sanusi, I. T., Dai, Y., & Chai, C. S. (2022). A self-determination theory (SDT) design approach for inclusive and diverse artificial intelligence (AI) education. *Computers & Education*, *189*, 104582.

Xiao, R., Schwarz, J., Throm, N., Wilson, A. D., & Benko, H. (2018). MRTouch: Adding touch input to head-mounted mixed reality. *IEEE transactions on visualization and computer graphics*, *24*(4), 1653–1660.

XReal. (2024). *XReal Air 2*. Retrieved April 12, 2024, from https://www.xreal.com/us/air2/

Xu, W., Liang, H.-N., Chen, Y., Li, X., & Yu, K. (2020). Exploring visual techniques for boundary awareness during interaction in augmented reality head-mounted displays. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (pp. 204–211). IEEE.

Xu, W., Liang, H.-N., He, A., & Wang, Z. (2019a). Pointing and selection methods for text entry in augmented reality head mounted displays. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (pp. 279–288). IEEE.

Xu, W., Liang, H.-N., Zhao, Y., Yu, D., & Monteiro, D. (2019b). DMove: Directional motion-based interaction for augmented reality head-mounted displays. In *Proceedings of the 2019 CHI conference on human factors in computing systems* (pp. 1–14).

Xu, W., Liang, H.-N., Zhao, Y., Zhang, T., Yu, D., & Monteiro, D. (2019c). Ringtext: Dwell-free and hands-free text entry for mobile head-mounted displays using head motions. *IEEE transactions on visualization and computer graphics*, *25*(5), 1991–2001.

Yang, X., & Aurisicchio, M. (2021). Designing conversational agents: A self-determination theory approach. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (p. 116).

Yao, Y., Huang, L., He, Y., Ma, Z., Xu, X., & Mi, H. (2023). Reviewing and Reflecting on Smart Home Research from the Human-Centered Perspective. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (p. 121).

Yeo, H.-S., Phang, X.-S., Ha, T., Woo, W., & Quigley, A. (2017). TiTAN: Exploring midair text entry using freehand input. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 3041–3049).

Yi, X., Yu, C., Zhang, M., Gao, S., Sun, K., & Shi, Y. (2015). Atk: Enabling ten-finger freehand typing in air based on 3d hand tracking data. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology* (pp. 539–548).

Yin, K., He, Z., Xiong, J., Zou, J., Li, K., & Wu, S.-T. (2021). Virtual reality and augmented reality displays: Advances and future perspectives. *Journal of Physics: Photonics*, *3*(2), 022010.

Yoo, S.-J., & Choi, S.-H. (2022). Indoor AR Navigation and Emergency Evacuation System Based on Machine Learning and IoT Technologies. *IEEE Internet of Things Journal*.

Yu, C., Gu, Y., Yang, Z., Yi, X., Luo, H., & Shi, Y. (2017). Tap, dwell or gesture? exploring head-based text entry techniques for hmds. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 4479–4488).

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 transactions on visualization and computer graphics*, *24*(11), 2927–2935. http://www.difeng.me/papers/18_PizzaText.pdf

Zafari, F., Gkelias, A., & Leung, K. K. (2019). A survey of indoor localization systems and technologies. *IEEE Communications Surveys & Tutorials*, *21*(3), 2568–2599.

Zdravevski, E., Lameski, P., Trajkovik, V., Kulakov, A., Chorbev, I., Goleva, R., Pombo, N., & Garcia, N. (2017). Improving activity recognition accuracy in ambient-assisted living systems by automated feature engineering. *Ieee Access*, *5*, 5262–5280. Retrieved March 28, 2024, from https://ieeexplore.ieee.org/abstract/document/7883880/

Zha, Z.-J., Yang, L., Mei, T., Wang, M., Wang, Z., Chua, T.-S., & Hua, X.-S. (Eds.). (2010). *Visual Query Suggestion: Towards capturing user intent in internet image search* (Vol. 6).

Zhang, M. R., Zhai, S., & Wobbrock, J. O. (2022). TypeAnywhere: A QWERTY-Based Text Entry Solution for Ubiquitous Computing. In *CHI Conference on Human Factors in Computing Systems* (pp. 1–16).

Zhang, X., Kulkarni, H., & Morris, M. R. (2017). Smartphone-based gaze gesture communication for people with motor disabilities. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 2878–2889).

Zhao, M., Pierce, A. M., Tan, R., Zhang, T., Wang, T., Jonker, T. R., Benko, H., & Gupta, A. (2023). Gaze Speedup: Eye Gaze Assisted Gesture Typing in Virtual Reality. In *Proceedings of the 28th International Conference on Intelligent User Interfaces* (pp. 595–606).

Zheng, M., Pan, X., Bermeo, N. V., Thomas, R. J., Coyle, D., O'hare, G. M., & Campbell, A. G. (2022). Stare: Augmented reality data visualization for explainable decision support in smart environments. *IEEE Access*, *10*, 29543–29557. Retrieved March 28, 2024, from https://ieeexplore.ieee.org/abstract/document/9727109/

Zheng, S., Apthorpe, N., Chetty, M., & Feamster, N. (2018). User perceptions of smart home IoT privacy. *Proceedings of the ACM on Human-Computer Interaction*, *2*, 120.

Zhou, B., Li, W., Chan, K. W., Cao, Y., Kuang, Y., Liu, X., & Wang, X. (2016). Smart home energy management systems: Concept, configurations, and scheduling strategies. *Renewable and Sustainable Energy Reviews*, *61*, 3040.

Zhu, M., Sun, Z., Zhang, Z., Shi, Q., He, T., Liu, H., Chen, T., & Lee, C. (2020). Haptic-feedback smart glove as a creative human-machine interface (HMI) for virtual/augmented reality applications. *Science Advances*, *6*(19), eaaz8693.

Ziefle, M., & Bay, S. (2005). How older adults meet complexity: Aging effects on the usability of different mobile phones. *Behaviour and Information Technology*, *24*(5), 375–389.

Zou, H., Zhou, Y., Yang, J., & Spanos, C. J. (2018). Device-free occupancy detection and crowd counting in smart buildings with WiFi-enabled IoT. *Energy and Buildings*, *174*, 309–322.

Zuckerberg, M. (2015). *VR / AR strategy and One*. Retrieved April 12, 2024, from https://sriramk.com/memos/zuck-unity.pdf

Declarations

Eidesstattliche Versicherung

gemäß §13 Absatz 2 Ziffer 3 der Promotionsordnung des Karlsruher Instituts für Technologie für die KIT-Fakultät für Wirtschaftswissenschaften

1. Bei der eingereichten Dissertation zu dem Thema *"Augmented Everything: Engineering Compelling Ubiquitous AR Experiences"* handelt es sich um meine eigenständig erbrachte Leistung.

2. Ich habe nur die angegebenen Quellen und Hilfsmittel benutzt und mich keiner unzulässigen Hilfe Dritter bedient. Insbesondere habe ich wörtlich oder sinngemäß aus anderen Werken übernommene Inhalte als solche kenntlich gemacht.

3. Die Arbeit oder Teile davon habe ich bislang nicht an einer Hochschule des In- oder Auslands als Bestandteil einer Prüfungs- oder Qualifikationsleistung vorgelegt.

4. Die Richtigkeit der vorstehenden Erklärungen bestätige ich.

5. Die Bedeutung der eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Versicherung sind mir bekannt.

Ich versichere an Eides statt, dass ich nach bestem Wissen die reine Wahrheit erklärt und nichts verschwiegen habe.

Karlsruhe, den 15.04.2024

Marius Franz Schenkluhn
(M.Sc.)