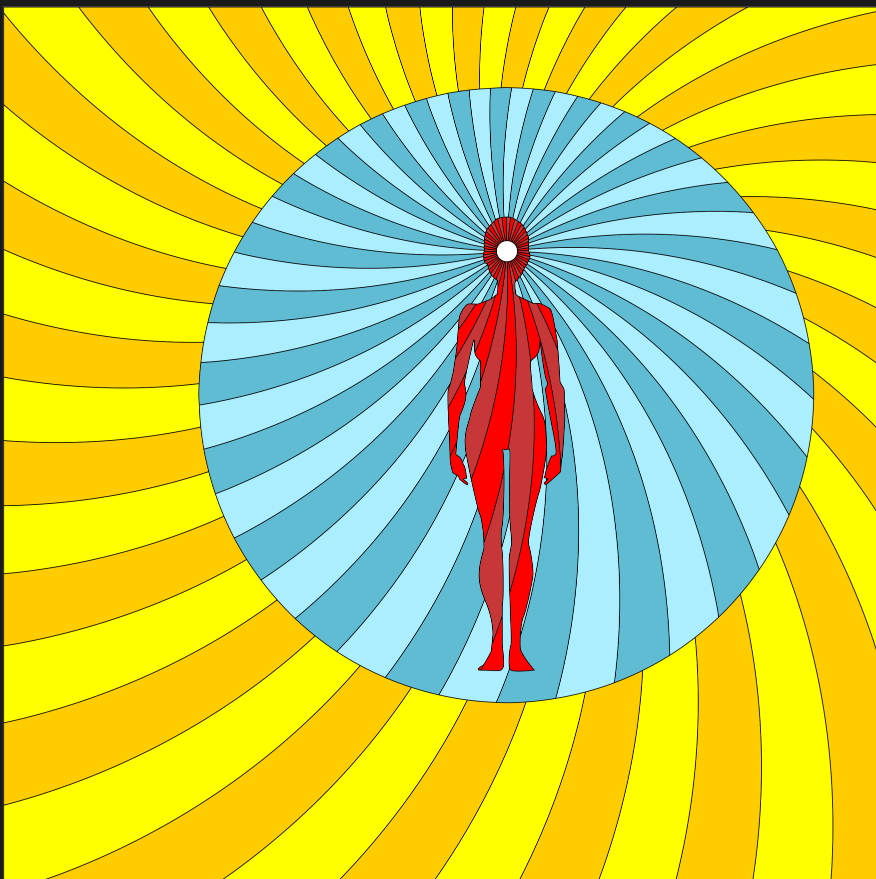
Reshaping Ubiquitous Interaction Through Sensory Augmentation



Francisco Kiss

RESHAPING UBIQUITOUS INTERACTION THROUGH SENSORY AUGMENTATION

Von der Fakultät für Informatik, Elektrotechnik und Informationstechnik der Universität Stuttgart zur Erlangung der Würde eines Doktors der Naturwissenschaften (Dr. rer. nat.) genehmigte Abhandlung

vorgelegt von

FRANCISCO KISS

aus Buenos Aires

Hauptberichter:	Prof. Dr. Albrecht Schmidt
Mitberichter:	Prof. Dr. Kristof Van Laerhoven

Tag der mündlichen Prüfung: 19. August 2020

Institut für Visualisierung und Interaktive Systeme
der Universität Stuttgart

2020

ABSTRACT

Current technological advances enable unprecedented approaches for aiding human perception through digital technologies. Powerful *Sensory Augmentation* becomes feasible. Our concept aims to improve the natural sensory capabilities of users while transcending the traditional interaction concept of tools. This approach promises to facilitate a more natural interaction with ubiquitous computers and support the enhancement of interaction with reality through computing machines. The goal of this research is to investigate the possibilities of this emerging paradigm and to provide a formal structure for future efforts in this field.

In this thesis, we propose a definition for *Sensory Augmentation* and a design space to structure its study. We present a vision for its application to human activities and identify opportunities and challenges for the incorporation of *Sensory Augmentation* to human activities. Further, we report on three research probes that help investigate these opportunities and challenges, as well as users' experiences and preferences for this technology. Each of these probes explores a specific area of augmentation, both assessing specific theoretical aspects of *Sensory Augmentation* and providing practical insights in the technical challenges posed by the development of applications.

Additionally, we investigate and discuss the implications of augmenting the senses from the perspectives of users and society, focusing on benefits but also addressing the main foreseeable negative effects and possible strategies to minimize them. Finally, we present a design process for *Sensory Augmentation* applications alongside with practical recommendations and discuss future research directions of this new field.

ZUSAMMENFASSUNG

Die aktuellen technologischen Fortschritte ermöglichen eine nie zuvor mögliche Unterstützung der menschlichen Wahrnehmung mittels digitaler Technologie. Eine umfassende und weitreichende *sensorische Augmentierung* wird möglich. Unser Konzept zielt darauf ab, die natürlichen sensorischen Fähigkeiten der Benutzer zu verbessern und gleichzeitig das traditionelle Interaktionskonzept von Werkzeugen zu transzendieren. Dieser Ansatz verspricht, eine natürlichere Interaktion mit allgegenwärtigen Computern zu ermöglichen und die Verbesserung der Interaktion mit der Realität durch Computermaschinen zu unterstützen. Das Ziel dieser Forschung ist es, die Möglichkeiten dieses aufkommenden Paradigmas zu untersuchen und eine formale Struktur für zukünftige Bemühungen auf diesem Gebiet zu schaffen.

In dieser Arbeit schlagen wir eine Definition für die *sensorische Augmentierung* und einen Designraum zur Strukturierung seiner Untersuchung vor. Wir präsentieren eine Vision für ihre Anwendung auf menschliche Aktivitäten und identifizieren Möglichkeiten und Herausforderungen für die Integration von der *sensorischen Augmentierung* in menschliche Aktivitäten. Darüber hinaus berichten wir über drei Forschungsproben, die dazu beitragen, diese Chancen und Herausforderungen sowie die Erfahrungen und Präferenzen der Anwender für diese Technologie zu untersuchen. Jede dieser Proben erforscht einen bestimmten Bereich der Augmentierung, wobei sowohl spezifische theoretische Aspekte der *sensorischen Augmentierung* bewertet werden als auch praktische Einblicke in die technischen Herausforderungen, die die Entwicklung von Anwendungen mit sich bringt, gegeben werden.

Außerdem untersuchen und diskutieren wir die Auswirkungen der Sinneserweiterung aus der Perspektive der Anwender und der Gesellschaft, wobei wir uns sowohl auf die Chancen, wie auch auf die wichtigsten vorhersehbaren negativen Auswirkungen und mögliche Strategien zu deren Minimierung konzentrieren. Schließlich stellen wir einen Designprozess für *sensorische Augmentierung*-Anwendungen zusammen mit praktischen Empfehlungen vor und diskutieren zukünftige Forschungsrichtungen dieses neuen Feldes.

PREFACE

In this thesis, I present work conducted during the last years at the University of Stuttgart. Given the range of scientific disciplines involved in the study of Human-Computer Interaction, this work was done in close collaboration with experts from the University of Stuttgart, the Ludwig Maximilian University of Munich, the University of Utrecht, and the Łodź University of Technology, who contributed with their knowledge from their respective research fields. These collaborations resulted in multiple publications that largely overlap with the research presented in this thesis. The contribution of individual researchers is recognized at the beginning of each chapter, together with a reference to the scientific publication that resulted from the collaboration. Through this thesis, I use the scientific plural ("we"), highlighting the collaborative nature of this work.

Acknowledgements

The work presented in this thesis could have never come to fruition without the support and contributions of a vast amount of people. First and foremost, I want to thank my supervisor Albrecht Schmidt for his trust, support, and guidance, and for allowing me to roam freely the path towards becoming a researcher, only nudging me when I wandered off too far. I also want to thank Paweł Woźniak, a true friend, resourceful mentor, and solid guide in this adventure. I am also in debt to Niels Henze, my favorite Devil's advocate, and living proof that it is possible to have fun while being a proper scientist. I am very grateful for Tonja Machulla's mentoring in how to supervise students with fairness without forgoing kindness.

I thank my great friend Tilman Dingler for inspiring me to venture into the realms of academia and also for all the fun times we had together sharing the things we love to do the most. I thank Sven Mayer for his friendship and guidance in navigating the world of German academic bureaucracy and student supervision, and for his amazing and not-weird-at-all back massages. I also feel thankful for my early mentors, Thomas Kubitza and Stefan Schneegass, who kept me busy with fun projects during my time as a student assistant.

I am grateful for the friendship and support of all my wonderful colleagues, who made me feel part of this amazing research group and helped me in so many ways during these years: Romina Poguntke, Alex Voit, Pascal Knierim, Valentin Schwind, Huy Le, Thomas Kosch, Jakob Karolus, Markus Funk, Rufat Rzayev, Passant ElAgroudy, Yomna Abdelrahman, Céline Coutrix, Hyunyoung Kim, Bastian Pfleging, Lars Lischke, Mauro Ávila, Norman Pohl, Patrick Bader, Dominik Weber, Anja Mebus, and Murielle Naud-Barthelmeß. It is often said that the path is more important than the goal. I believe that the people who help you and care for you along that path are actually the true treasure.

Last but not least, I thank my very supportive family. I am lucky to have parents such as Esteban and Patricia, who made me an independent thinker, curious about Reality, and an avid reader of books. I am deeply grateful to my sisters Luisa and Teresa, whose support and advice helped me gain confidence in my abilities. I thank my amazing wife Sandra for her constant support and patience, and particularly for tolerating my grumpy moods caused by deadlines during family vacations. Finally, I want to also thank my little Felipe and Amalia for giving a true meaning to everything I do and putting my priorities in order.

TABLE OF CONTENTS

List of Figures

List of Tables

List of Acronyms

ANC Active Noise Cancelling

AR Augmented Reality

BCI Brain-Computer Interface

CMP Computer Mediated Perception

CMR Computer Mediated Reality

DPSA Design Procedure for Sensory Augmentation

ESH Enhanced Selective Hearing

FTW Future Technology Workshop

HA Hearing Augmentation

HCI Human-Computer Interaction

HMD Head Mounted Display

ML Machine Learning

MR Mediated Reality

RQ Research Question

SA Sensory Augmentation

SNR Signal-to-Noise Ratio

UCD User-Centered Design

UX User Experience

VR Virtual Reality

VUM Virtual Ubiquitous Microscope

VZT Visual Zoom Technology

Chapter 1

Introduction

> First were mainframes, each shared by lots of people. Now we are in the personal computing era, person and machine staring uneasily at each other across the desktop. Next comes ubiquitous computing, or the age of calm technology, when technology recedes into the background of our lives.
>
> — Mark David Weiser

The advances of technology are driving computing machines away from the simple task of managing information and into the role of mediators between humans and reality. Modern mobile devices play a large part in communication between humans, not only facilitating it but also shaping how humans interact with each other [216, 242]. The advent of ubiquitous Augmented Reality (AR) promises to extend mediation to potentially everything we see and hear [195]. In this thesis, we investigate the concept of Sensory Augmentation (SA), which describes the technological intervention of the human senses particularly to increment or modify what humans naturally perceive. But before introducing the concept of SA, we will first present a summary of the evolution of human-computer interfaces to provide adequate context to our research topic.

1.1 The Evolution of Interfaces

Since the invention of digital computing machines, several revolutions shaped the way we interact with the world and particularly information. While these revolutions arose mostly from technological development, their ultimate motivation and effect belong in the realm of Human-Computer Interaction (HCI). The evolution of human-computer interfaces is strongly linked to these technical advances, which enabled interaction designs increasingly centered on the user.

In the initial stage of this historical process, users required considerable training to input information into a computing machine and to interpret the output produced by the computer. Interfaces of early computers revolved around punch cards (see Figure 1.1), light-bulbs and bells. Interaction with such machinery was limited to loading a program and data, letting the program run until it terminated, and retrieving punched or printed output. The introduction of computer screens and keyboards improved both input and output possibilities, enabling users to use text and numbers. This new kind of interfaces allowed interactions to be more dynamic, and considerably closer to human communication media (see Figure 1.2). The evolution from binary input, to low-level programming languages, to terminal prompts, and finally graphics interfaces (see Figure 1.3) accommodated an increasing number of users, while reducing the overhead expertise required to accomplish specific tasks [160]. This ease of use, along with the ever-growing versatility of machines, is arguably a critical reason for today's popularity and ubiquity of personal computers, both at home and work.

A more recent revolution in HCI, maybe the greatest in terms of popularity and impact, came at the hand of handheld devices, particularly smartphones. These are the epitome of ease of use, portability, and versatility, at least as of the time of writing this thesis. Touch interaction and mobile interfaces made interaction even friendlier to users (see Figure 1.4). Handheld devices empower billions[1] of users worldwide to communicate and access vasts amounts of information. Further, these devices perform multiple tasks that previously required dedicated tools such as photographic cameras, audio recorders, calculators, calendars, flashlights, chronographs, GPS navigation systems, and many more. Arguably, these devices incorporated a new layer of information on top of our daily lives, enabling the constant and ubiquitous access to a massive and ever-growing pool of data.

Current trends and emerging technologies strongly suggest further advances in portability, power, and sophistication of the devices we use to interact with

[1] https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/

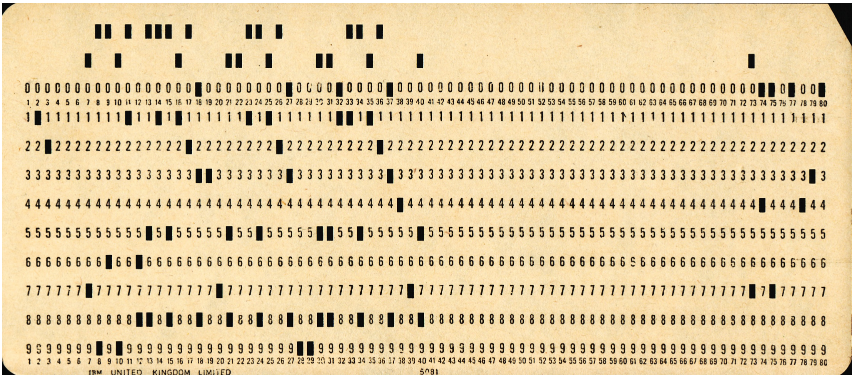


Figure 1.1: Punched cards were used for data storage, input, and output for early computing machines. Picture by Peter Birkinshaw from Manchester, UK (6 November 2010) Wikipedia. Licensed under CC by 2.0.

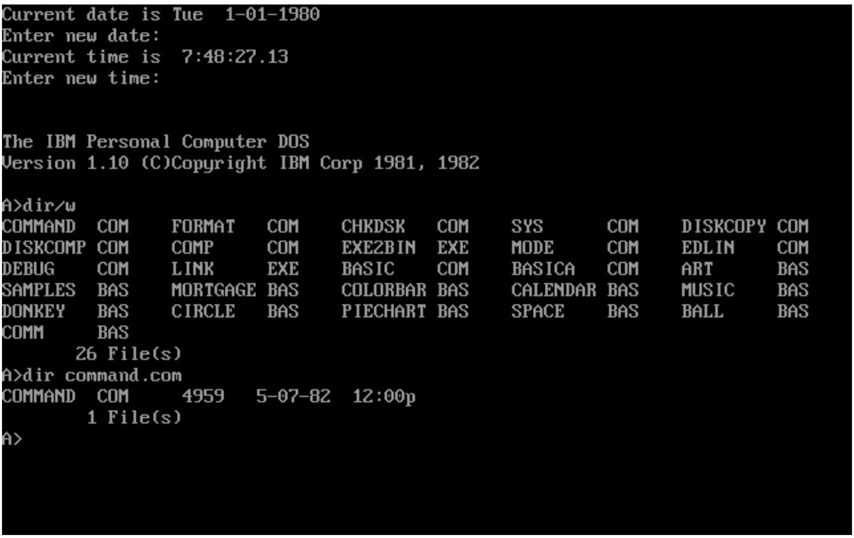


Figure 1.2: PC DOS, a popular operative system for x86 computers, used a text-based interface. Users could input commands and data using a keyboard and then read the output on a screen. This picture belongs to the public domain. Source: Wikipedia.

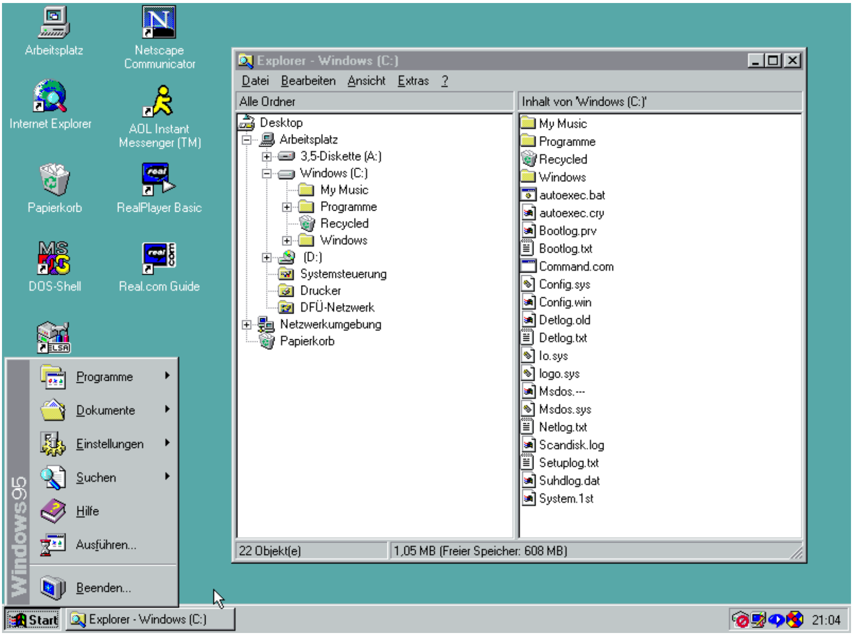


Figure 1.3: Windows 95 used a graphic interface design known as WIMP (Windows, Icons, Menus, and Pointer). The interaction was characterized by the addition of a computer mouse. This picture belongs to the public domain. Source: Wikipedia.

digital information. Ubiquitous computing, AR, and Virtual Reality (VR) are promising possibilities that will likely precipitate the next interaction paradigm (see Figure 1.5). As pointed out by Harper et al., the boundary between humans and their machines is moving inwards, towards a more intimate relationship [85]. New interfaces are closer to us, in some cases even directly attached to our bodies. Computers now collect vasts amounts of information about us, but also are turning into effective mediators between us and reality.

Decades ago, Mark Weiser envisioned a world where computers would become invisible assistants, helping users as needed without requiring cumbersome interfaces [243]. Computers are ultimately mere tools to perform a given set of functions, being these functions the only reason for their existence and our interactions with them. It follows from this premise that the less interaction is needed to use the function, then better is the design. Thus, the best possible interface is *no interface at all*. Current trends in interaction design suggest we might see this vision come to fruition, particularly at the hand of technologies such as ubiquitous

Figure 1.4: iPadOS 13 on an iPad Pro. User input is based on direct touch of the visual elements, although speech can be used to perform many actions.

AR. Especially in the functions related to retrieving information and presenting it to the user, these technologies can act as a seamless mediator between reality and what the user perceives, enriching natural perception with additional information. In this context, the prospect of augmenting the human senses and thus enabling users to perceive a larger part of reality becomes a promising field of research. However, this opportunity also prompts multiple questions, such as how this goal can be achieved, how interfaces for augmenting the senses should be designed, and what implications for people would the proliferation of such technologies have.

These questions resonate with the wider framework of inquiry proposed by Harper et al. [85], highlighting the need for a better understanding of these emerging technologies and their potentially high impact in human society. In particular, Sensory Augmentation (SA) poses a series of challenges to be addressed promptly and failing to do so can result in ill-devised interactions that, when in place, might be difficult to improve and could produce detrimental effects.

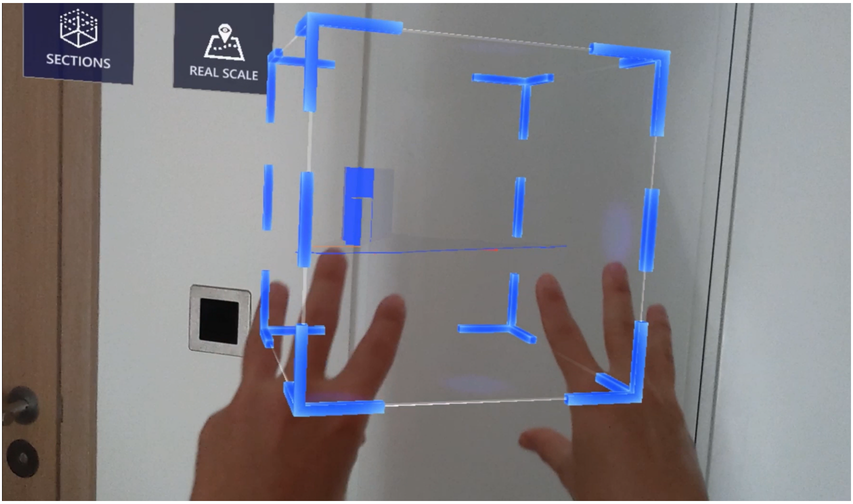


Figure 1.5: Perspective of a user wearing the Microsoft HoloLens 2. The user can interact with digital contents through direct manipulation and voice commands.

In this thesis, we investigate the fundamental aspects of SA, starting by proposing a formal definition of this concept and derive a design space from the existing literature. Through research probes, we explore the application possibilities of SA in terms of feasibility and potential benefit, and focus particularly on the interactive aspects of these technologies. We conclude this work with an analysis of the implications of these technologies for individual users and society as a whole, and provide a set of general recommendations for the design of SA applications.

1.2 Methodology

The methodology used in this thesis is based on User-Centered Design methods [106], supported by a systematic review of relevant literature. We structure and guide the work in this thesis by formulating a set of overarching Research Questions (RQs). Next, we provide general background information about the human senses, followed by a definition of SA in the context of HCI. Using this definition, we report on a review of scientific work on sensory amplification and augmentation in HCI, which we use to define a design space for SA. Once having defined the context and scope of SA, we proceed to present a vision on how this

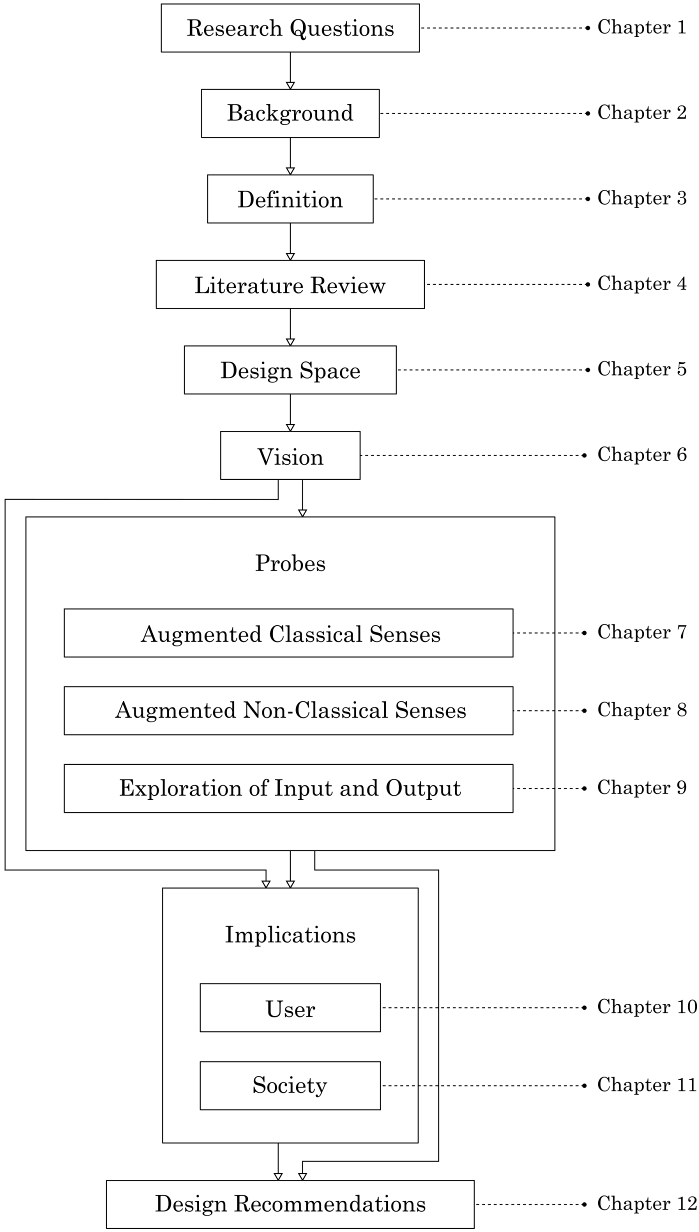


Figure 1.6: Methodology and structure of the work presented in this thesis.

technological and design concept can be applied to human activities and which benefits and drawbacks it carries.

We explore this vision through research probes, assessing the feasibility, design challenges, usability, and benefits of SA. We propose systems to augment a classical sense (hearing) and a non-classical sense (orientation), as well as a design to investigate the input and output possibilities for SA applications. The systems were conceived through user-centered design processes, enabling us to consider users' requirements and expectations already in the respective ideation phases of our research probes. To validate our designs, we implemented physical prototypes using iterative methods, and then evaluated them through user studies.

Next, we investigate the implications identified in the vision and the research probes for both individual users and general society. For this, conducted user studies which comprised interviews and a survey, as well as analyzed the relevant literature.

Finally, based on the knowledge throughout the work presented on this thesis, we offer a set of design recommendations to aid future research and development of SA applications.

1.2.1 Literature Analysis and Preliminary Investigation

Our work is based on literature from diverse fields, mostly HCI and Psychology, and to a minor extent, Information Science. In the first stage, we focused our efforts on reviewing HCI work on SA, particularly on systems that enhance human perception, even in cases where SA is not the main focus of research. Having a broad overview of past applications and designs helped us formulate a definition for SA that helped classify existing work and further refine our review. The review was continued throughout the work described in this thesis, thus allowing the incorporation of new designs and systems.

The outcome of the literature review is a set of interactive designs that embody the identity of SA. Further, these works enable the formulation of a design space for SA, that describes the main design dimensions for systems that augment the human senses.

Additionally, our design probes and the assessment of the implications of SA were based on literature, which is reviewed and presented in each respective chapter.

1.2.2 Research Probes

Artifact-driven research is a widely used approach in HCI [44, 122, 133]. Research probes belong to this research method, offering practical and empirical insights about a topic, otherwise unavailable through purely scholarly work. To evaluate the concept of SA, we conducted two main research probes, focusing on the augmentation of a classical sense, and a non-classical sense.

Each probe followed a user-centered design approach, starting with the involvement of users in the conception and first stages of the design process. Through interviews, workshops, and mediated group activities we gained knowledge about users' expectations, requirements, and preferences for systems that improve sensing and perception.

These insights were the foundation of our initial designs, which we presented again to users in the form of low-fidelity prototypes to gain feedback and improve our design. Through an iterative process, we refined our interaction concepts and physical prototypes until definitive designs were evaluated in controlled experiments. These user studies provided the scientific validation of our designs in terms of usability, effectiveness, and feasibility, as well as produced valuable insights about the users' impressions and reception of our work.

In a third probe, we investigated input and output mechanisms for visual augmentations. We used a SA that enabled users to ubiquitously magnify objects to microscopic levels as a case study. We implemented a prototype capable of simulating this functionality for a given set of objects and used this system to evaluate four different control methods. These methods were derived from the literature and represent the main approaches used in current AR interaction. Further, we compared two output methods, also reflecting current approaches and their subsequent challenges. We assessed these input and output variations in a controlled user study, evaluating the performance and preferences of participants.

1.2.3 Analysis of Implications

The incorporation of SA to human activities will likely have an impact on multiple aspects of our way of life. We investigated what implications can be expected, both for the individual user and for society as a whole. To assess this, we took a mixed-methods approach, combining informed speculative thinking with scientific literature and user studies. Particularly regarding the user studies, we conducted interviews and surveys, proposing different hypothetical scenarios to collect representative reactions and opinions.

1.3 Research Context

The work presented in this thesis was conducted during the last three years in the Human-Computer Interaction Group at the University of Stuttgart. This research is partially the product of collaboration with other researchers, both within the Human-Computer Interaction Group and from other institutions. The work presented here took place within the larger context of two major projects, AMPLIFY and Be-Greifen.

1.3.1 AMPLIFY

The main project leading the work presented in this thesis was AMPLIFY[2]. This project was funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (ERC Consolidator Grant no.683008). The main goal of this project was to assess the feasibility of creating new senses or enhancing existing ones, without increasing the cognitive load of users but enabling the intake of larger quantities of information. My role in this project was the exploration of sensory augmentation through both theoretical and empirical research, particularly through the experimental evaluation of prototypical applications.

1.3.2 Be-Greifen

Some of the work presented in this thesis was done in the context of Be-Greifen[3], a project funded by the German Federal Ministry of Education and Research under the grant number 16SV752. The main objective of this project was the development of novel interaction methods to incorporate digital information to augment didactic experiments. My contribution to this project was the formulation, development, and evaluation of prototypical systems to embody novel interaction concepts.

[2] http://amp.ubicomp.net/

[3] http://begreifen.dfki.de/

1.4 Contributions

The work presented in this thesis produced eight contributions to the field of HCI. The contributions can be summarized as:

1. a definition of SA and a revised theoretical framework for SA in HCI.
2. a general design space for SA.
3. a specific design space for Hearing Augmentation (HA).
4. the design and evaluation of a system to improve hearing selectivity.
5. the design and evaluation of a system to improve orientation for swimmers.
6. the comparison of different methods to control ubiquitous visual magnification.
7. an analysis of implications of SA for individual users and society.
8. a set of design recommendations for SA applications.

In the following sections, we present more details for each of the main contributions.

1.4.1 Definition of Sensory Augmentation

We formulated a formal definition for SA and placed this area of investigation in the wider context of AR and Computer Mediated Perception (CMP). This aids in the classification of existing work, the ideation of novel applications, and the research on the field as a whole. Further, this offers a richer conception of the interaction space in the realm of HCI.

1.4.2 A Design Space for Sensory Augmentation

We derived a design space for general SA, defining five different classes of augmentation that we believe will encompass all existing SA applications. We characterized the design space through a literature review of the most relevant HCI work published until the date of the elaboration of this thesis. Based on the proposed definition, we identified a corpus of work on SA and used it to discover the dimensions of the design space. The design space will serve as a framework for the classification of existing work and as a tool for the exploration of novel applications in this thesis.

1.4.3 A Design Space for Hearing Augmentation

We propose a design space for HA based on iterative investigations on a HA application. This empirical knowledge expands one dimension of the proposed general design space for SA, offering more depth and detail in the characteristics of such augmentations. Further, this design space serves as a guide to expand other dimensions of the design space for general SA.

1.4.4 A System to Improve Hearing Selectivity

We designed a SA to enable users to increase the selectivity of their hearing attention. This design amplifies the loudness of sounds generated by sources in the direction the user is facing. Through an iterative process, we evaluated multiple aspects of this design in controlled experiments, assessing it in terms of feasibility, usability, and effectiveness. This system demonstrates the use of the design space as a tool to augment a human sense. Further, its development highlighted some challenges and opportunities that result in design recommendations.

1.4.5 A System to Improve Orientation for Swimmers

We designed a SA to improve the sense of orientation for swimmers in open waters. The system provides visual feedback, compensating the lack of visual cues typical from subaquatic environments. We evaluated the design through controlled experiments in terms of feasibility, usability, and effectiveness. Further, we compared the difference in presenting different types of feedback to users. This system illustrates the use of the design space as a research tool. Additionally, it prompts a series of design and evaluation challenges that result in design recommendations.

1.4.6 Analysis of Control Methods for Visual Magnification

We conducted an experimental comparison of control methods for visual magnification. For this purpose, we created a visual augmentation capable of simulating a ubiquitous microscope that allows users to magnify objects to microscopic levels. We proposed four different control methods that encompass the most representative control mechanisms for AR technologies. These methods were contrasted in a user study, providing the basis for a design guideline for controlling SA.

1.4.7 Implications of SA for Users and Society

We investigated the implications of deploying SA applications, both for individual users and third parties. We focused on the mental health consequences that increasing the amount of information input might have on users based on existing literature. Based on this, we proposed a series of design recommendations to avoid foreseeable problems with future interaction designs. In terms of social implications, we conducted a study to assess the perception of the public regarding privacy threats posed by SA. The observed results provide insights on what design approaches might improve the acceptability of SA applications.

1.4.8 Design Recommendation for SA Applications

We proposed a series of design recommendations based on the knowledge gained during the work that resulted in this thesis. The recommendations encompass different aspects of design, ranging from conception, to prototyping, to evaluation, to user acceptability. We believe this practical contribution can aid future research and applications of SA.

1.5 Thesis Overview

This thesis consists of fourteen chapters, followed by the bibliography and an appendix.

The first six chapters can be described as a theoretical and scholarly approach to the subject. The first chapter is this *Introduction*, which presents the *Research Questions* that guide the work discussed in this thesis. The second chapter,*Background*, reviews the fundamental concepts of human sensation from the perspective of Psychology. This is followed by a *Concept* chapter, which offers a definition of SA within the context of HCI. The *Literature Review and Design Space* chapter comprises a systematic literature review of work in HCI, from which a definition and design space for SA arise. This chapter is followed by a *Vision*, where we discuss the potential and possibilities of SA.

The next three chapters present the evaluation process and the results of three research probes. In the chapter *Augmenting the Sense of Hearing*, we propose a design concept to enable users to increase their ability to focus hearing on individual sound sources. This design concept is then studied through multiple iterations of prototypes and user studies, from which we derive a design space for HA. The chapter *Augmenting the Sense of Orientation* describes a system to improve the orientation of swimmers in open waters. This system is evaluated in a controlled experiment in terms of feasibility and usability. The chapter *Controlling Sensory Augmentation* presents our research in control methods for SA based on a study case. We evaluate and compare different interaction concepts for an AR microscope in a user study and derive design recommendations from the observed interaction.

Chapters ten and eleven discuss the implications of SA, respectively for individual users and from a social perspective.

Finally, the last three chapters provide design recommendations for SA, discuss the presented work in terms of the research questions, presents conclusions, and proposes future research that build upon the knowledge we gained through the work that constitutes this thesis. A more detailed summary of each chapter is presented hereunder.

Chapter 1: Introduction

In this chapter, we present the theme of this thesis and explain the motivation and challenge we address. Further, we give an overview of the context in which the presented research was conducted. We list the main contributions of this thesis and provide an overview of its structure and contents.

Here we also formulate the research questions that guide our efforts and describe the methodology through which we aim to answer them.

Chapter 2: Background

This chapter provides the definitions used in this thesis for the human senses. We distinguish between the so-called *classical senses* and the *non-classical senses*. Each category is explained and a summary of the natural limitations of human sensation is provided.

Chapter 3: Concept

In this chapter we provide a definition of SA and place this concept within the theoretical framework from the literature.

Chapter 4: Related Work

In this chapter we present the methodology and outcome of a literature review on the topic of SA. In this study of previous work, we recognize a corpus that embodies the concept of SA.

Chapter 5: Design Space

We derive a design space from the literature review, encompassing both the existing literature and future applications.

Chapter 6: Vision

This chapter describes our vision on how SA can be implemented and deployed. Further, we propose a series of applications and use case scenarios for augmenting the human senses and derive requirements, challenges, and risks for these emerging technologies.

Chapter 7: Research Probe I - Enhanced Selective Hearing

In this chapter, we present a system that augments the sense of hearing. Our design enables users to increase the spatial selectivity of their hearing attention by amplifying the sound generated by sources in the direction they face. We present the iterative design process, prototype implementation and multiple evaluations of this system. Further, we present a design space for hearing augmentations derived from the insights gained through multiple user studies.

Chapter 8: Research Probe II - Clairbuoyance

This chapter presents a research probe for augmenting non-classical senses: a system to improve the sense of orientation for swimmers in open waters. The design is based on visual feedback to compensate the lack of visual cues typical from subaquatic environments, such as the sea or lakes. Further, we present a prototypical implementation, which we evaluated in a controlled experiment. Based on the observed results, we propose a series of design recommendations.

Chapter 9: Research Probe III - Virtual Ubiquitous Microscope

In this chapter, we report on our research on control methods for SA through a case study. We propose a system that enables users to ubiquitously magnify what they see to microscopic levels. Through a functional prototype, we compare four different control methods typically used in AR interaction. Based on our findings, we suggest a series of design recommendations for controlling SA.

Chapter 10: Implications for Users of Sensory Augmentation

This chapter investigates the implications for users in terms of cognitive load and stress. Based on previous work, we identify likely problems and their possible causes. Further, we propose design recommendations for avoiding the causes, or when impossible, minimizing the effects on users.

Chapter 11: Social Implications of Sensory Augmentation

This chapter explores the social implications of SA in terms of social acceptability. Through interviews and a survey, we investigate the perceptions of potential users of hypothetical technologies that pose a threat to privacy. Further, we derive from our findings recommendations for reducing the negative perception of SA technologies.

Chapter 12: Designing Sensory Augmentation

In this chapter, we summarize the knowledge gained throughout the conducted research in the form of a set of design recommendations structured as a workflow.

Chapter 13: Conclusions

This chapter reviews the main findings and contributions of this thesis. Further, it explicitly addresses the research questions posed in Chapter 2.

Chapter 14: Future work

In this chapter, we review the limitations of the research conducted in this thesis and suggest how our research can be extended. We propose future research directions and opportunities, both for theoretical and practical work. Additionally, we propose a set of design concepts as future research goals.

1.6 Research Questions

In this section, we present the Research Questions (RQs) that guide our research and explain them in detail.

One goal of this thesis is to define SA as a research field within HCI, and propose it as a tool for the development of future interaction design. Thus, our research questions must first seek a definition of SA, then identify challenges, advantages, and opportunities of such technology, and finally recognize its implications for users and society. Hence, our investigation of SA focuses on six overarching RQs, summarized in Table 1.1.

The first question (RQ1) seeks implicitly to enunciate a formal definition of SA and explicitly to derive a design space. These two concepts are fundamental for studying SA since they define the scope of the proposed research focus. Further, the proposed definition of SA should be organic to the theoretical framework for the reality continuum, as proposed by Milgram et al. [161].

Using the proposed theoretical background, we then proceed to investigate SA empirically, thus assessing the feasibility of SA applications to identify difficulties and challenges inherent to the augmentation of the human senses. We differentiate between the so-called *classical* senses, and the *non-classical* ones (see the

Table 1.1: The six overarching research questions addressed in this thesis.

	Research Question (RQ)	**Chapter**
RQ1	How to structure a design space for SA?	Chapter 5
RQ2	Can SA improve perception through the classical senses?	Chapter 7
RQ3	Can SA improve perception through the non-classical senses?	Chapter 8
RQ4	What type of interfaces are suitable for SA?	Chapter 9
RQ5	What implications presents SA for users?	Chapter 10
RQ6	What are the social implications of SA?	Chapter 11

Background chapter), so both inquiries are conducted independently as RQ2 and RQ3, asking if SA can improve perception through classical and non-classical senses respectively.

A particularly challenging aspect of SA design is interaction design since conventional interactive elements might not be adequate or convenient for applications that augmented the senses. Most current interfaces were designed originally for stationary or handheld screens and are not necessarily suitable for SA. Thus, RQ4 seeks to assess which interfaces are preferred by users and result in higher performance.

Finally, SA can have stark implications for human society. Enabling users to modify what and how they perceive can affect their interaction with reality and other people. We investigate this in terms of implications for the user of such technologies in RQ5, and from a social perspective in RQ6.

We acknowledge the limitations in the scope of the formulated research questions. Topics that escape our area of competence, such as physiology or psychology of perception, have been taking into account for our research but are not explicit focuses of investigation. Thus, a number of possible research objectives, which are related to understanding SA, have been intentionally excluded from this thesis and remain as questions to motivate further research. These open topics are suggested in more detail in the Future Work chapter.

Chapter 2

Background: Human Perception and Sensation

> The blind say that eyes have no sense of smell.
>
> — Nigerian proverb

To augment the human senses, it is essential to first understand their functions and limitations. In the following, we will provide definitions of what we call the senses, reviewing some fundamental concepts. Starting from the five classical senses and their non-classical siblings, we summarize for each of them their key mechanisms and natural limitations. In conjunction with previous work from the field of human-computer interaction illustrating some examples of how to augment human perception, we conclude this passage by clarifying the terms sensory augmentation, enhancement, etc. By the explanation we aim to provide a common ground on what we will focalize in this chapter. We based the following summarized explanations of the senses on Goldstein's Sensation and Perception [76], which we recommend for a deeper understanding of how senses and cognition work from a physiological point of view.

Parts of this chapter are planned to be published in the following form:

- F. Kiss and R. Poguntke. Augmented Senses: Evaluating Sensory Enhancement Applications. In Technology-Augmented Perception and Cognition. Springer, Cham.

2.1 Five Classical Senses

We call senses the physiological mechanisms that enable organisms to obtain information from reality. This information is some physical effect generated by *observable* physical phenomena, which interacts with the sensory organs [76]. Sensory organs are sensitive to variations in some aspects of reality, producing the physiological signals that in the case of higher vertebrates are decoded during cognition. Since sensory organs are sensitive to a particular aspect of reality, they can perceive a limited type and range of phenomena. This classification of senses has been traditionally circumscribed to the commonly called *five senses*[4]: audition, taste, smell, touch, and sight.

The physiology and cognition behind each sense are a defining characteristic of the way the sense can be enhanced: natural limitations suggest challenges to be overcome and closed domains define frontiers to be explored. Thus, alone the definition of each individual sense suggests a starting point for its own sensory enhancement research.

Apart from what we are often calling the five classical senses, there is more to consider, particularly when taking a closer at how we perceive smells and flavor. In this context, we have to differentiate between taste or gustation and the overall term for gustatory and olfactory inputs, which is flavor [222]. Only the triad of the senses of taste, olfaction and the trigeminal perception enables us to distinguish flavors. Often, the trigeminal chemosensory system mainly consisting of the trigeminal nerve and adjacent axons is being neglected when talking about the classical senses. Nevertheless, its existence facilitates chemical processes which build the foundation for our perception. To clarify their differences we start with the gustatory sense before we explain what the olfactory sense is. We then refer to the tactile and visual sense before we conclude with the auditory sense.

---

[4] Even if sometimes the five classical senses are called Aristotelian senses, Aristotle defined in *De Sensu et Sensibilibus* only four senses, describing taste as a form of touch [17]. However, by the 11th century multiple sources start counting the senses as five [9]

2.1.1 The Gustatory Sense

The sense of taste provides information about what we introduce in our mouthes, which has the potential benefit of recognizing both sources of nutrients and substances that should not be ingested. Humans perceive taste through gustatory *calyculi* (or taste buds), which are concentrated on the upper side of the tongue [76].

Thorugh chemical reactions, gustations enables us to detect specific, albeit limited, properties of matter we put in our mouthes. Namely these are sensations of salty, sour, bitter, sweet and umami. Interestingly, recent scientific work argues that *oleogustus* [111, 200] and *kokumi* [132] need to be considered separately when speaking of our different tastes.

The gustatory sensation can be also achieved by electrical stimulation of our gustatory muscle, i.e. the tongue. Although this method has been explored already about 80 years ago [34], researchers still take advantage of the artificial elicitation of taste using electrical muscle stimulation [172, 196], thermal stimuli actuation [202], or relying on chemical substances [171]. For a detailed literature review please refer to Spence et al. [222].

2.1.2 The Olfactory Sense

Similarly to the sense of taste, smelling occurs through chemical reactions between inhaled substances and sensory cells. Olfaction occurs in humans in a small surface of tissue inside the nose, which is populated by cells specialized in perceiving smells [76]. The sense of smell enables us to detect the presence of chemical compounds in the air, which might be beneficial by indicating the quality of food, the presence of a threat (such as fire) or even help infants recognize their parents [147]. Although humans are microsmatic (our sense of smell is comparatively weak and not crucial to our survival), we can tell apart about 100,000 different odours [70], despite we may have difficulties identifying particular scents.

The sense of smell plays an important role in our culture, a phenomenon clearly exemplified by the use of deodorants, perfumes or ambient fragrancies. Thus, enhancing the sense of smell is an area of interest for researchers, although it is a challenging interaction space since a technology to emulate the odorous stimuli is yet to be invented.

However, some researchers investigated different enhancements of the olfactory sense with the release of individual scents. For example, Brewster et al. [30] suggested to use smell-based interaction for tagging photos and Amores et al. [8] investigated the effect of smell as an interface to influence mood and cognitive states.

2.1.3 The Tactile Sense

The sense of touch resides in the skin, the hairs and some inner surfaces of the body (such as the tongue and throat). Four types of mechanoreceptors located in the skin allow to detect when pressure is applied to, and removed from, their effect area, allowing to perceive touch, pressure, vibrations, and textures [76]. This enables us to perceive a wide range of individual phenomena, such as the tickling caused by an insect moving on our skin, or the strong pressure of the ground on the soles of our feet, or even the direction, speed and intensity of a caress. The sense of touch plays an important role in preventing harm to our bodies, as well as manipulating physical objects. It is also an important aspect of social interaction and fundamental to sexual activity.

The sense of touch has been widely investigated in technology applications as a means of conveying information. Tactile feedback, in particular vibration, has become ubiquitous through the use of mobile phones to replace sound stimuli. More sophisticated applications explore touch as a modality to convey complex information. For example, Blum and Cooperstock [25] mapped user state information, such as activity level or distance, to tactile stimulation patterns as a mean for remote implicit communication. Brewster and Brown [29] proposed a more formal method of communication-based on vibration: Tactons, or tactile icons, present the possibility of encoding symbols by using frequency, amplitude, rhythm and duration of tactile pulses.

2.1.4 The Visual Sense

The sense of sight is arguably the one we depend on the most. The photons coming from light sources bounce on the matter and are dispersed, and some of them are captured by photoreceptors within our eyes. Two different kinds of photoreceptors enable us to distinguish shapes and colors, and thus we are able to see our surroundings. We sensitive only to a small range of light frequency, and are capable of perceiving a limited dimensional granularity for both space and time [76]. One of the earliest sensory enhancements was the use of transparent lenses to enlarge tiny objects or bring closer objects that are far away. But other limitations have also been overcome in more recent times IR vision.

Additionally, having two eyes enables us to perceive distances and spaces, adding the so-called third dimension to our reality. This has been proven to add realism to augmentations such as virtual reality.

Likewise, we are able to perceive and distinguish colors and shapes enabling us to connect different pieces of information what results in the ability of mapping colors to information. An interesting example of how research took advantage of this to facilitate learning, is the "Motion Echo Snowboard" [185]. Hereby, LEDs represent the weight distribution on a snowboard, so that the natural feeling of balance is augmented by visual information resulting is real-time visual feedback that supports perception. Moreover, vision can be manipulated using a head.mounted display and an eye-tracking system in a way, that the viewing perspective can be changed. Kasahara et al. [110] showed that changing visual input could result in differences for spatial perception when e.g. the first person view is shared among a group of four persons.

2.1.5 The Auditory Sense

Hearing is the perception of sound, which is a mechanic vibration transmitted over a medium, typically air. Vibrations caused by physical phenomena propagate in form of waves, and when these waves reach our ears, they exert pressure along the auditory canal, activating mechanoreceptors that translate vibrations to the electrical signals that are transmitted to the brain. Humans can normally perceive sounds within the spectrum of 20Hz to 20kHz, being most sensitive to the frequencies between 2000 and 4000 Hz, the most important range for understanding speech [76].

Additionally, having two ears enables humans to perceive spatial information about what we hear, such as the relative direction of a sound source, or geometrical and mechanical properties of our environment. Since sound propagates in the form of waves, a given signal may reach the two ears with a time delay. The difference in these two signals is used by the brain to calculate direction and distance of the position in space where the sound is coming from. Furthermore, the sound waves bounce on environmental features and objects (e.g. walls) and reach the ears in form of repetitions of the original signal, although with lesser intensity and affected by the materials with which they interact. These additional waves are used to gain spatial and material insights about these objects [197, 198].

This ability has been used in multiple spatial audio applications, and is particularly useful in adding realism to virtual environments. Rendering spatial audio in realistic virtual environments, adding a virtual soundscape on top of reality with an office PC was already feasible in 2004, as shown by Zotkin et al. [260], with increasingly better consumer applications in modern computer games.

The underlying cognitive mechanism of binaural hearing has been also explored as a tool to affect mood and cognition, showing promising possibilities but yielding yet no concrete results [43]. If claims about binaural beats being capable of driving brain activity result to be true, which current research suggests to be the case[105], a whole new interaction design space based on audio would become available.

Hearing also enabled the development of language and is still fundamental for spoken communication. Improving communications over distance, as with telephones, or over time, as with audio recording and reproduction, predates digital electronics, but can be seen as a precursor of sensory augmentation. More recently, research efforts have resulted in many accessibility applications, such as hearing aids to improve audition for the hearing impaired, or the use of audio to convey environmental information to the visually impaired[164, 83, 259]. Apart from this target group, auditory information has been used to navigate pedestrians by adding directional cues to music [252].

2.2 Non-classical Senses

Human perception is not limited to the five classical senses we have presented [76, 150]. We are able to perceive both external and internal phenomena beyond taste, smell, sight, hearing, and touch. We perceive time passing, even independently from external stimuli. We can feel pain, allowing us to react to situations that pose a danger to our well-being. We recognize our posture and the position of our limbs through proprioception, and sense variations in our bodies such as hunger, pulmonary stretch, blushing or sexual arousal. We perceive our position and orientation in space, relative to the gravitational pull, as well as movement and acceleration. We can perceive the internal somatic effects of fear, anxiety, or happiness. We can become aware of our heartbeat and other physiological phenomena taking place in our bodies. Some of these senses can be related to specific organs or structures, while others, such as the sense of direction, are constructs made at a cognitive level from information gathered by physiological sensors.

Despite the value we recognize in Macherson's consolidating view of sensation as a multidimensional space, we opted in this thesis for a simplified classification of the human senses. We use these two categories of senses and assume that the classical senses provide information about external phenomena, while the non-classical senses enable the perception of internal phenomena, as well as the situation of the observer with respect to the external reality (see Figure 2.1).

This classification enables drawing a parallel between sensation and computing. Classical computing is analogous to the classical senses, in terms of focusing on specific and clearly delimited aspects of reality. In contrast, ubiquitous computing and non-classical senses can provide insights about contextual and situational characteristics.

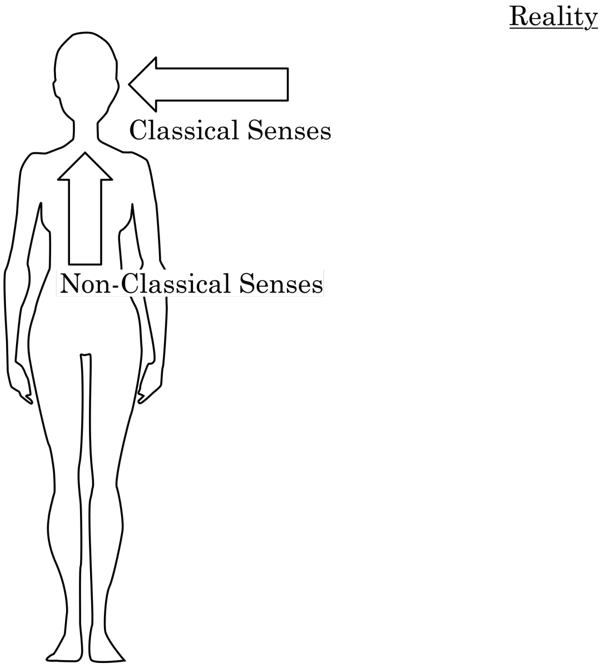


Figure 2.1: A simplified model for human sensation of reality. The classical senses provide information about external phenomena, while the non-classical senses provide information about internal phenomena or situations of the observer respect to the external world.

Chapter 3

Sensory Augmentation

> Reality is that which, when you stop believing in it, doesn't go away.
>
> — Philip K. Dick

This chapter presents the context and general definition of SA within the theoretical framework provided by the literature. We first review the existing taxonomy for computer systems that mediate the human perception of reality. This is followed by a general definition for SA, based on the literature and as a refinement of the existing theoretical framework. Finally, we highlight the conceptual difference between SA and *accesible technologies*.

Parts of this chapter are currently ***under submission*** *and planned to be published in the following form:*

- F. Kiss, P. W. Woźniak, P. ElAgroudy, R. Rzayev, R. Poguntke, C. Eghtebas, T. K. Machulla, and A. Schmidt. Designing Sensory Augmentation: Literature Review and Design Space. ACM International Conference on Interactive Surfaces and Spaces (ISS).

3.1 Theoretical Context

An early approach to this subject can be traced back to Douglas Engelbart's conceptual framework on augmenting the human's intellect [67]. In this work, the author proposed to improve the intellectual effectiveness of human beings with help of computers. Engelbart envisioned that computers would amplify the human natural capabilities to match the challenges posed by highly complex problems and formulated the founding concepts of computer-based augmentation systems.

Steve Mann proposed a taxonomy for Mediated Reality (MR) and explores Computer Mediated Reality (CMR) for the particular case of mediated vision [152]. Extending Milgram's *augmented reality continuum* [161], Mann added an orthogonal dimension to represent *mediality* (see Figure 3.1). While virtuality describes the extent to which the computer modifies what the user perceives, mediality describes the extent to which the computer can modify reality. Thus, while AR elements provide digital information on top of what users perceive, MR elements offer the additional capability of modifying reality.

Mann defined MR as the union of modulated reality and mixed reality, which includes AR and VR. This definition aims to encompass all modalities in which systems mediate between users and reality, categorizing them according to how they intervene in the users' perception of reality, as well as the extent to which they can affect the real world.

Rekimoto and Nagao proposed the concept of *augmented interaction* in terms of the interaction between humans, computers and the environment [199]. The authors identified four main styles of HCI in terms of the role of computers in the interaction between humans and reality (see Figure 3.2). Interactions with desktop computers are typically isolated from the real world: the computer is not aware of the environment of the user, nor has any direct effect on it. Virtual Realty (VR) overrides the natural perception of the environment, replacing the perception of visual and auditory reality with digital simulation, and thus hindering interaction with reality. With the emergence of ubiquitous computing, users interact alternatively with real-world phenomena and digital information through computers embedded in their environments. Interaction with these computers is often intertwined with interactions of what we perceive as the real world, such as traffic lights, ATMs, navigation systems, or microwave ovens. In contrast to the ubiquitous computing approach of creating a reality augmented by computer, Rekimoto and Nagao proposed to augment the interaction with

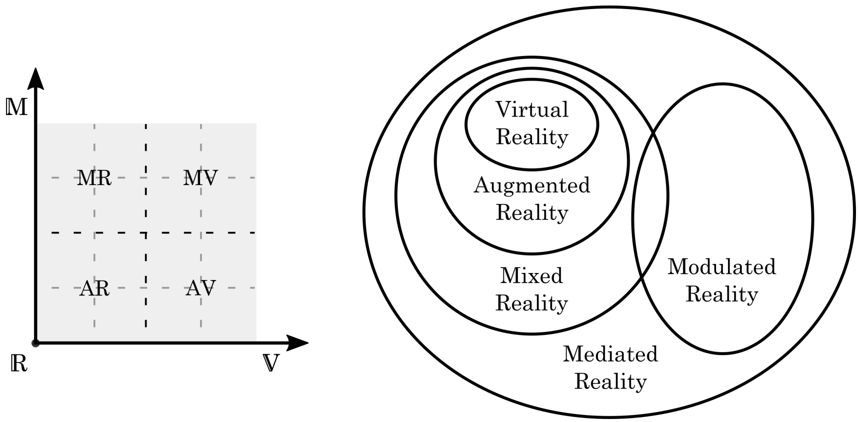


Figure 3.1: Left: Taxonomy of Reality, Virtuality, Mediality, as proposed by Mann. The origin R indicates unaltered perception. The V axis represents Virtuality, as in Milgram's Augmented Reality Continuum [161]. The M axis represents Mediality, which corresponds to the extent of modification of reality.
Right: Venn's Diagram representing Mediated Reality [152]. Mann suggested Augmented Reality to contain Virtual Reality, both being a subset of Mixed Reality, but being distinct from the intersection between Mixed Reality and Modulated Reality. Mediated Reality with implementations for everyday life. Steve Mann [152].

reality with just one computer. This device should be context-aware and intervene in the users' interactions with reality accordingly.

3.2 Sensory Augmentation

The concept of *augmented interaction* can be further categorized by distinguishing *augmented perception* from *augmented action* (see Figure 3.3). Through action, users exert agency over reality or the computer. Thus *augmented action* empowers users to act in ways typically unavailable to them. In counterpart, information about reality flows to the user through perception, and thus *augmented perception* enhances what users can normally perceive.

Augmented action strongly relies in ubiquitous computing since the ubiquitous indirect control of phenomena would require either an infrastructure in place to actuate real-world elements, or rely on automatons and robot servants.

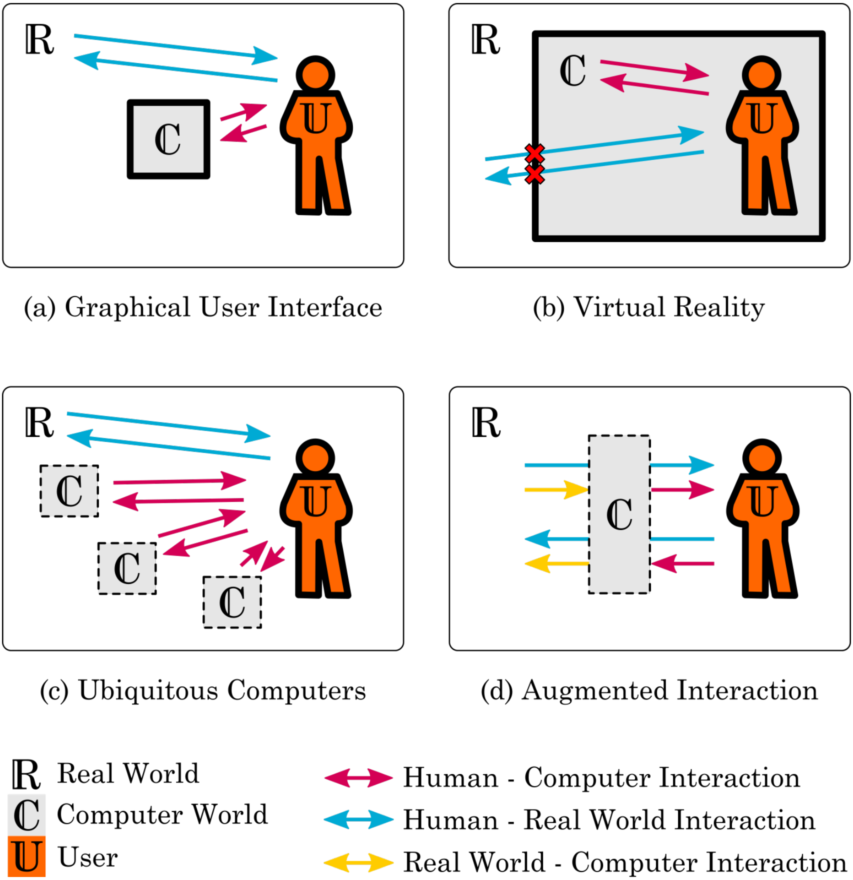
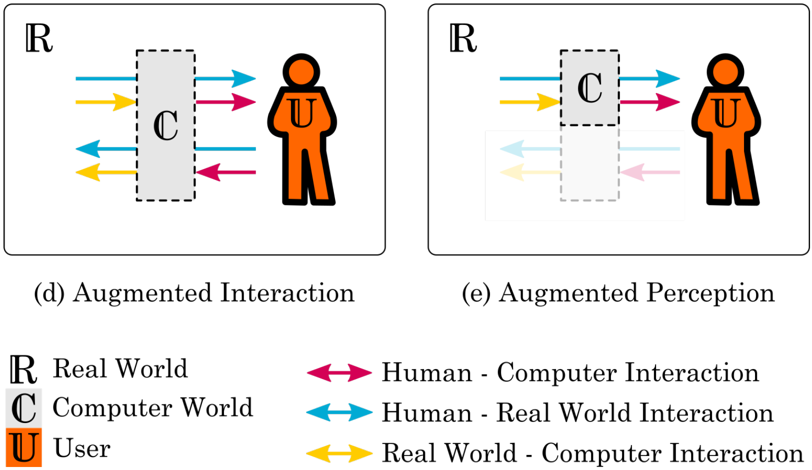


Figure 3.2: HCI styles: (a) interactions with desktop-computers is isolated from interactions between users and the real world, (b) virtual reality impedes the interactions between the user and the real world, (c) ubiquitous computing allows users to alternatively interact with the real world, or with computers embedded in it, and (d) augmented interaction enhances the interaction of the user with the world by interventions of computers. The difference between (c) and (d) is the number of computers. Adapted from: The World through the Computer: Computer Augmented Interaction with Real World Environments. Jun Rekimoto and Katashi Nagao [199].



(d) Augmented Interaction (e) Augmented Perception

Figure 3.3: Rekimoto and Nagao's Augmented Interaction (d) can be further subdivided by the direction in which information flows. Sensory Augmentation can be interpreted as Augmented Perception, encompassing the information flowing to the user from the real world (e).

In contrast, *augmented perception* can be achieved with a single computer, an approach more aligned with Rekimoto and Nagao's vision. Admittedly, current technology requires the decentralized support of information management systems to deliver a large portion of the information available to users. However, this limitation is mainly technical, and would disappear once processors, storage, and sensors become good enough.

Perception takes place at a physiological level. Thus the literal augmentation of perception would require interventions on the sensory organs. An alternative to this approach would be the augmentation of the sensory information just before it reaches the sensory organs. Thus, the interface between users and computers lies very close to the sensory organs, but still outside the human body. We describe this approach as Sensory Augmentation (SA), which can be classified as a subset of Rekimoto and Nagao's *augmented interaction* and can be included in the *mixed reality* subset in Mann's MR.

We propose a definition of SA based on Schmidt's vision [208]:

> A system provides Sensory Augmentation if it enhances the capabilities of a given human sense to perceive a quantity of information about reality beyond the sense's natural capabilities.

3.3 Sensory Augmentation is not Accessibility

Despite having the common goal of enhancing human perception, SA and accessibility are intrinsically different. At a theoretical level, the goal of accessibility is to tackle an atypical limitation in perception affecting a subgroup of the general population. SA seeks instead to augment perception beyond the typical limits for everyone. On a practical level, despite presenting information of the same nature, accessibility and SA applications have different technical needs and a technology that tackles a disability cannot augment perception beyond the natural limits. For instance, a system that compensates a lower sensitivity for a part of the hearable spectrum will achieve this by amplifying sounds with the affected frequencies. A SA application that enables users to hear beyond the hearable auditory spectrum cannot achieve its goal with the same method since the human ear cannot physiologically achieve this task.

Figure 3.4 illustrates the difference using three exemplary cases: Person A can hear a little below the typical highest pitch people can perceive, Person B has a stronger disability, being able to perceive only the lower half of the auditory spectrum, and Person C can perceive higher frequencies than most. Accessible technology applications can greatly benefit Person B, while Person A sees only a little improvement. In contrast, SA can for instance scale down the pitch of sounds allowing to perceive frequencies way beyond the natural human ability. An additional example that clarifies the difference between SA and accessible technologies is the maximal distance in which we can perceive objects. While corrective lenses can reduce the effects of short-sightedness in persons affected by *myopia*, they provide no benefit to the rest of people. In contrast, binoculars help focus vision on objects well beyond the typical natural limitations for human sight.

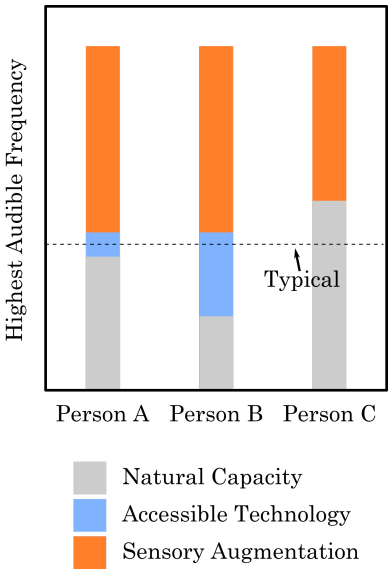


Figure 3.4: The difference between accessibility and augmentation: Persons A, B, and C can naturally perceive different maximal frequencies. Accessible technologies can increment the hearing spectrum up to the natural physiological limits, but Sensory Augmentation can trascend the natural limits.

Chapter 4

Related Work

> Without the library, you have no civilization.
>
> Ray Bradbury

To identify and analyze previous work relevant to our research, we conducted a literature review on SA based on the definition proposed in the previous chapter. In this chapter, we describe the applied methodology and our findings.

Parts of this chapter are currently ***under submission*** *and planned to be published in the following form:*

- F. Kiss, P. W. Woźniak, P. ElAgroudy, R. Rzayev, R. Poguntke, C. Eghtebas, T. K. Machulla, and A. Schmidt. Designing Sensory Augmentation: Literature Review and Design Space. ACM International Conference on Interactive Surfaces and Spaces (ISS).

4.1 Introduction

Given the novel nature of SA as a specific field of research, no systematic effort has been yet realized to identify a corpus of related work. This results in the uncertainty about the extent to which SA is present in the literature, even if not portrayed as such. To address this gap, we conducted a literature review of SA to identify research and applications in HCI. Our approach shares principles and goals with past literature reviews in HCI. As our primary goal is to delineate the research area of SA, we were inspired by past reviews that also sought conceptual clarity. Hence, we were inspired by work by Speicher et al. [221] who clarified the design space of mixed reality and Frich et al. [74] who presented a retrospective of creativity support tools in HCI.

4.2 Methodology

Our reporting follows the guideline proposed in the PRISMA statement [162], which is a structured review methodology, originally developed for the medical sciences. This approach is based on four stages: *identification*, *screening*, *eligibility* and *inclusion*. In the *identification* stage, sources of scholarly articles and publications are identified for the target topic. The *screening* is implemented by removing publications which are thematically irrelevant for the corpus. In the *eligibility* stage, the corpus is further reduced by excluding publications using predefined criteria. Finally, in the last stage, the remaining publications are analyzed in detail and their final *inclusion* is decided individually.

4.2.1 Strategy

We followed the PRISMA approach, performing each of the four stages. The first step was to characterize the focus and scope of our search. Based on the proposed definition for SA, we decided to use the following search criteria: we wanted to find publications describing a design or artefact, which enhances (at least) one human sense, so the user perceives a larger quantity of information in the immediate reality. With immediate reality, we mean the real physical world directly around the participant, which excludes static media. In line with past reviews in HCI, e.g. [228], we focused solely on archival material from HCI publication venues. These criteria result in a considerable part of AR applications being excluded from SA, since in most cases, graphic or textual information does not increment human perception, and merely provides symbolic information. Further, as stated in the previous section, the requirement for an increase in perception beyond typical human capacity excludes applications of accessibility technologies.

4.2.2 Information Sources and Eligibility Criteria

We considered conference papers and journal articles for a broad range of HCI publications. We chose a representative set of publication venues that were likely to contain AS publications through a discussion among the authors.

Having defined our search goals, we proceeded with the *identification* stage and defined a corpus of work in HCI to review. We included in our search the complete proceedings and volumes of the most recognized publication venues in the field: Conference on Human Factors in Computing Systems (CHI), Augmented Human International Conference (AH), International Joint Conference on Pervasive and Ubiquitous Computing (ubicomp, including predecessor conferences and the current journal version IMWUT), International Symposium on Wearable Computers (ISWC), Conference on Designing Interactive Systems (DIS), Personal and Ubiquitous Computing, and Transactions on Computer-Human Interaction (TOCHI). Our initial corpus included all volumes of these publications published until August 2019, yielding a total of 13231 publications.

Here, it is worth noting that we considered a keyword-based approach instead of browsing all papers in the proceedings listed above. An investigation of the papers known to the authors prior to the review revealed that no consistent terminology was used for SA systems. Terms such as augmented perception [98], amplified perception [208], augmented senses [116] were used to describe systems that features SA. Consequently, we were unable to formulate a set of keywords that would comprehensively search databases for papers that addressed SA. Consequently, comprehensively mining the entirety of the volumes selected was our option of choice.

4.2.3 Procedure

We reviewed the complete corpus of work in the publications above in the three stages, as required by PRISMA. The PRISMA Flow Diagram [162], as depicted in Figure 4.1, provides an overview of the process.

In the *screening* stage, all the publication titles were split between the authors of this paper. We removed from the corpus the publications with titles and abstracts that indicated unambiguously that the publication was not related to human perception or any sensing technology, nor presented any possible relation to those topics. Ambiguous or borderline cases were included, to allow deeper scrutiny at the next stage.

In the *eligibility* stage, we first established eligibility criteria for our corpus in an iterative discussion. We defined the following requirements for a paper to be included: only archival material (full papers or articles) would be included, which presented the evaluation of the system providing a SA. Again, unclear or borderline cases should be included, to allow for further scrutiny at the next stage.

Next, the intermediate corpus was divided between the authors in a way that each paper was reviewed by two authors. Based on the abstract, each author assessed the papers assigned using the eligibility criteria and suggested inclusion and rejection. In cases were authors disagreed, a discussion among all the authors resolved the difference. This step resulted in an intermediate corpus of 122 papers. A large share of the papers excluded in this phase were accessibility papers. While we did recognize that there is an extensive work on manipulating the senses in the accessibility community, we decided that an accessibility-focused review would be more appropriate for such systems.

The authors then read all the remaining papers to commence the *inclusion* stage. After an initial reading round, we decided that two authors will assess the relevance of each paper at a 3-point scale: 'low', 'medium', and 'high'. We only included papers judged to be of high relevance for the review. In cases where only one author considered a paper highly relevant, we conducted a discussion among all the authors to reach a final decision. The final corpus consisted of 28 papers, shown in Table 4.1.

4.3 Findings

This section presents an overview of the findings of our review. We first look at the dynamics of publication in the SA area. Then, we present our conceptual understanding of past work on SA in the form of a classification.

4.3.1 Sensory Augmentation as Emerging Topic

The number of yearly publications about SA increased in time, particularly in recent years (see Figure 4.2). What started as sporadic technical explorations in the CHI conference, gain sustained traction from the AH community, starting in 2011. The popularity of this SA transpired to all major HCI venues since 2015. This behavior suggests an increasing relevance of SA as a focus of research.

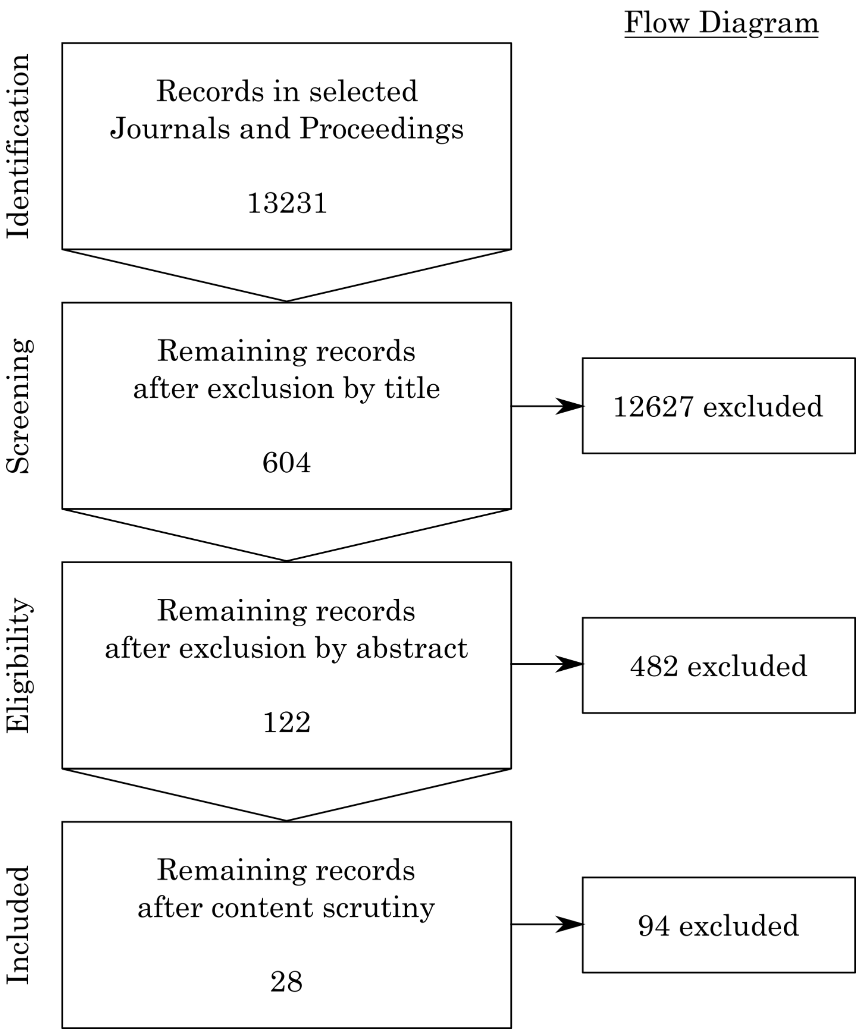


Figure 4.1: Flow diagram of the literature review, based on the PRISMA methodology. In each stage, the number of eligible publications is reduced, until the final corpus of our review Sensory Augmentation is established.

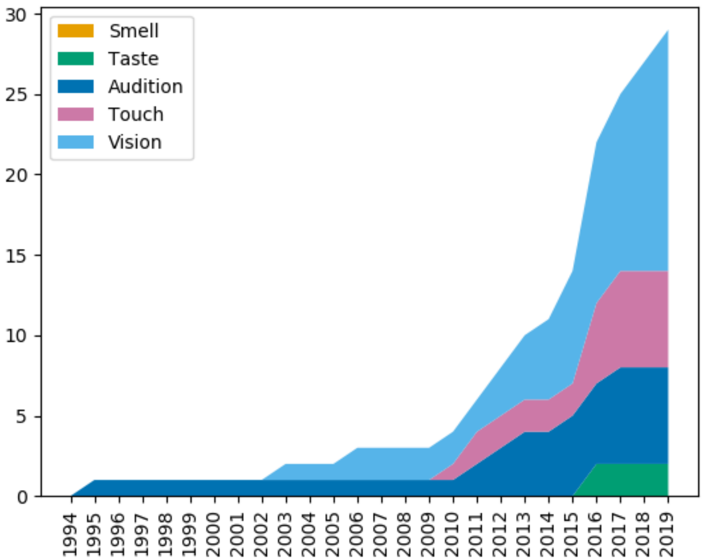


Figure 4.2: Accumulative timeline of artefact publication in the included literature review, categorized by sense.

Additionally, the growth of the yearly publication number in the topic of SA resonates with the growth of research focus in HCI, if we take the CHI as reference. This hints that the general research in HCI and the development of computing technology furthers the work in SA.

4.3.2 Coding

Two authors coded the entire paper corpus to establish the dimensions of the review. We coded all the bodily senses that each of the papers addressed (taste, vision, touch, smell, hearing). Further, we used open coding for the type of augmentation implemented, the effect of the augmentation and the methods used to evaluate the system. Next, we coordinated coding trees in an iterative discussion and conducted a second coding of the material. We used thematic analysis [23] to build meta-level themes based on the codes in each of the coding dimensions. Finally, we completed a final coding of the publication set ensuring that both coders were certain of all the papers being assigned to the correct categories for each dimension.

4.3.3 Senses Used

We first study which senses were the target of augmentations in past research. The sense most frequently augmented was sight (15 out of 28 publications). Next was hearing (6 publications), touch (6 publications), and taste (2 publications). We found no SA for the sense of smell that fulfilled our inclusion criteria. Figure 2.1 shows how the dynamics of publication in SA for different senses evolved over the years included in the corpus.

4.3.4 Evaluation Methods

The evaluation of SA systems is a predominant theme in the identified literature. In 19 publications, the SA design was evaluated with a user study, five with preliminary evaluations, one in a technical feasibility study, and two reported no evaluation at all (see Figure 4.3). The user studies of 14 publications were based on quantitative evaluations, and 13 on qualitative evaluations, with eight using both methods.

Quantitative experiments were diverse, holding in common the use of numeric metrics to measure performance and other aspects of designs. Qualitative research, both in main and support evaluation roles, consisted mostly of interviews, with a few exceptional workshops. Preliminary studies were either quantitative or a combination of quantitative and qualitative measures and aimed to provide a general idea of the performance of some aspect of the design with a low number of participants. Technical or feasibility evaluations assess the technical capabilities of the system, but no user study is involved in the evaluation.

4.3.5 Type of augmentation

Here, we present five types of augmentation that define SA. These types describe how the perception of the SA's user is changed when the system is active. Given that humans perceive reality exclusively through their senses, there is a finite number of ways information about physical phenomena can be conveyed to the

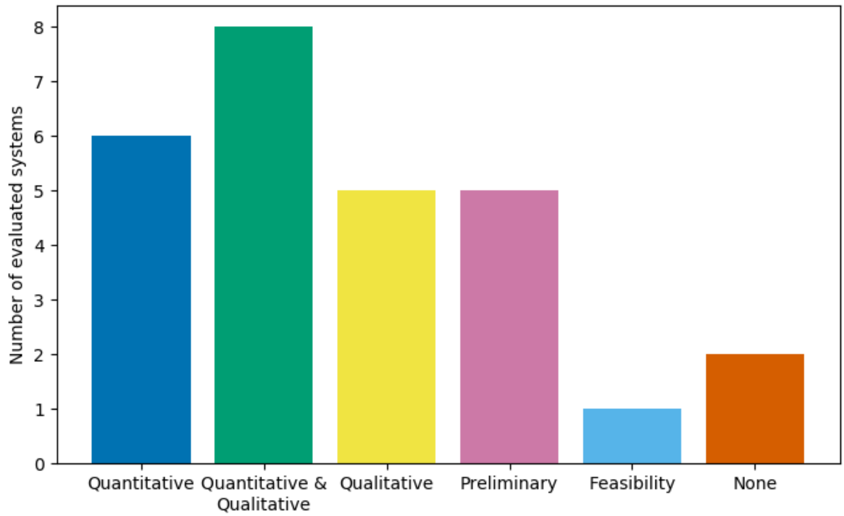


Figure 4.3: Evaluation methods of the Sensory Augmentation corpus.

user. We identified four recurring augmentation types in the literature through thematic analysis. Further, we postulate a fifth theoretical type, which was not yet observed in SA literature in HCI, but it is logically possible. While the terms we use to name the augmentation types are only suggestions from the authors, we believe that they provide an opportunity to disambiguate the SA space.

Amplification

Sensory amplification allows users to perceive circumstantially unavailable stimuli, but within the naturally perceivable range for humans. This augmentation makes a larger volume of information available than can be perceived by senses without the amplification. In other words, the strength (or magnitude) of a given sense is increased, providing a wider range of perceptual stimuli to be processed by the user. Examples of amplification include the magnification of distant objects through AR systems or increasing the volume of very soft sounds through sensitive microphones.

Sensory amplification can be illustrated by the work of Tao et al. [227], which consists of a system that enables users to control the angle and speed of what they see. Both the perception from different perspectives and the perception of movement belong to natural human capabilities, although the proposed amplification enables perception beyond the natural limits. The work of Duarte et al. [59] presents another clear example of amplification. In this case, the design amplified hearing to enable users to perceive the heartbeat of others remotely, a feat naturally possible only through physical contact.

Extension

Sensory extension allows to perceive stimuli that human sensory organs are not sensitive enough to perceive, but belong to the domain of perception of said sense. The augmentation consists on extending one magnitude of sensation to include a wider range of a given physical part of reality, e.g. UV light, high-frequency sound, microscopic objects.

Augmentation through extension was particularly evident in the work by Lilija et al. [139]. The system enables to see occluded objects, thus portraying stimuli from the same domain as the augmented sense but unperceivable due to natural constraints. Alternatively, the concept of sensory extension can be illustrated by My Eyes [183], a system that enables to simultaneously see naturally and from the point of view of a different person.

Enhancement

Sensory enhancements use complementary abstract stimuli to *enhance* or accentuate a dimension of perceived reality. The increment in perception does not necessarily occur using the sense that is being augmented. Thus, this augmentation enables perceiving something beyond human perception through its representation using additional sensory stimuli. An example would be using the pitch of a tone to indicate the proximity of ferromagnetic objects, as metal detectors do, or using haptic feedback to convey distance to an object.

Examples of enhancement in the corpus include Clairbuoyance [121], a system that improves the sense of direction through visual stimuli, and the work by Hasegawa et al. [89] on sonification of the gravity center of the user's body. Both designs increase the perception of the user by using a sense that normally cannot perceive the target dimension of reality.

Annotation

Sensory annotation is based on symbols, such as text or graphic representations, to convey information based on an established encoding scheme or language. This concept is different for AR or head-worn displays. In the case of augmentation through annotation, information is not an attribute of observed objects, but of reality as experienced by the user. The augmentation thus improves knowledge about perceived physical phenomena using symbols. Such symbols could be provided to the user through any sense. An example of annotation is an AR system that provides a text overlay indicating the distance to an object while walking or a digital voice assistant describing what happens behind the user.

Annotation augmentation designs in the corpus include BodyVis [178], which uses pictographic representations of bodily organs to display inner biological processes of the user, and GrabAmps [65], a system that helps visualize the current intensity flowing through touched cables. Both augmentations make use of symbology to visually convey information from different domains of reality.

Addition

Finally, we discuss a, so far, hypothetical type of augmentation. Sensory addition dispenses of sensory organs and uses electrical stimuli applied directly to the central nervous system to build new sensory experiences. Given that the human brain receives electrical signals from the peripheral neurons and interprets them into what we call perception, intervening or generating such signals can enable a direct interface to feed information to the human brain and thus create new senses. While this possibility is currently theoretical, we decided to add it to our augmentation types to cover the full spectrum of possible augmentations.

4.4 Analysis

During a comparative analysis of the included literature, we observed a set of common characteristics across all publications, as well as features that allow classifying the papers into sub-categories.

The next categorization variable we identified was the type of augmentation: some work focuses on improving the performance of a sense in its natural domain (e.g. Stop Motion Goggles [129]), while other adds new features to what a sense normally can perceive (e.g. Clairbuoyance [121]). These different types of augmentation are a finite set, limited by the nature of perception and existential constraints.

Finally, the last defining characteristic of a sensory augmentation is the nature of the additional information that users gain through the use of this technology. This depends starkly from the type of augmentation and the targeted sense. However, we observed that some themes are also repeated or revised, and thus this variable may serve also as a categorization parameter for (future) larger collections of artefacts.

Artefact	Type	Sense	Effect	Authors	Year
AudioStreamer §	extends	hearing	to spatialize monoaural audio signals	Schmandt et al.	1995
Stop Motion Goggle §	extends	vision	to perceive selective by controlling shutter speed	Koizumi et al.	2003
Motion Echo Snowboard †	enhances	vision	to improve proprioception and movement awareness	Park et al.	2006
SpiderSense §	enhances	touch	to sense surrounding objects	Mateevitsi et al.	2010
Audiolizing of body movement	enhances	hearing	to reinforce proprioception and sense of embodiement	Houri et al.	2011
Sensible Shadow ‡	enhance	touch	to sense what the shadow of the user touches	Hiraki et al.	2011
Augmento §	amplifies	hearing	to perceive the pulse of somebody else	Duarte et al.	2012
Laplacian Vision ○ †	enhance	vision	to predict trajectory of moving objects	Itoh et al.	2012
Sonification of gravity center §	enhances	hearing	to improve proprioception and movement awareness	Hasegawa et al.	2013
SpiderVision ○ †	extends	vision	to perceive back view when movement is detected	Fan et al.	2013
My Eyes ○ †	extends	vision	to simultaneously view from the point of view of partner	Pan et al.	2014
The Mad Hatter's Cocktail Party †	amplifies	hearing	to increase auditory attention selectivity	Aoki et al.	2015
BodyVis †	annotates	vision	to perceive representations of the inner body	Norooz et al.	2015
Afterimage ○	extends	vision	to show an afterimage when the user blinks / close eyes	Nojiri et al.	2015
Electric Taste §	extends	taste	to perceive digital flavors	Nakamura et al.	2012
Virtual Food Texture †	extends	taste	to perceive digital texture	Nijima et al.	2016
ProximityHat ○ †	amplifies	touch	to sense the distance to surrounding objects	Berning et al.	2016
Thermal feedback ○ †	enhance	touch	to sense the distance to surrounding objects	Wilson et al.	2016
Tactile Distance Feedback ○ †	extends	touch	to sense surrounding objects	Carton et al.	2016
GrabAmps ○	annotates	vision	to visualize current intensity of handheld cable	Elvitigala et al.	2016
Parallel Eyes ○	extends	vision	to simultaneously view from the point of view of others	Kasahara et al.	2016
Looming Silhouette ○	extends	vision	to perceive occluded incoming people	Yokoyama et al.	2016
3DAATIS †	enhances	hearing	to identify direction of causes of alerts and warnings	Wang et al.	2017
FeltRadio ○	extends	vision, touch	to perceive radio signals	Grönvall et al.	2017
Unconstrained Neck	amplifies	vision	to overcome physiological limitations of the human neck	Shen et al.	2018
4D Visualization on HMD ○	amplifies	vision	to control the speed of time and point of view of sight	Tao et al.	2018
Clairbuoyance ○ †	enhances	vision	to improve the sense of orientation	Kiss et al.	2019
AR Views ○ †	extends	vision	to perceive occluded objects	Lilija et al.	2019

Table 4.1: Projects included in the literature review, classified into the dimensions of the review. The types of evaluation of each project are coded with the following notation: (†) qualitative, (○) quantitative, (§) technical, and (‡) preliminary.

4.4.1 Effect

Our analysis revealed high diversity in the analyzed corpus in terms of the aims of the presented SA systems. An exhaustive enumeration of possible augmentation effects is unpractical, given the considerable diversity and the fact that some works were simply unique, e.g. Audiostreamer spatializing monoaural audio signals to facilitate selective attention [207]. Table 4.2 presents the clusters of augmentation effects which we derived through thematic analysis. These effects represent the range of usage scenarios for SA that were addressed in HCI literature to date. They can be used by future researchers to identify opportunities for new scenarios or comparisons with past systems. We identified as clusters groups of at least two publications presenting the same or similar effects.

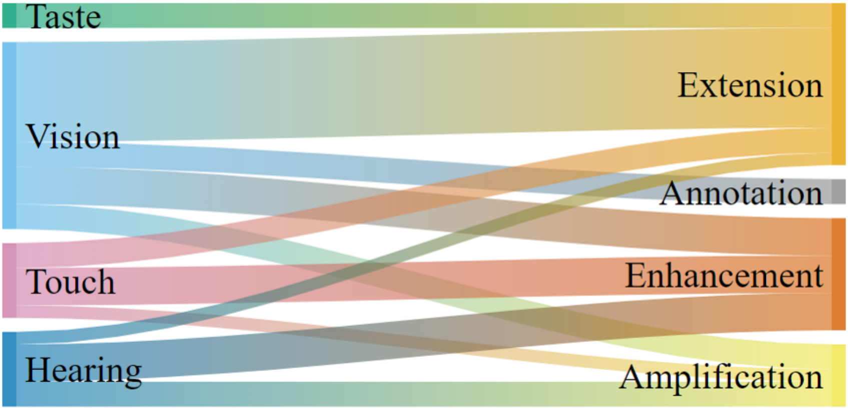


Figure 4.4: A Sankey diagram showing the relation between senses used and augmentation types offered. The artefacts included in the literature were designed for different senses and implemented different augmentation types.

Augmentation effect	Papers included	Comment
More information about others	[59, 254, 110, 183]	These systems allow sensing more information about other users, e.g their heart rate or point of view.
More informatin about oneself	[89, 97, 178]	These systems provide additional perception of the user's own body.
More information about surrounding objects	[21, 41, 157, 245]	The systems allow to perceive spatial characteristics of surrounding objects, such as presence and distance.
Overcome a specific sensory constraint	[69, 129, 213, 227]	These designs target specific ergonomic or physiological limitations of human perception.
Improve attention selectiveness	[10, 207]	These systems enable users to better focus their attention in desired targets.

Table 4.2: A summary of the augmentation effects which we identified in the corpus.

Chapter 5

Design Space

> Design is not just what it looks like and feels like. Design is how it works.
>
> — Steve Jobs

In the previous chapter, we presented the methodology and findings of a literature review on SA. In this chapter, we articulate the gained knowledge to propose a design space for SA which we use to provide insights on this research topic.

*Parts of this chapter are currently **under submission** and planned to be published in the following form:*

- F. Kiss, P. W. Woźniak, P. ElAgroudy, R. Rzayev, R. Poguntke, C. Eghtebas, T. K. Machulla, and A. Schmidt. Designing Sensory Augmentation: Literature Review and Design Space. ACM International Conference on Interactive Surfaces and Spaces (ISS).

5.1 A Design Space for Sensory Augmentation

Based on the findings of our review, we present a design space to classify past and future SA systems. Such a classification enables understanding the field of SA. Further, given that our review showed that work focused primarily on using vision for expansion and enhancement, cf. Figure 5.4, our design space intends to sensitize researchers to alternative design choices in SA. Figure 5.1 shows the design space. The space has two main dimensions to categorize a SA system: the sense(s) targeted and the type of augmentation provided. We postulate that any systems that address both dimensions should be considered SA systems. This may reduce the terminological ambiguity currently present in the literature.

	Vision	Hearing	Touch	Taste	Smell
Amplification					
Extension					
Enhancement					
Annotation					
Addition					

Figure 5.1: Design Space for Sensory Augmentation. A system can be considered an SA system if it uses at least one of the basic senses to deliver information to the user (horizontal) for one of the augmentation types (vertical).

The design space summarises the results of our literature review. To ensure the classification is comprehensive, we decided to include all basic human senses in it. The sense targeted by the augmentation is thus defined by the modality of the augmenting system's output. It is important to clarify that the targeted sense is not necessarily the sense to be augmented, but the sense used to deliver information to the user. These two senses are matched for *amplification* and *extension* types of augmentation, but may be different for *enhancement* and *annotation*.

We observed that despite the existence of the so-called *non-classical human senses* [150], all reviewed work addressed exclusively one (or more) of the five *classical senses*: vision, audition, touch, taste, and smell. Further, despite the possibility of presenting information beyond the natural perceptual domains, this information will always be presented to the user through one of the classical senses. Consequently, we adopt this more conservative taxonomy for our classification. Additionally, we do not exclude the possibility of future work on SA exploring the augmentation of olfaction and thus, we include smell in our categorization. Figure 5.2 shows a heatmap of the corpus from the literature review in the design space.

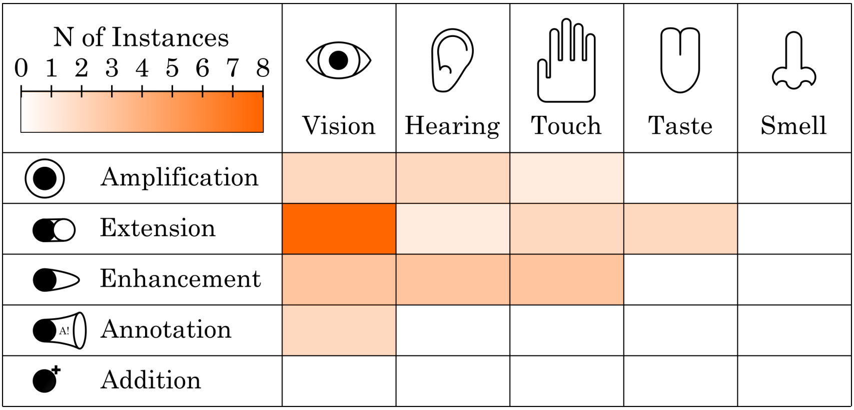


Figure 5.2: Heatmap of reviewed work in the proposed design space. Vision is the most frequently augmented sense.

Each of human senses has its own domain of perception and delimited ranges of sensitivity and areas of effect. The physical propagation of the information we perceive through each sense is different in kind and form. Hence, the information we perceive, its nature, coding, content, and constraints are unique to each sense: we can smell what we cannot see, we can hear what we cannot touch, or taste what we cannot smell. Sometimes we perceive phenomena through multiple senses simultaneously and our brains combine the information gathered by the participating senses into a richer experience. For instance, when we experience swimming in the sea and we see it, hear it, touch it, taste it, and smell it, and these are all different kinds of information, but complementary. Thus, the individual features and constraints of each sense dictate both how a system can to convey information to a user, and what aspects of perception can potentially profit from augmentation.

This plays a role in the selection of target sense, as seen in most designs within the SA corpus: the nature of the information often dictates which channel is adequate, or even optimal. Thus, the selection of a given sense is the product of a system driven constraint or convenience, as seen frequently in systems with *enhancement*-type augmentations. However, in some cases, and especially for *amplification*-type and *extension*-type applications, the selection of a sense is trivial, since the goal of the augmentation is improving said sense (see Figure 4.4).

Consequently, the two dimensions in the design space are inherently intertwined. It is up to the designer to decide which augmentation type may offer the most benefit given their application scenario and how this augmentation is to be provided for chosen senses. Our design space highlights the variety of possible SA systems that can be built and enables comparing future systems.

5.2 Discussion

Based on the proposed designed space and the knowledge gained through the literature review, we gained a series of insights about SA. These can be classified in *constraints* and *opportunities* for the design of systems that augment the human senses. Finally, we discuss the limitations of the design space, both in terms of the proposed design and of the literature review on which it was based.

5.2.1 Constraints

Augmenting the senses enables providing humans with more information about reality through increasing their perception, which in turn implies that the volume of information processed increases. However, the presented approaches to SA do not involve any modification of sensory organs. Hence, the baseline throughput of sensory information remains the same, since its physiological limitations remain unchanged. Naturally, this results in some sensory information being lost and replaced by the augmentation. This constraint is not only true for the sensory throughput, but also for the cognitive capacity of the human brain. Even in the hypothetical case of *addition*-type augmentation, where information is input to the brain on top of the existing natural channels, the human capacity to process and make sense of perception remains the same. Consequently, future SA systems will have to ensure that the cognitive capacities of human brains are not strained by SA and the sum of augmentations possible for a single user is limited by human physiology.

Another obstacle for SA is the granularity of perception. Conveying information about continuous variables (such as spatial information) may be achieved only to a certain grade of accuracy. For instance, representing the distance to an object with the pitch of a sound may give a notion of the object being close or far away, but not convey an exact distance in millimeters. The complexity of represented information may also be a constraint for SA designs. Even for the natural senses, a high degree of complexity poses a challenge to perception. The difficulty in fining Wally in the famous children's puzzle book is a prime example of a single sense being strained by complexity. This is also true for SA. Limiting the complexity of what a user can understand from the interaction. Most reviewed SA succeeded in conveying simple information, mostly boolean or unidimensional variables. The communication of more complex information can be challenging, particularly when *annotation* is not desired or possible.

Finally, some technical constraints limit the design space of SA. Most works in the SA corpus was implemented in the form of prototypes and only functioned in controlled conditions. The feasibility of introducing these technologies to the daily activities of people in an unobtrusive, seamless way is still hypothetical and extensive research and development are necessary before deployment in the wild.

5.2.2 Opportunities

The field of SA is a promising one, and having a design space enables the systematic exploration of possible applications. Existing systems can be used as brainstorming material by changing the type or sense of augmentation while attempting to produce the same effect. Additionally, the combination of each augmentation type and sense can be used to explore which natural constraints of the sense can be overcome, and what possible benefits would this provide.

Further, the field of SA can drastically change the way we live. Modern technologies enable extensive modification of perception, with both VR and AR technologies achieving highly realistic simulations in consumer-grade products. The incorporation of these technologies to daily life can result in making life easier for many people. SA will likely be an important part of this future of interaction, as suggested by the trend depicted on Figure 2.1. The work presented here aims to provides the opportunity to explore and elaborate the design space of SA from early on.

5.2.3 Limitations

We reviewed a corpus of work confined to the extent of the most relevant publication venues in the field. However, the existence of further SA publications scattered in other journals and conferences is possible and likely, especially given terminological issues. Thus, some SA systems may have been omitted. However, we believe that our work shows a comprehensive view of SA as reported in the HCI field. However, as out classification is comprehensive, we believe it can be extended to more systems.

Further, we recognize that we addressed the categorization and classification of artefacts on a high level of abstraction, in order to identify trends and lines of research in SA. However, alternative approaches could have been used. We find it particularly relevant for future research to address SA systems using knowledge from perceptual psychology, analyzing the use of sensory channels in SA systems. Further, a design-oriented analysis is also possible, focusing on the design qualities of the artefacts and the values they embody.

Finally, we settled for two principal dimensions for our design space. However, additional dimensions are not only possible, but desirable, particularly once the field reaches a higher count of publications and artefacts. We recognized some further possible domains, particularly in the methods to control the systems, the level of explicit involvement of the user in the interaction, or to which degree the augmentation excludes information otherwise perceived normally by the user. At the current stage, these deeper levels of classification add unnecessary complexity to the design space, but we believe that these dimensions could be useful in future work.

Chapter 6

Vision

> If the doors of perception were cleansed everything would appear to man as it is, infinite.
>
> William Blake

In previous chapters, we have reviewed the evolution of interfaces up to their present state and hinted at possible future developments. We followed by providing a theoretical and practical background on the human senses and mediated perception. Further, we proposed a definition for SA, which was used to conduct a literature review aiming to identify previous work that fits this category. Finally, we proposed a design space for SA. However, we have, so far, not discussed in detail what we expect SA to achieve, either from a technical nor an interaction perspective. We believe that the large scale implementation of SA applications has the potential to benefit society in many ways by empowering individuals beyond natural limitations. These themes will be discussed in this chapter, where we present a vision for SA and the effects in individuals and society of its incorporation into human activities.

Parts of this chapter are based on the following publication:

- F. Kiss and A. Schmidt. 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*, PervasiveHealth'19, pages 377–382, New York, NY, USA, 2019. ACM

6.1 Technical Development

Humans develop technology to harness the power of natural laws and use it for their advantage. Technology increases our power over reality and our ability to perceive it. The later goal, namely the empowerment of the human senses to perceive more of reality, has been an object of research since very early technological developments [217]. This has been a steady trend across human history, resulting in today's technological artefacts being deeply ingrained in human activities; such is the case for telephones, telescopes, radiographs, or microscopes.

Modern break-throughs unlocked and will keep on unlocking novel possibilities to increase the portion of reality we can perceive, thus enabling users to augment their sensation and perception. Further, the development of ubiquitous computing now allows rethinking how we interact with these technologies, transcending the concept of tools and entering into the realm of human augmentation as suggested by Schmidt [208].

Transparent Head Mounted Displays (HMDs) and hear-though headphones can deliver information collected over ubiquitous distributed sensors and information structures, effectively providing an enriched experience of reality that transcends the naturally perceivable phenomena. This technological context provides the fundamental infrastructure to make Weiser's dream real [243]. Our vision for SA is aligned with Weiser's, both seeking to eliminate the operation of technology from the interaction and in consequence enabling a *truly augmented* experience of reality.

6.2 Weiser's dream

"The most profound technologies are those that disappear." - so started Mark Weiser his visionary article from 1991 "The Computer for the 21st Century" [243]. The author made an educated prediction of how interaction with computers would escape the constraints of workstations and become embedded in the users' environment through *embodied virtuality*, while *"vanishing into the background"*.

Computers empower users, giving them access to vasts amounts of information or remote access to distributed systems, but at a cost: the interaction. Operating computers requires time, attention, and training, and for a long time also required users to physically be in front of a workstation. Advances in technology and HCI enabled luggable and portable computers, and wearable computers are currently making their way into the public's lives. These advances result also in the operation of computers requiring shorter times, less attention, and in some cases also less training. This contributes to reducing the part of the interaction that consists in the operation of a system and thus brings Weiser's vision closer to fruition.

Some work reviews Weiser's vision and reinterprets it from a more actual perspective, resulting in new fields of research and application, such as collective computing and pervasive computing [2, 201]. Our vision builds upon this work, adding a new layer to interaction design. As pointed out by Bell and Dourish in the revision of Weiser's work, ubiquitous computing is already an intrinsical aspect of human society [19]. A vast amount of computers are embedded in our environments. They are interconnected, forming a global web of information collection, storage, and distribution. Currently, we access this information through classical interfaces, most of them consisting of screens and buttons. These interfaces have been the product of human creativity overcoming technical limitations over decades, as reviewed in Chapter 1. These limitations are no longer constraints for interface design, and thus we believe now is the time to rethink interaction. For this, we argue not to focus on the architecture of the interaction, but on its ultimate goals.

6.3 Augmenting the Human Senses

The main objective of SA is empowering users to perceive more of reality. Like any other interactive computer systems, SA applications need an interface to allow users to control them and to deliver information to the users. So far, this kind of interaction has been based on the construct of tools: artefacts that perform a function.

For these applications to be true extensions of the natural human senses, what we call *true augmentation*, their interfaces must be minimized out of existence. The information presented to the user should not obstruct natural perception. The operation of the augmentation should not require active engagement, thus rendering traditional interfaces hopelessly inadequate. Natural perception occurs almost automatically, with some particular cases allowing control through intent, such as how humans can control the focus of their gaze. These requirements pose a challenge to current interaction design, which can be seen as a research opportunity.

6.4 Usage Scenarios

SA can be beneficially applied to a wide spectrum of human activities. Sight magnification has advantages for diverse fields, ranging from surgery, fine mechanics, biology, electronics, while also being useful for trivial activities, such as looking for a small item under the table or reading the small print of an advertisement. Similarly, infrared vision, auditory selectivity, or a refined sense of smell can have multiple fields of application, both in work and leisure activities. In general, it can be argued that most human activities significantly depend on perception. An increment and refinement of perception can result in quantifiable benefits in many of these activities without the tradeoffs typical from tool usage.

If we disregard current technical limitations, it is easy to imagine how SA can empower humans to know everything only by wishing to do so. To better illustrate our vision, we describe how a fully deployed multimodal SA system could be experienced in the following fictional testimonial:

> I am sitting in a coffee shop, looking outside through the window. I look at a tree on the other side of the street. I try to *recall* the species and instantly *remember* is a “pinus pinea”, typical from the Mediterranean region. There is a tiny bird on the tree. I *squint my eyes*, and I *zoom in* to the bird for a few seconds. I try to *pay attention* and can *hear* through the glass and at a long distance how it sings. I look down at my cup of coffee on my table, and *wonder* if it is still hot enough to drink. I *know* it is 30° Centigrade, so I desist.

> Suddenly, I *become aware* that my friend Joe wants to talk to me, so I *decide* to talk to him. I *hear* his voice in my head, asking if I want to meet and hang out, so I *form* the words in my mind, saying he should come to meet me at the coffee shop. I *think* of where I am, remembering the name and address of the coffee shop, *knowing* Joe will then also know. He *replies* "see you soon", which I *know* it means he will be here in approximately 17 minutes, so I *look* at the waiter, and he nods and starts preparing another espresso.

The example above illustrates a series of interactions with SA intertwined with common activities of a user's daily life. These interactions are highlighted with *italic* type. In this example, a SA system detects and interprets the user's intent, retrieves the relevant information from a network of sensors and stored information, and conveys it to the user. Further, using the same principles, it enables *telepathic*-like communication with other users. This technology thus enables to access information about the immediate physical environment, to retrieve information stored in distributed systems, and to communicate remotely with others effortlessly and with an unmatched level of privacy. Naturally, some aspects of the interaction described above are technically unfeasible with existing technology. However, the further development of Brain-Computer Interfaces (BCIs), context-aware systems, Machine Learning (ML), and other technologies will likely support these interactions in the mid-term.

6.5 Requirements and Challenges

Multiple aspects of our vision for SA need extensive work before becoming viable. As mentioned in the previous section, some fundamental technical aspects are yet unresolved, albeit these are beyond the scope of this thesis. Our work is focused on the HCI perspective of SA and thus we are only concerned with such requirements and challenges for the augmentation of the senses.

The main challenges we face for creating SA technologies are how to control such applications and how is information delivered to users. The need to identify possible collateral effects of SA on users requires particular attention since it is unclear if these systems can have detrimental effects. Further, it is important to understand how these technologies will be received by society, given the possibility that augmenting the perception of individuals can be perceived as a threat by third parties. These questions, previously formulated in Chapter 2, guide and structure the rest of this thesis. We believe that answering them, even if only partially, is an important step towards bringing our vision, and Weiser's, to become reality.

However, answering these questions from a purely theoretical stance is impossible. Understanding SA requires concrete applications to be designed and studied. This approach is aligned with conventional HCI research strategies and it is the methodology applied in this thesis. Thus, we use the proven methods of HCI to explore and understand SA.

Chapter 7

Research Probe: Enhanced Selective Hearing

> The things that really matter don't mix with idle chatter.
>
> Mose Allison

In this chapter, we present a research probe aimed at exploring the possibilities for augmenting the classical sense of hearing. Despite the existence of multiple instances of work in this topic, the lack of a targeted and systematic investigation of augmented hearing prompts to examine this particular direction in detail. The work discussed in this focused in the *amplification* of the sense of hearing, resulted in an artefact which was evaluated in five users studies, and produced the specific design space for HA.

Parts of this chapter are currently ***under submission*** *and planned to be published in the following form:*

- F. Kiss, P. W. Woźniak, and A. Schmidt. An Empirically Informed Design Space for Auditory Augmentation. In ACM Transactions on Computer-Human Interaction (TOCHI).

7.1 Introduction

Computer mediation reshapes how we experience reality [209]. The amplification and augmentation of the human senses have the potential of enriching our experience in as-yet unimaginable ways, but only in the measure in which it does not overexert human attention and cognition. We live in an *information society* where we are continually exposed to large volumes of sensory stimuli [92]. As we further develop technology, the information we can ubiquitously access grows in both quantity and variety [93].

The amount of information available and presented to us results in high demands from our cognitive capacity. Information consumers are prone to keep their stimulation at a high level, often at the edge of overloading their cognitive capacities, and stopping consumption barely before limits are reached [103]. This increasing amount and variety of information poses a threat to the quality of life, even to the health, of users and already takes a toll on their cognitive abilities [16].

As exposure to information increases, strategies to optimize our information consumption become necessary. Not all information we perceive is relevant and its relevance depends on context and situation, so dynamically filtering irrelevant information can aid to avoid overwhelming users [20, 28]. In practice, this is strongly suggested by the popularity of active Active Noise Cancelling (ANC) headphones: users ascribe an increasing value to quietness and personal space [82]. Being able to limit environmental noises and shaping the surrounding soundscape could empower users to silence background auditory information streams when deemed convenient [80]. Consequently, understanding how the auditory environment can be augmented to benefit the users' interactions with the world around them emerges as a challenge for HCI.

To answer this challenge, here we investigate the use of HA for selective auditory attention and present a twofold contribution: first, we propose a Enhanced Selective Hearing (ESH) — an HA that tailors soundscapes by amplifying a selected sound source while muting other sounds. Through an incremental and iterative design and evaluation process, we evaluated ESH in terms of usability, performance, and acceptability through five different formative studies. A significant body of past work has explored soundscape curation and selective attention, both in Psychology (e.g. the cocktail party effect [249]) and HCI [80, 143, 239]. However, the design requirements and constraints involved in building HA systems were not explored in these.

We have derived a design space for HA based on the findings and insights gained during the evaluation of ESH. This design space classifies HA in terms of four dimensions: sound properties, control properties, information properties and social properties, facilitating the categorization and analysis of existing systems, and aiding in the identification of research opportunities and ideation of novel HA interfaces.

We contribute the following: (1) the development and evaluation of proof-of-concept prototypes of ESH, a HA system; (2) five formative studies that explore the feasibility, control and spatial properties in ESH; and (3) a design space for HA based on the five studies and related literature. In this paper, we first review relevant related work. We then report on the five studies conducted, discussing key insights from each study. Finally, we introduce the design space for HA.

7.2 Related Work

The augmentation of the sense of hearing to increase selective auditory attention requires, as a first step, understanding the cognitive basis of this challenge. Once the underlying problem is identified, a technical solution needs to be developed, and an interaction design proposed.

7.2.1 Selective auditory attention

Humans can focus perception and cognition on individual parts of reality by controlling attention. In particular, this is also true to a certain extent for the sense of hearing. This ability was investigated by Cherry [47] through a series of experiments in which he presented participants with different spoken messages in each ear. Participants were instructed to listen to the message in one ear and ignore the other one. Cherry observed that participants could successfully retrieve the required information, but they were unable to report on the other message beyond some limited features, such as gender or tone of the voice. This line of work was continued by Broadbent [32, 31] with the proposition of a model to describe selective attention. He suggested that the human brain completely filtered information that is identified as irrelevant for the task at hand. This theory was questioned by Moray [167] with data obtained in experiments using the same Cherry's apparatus. He reported that participants were able to recognize their name being said in the ignored channel, indicating that their brains were not completely blocking the perceptual information identified as unnecessary. The authors likened this effect to what sometimes happens in social gatherings, coining the term “cocktail party phenomenon”. This effect was not universal but present in only a third of the study participants. Moray’s experiment was replicated in 1995 by Wood and Cowan [249], who reported that 34.6% of participants were able to hear their names in the irrelevant message.

Conway et al. found a correlation between the “cocktail party phenomenon” and the working memory capacity of the people who experience it [49]. Experiment participants with lower working memory capacity were significantly more likely to report hearing their names in the irrelevant message. The authors thus identified working memory capacity as a critical in the cognitive process, having a major influence in the performance of cognition, particularly in selective attention tasks, although they could not establish the causality between the two observable metrics. The correlation of working memory ca pacity and auditory attention might be explained by the human capacity to isolate and identify sound sources and track them in time. Griffiths and Warren [78] suggest that *auditory objects* are mental representations of sound sources, with individual properties that identify them in time and space. In the case of complex auditory scenes, the identity of *auditory objects* is constructed over time, as shown by Best et al. [22]. In particular, spatial continuity over time enhances spatial selective attention. This means that subjects perform increasingly better at focusing auditory attention in particular objects as we track their position in time. Furthermore, the authors suggest strong parallelism between selective attention in complex visual and auditory scenes, especially when information perception demands sustained attention. Not only the individual position of sound sources helps humans identify them, but the relative position in an array of sounds plays a role in their audition. Wood and McDermott [250] explored the use of schemas to recognize sources in complex soundscapes. These schemas consist of features that are common to some sounds and enable listeners to identify them as being produced in the same source.

Research on the role of space and position in auditory perception precedes the “cocktail party effect”. Kock’s work [126] presented theoretical foundations of how human cognition decodes the direction of incoming sound using the principle of *binaurality*. Sound propagates at a finite speed and may thus the sound generated by a particular source reach each of our ears at different times, as the source is generally closer to one ear than the other. This time delay between the two auditory channels, combined with the echoes produced by the sound bouncing off our shoulders and surrounding objects, enables our brains to determine where the sound originates. Kock reported on experiments confirming this theory and discussed the human ability to control listening attention, explaining the ability to perceive weak sounds in the presence of noise as a function of the direction of the targeted sound and noise, and suggesting that increasing the volume of sound coming from a particular direction might improve attention. This observation serves as a key inspiration for designing our system for facilitating augmented hearing.

Ebata et al. [62] researched the possibilities for increasing spatial auditory attention. The authors used the concept of "attention" to describe the ability to selectively hear a particular signal within a noisy soundscape. The authors used a circular array of speakers to assess the effect of the difference in direction between signals and noise. Their results indicated a proportional relation between the difference in direction and the performance of listeners in detecting the desired signal. In other words, signals were easier to perceive when the noise came from different directions. The authors also observed that participants actively focusing "attention" to the signal performed better than those who did not. A more systematic analysis was carried out by Carlile et al.[40], who investigated the nature and distribution of errors in sound localization. The elevation error was its minimum on the horizontal plane at the height of the ears of the listener. These works showed that while the human hearing was capable of sensing direction efficiently, it was prone to certain limitations that potentially affected the ability to extract information from the environment.

This corpus of work provides insights on the challenge of focusing auditory attention at will as well as highlights its importance and difficulty from a cognitive Psychology perspective. Further, it serves as a framework to plan and evaluate applications to address this problem.

7.2.2 HCI Applications to Improve Auditory Attention

Work in the field of HCI addressed a variety of problems related to auditory attention, particularly investigating the relationship between auditory attention and spatial perception. A few HCI applications focused on the "cocktail party effect". For example, Kobayashi and Schmandt [125] mapped time to space for audio browsing, creating a user interface for temporal navigation of audio files. This work showed how the auditory perception of direction could be used effectively to manipulate complex information. In terms of spatial audio applications, AR addressed directional audio extensively. Dobler et al. [56] proposed ASR (Augmented Sound Reality), sketching a mixed reality overlay that combined 3D sounds and virtual objects. Albrecht et al. [6, 5] explored the use of audio in AR, using binaural microphones to preserve the spatiality and direction of sound in mixed or augmented environments and exploring the effect of distance in these auditory environments. Further, Naseh Hussaini [173] created a wearable prototype based on a smartphone that enabled ASR. The prototype used the accelerometer and gyroscope in an iPhone 4 to track head movements and provided the basis for audio virtual environments. Similarly, Heller and Borchers [91, 90] used the sensors embedded in smartphones to calculate the relative position and orientation of virtual sound sources and render sounds coming from specific directions. These examples show that directional audio can be effectively used to enhance the auditory experience using already existing technologies such as smartphone sensors. Our work is interestingly different as it explores how the direction of sound can be used to better perceive real life.

Another recurrent use of directional audio in HCI was the visualizing of the location of audio sources, or their direction, primarily for accessibility. VisAural [77] was a system for the hearing impaired that converted audio signals into visual cues, providing luminous feedback about the direction from which the sound was being emitted. This approach was later extended by Jain et al. [104] who developed and evaluated a head-mounted display to help visually localize sound sources. The purpose of their design was to facilitate communication for the hearing impaired, indicating when and who was talking so that the user could focus their attention on the right person. Similar work was performed by Shen et al. [214], who designed an AR system for visualizing sound sources that overlaid a visual symbol on the source in a head-mounted display. The majority of work in this area in HCI is aimed at tackling impairments. The vision of Schmidt of transcending our natural sensory limitations is rarely pursued in current research citeschmidt2017augmenting. We see an opportunity to extend the sense of hearing for all users, profiting from an enhanced capacity of understanding individual sound signals in noisy or loud environments.

A few HCI applications address the focus of auditory attention. Kidd and Favrot [115] presented a laboratory-based prototype of a visually guided hearing aid. This system was capable of steering the focus of sound amplification perceived by users by tracking their gaze. Kidd [114] evaluated the system in a controlled experiment, recruiting as participants both listeners with normal hearing and with sensorineural hearing loss. The collected data strongly suggested that the system helps users with sensorineural hearing loss to isolate and attend to single sound sources in noisy environments, particularly those closer to the users.

To our knowledge, no other research so far proposes an augmentation to increase the hearing attention of users. Further, there is no systematic exploration of hearing augmentation, nor extensive HCI research into auditory applications. The reviewed literature highlights the importance of a deeper understanding of hearing augmentation from the perspective of HCI, which is necessary to chart future research directions. Given the lack of a defined design space for hearing augmentation, we propose to formulate one based on the particular case of augmenting auditory attention.

7.3 Enhanced Selective Hearing

The goal of our work is twofold: first, to propose an interactive design to enhance the selective auditory attention, and second, to propose a design space for HA. The second goal can be effectively achieved through an analysis of the artefact created for the first goal, as recommended by MacLean et al. [148]. For the first goal, namely the design of a system to enhance selective auditory attention, we take a classic approach: we first formulate Research Questions (RQs) as guiding criteria and then propose a design capable of addressing these RQs. On a meta level, these RQs can be regarded as part of RQ2 in the larger context of this thesis since they address the main aspects of the general problem of augmenting a classical sense.

7.3.1 Research Questions

Given the exploratory nature of our research, the RQs were extended, reformulated, and refined throughout the duration of this research endeavor. For the sake of clarity, we show these questions in their final form on Table 7.1.

7.3.2 General Concept

We propose a HA to enhance the capacity of humans for focusing on individual sound sources, which we name ESH. Since humans find it easier to focus on the louder signals within a soundscape, we argue that controlling the loudness of individual sources can aid humans to better focus on particular sounds. In this analysis, we consider the sound emitted by the target sound source as a

Table 7.1: Research Questions for ESH to be addressed in Chapter 8. The questions can be regarded as sub-questions of RQ2 and are thus numbered correspondingly.

	Research Question (RQ)
RQ2.1	What are the users' expectations and requirements for HA?
RQ2.2	How can HA be controlled?
RQ2.3	Does the nature of auditory information effect the effectivity of HA?
RQ2.4	Does HA improve the users' auditory attention?
RQ2.5	Does HA have an effect in the cognitive performance of users?

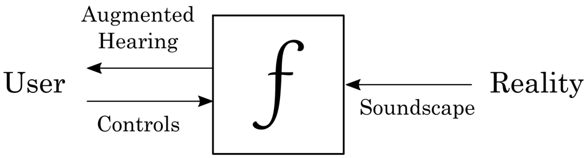


Figure 7.1: Model of a system to enhance auditory selectivity. The user controls the system to sense reality in an augmented fashion, where the augmentation is described as a function, the soundscape and the user commands as input, and augmented hearing the output of the system.

signal, while considering the rest of the soundscape as noise. Thus, we propose facilitating the focus on individual sounds by the selective increment of its Signal-to-Noise Ratio (SNR). This can be achieved by amplifying the target signal, curtailing the auditory intensity of the noise, or a combination of both.

We designed the interaction with ESH based on the following premises: (i) users should be able to toggle the augmentation on and off at will, (ii) users should be capable of easily steering the focus of the augmentation, and (iii) the main goal of the HA is to improve the reception of the information contained in a given sound signal. The considerations (i) and (ii) result in requirements for the control of the system, thus responding to user input and resulting in a modification of the output provided by the system. The consideration (iii) defines the function performed by the system, both regarding the augmentation of auditory features and the reaction to the users as considered in (i) and (ii). Consequently, we model our system as illustrated in Figure 7.1: the augmentation is the function performed by the system. Its inputs are the soundscape it senses from the real world and the user controls. Its output is the soundscape processed in a manner that enhances the ability of the user to perceive information from individual sound sources.

We identified two main challenges in the creation of ESH: a technical problem, namely the creation of a system capable of dynamically increasing the SNR for specific elements in a soundscape and obeying user input, and a design problem, concerning the user interface and interaction. Addressing the second problem is the main objective of this work since it enables answering the RQs and extrapolating a design space.

7.3.3 Design

We focus on exploring the feasibility of HA for the enhancement of auditory selectivity. In consequence, we chose to limit the interaction space to the bare minimum functionality and controls and thus reduce *tertium quid* effects when evaluating the system.

The user should be able to toggle the augmentation on and off, and so choose to listen normally or to focus on a particular source of sound. Thus, we defined a boolean input variable for the system as its activation.

In contrast, the selection of sound sources is a more complex problem since the spectrum of possibilities is large and variable. Following the relationship between sight and hearing attention suggested by O'Sullivan et al. [181], we chose to use the direction in which the user faces as a spatial selector. This decision is aligned with previous work by Kidd et al. [115, 114]. Arguably, gaze would allow finer control at the cost of triggering unwanted targeting more often. However, humans are naturally accustomed to move and turn their heads to improve hearing [244]. Other methods were considered and eliminated due to their impracticality, both from the perspectives of User Experience (UX) and technical implementation. The possibility of using alternative approaches is further discussed in the limitations section.

The output of the system, while active, should be the targeted sound signal, amplified to a level that is a compromise between a high SNR and both safe and comfortable for the user. Since the selected sound source should be in the exact direction faces by the user, the output signal will be identical for both ears. We assume that neglecting environmental effects of sound propagation and reflection has no effect in ESH, and if this was the case, their effect would likely be cognitively detrimental. While the system is not active, the user should be able to listen normally to the environmental sound.

7.3.4 Functionality

Providing the functionality of ESH can be achieved by diverse methods. To minimize the complexity of the system, we decided to use a directional microphone fixed to the head of the user, aligned with the direction the user is facing. The sound signal captured by the directional microphone is amplified and reproduced for the user through headphones. The SNR of this signal can be further increased using ANC headphones, thus diminishing the noise level.

7.3.5 Method: Five Studies of Hearing Augmentation

Addressing the RQs and augmenting audition to produce ESH presents multiple challenges: the critical factors in the interaction to be evaluated must be identified, the adequate experiments to evaluate those factors must be designed, and a functional prototype to support the interaction must be implemented. We chose a user-centered design approach. The involvement of users in the design process and evaluation is necessary to guarantee a positive interaction and UX. Finally, given the exploratory nature of our work, the development and evaluation of the augmentation can greatly benefit from an incremental and iterative approach.

To address the challenges described above, we based our approach on a user-centered design process [100], thus ensuring the involvement of potential users both in the design and evaluation of the ESH. During the initial phase, we established preliminary criteria both for the concept and its evaluation. Additionally, we implemented a basic physical prototype that embodies the functionality and primitive interface of the interaction. This first stage was conducted based on the literature and the goals established in the Concept section, producing requirements for the first design concept. The second stage was the iterative process, which involved potential users in the development and evaluation of the design. The requirements for the design are used to implement a prototype, which is evaluated in a user study. The results of the study were analyzed to produce new requirements, aiming to improve the interaction both in terms of interface and functionality of the system (see Figure 7.2).

This process resulted in five iterations of the interaction design with five studies, each one with prototypical implementations of incrementing fidelity. Each of these iterations builds upon the previously gained knowledge, refining the design and focusing on arising characteristics of the interaction. Further, each study contributes to answering the RQs and advances the design of ESH. These studies

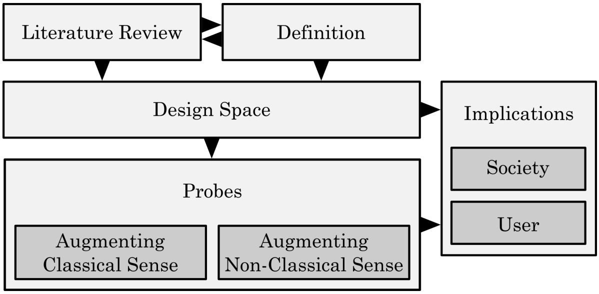


Figure 7.2: Iterative Design: at a preliminary stage, we formulated the initial requirements for the system based on its goals and literature. Based on this, the system was developed in an iterative, user-centred process.

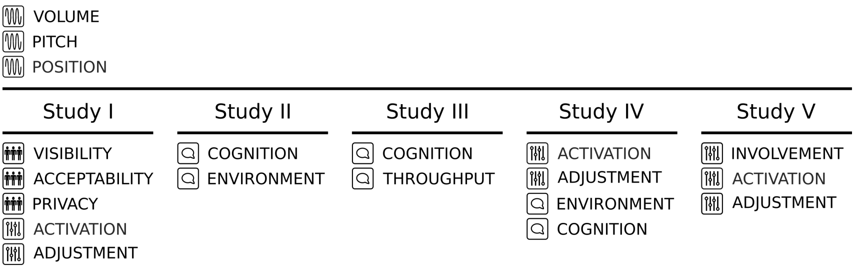


Figure 7.3: Overview of the five studies conducted to explore ESH. Each study evaluated specific dimensions that together define a Design Space for Hearing Augmentation. The dimensions can be grouped in four general categories: sound properties, control properties, information properties and social properties.

are explained in detail in the next sections, followed by a design spaced based on the principal dimensions we identified during the design and evaluation process (see Figure 7.3).

7.4 Study I: Usage Scenarios

The first study consisted of a preliminary design, based on the defined goals and literature on the topic and a *future technology workshop*. The focus of this study was to gain a fundamental understanding of users' expectations, requirements, and preferences for a system providing ESH, as well as establishing the usage scenarios that potential users consider of importance.

7.4.1 Preliminary Design

Recognizing the difficulties of envisioning a hypothetical system, the first approach to the subject poses the challenge of ensuring that participants understand the system's characteristics without priming their views and expectations.

We opted for merely illustrating the fundamental functionality of ESH in a video clip. The video display a scene with three actors speaking simultaneously. For the first 10 seconds, the voices of all three actors were reproduced simultaneously, making difficult to understand what any of the actors said. Next, the ESH augmentation is activated and only the voice of one actor remains audible while the other two are muted. A text overlay indicates the change and which actors are being heard.

The interaction with the system was sketched based on the features established in the concept section. The user hears the augmented soundscape through headphones, can toggle this function on and off, and can select individual sources. The choice of method for selecting is based on past work. A strong correlation between visual and auditory selective attention has been highlight by multiple authors [22, 181, 250]. Further, a number of applications have bound gaze or visual targeting to sound localization [5, 6, 77, 104, 114, 115, 125, 214]. This approach takes advantage of the intuitive mechanism of focusing attention visually, as typically experienced when reacting to sudden unexpected sounds by looking in their perceived direction of origin. Thus, we propose to control ESH by targeting the sound sources with the direction of the field-of-view of users.

The method for controlling the ESH is not explicitly indicated in the video and thus encourage the proposal of alternative control methods by participants of the *future technology workshop*.

7.4.2 Future Technology Workshop

We assessed the users' expectations, requirements, and concerns for ESH in a user study. We chose to achieve this with a *future technology workshop* (FTW) since this validated method facilitates a collaborative exploration and design of interactions with yet unexisting technologies [238].

Apparatus

The workshop was conducted in a meeting room with a projector used to show the video clip to participants. Participants were provided a flip-chart, paper, sticky notes, and writing material to aid them in sketching and describing their ideas. The complete duration was recorded with a video camera positioned to simultaneously capture all participants and the flip-chart. The session was conducted by two experiments, one leading the session and the other taking notes.

Procedure

Participants were first briefed about the nature of the study, their role in the workshop and their legal rights as participants. Next, they were asked for consent to participate and record the session, after which the video recording was initiated.

The workshop started with a short presentation about augmentation through technology and a brief group ice-breaking activity to encourage participation. Then participants were shown the video illustrating ESH.

In three separate ideation sessions, participants were asked to imagine control methods and implementation approaches, use case scenarios, and social and individual implications for the deployment of ESH. Each of these topics was asked individually, giving the participants five minutes to visualize and sketch their answers, and five minutes to present them. After each participant presented her answer, a group discussion was motivated. The experimenters only participated in the discussion as moderators and limited their input to suggesting topics once it was clear that participants would not suggest them themselves.

Once the three ideation sessions were concluded, the participants were asked to provide any additional feedback or suggestion. The total duration of each FTW session was of two hours.

Participants

We conducted two sessions of the FTW, each with six participants. Participants were recruited using mailing lists and posters in public spaces. Seven of the participants declared being female and 5 male. Their ages ranged between 20 and 25 years. Participants received compensation of 20€each.

Results

During the FTW, we discussed social and individual implications, use cases, control methods and different implementation approaches. Two researchers analyzed the recordings of the sessions, generating notes independently to later structure them using affinity diagrams. The collected feedback suggested a preference for gesture-based controls. In particular, participants mentioned raising one's open palm to the ear — a gesture known to signify that one is trying to hear better. There was a consensus among participants on the importance of allowing the user to turn the augmentation on and off at will. Interestingly, most suggested application scenarios concerned accessibility, and privacy concerns were only discussed when the participants were explicitly prompted to think about them. Finally, the participants wondered if they would be able to react to stimuli by changing hearing modes quickly enough to extract information from the environment efficiently. They related this problem to the different nature of sounds, particularly continuous signals versus short individual sounds.

Discussion

The FTW showed that the ESH concept was appealing to users and they could easily identify its possible benefits in real life.

It also highlighted how control was a key design consideration for AH systems. The main findings of this study come as challenges for the next iteration: the nature of the perceived signals can have a determining effect in the effectiveness of the ESH. Further, the complexity of the task will also likely influence to which extent ESH can improve hearing. Finally, participants of the FTW suggested a defined gesture to toggle ESH.

7.5 Study II: Proof of Concept

Based on the insights collected during the first study, we proceeded by conducting a proof-of-concept. We focused on the evaluation of two main aspects highlighted by participants of the FTW: (1) comparison of benefits of directional hearing for listening to continuous signals versus transient bursts of auditory information, and (2) performance in both simple tasks and complex realistic scenarios. Further, the additional knowledge about user preferences gained through the FTW enabled the completion of a design for ESH, and the implementation of a prototype.

In this section, we present the second iteration of the design for ESH and its experimental evaluation.

7.5.1 Design

The new design was based on the knowledge gained in Study I. Based on the consensus from the FTW participants, the augmentation should be activated by positioning a hand next to the year, performing a typical mimic for hearing. While holding this position, the user clearly hears the selected sound signal and the remaining environmental sound is silenced. The user selects the sound of the source by facing its direction. The interaction is terminated when the user removes the hand from the position.

7.5.2 Evaluation

To evaluate the design, a physical prototype was implemented. Further, it was necessary to construct an adequate apparatus, capable of providing complete control over the soundscape experience by the user to the experimenters, thus ensuring consistency across trials. To assess the effect of the nature of signals, a suitable procedure and an appropriate task had to be designed. All these aspects of the experiment, as well as the collected data, are presented in the following.

Prototype

We implemented a physical prototype capable of providing the functionality and interface proposed in the design. The prototype consisted of a wearable system, comprising a directional microphone for capturing the auditory signal, a set of headphones for reproducing the signal to the user, a smartphone for amplification, and a proximity sensor to detect the gesture (see Figure 7.4, left). We used a RØDE NTG-2 [5] directional microphone, which offers intermediate directionality and thus requires less precision from the users' targeting. The microphone was fixated on the top of a reinforced hat, pointing in the direction perpendicular to the user's face. The signal picked by the microphone was amplified by an Android phone running a custom application. The activation gesture was detected by a proximity sensor (see Figure 7.4, right). This was implemented using an infrared emitter and detector module attached to an RFduino[6] board powered by a coin-cell battery. The sensor was attached to the hat and over the ear of the user, on the side of the dominant hand. When the user performed the gesture, the proximity of the hand triggered the sensor, which wirelessly communicated

[5] http://www.rode.com/microphones/ntg-2

[6] https://github.com/RFduino/RFduino

with an Android application on the smartphone, activating the augmentation. As soon as the hand was no longer detected by the sensor, the augmentation was deactivated.

Apparatus

In addition to the prototypical system, the evaluation of the design requires an experimental setup capable of providing a controlled soundscape. Drawing from previous work by Ebata et al. [62] and Carlile et al. [40], we positioned a set of eight audio speakers forming a regular octagon (see Figure **??**) in a quiet room. The speakers were positioned at a height of 130 cm, roughly the height of the directional microphone for a sitting user. The speakers were connected to an audio interface (Behringer FIREPOWER FCA610[7]), capable of reproducing audio simultaneously through eight channels, each assigned to one speaker. Using Reaper v5.40[8] to control the interface from a computer, we were able to reproduce an individual track of audio in each speaker, simultaneously. With all the speakers pointing towards the center of the array, a person in this position would perceive all signals as coming from all around. We marked the speakers from 1 to 8 in clearly visible signs, positioned a rotating chair in the center of the array, and adjusted the volume of the speakers so the sound measured at the center of the array was below $60dB$.

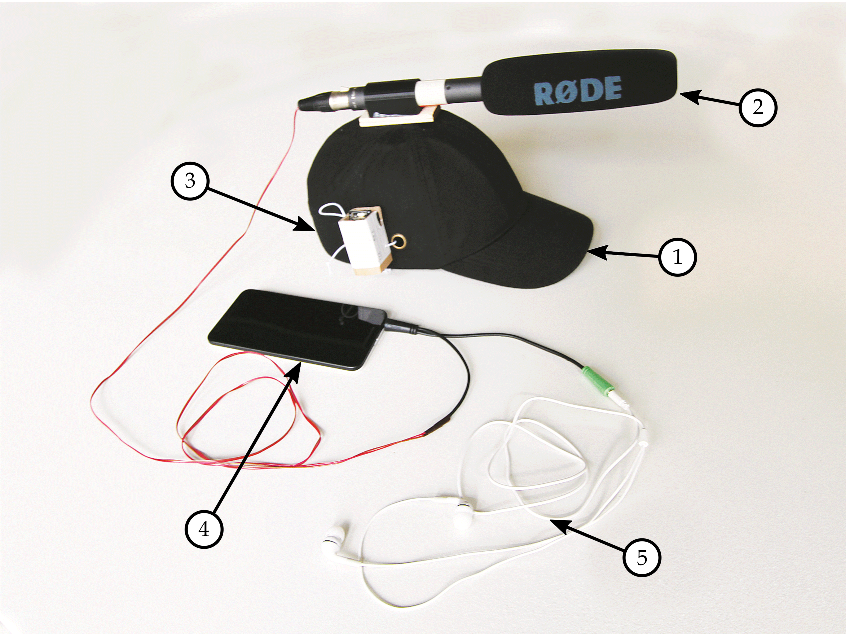
Tasks

The user experiment consisted of two tasks. In the first task, we evaluated the effectiveness of the system for obtaining information from periodic signals. To this end, the audio system repeated a different series of four digits on each audio channel, preceding each one with the command “ignore” (e.g. “Ignore one, ignore five, ignore three, ignore eight"). All speakers played each number simultaneously, at a frequency of 0.5 Hz. After a minute, one of the speakers started preceding the digits with the command “write” (e.g. “write one, write five, write three, write eight"), signaling the participant to input the numbers on a form on a tablet. Once the participant had all digits, she signaled the experimenter and concluded the trial. The task was completed correctly if all the numbers in the sequence were correctly written, irrespective of the starting digit of the series.

---

[7] `https://www.music-group.com/Categories/Behringer/Computer-Audio/Audio-Interfaces/FCA610/p/P0A3B`

[8] `https://www.reaper.fm/`



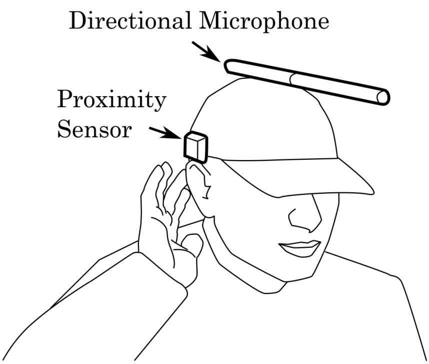


Figure 7.4: Physical implementation of the design for Study II: the prototype (top) consists of: (1) a hat, (2) a directional microphone, (3) a wireless proximity sensor, (4) a smartphone, and (5) headphones. While performing the activation gesture (bottom), the user can select sound sources with the direction of her head.

The second task aimed at assessing the effectiveness of the system in obtaining transient information in short bursts within a complex soundscape. We chose a use case scenario fitting this task: the audio system simulated the soundscape of a large train station. We recorded environmental sound from a local hub, which we mixed with computer-generated train announcements that we played once at irregular moments, in different speakers. The task of the participant was to note down the platform and delay of a particular train, having been provided with the train number and destination.

For both tasks we measured the amount of errors and the task completion time (TCT), as well as asking participants to fill a NASA Task Load Index (TLX) questionnaire [88] for each task and condition. Both tasks were performed twice; once using the prototype as aid, and once without the augmentation activated but still wearing the device. Before performing the tasks, participants were allowed time to familiarize themselves with the task prototype in two soundscapes in trial tasks.

Participants

We recruited 16 participants (5 females and 11 males) with ages between 22 and 31 years ($M = 24.93$, $SD = 2.57$) and without any hearing, sight or cognition impairments. Participants received 10€as compensation.

Results

All participants provided the correct answer to the first task, within the minimal time possible. No participant was able to gather any correct information during the second task. Participants stated in post-participation interviews that the first task was too easy, while the second was too difficult, which is consistent with the observed results. All participants described the system as attractive, novel and potentially useful. Some found the control method problematic, suggesting the use of tactile buttons. One participant stated that she did not feel sure about performing the gesture correctly, but “there is no wrong way to press a button”. Further, some participants stated to be unsure if the augmentation was active, suggesting to provide additional feedback, e.g. a LED. In terms of the hardware used, participants reported perceiving a delay between the sound played on the speakers and the signal amplified by the prototype. Some of them claimed this was distracting and had a detrimental effect on their performance. All results of the NASA TLX were non-significant. During post-participant interviews, participants recognized the novelty and usefulness of the device. When asked about possible concerns, participants only mentioned wearability and social acceptability due to its appearance, but privacy issues were not mentioned. When asked about possible use case scenarios, participants only suggested accessibility applications.

Findings

The feasibility study suggested that locating the sound source using directional sound required extra time. This implied that the utility of augmented hearing might be limited when listening to transient short signals. On the other hand, we observed that using the prototype was effective for information that was in the environment for a certain amount of time. Further, the gesture-based gesture proposed in the workshops was not well received by participants and, in practice, hindered the usage of the prototype. These results suggest that the properties of the sound to be augmented and the augmentation itself require finer control in order to provide user benefit. Finally, we also identified a need to further study the control method for the mechanism.

7.6 Study III: Refined prototype

Study II allowed us to further improve our prototype and interaction design. Adapting our strategy to the lessons learned, we excluded the evaluation of the activation and control methods from the study, focusing exclusively on the cognitive implications for ESH.

The following sections describe the incremental refinement of the prototype and the evaluation for this third iteration.

7.6.1 Design

The interaction designed remained mostly unchanged. The only modification consisted on the removal of the gesture activation, which was replaced by a simple press button. This way, the augmentation is active as long as the user is pressing the button and terminated upon release.

7.6.2 Evaluation

Building upon the knowledge gained in the first two stages, we identified two main activities that users perform while focusing their auditory attention: determining the location of the source and obtaining the information. This separation is congruent with the Psychology literature on the topic [180, 12]. We designed an experiment to evaluate the improved prototype in terms of these two activities, individually and combined, resulting in three tasks. The tasks, and especially the spatial aspects, are based on previous work from the field of Cognitive Psychology [57, 73].

Prototype

In contrast with the minimal modifications of the design, the physical prototype was largely revamped (see Figure 7.5). The smartphone was replaced by a custom made amplifier, miniaturized to be directly attached to the microphone. The amplifier was based on the LM386 Low Voltage Audio Power Amplifier[9], achieving nominal gains of up to 200 (approx. $45dB$). The augmentation was activated by holding down a physical button, with the active status being indicated by a red LED next to the button. Finally, to further increase the SNR, the standard headphones were replaced by Active Noise Cancelling (ANC) headphones (AKG N60NC[10]), toggled simultaneously with the augmentation.

Apparatus

In addition to the improved physical prototype, we reutilized the experimental setup from Study II.

Tasks

To investigate the aspects of interaction identified above, we designed three tasks. Each task consisted of five repetitions of the same assignment, but with increasing levels of difficulty. For each repetition, the participants would receive a card with indications for that given assignment and difficulty level. Each card also provided space for the participant to write down the answer to the task. For the complete duration of each task, all speakers were simultaneously repeating their respective number series.

[9] `www.ti.com/lit/ds/symlink/lm386.pdf`

[10] `https://www.akg.com/Headphones/Over-ear%20%26%20On-ear/N60+NC.html`

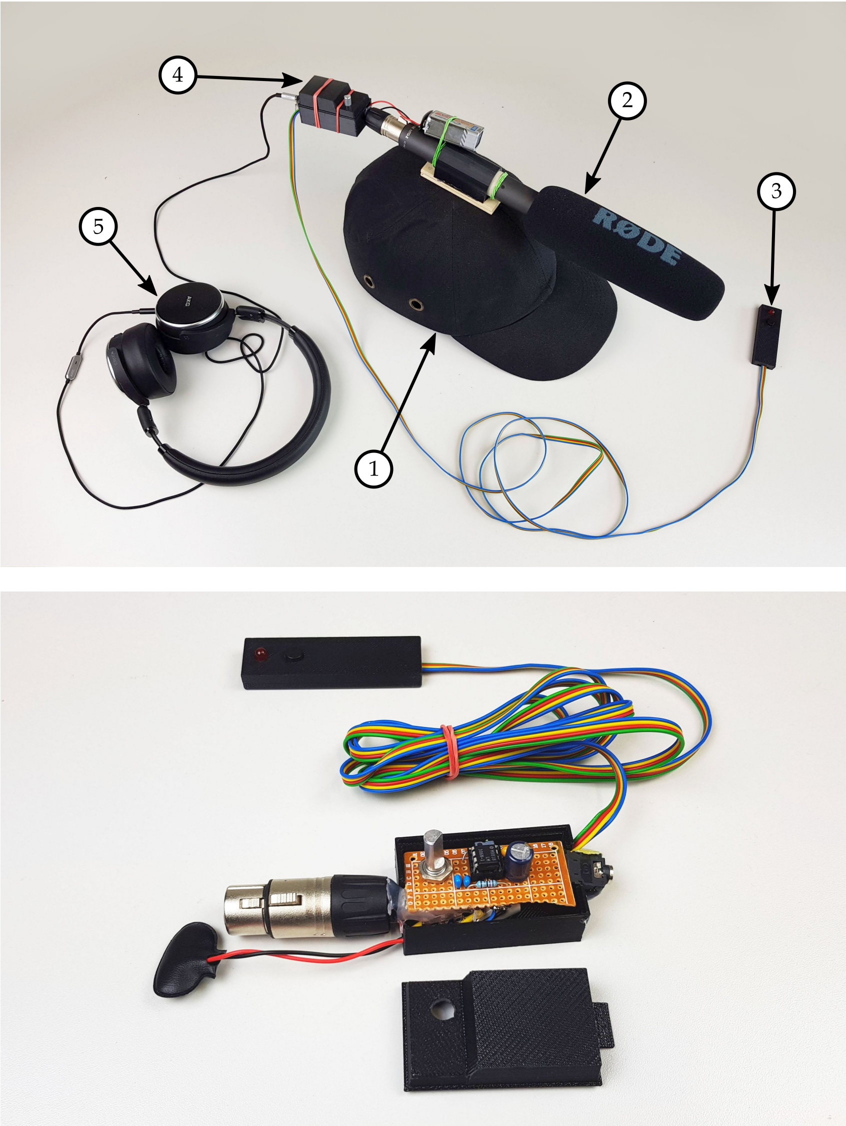


Figure 7.5: Physical implementation of the design for Study III: the prototype (top) consists of: (1) a hat, (2) a directional microphone, (3) a press button, (4) an audio amplifier, and (5) noise-cancelling headphones. A battery powered custom audio amplifier (bottom) was designed for the augmentation.

The first task (*source identification*) aimed to measure the effect of the hearing augmentation in identifying the source of audio information. A speaker was continuously repeating a number printed on a card given to the participant while all other speakers were broadcasting a series of other numbers. Once the speaker was identified, the participant wrote down its number on the card. For the first card, the number had just one digit, increasing by one with each level of difficulty up to five. The second task (*information retrieval*) aimed to measure the effect of the hearing augmentation in obtaining information from a given speaker. The card given to the participant indicated the speaker number, and the participant had to write down the number series, which had just one digit for the easiest repetition and was incremented by one with each card, reaching a maximal difficulty of five digits. The third task (*compound task*) was targeted at measuring both spatially finding sound sources and obtaining information, all while under cognitive load. During this task, each speaker continuously repeated "Go to.." and the number of a different speaker, directing the user to a different sound source. An increasing number of speaker switches (starting with one for the first card and up to five for the last one) and a starting speaker number were printed on the cards. The task of the participant was to start listening to the speaker indicated on the card and then following the directions to different speakers as many times as indicated on the card. This task can be illustrated with the following example:

1. Card indicates "Speaker 3 - 2 times".
2. The participant finds speaker 3 and listens to its message: "Go to 5".
3. The participant then follows the instruction for the first time and thus focuses on speaker 5. Speaker 5 repeats "Go to 1".
4. The participant follows an instruction for the second time and thus focuses on speaker 1. Speaker 1 repeats "Go to 7".
5. Having followed speaker instructions two times, the participant writes "Go to 7" on the card and says "Done".

For every task and assignment, we measured task completion time (TCT) and the error rate (the ratio of correctly reported digits to all digits in the task). Additionally, after each task and condition, participants were asked to answer a NASA Task Load Index questionnaire [88], and for each condition a System Usability Questionnaire [14], to respectively assess the demand of the task on participants, and how they perceived the usability of the system.

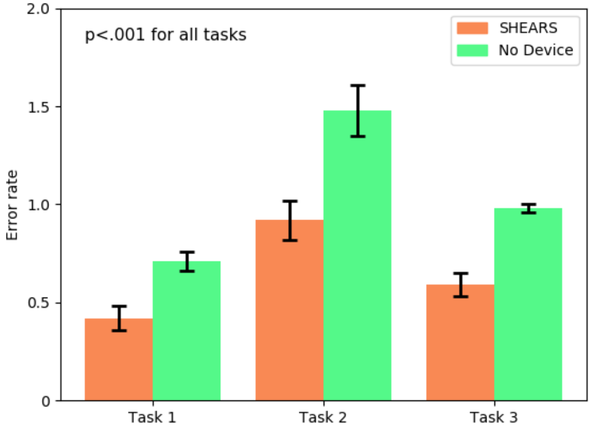


Figure 7.6: Study III: Mean value and standard error for error rates among participants, for Tasks 1, 2 and 3.

All tasks were performed twice, with and without hearing augmentation. To avoid learning effects, we generated two sets of audio tracks for each task and assignment. The order of the conditions, tasks and their matching sets of audio tracks was counterbalanced across the participants using Latin squares.

Participants

We recruited 16 participants (11 male and 5 female) aged between 22 and 31 years ($M = 24.93$, $SD = 2.57$). None of the participants had any hearing, sight or cognitive impairments. Participants received 10€as compensation.

Results

We performed repeated-measures one-way ANCOVAs to investigate the effect of using the augmented hearing device on task completion time (TCT) and error rates, controlling for the length of number sequence in the task for all tasks. Table 7.2 shows the results of the analysis. We found that using the HA resulted in a significant decrease in error rate for all tasks. In the *compound task*, we also observed a significant increase in TCT.

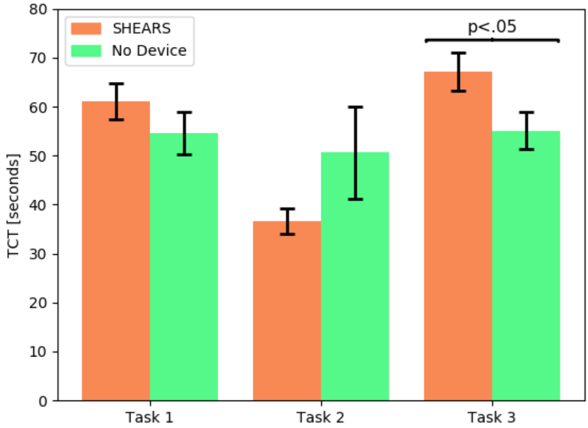


Figure 7.7: Study III: Mean value and standard error for TCT participants, for Tasks 1, 2 and 3.

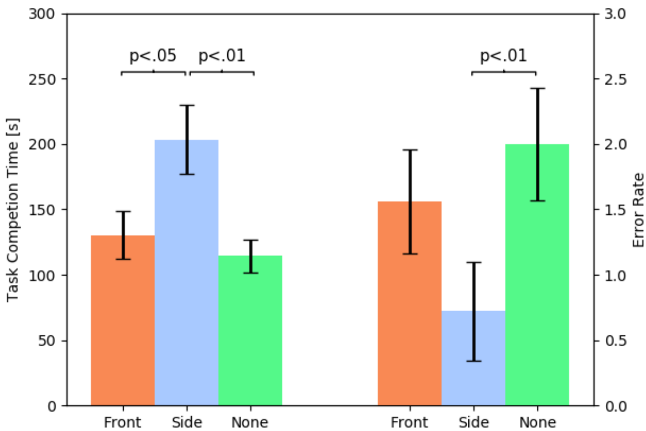


Figure 7.8: Study IV: Mean value and standard error for TCT and Error rates for the three evaluated conditions.

	TCT				Error			
	$M_{NoDevice}$	M_{ESH}	$F_{1,154}$	p	$M_{NoDevice}$	M_{ESH}	$F_{1,154}$	p
Task 1	54.62	61.06	1.50	0.22	0.71	0.42	16.81	<0.001
Task 2	50.64	36.66	2.13	0.15	1.48	0.92	17.32	<0.001
Task 3	55.12	67.12	5.14	<0.05	0.98	0.59	47.68	<0.001

Table 7.2: Mean values of task completion time and error rate, with and without assistance of the prototype, for Tasks 1, 2 and 3.

The mean values for the NASA–TLX were an average of 63.90 without using ESH and of 64.44 while using our prototype. We performed a t-test on the collected data, which showed no statistical significance: $t(93.73) = -0.15$, $p = 0.89$. The SUS score across participants was $M = 65.71$ ($SD = 9.08$).

Findings

We observed a consistent improvement in the performance of participants across tasks when using the HA, in terms of errors made. Participants recognized the presented information more accurately in all three tasks. The increased accuracy may come at a cost since TCT showed a significant increment in Task 3, which could be attributed to the difference in strategy for locating the sound sources, namely linear scanning instead of the neural representation of the head-related transfer function [12]. However, this interpretation remains inconclusive, since the behavior was not observed for the two other tasks. Overall, the data suggests that using the HA results in an increment in performance at a relatively low increment in time demand.

NASA-TLX results did not show a significant effect on the cognitive load perceived by participants during the experiment when comparing the control condition with the usage of the HA. This suggests that the use of our prototype provides a performance improvement without a significant cost in cognitive load.

The use of a push-button provided an intuitive control on the activation of the sense, but it is arguably inconvenient for use case scenarios outside a controlled environment. The proposed control method (i.e. pointing at sound sources with the head) showed an improvement in performance, but feedback collected in post-participation interviews suggested that this method is far from optimal, since we are used to hearing to the sides, and not to the front. Despite this caveat, the average SUS score for the HA can be interpreted as the device having an average usability, according to the literature [14].

7.7 Study IV: Spatial properties

Study III showed us that an ESH prototype could be successfully deployed and users could benefit from augmented hearing in specific tasks. An aspect that remained unexplored was the directionality of augmented hearing. In the previous studies, users appreciated activating ESH both through a button (central control) and touching their ear (side control). This poses the design question of whether an AH usage scenario should enhance the sound in front of the user—follow their gaze—or from the side of the user—following the user's ears. To explore this question, we conducted another experiment.

7.7.1 Design

The interaction design remained mostly unchanged respect Study III, with the only variation being the subject under investigation: the direction in which the microphone points. To investigate the effect of this factor, we propose an alternative design with the directional microphone positioned to amplify signals coming from the direction perpendicular to the dominant ear.

7.7.2 Evaluation

To investigate the effect of the direction of amplification, we designed and conducted a new experiment.

Prototype

We modified the prototype from Study III by adding a hinge, thus allowing the microphone to be rotated to both sides. The prototype remained otherwise unmodified.

Apparatus

For this study we used the same apparatus from Study III, with a modification: we added a hinge to the fixation holding the microphone on the hat. This way, we could rotate the microphone on the horizontal plane, enabling us to use the prototype focusing forward, or laterally in the direction of the dominant ear. To assess which ear was dominant, each participant was asked to pay attention to a very low sound played on a speaker in front of the participant. The participant then rotated the head to point with one ear to the speaker, revealing which ear was dominant.

Task

To compare the performance of users across conditions, namely forward focus (FRONT), lateral focus (SIDE) and control (NO DEVICE), we created three sets of eight spoken phrases using Amazon Polly[11]. Each phrase contained five individual pieces of information that the participant had to retrieve. All phrases had the same length and level of complexity. These phrases were repeated, one on each speaker, and started each time simultaneously (minimal track duration differences were compensated with silence).

The participants were presented with a card on which one of the phrases was printed. Three of the five information pieces in the phrase were missing, replaced with a horizontal line. The task of the participants was to retrieve the missing information and complete the phrase.

This task was repeated three times, once with each set and for each condition. The assignment of track sets to conditions was counterbalanced across participants, as well as the order of the conditions.

We recorded for each trial the TCT, error rate, and subjective task load using the NASA TLX. The error rate was calculated using a score point system: wrong or missing information resulted in zero points, partially correct information resulted in one point, and totally correct information resulted in two points. Thus, each task produced an error rate between zero and six for each condition.

Participants

We recruited 18 participants, four of them female and the rest male. Their ages ranged from 18 to 25 years ($M = 24.22$, $SD = 3.75$) and none of them presented any sight or hearing problems. Participants received 10€for completing the study.

7.7.3 Results

We conducted a one-way ANOVA to investigate the effect of the device version used on TCT. A significant effect was observed, $F_{2,51} = 5.76$, $p < .01$. The task

[11] https://aws.amazon.com/en/polly/

was completed in the shortest time in the NO DEVICE condition. Post-hoc tests with Tukey HSD revealed significant differences at $p < .05$ for the pairs SIDE–FRONT and SIDE–NO DEVICE.

Similarly, a one-way ANOVA examined the effect of the experimental condition on the error rate and showed a significant result, $F_{2,51} = 5.30$, $p < .01$. Here, the FRONT condition performed best. Post-hoc testing showed a significant difference between the SIDE–NO DEVICE conditions, $p < .01$. Figure 7.8 shows TCT and error rate results.

Another one-way ANOVA analysed raw NASA TLX scores and found no significant effect, $p > .05$. We then investigated TLX performance scores to investigate the participants' subjective impression of performance achieved. A significant effect was found, $F_{2,51} = 3.31$, $p < .05$ with post-hoc analysis revealing no significant differences. The SIDE was ranked as enabling the best performance, $M = 15.78$ with NO DEVICE ranked second, $M = 15.44$. In the FRONT condition, the participants assessed their performance most negatively, $M = 9.11$

Findings

The observed performance in terms of error rate confirms our findings from Study III, suggesting that ESH improves the user's ability to understand information. Participants solved the task fastest without the aid of ESH, and when using the prototype, needed less time for the FRONT facing prototype. This result is congruent with our observations from Study III in terms of TCT. From the cognitive load reported by the users, we found no evidence of a significant difference between conditions. However, participants perceived their performance to be better under the SIDE condition, which resonates with comments from participants of Study III highlighting the fact that human ears point to the sides and therefore focusing forwards feels unnatural. Consistent with this idea, the FRONT condition was perceived as worse-performing than NO DEVICE. From the findings discussed above, we derive that ESH enables a better perception of information in exchange of an increment in time to do it. A front-facing focus of the HA performs better than a side-facing one in terms of error rate and TCT. However, this does not match the users' perception.

7.8 Study V: Control

Based on the findings of the previous studies, we decided to address in detail how a possible ESH system would be controlled. To that end, we conducted a survey to identify control methods for the system. Based on the results, we conducted a controlled experiment to compare the three preferred control methods in terms of subjective workload and usability.

7.8.1 Survey

We conducted a survey with 57 participants (35 female and 22 male), with ages from 18 to 56 years ($M = 23.6$, $SD = 6.25$) to assess preferences for control modalities. The proposed modalities were: a physical button, voice commands, gestures, gaze and a smartphone (as suggested by users in Study I). Survey participants were presented with a video illustrating the cocktail party problem and possible use of ESH, along with a scenario description. We then asked them to order the methods from most to least preferred. Both in terms of most preferred and least preferred valuation, the smartphone was the most frequent choice (17 times most preferred, 5 times least preferred method), with gaze (16 most preferred, 11 least) and button (15 most, 7 least) being second and third in preference. Voice and gesture control methods were the least preferred method 20 and 14 times respectively, and most preferred method 5 and 4 times. Based on the results of the survey, we conducted a controlled experiment to compare the three preferred control methods in terms of subjective workload and usability.

7.8.2 Design

For the last iteration of the design, we compared three control methods: a smartphone, a physical button, and gaze. In all cases, the selection of the signal was controlled as in the previous studies. Smartphone control was realized through an Android application, displaying a button. The augmentation was thus active as long as the user kept a finger on the virtual button. The control through a physical button resembled the interaction from Studies III and IV: the augmentation remained active while the user held the button pressed down. Finally, the gaze controlled augmentation was continuously active as long as the user-focused the gaze on a particular sound source.

7.8.3 Evaluation

The evaluation of the different control methods posed a technical challenge beyond the scope of this work. Thus, we opted for a *wizard of Oz* approach. In consequence, we did not implement a prototype to embody the ESH and instead produced an apparatus capable of simulating the experience.

Apparatus

We used four speakers on orthogonal axes, and labeled them with numbers from 1 to 4. This setup can be seen as a simplified version of the system described for Studies II, III and IV, using only every second speaker. Each speaker repeated the number of a different speaker, with a frequency of $1Hz$, on top of a noisy background recorded from an urban environment. In this study, HA and gaze control were achieved with a Wizard of Oz approach [51, 87]. This enabled us to reduce the technical complexity of the experiment and focus solely on the question of controlling HA, eliminating factors investigated in studies I–IV. Four speakers played audio simultaneously by default. ESH was simulated by turning the volume of the other speakers when a particular speaker was activated. For the button condition, the HA would be active as long as the participant held the button pressed. For the smartphone condition, this would be controlled by a toggle button on a dedicated app. Gaze activation was achieved by activating a speaker when the user was looking at it for at least one second. These designs were chosen based on the survey and study I.

Task

The task was identical to the *compound task* in Study III. We used a within-subjects design, with the input method as the independent variable. Three conditions were evaluated: using a press-button, a dedicated smartphone application (ESH activated by pressing a software button) and gaze tracking, by performing a task three times on each condition. We measured error rate for each task and condition, and after completing the task, participants filled out a NASA Task Load Index questionnaire, and a SUS questionnaire [88, 14].

Participants

A total of 24 participants (16 females and 8 males) took part in the experiment. They were from 19 to 65 years ($M = 27.1$, $SD = 12.71$). All participants reported that they did not suffer from any physical or mental impairments, especially any hearing impairment.

Condition	TLX*	SUS	Error*	Preference
BUTTON	29.83*	73.12	0.15	7
EYE TRACKING	27.26*	80.42	0.11*	12
SMARTPHONE	29.31	78.23	0.20*	2

Table 7.3: The results of study V. Asterisks show significant main effects and significant pairs in post-hoc comparisons.

Results

We used a one-way ANOVA and found no difference in NASA TLX scores ($F(2,69) = 0.238$, $p = .789$). Aligned-rank transformed[4] SUS scores were analysed with a one-way ANOVA. We found a significant effect ($F(2,69) = 3.60$, $p < .05$), with all scores suggesting a good usability[14]. BUTTON was rated as the least usable and post-hoc tests with the Tukey method revealed a significant difference in the BUTTON–EYE TRACKING pair, $p < .05$. All other comparisons were not significant, $p > .05$ A significant difference was observed for the error rates ($F(2,69) = 4.68$, $p < .05$). SMARTPHONE produced the most errors while EYE TRACKING led to the least mistakes. Tukey HSD analysis showed a significant difference between SMARTPHONE and EYE TRACKING, at $p < .05$. Table 7.3 shows the results of the study.

Findings

Gaze-controlled HA was perceived as the most usable and resulted in the least errors. This suggests that that users prefer implicit interaction when controlling the augmentation. Interestingly, smartphone control resulted in users committing significantly more errors. While controlling HA with devices that users already possess has a practical appeal, our results suggest that using a secondary device to control HA may have detrimental effects on performance. Consequently, dedicated HA devices are likely to offer increased performance.

7.9 A Design Space for Augmented Hearing

The five studies presented in this paper constitute an analytical inquiry into the design of HA systems. We decided to study different aspects of the system in separate studies as past literature offered limited guidelines. Through engaging in the iterative design of ESH, we gained practical knowledge designing HA.

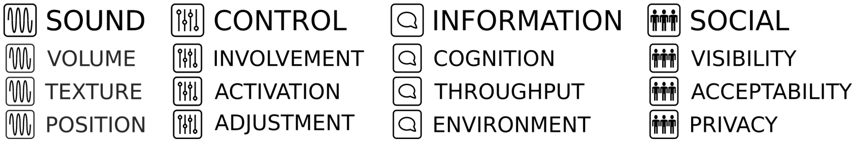


Figure 7.9: Design Space for Hearing Augmentations

Early prototypes of ESH had many shortcomings and required improvements in many dimensions. Through the five studies, we gained an understanding of the requirements and constraints to which a system like ESH is subjected. For instance, decoupling the control method from the sound properties (as in study V) allowed us to focus on the control aspect. The social dimension was explored mainly in Study I, with feedback from later studies only corroborating the importance of the topics mentioned during the FTW: privacy, social acceptability and visibility.

In order to stimulate further research in HA and allow others to learn from our mistakes, we propose a design space for HA. The design space is intended to serve as a starting point for designing new HA systems. Further, it enables the comparison between HA systems and provides a framework for the analysis of existing interactions for HA. Finally, we hope that the design space can serve as an ideation tool for creating novel HA applications. Figure 7.9 shows an overview of the design space. Based on the experience and knowledge we gained through our research, we identify four dimensions of the HA design space: sound, control, information and social.

7.9.1 Sound properties

The physical properties of sound and the subsequent need for processing are simultaneously design possibilities and constraints that need to be taken into account. Our studies showed that designing for HA requires refocusing from the traditional signal-oriented understanding of sound to properties that describe sound like users do. Consequently, the design space includes five properties: volume, range, pitch, position, and content.

Volume

The volume, or loudness, of a sound describes its intensity as perceived by the user [7]. It is equivalent to the amplitude of the signal, when measured exactly at the ears of the user. The volume of sound is a continuous dimension, ranging from zero (below the lower limit of what is audible) to one hundred, which should be chosen as the maximal possible value that does not endanger the users' auditory system. Volume for HA is not only defined by sound intensity, but also as a function of position in space. This dimension describes the loudness of sound in space, which enables the user to locate sounds. As we observed in studies III and IV, volume amplification can help users locate sources and extract information from the environment. This dimension is best described with a mathematical function in the three-dimensional space, but it can also be defined as "natural", when the behavior is the same as in the real world (inverse square law) or "constant", when the volume is independent from space coordinates [258].

Texture

The texture of sound describes the cognitive separation and interaction of auditory objects [39]. Even within individual sources of sound, the texture of a sound aids the user to recognize its originating phenomenon and classify the sound within known categories [60]. Further, the texture of a sound is a necessary characteristic of music, since harmony is precisely the interaction and interrelation between sounds [109]. Texture can convey meaning, emotions, environmental cues, and further information.

Position

The position of a sound describes the location of its source in 3D space. Studies II and III have shown that users are eager to use augmentations to better perceive the position of the sound. There are several possibilities for a HA system to provide this ability. Realistic binaural sound can be described with coordinates relative to the user [24]. Simple stereo sound must be described in terms of panning, or being present in just one channel. Mono signals, presented to both ears, have no defined position and are usually perceived as *inside the user's head*. The HA system can change the perceived position of a sound, indicating the motion of its source, simulating a stationary source where the user is moving, or a combination of both.

This dimension is best described with a mathematical function in space and time. Additionally, we propose to distinguish frames of reference between "relative" behavior, in which the position of sound sources is referenced to the user's position; "absolute", using a coordinate system fixed to the real world; and "void", with sounds that are not linked to any position in space, such as typical *stereo* or *mono* signals in headphones.

7.9.2 Control properties

Although humans possess a limited control of the natural sense of hearing, auditory augmentations can allow users to gain control on some aspects of this sense. When designing an HA system, providing effective control is a key requirement. We observed that in studies II and III where participants commented extensively on control methods, which led us to a focused exploration in study V. This, in turn, showed that HA performance and perceived usability were heavily dependent on the control methods. We define three properties of control in HA: Involvement, Activation, and Adjustment.

Involvement

The involvement of the user in controlling an augmentation can be *explicit*, which means that the user actively decides on what they hear through controlling individual sound properties. Conversely, the control of the augmentation can be *implicit*. In this case, the system supporting the augmentation is in charge of deciding what to present to the user, based on contextual information and the user's actions. These design possibilities generate additional constraints. For example, while study V showed a preference for implicit interaction, implementing such a system is challenging.

Activation

The activation property of an AH system defines how and when this enhancement is active. For some applications it might be convenient that the HA is constantly active. In others, users might prefer activating HA on demand so that the user is not overstimulated or distracted. Finally, the activation can depend on the situation, activity, time or place, suggesting opportunities for context activation. This property offers considerable opportunity for future exploration. While users in study I expressed an explicit need to turn the HA system on and off, study V showed that an always-on solution offered significantly better performance in a specific task.

Adjustment

This property is closely linked to the users' perception of the position of the sound. Our studies showed that the way users navigate the soundscape is inherently connected to physical references, both with relation to the environment and the user's body. In our work, especially in study IV, we explored two approaches to this property: use of gaze or head movement (vision-bound control); or the direction of the ears (hearing-bound control). Our work showed a preference for vision-bound control, but this may be specific to the ESH system. Further possibilities within this property involve the use of spoken commands (speech-bound [84]), the use of BCI (thought-bound [159]), and the use of the hands or hand-held objects (tool-bound [240, 174]).

7.9.3 Information properties

Another design dimension we identified related to the information that an HA system conveys to the user. Our studies featured different tasks which resulted in varying performance. This dimension helps designers navigate how, which and how much information can be processed in an HA system.

Cognition

Depending on the content carried and sound properties, sound signals cause cognitive load. Our studies echo the Bray's [28] remarks and suggest attention in HA systems should be regarded as a scarce resource. Thus, the cognitive cost of information is a highly relevant property of sounds in HA. In our studies, we observed both relatively high (study III) and low (study V) levels of perceived cognitive load. This shows that measuring cognitive load is an important consideration in HA systems. Further, our results confirm the possibility of building HA systems that do not cause excessive cognitive load.

Past work differs on how cognitive load is to be measured [55], which may be a challenge for developing new HA systems. Our work suggests that the use of application-specific metrics, or ratings based on self-reporting may provide a reliable estimate. However, the high differences in cognitive load we observe suggest that future HAs should be developed to be cognition-aware systems [35].

Throughput

The amount of information transmitted by sound emerged as a key design consideration for ESH. In Study III, we observed that an excessive amount of information presented may render the HA system useless. In contrast, if the HA does not augment a sufficient amount of information, the user may not see any benefit in using it. An analytical approach to this property of HA would include Information Theory with information rate and bandwidth as key parameters. However, these concepts do not capture the users' perception of the complexity of information shown in our work, especially studies II and III. Consequently, we propose a user-centred categorization for types of information that can be augmented, listed by increasing throughput:

Auditory icons and earcons: short bursts of sound that alert users of events or situations (e.g. notifications), or provide simple information, such as confirmation of an action.

Alarms: persistent sound signals, that make the user aware of a situation that require immediate attention, such as an incoming phone call.

Tones: persistent sound signals consisting of a modulated monotone or continuous periodic signal. Tones describe the variation in time of a given measure (e.g. heartrate, the temperature of a kettle) by matching a property of the signal, such as frequency or amplitude.

Speech: human auditory communication with a wide range of levels of complexity, depending on the content and context. It has the highest potential throughput, since it communicates information in a finite time that no other category can.

Environment

The context of the user experience is an integral part of the interaction with a hearing augmentation. The individual characteristics of the location where the interaction takes place, the background and mood of the user as well as several other factors play a role in how auditory information is perceived. As we observed in Study II, effective HA is easier to achieve in less noisy environments. On the other hand, HA may be most welcome when noise rates are high as suggested by study I and the wide popularity of noise-cancelling headphones. Consequently, the environmental context of where the HA system is to be used is a key design constraint.

7.9.4 Social properties

The social properties of the HA system are the final design dimension in our design space. The insights in this dimension were built primarily on study I and debriefing interviews in all our studies. We identified three main themes, which correspond with key research topics in HCI: visibility [146, 156], social acceptability [112, 128], and privacy [3].

Visibility

This dimension describes how likely other people are to become aware of someone using an HA system. In studies II-IV, we used a highly visible system which could be identified from a distance. This choice was dictated by the fact that a large-sized microphone was required for the sound fidelity needed for the studies. However, future HA systems could use smaller devices, even ones that may be invisible. In Study I, participants stated explicitly that they would have liked to know if someone else were using an HA system. Future systems could hide or actively show if they are currently active, resembling existing conventions such as the blinking red light on a recording video camera. Whether and how active HA is shown to other users is a key property of an HA system.

Acceptability

HA can be accepted by people to different degrees and for different reasons. Devices with unusual appearances may trigger unwanted reactions or become trendy. People might only feel comfortable using a hearing augmentation alone or at specific locations. Study I showed that users were eager to supply usage scenarios where they found an HA system to be acceptable. However, users in all our studies also remarked on HA technologies being potentially socially disruptive. Thus, understanding in which contexts a new HA can be acceptable is a key consideration for future systems.

Privacy

HA can potentially violate the privacy of other users. Designs that enable eavesdropping or tapping communications pose ethical problems. These potential issues were mentioned by users in Study I. HA can potentially provide advantages over others that could be perceived as unfair. Further, the novelty of such designs makes it hard for third parties to detect, prevent or combat privacy violations. Designers of future HA systems should carefully consider the potential misuse of HAs and discuss the implications. Designing privacy-respecting HA systems emerges as a design challenge.

7.10 Discussion

The conducted research and the evaluation of ESH provided the necessary insights to answer the RQs and understand the emerging topic of HA. In line with the approach suggested by MacLean et al. [148], we derived a design space from our artefact that describes the main characteristics of HA applications. In the following subsections we offer answers for the RQs based on the reported studies and discuss the implications of the proposed design space. Further, we comment on additional insights gained through our work.

RQ2.1: What are the users' expectations and requirements for HA?

Participants of all studies emphasized the importance of ergonomics and user-friendliness of the interaction, particularly regarding the activation of the augmentation. Participants valued the simplicity of usage and considered necessary to provide multimodal feedback about the status of the ESH (Study II).

Sound distortions, especially the latency of the amplified signal respect the soundscape due to sound processing, had a negative impact on the interaction, both in terms of performance and UX (Study II). Thus, ensuring high-quality of audio amplification and no noticeable delays is mandatory for ESH.

When asked about possible drawbacks of deploying ESH as a consumer-product, participants did not suggest privacy threats to be a concern. When prompted about this specific issue, the general response was to recognize it as a problem, although no concrete strategies to deal with it were suggested by any participant (all studies).

RQ2.2: How can HA be controlled?

The preference of a physical button over hand gestures was suggested by participants, arguing for the simplicity of actuation (Study II). A survey suggested physical buttons, smartphone interfaces and eye-tracking to be the most popular technologies to control ESH (Study V). Among those three, gaze control turned to be the best choice in terms of performance, usability, and user preference.

RQ2.3: Does the nature of auditory information affect the effectivity of HA?

Some features of the soundscape have shown noticeable effects in the performance of participants during studies. While ESH performed significantly better for retrieving information contained in periodical signals, short bursts of information were far more challenging, especially when reproduced just once per trial (Study II). ESH performed similarly for tasks consisting of understanding numbers series and phrasal sentences. Although it is possible to conclude that the nature of the targeted auditory information has an effect in the ESH, further investigation is necessary to determine the extent of this effect.

RQ2.4: Does HA improve the users' auditory attention?

The use of ESH resulted in a better performance in terms of error rate and task completion time (Studies II, III, and IV). Admittedly, this improvement in performance respect a control condition was observed only for periodic signals. This is likely explained by the time required to locate the target sound source and activating the HA, which might result in the loss of critical information for brief and transient signals.

RQ2.5: Does HA affect the cognitive performance of users?

The results of the NASA Task Load Index for Studies II, III, and IV showed no significant effects for ESH respect normal hearing. However, actuating the system inherently requires more effort than simply listening. Since the questionnaire measures the total task load for each condition, we cautiously interpret the observed lack of difference as a reduction in the trial task load while using ESH of a similar magnitude to the additional task load resulting from actuating the system. This suggests that HA might improve the cognitive performance of users, which could result in net improvement given an optimal interface design. However, the available data and the used method provide insufficient support for this interpretation, prompting a further investigation to obtain a more definite answer.

7.10.1 Design Space

The proposed design space encompasses the main features and properties of HA systems identified both in the literature and through our design process. These dimensions are grouped into four main categories, each of them belonging to a different discipline (Acoustics, Interaction Design, Information Science, and Sociology). All categories and their dimensions are integral parts of HA designs and thus, their consideration is critical for HA applications, although admittedly, in particular cases a dimension may play no significant role in the design.

The number and nature of the designs space's dimensions difficult a graphical representation. Despite this limitation, the design space can be useful to researchers both for classification and analysis of existing systems, and for the conception, design, and implementation of novel applications. By using the space dimensions as a template, designers are forced to consider simultaneously all aspects of an application starting in the early stages. This provides a clear overview of design decisions and a consistent structure for the design concept. Further, applying the classification to existing work can help designers to find solutions in the literature for problems they face for particular dimensions in their applications. Finally, the design space serves as an exploration tool for HA. Designers can ideate new applications by classifying existing systems and then incorporating variations in given dimensions.

7.10.2 Additional Insights

Some general observations were made during the conduction of the experiments, which do not address the RQs, nor belong to the design space. First, participants of the experiments consistently perceived HA in general, and ESH in particular, as accessibility applications. We can only explain this perception as anchoring or confirmation bias since it is likely that most hearing amplification devices known to participants are hearing aids. Participants did offer other possible use cases for HA and ESH but accessibility applications were the first suggestion in most cases.

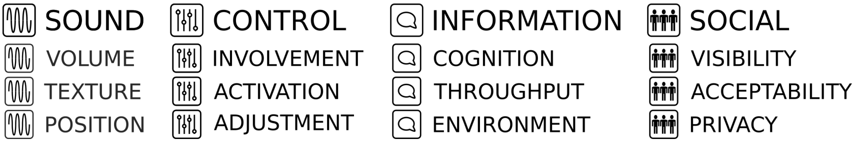


Figure 7.10: Design Space for Hearing Augmentations

7.11 Design Space for Augmented Hearing

The five studies presented in this paper constitute an analytical inquiry into the design of HA systems. We decided to study different aspects of the system in separate studies as past literature offered limited guidelines. Through engaging in the iterative design of ESH, we gained practical knowledge designing HA.

Early prototypes of ESH had many shortcomings and required improvements in many dimensions. Through the five studies, we gained an understanding of the requirements and constraints to which a system like ESH is subjected. For instance, decoupling the control method from the sound properties (as in study V) allowed us to focus on the control aspect. The social dimension was explored mainly in Study I, with feedback from later studies only corroborating the importance of the topics mentioned during the FTW: privacy, social acceptability and visibility.

In order to stimulate further research in HA and allow others to learn from our mistakes, we propose a design space for HA. The design space is intended to serve as a starting point for designing new HA systems. Further, it enables the comparison between HA systems and provides a framework for the analysis of existing interactions for HA. Finally, we hope that the design space can serve as an ideation tool for creating novel HA applications. Figure 7.10 shows an overview of the design space. Based on the experience and knowledge we gained through our research, we identify four dimensions of the HA design space: sound, control, information and social.

7.11.1 Sound properties

The physical properties of sound and the subsequent need for processing are simultaneously design possibilities and constraints that need to be taken into account. Our studies showed that designing for HA requires refocusing from the traditional signal-oriented understanding of sound to properties that describe sound like users do. Consequently, the design space includes five properties: volume, range, pitch, position, and content.

Volume

The volume, or loudness, of a sound describes its intensity as perceived by the user [7]. It is equivalent to the amplitude of the signal, when measured exactly at the ears of the user. The volume of sound is a continuous dimension, ranging from zero (below the lower limit of what is audible) to one hundred, which should be chosen as the maximal possible value that does not endanger the users' auditory system. Volume for HA is not only defined by sound intensity, but also as a function of position in space. This dimension describes the loudness of sound in space, which enables the user to locate sounds. As we observed in studies III and IV, volume amplification can help users locate sources and extract information from the environment. This dimension is best described with a mathematical function in the three-dimensional space, but it can also be defined as "natural", when the behavior is the same as in the real world (inverse square law) or "constant", when the volume is independent from space coordinates [258].

Texture

The texture of sound describes the cognitive separation and interaction of auditory objects [39]. Even within individual sources of sound, the texture of a sound aids the user to recognize its originating phenomenon and classify the sound within known categories [60]. Further, the texture of a sound is a necessary characteristic of music, since harmony is precisely the interaction and interrelation between sounds [109]. Texture can convey meaning, emotions, environmental cues, and further information.

Position

The position of a sound describes the location of its source in 3D space. Studies II and III have shown that users are eager to use augmentations to better perceive the position of the sound. There are several possibilities for a HA system to provide this ability. Realistic binaural sound can be described with coordinates relative to the user [24]. Simple stereo sound must be described in terms of panning, or being present in just one channel. Mono signals, presented to both ears, have no defined position and are usually perceived as *inside the user's head*. The HA system can change the perceived position of the sound, indicating the motion of its source, simulating a stationary source where the user is moving, or a combination of both.

This dimension is best described with a mathematical function in space and time. Additionally, we propose to distinguish frames of reference between "relative" behavior, in which the position of sound sources is referenced to the user's position; "absolute", using a coordinate system fixed to the real world; and "void", with sounds that are not linked to any position in space, such as typical *stereo* or *mono* signals in headphones.

7.11.2 Control properties

Although humans possess a limited control of the natural sense of hearing, auditory augmentations can allow users to gain control on some aspects of this sense. When designing an HA system, providing effective control is a key requirement. We observed that in studies II and III where participants commented extensively on control methods, which led us to a focused exploration in study V. This, in turn, showed that HA performance and perceived usability were heavily dependent on the control methods. We define three properties of control in HA: Involvement, Activation, and Adjustment.

Involvement

The involvement of the user in controlling an augmentation can be *explicit*, which means that the user actively decides on what they hear through controlling individual sound properties. Conversely, the control of the augmentation can be *implicit*. In this case, the system supporting the augmentation is in charge of deciding what to present to the user, based on contextual information and the user's actions. These design possibilities generate additional constraints. For example, while study V showed a preference for implicit interaction, implementing such a system is challenging.

Activation

The activation property of an AH system defines how and when this enhancement is active. For some applications it might be convenient that the HA is constantly active. In others, users might prefer activating HA on demand so that the user is not overstimulated or distracted. Finally, the activation can depend on the situation, activity, time or place, suggesting opportunities for context activation.

This property offers considerable opportunity for future exploration. While users in study I expressed an explicit need to turn the HA system on and off, study V showed that an always-on solution offered significantly better performance in a specific task.

Adjustment

This property is closely linked to the users' perception of the position of the sound. Our studies showed that the way users navigate the soundscape is inherently connected to physical references, both with relation to the environment and the user's body. In our work, especially in study IV, we explored two approaches to this property: use of gaze or head movement (vision-bound control); or the direction of the ears (hearing-bound control). Our work showed a preference for vision-bound control, but this may be specific to the ESH system. Further possibilities within this property involve the use of spoken commands (speech-bound [84]), the use of BCI (thought-bound [159]), and the use of the hands or hand-held objects (tool-bound [240, 174]).

7.11.3 Information properties

Another design dimension we identified related to the information that an HA system conveys to the user. Our studies featured different tasks which resulted in varying performance. This dimension helps designers navigate how, which and how much information can be processed in an HA system.

Cognition

Depending on the content carried and sound properties, sound signals cause cognitive load. Our studies echo the Bray's [28] remarks and suggest attention in HA systems should be regarded as a scarce resource. Thus, the cognitive cost of information is a highly relevant property of sounds in HA. In our studies, we observed both relatively high (study III) and low (study V) levels of perceived cognitive load. This shows that measuring cognitive load is an important consideration in HA systems. Further, our results confirm the possibility of building HA systems that do not cause excessive cognitive load.

Past work differs on how cognitive load is to be measured [55], which may be a challenge for developing new HA systems. Our work suggests that the use of application-specific metrics, or ratings based on self-reporting may provide a reliable estimate. However, the high differences in cognitive load we observe suggest that future HAs should be developed to be cognition-aware systems [35].

Throughput

The amount of information transmitted by sound emerged as a key design consideration for ESH. In Study III, we observed that an excessive amount of information presented may render the HA system useless. In contrast, if the HA does not augment a sufficient amount of information, the user may not see any benefit in using it. An analytical approach to this property of HA would include Information Theory with information rate and bandwidth as key parameters. However, these concepts do not capture the users' perception of the complexity of information shown in our work, especially studies II and III. Consequently, we propose a user-centred categorization for types of information that can be augmented, listed by increasing throughput:

Auditory icons and earcons: short bursts of sound that alert users of events or situations (e.g. notifications), or provide simple information, such as confirmation of an action.

Alarms: persistent sound signals, that make the user aware of a situation that require immediate attention, such as an incoming phone call.

Tones: persistent sound signals consisting of a modulated monotone or continuous periodic signal. Tones describe the variation in time of a given measure (e.g. heartrate, the temperature of a kettle) by matching a property of the signal, such as frequency or amplitude.

Speech: human auditory communication with a wide range of levels of complexity, depending on the content and context. It has the highest potential throughput, since it communicates information in a finite time that no other category can.

Environment

The context of the user experience is an integral part of the interaction with a hearing augmentation. The individual characteristics of the location where the interaction takes place, the background and mood of the user as well as several other factors play a role in how auditory information is perceived. As we observed in Study II, effective HA is easier to achieve in less noisy environments. On the other hand, HA may be most welcome when noise rates are high as suggested by study I and the wide popularity of noise-cancelling headphones. Consequently, the environmental context of where the HA system is to be used is a key design constraint.

7.11.4 Social properties

The social properties of the HA system are the final design dimension in our design space. The insights in this dimension were built primarily on study I and debriefing interviews in all our studies. We identified three main themes, which correspond with key research topics in HCI: visibility [146, 156], social acceptability [112, 128], and privacy [3].

Visibility

This dimension describes how likely other people are to become aware of someone using an HA system. In studies II-IV, we used a highly visible system which could be identified from a distance. This choice was dictated by the fact that a large-sized microphone was required for the sound fidelity needed for the studies. However, future HA systems could use smaller devices, even ones that may be invisible. In Study I, participants stated explicitly that they would have liked to know if someone else were using an HA system. Future systems could hide or actively show if they are currently active, resembling existing conventions such as the blinking red light on a recording video camera. Whether and how active HA is shown to other users is a key property of an HA system.

Acceptability

HA can be accepted by people to different degrees and for different reasons. Devices with unusual appearances may trigger unwanted reactions or become trendy. People might only feel comfortable using a hearing augmentation alone or at specific locations. Study I showed that users were eager to supply usage scenarios where they found an HA system to be acceptable. However, users in all our studies also remarked on HA technologies being potentially socially disruptive. Thus, understanding in which contexts a new HA can be acceptable is a key consideration for future systems.

Privacy

HA can potentially violate the privacy of other users. Designs that enable eavesdropping or tapping communications pose ethical problems. These potential issues were mentioned by users in Study I. HA can potentially provide advantages over others that could be perceived as unfair. Further, the novelty of such designs makes it hard for third parties to detect, prevent or combat privacy violations. Designers of future HA systems should carefully consider the potential misuse of HAs and discuss the implications. Designing privacy-respecting HA systems emerges as a design challenge.

In general, in this chapter we have explored the potential for HA, thus demonstrating the feasibility of augmentation of a classical sense to provide an enriched experience of reality. We have explored the methodology to design, implement, and evaluate a SA application though an informal User-Centered Design (UCD) process, gaining valuable knowledge about the opportunities and challenges that result from involving users. Further, we have proposed a design space for HA that enables to investigate this subject in a systematic and structured manner and expands the design space for SA presented in Chapter 5.

Chapter 8

Research Probe II: Clairbuoyance

> If you do not change direction, you may end up where you are heading.
>
> — Lao Tzu

In this chapter, we report on a second probe on SA. To explore the possibilities for augmenting non-classical senses, we focus our efforts on the sense of orientation. In the following sections, we present the study of a concrete challenge, namely orientation in open waters.

Having a sense of orientation is important for navigation tasks. This non-classical sense is critical to navigate the world and is the result of complex cognitive processes comprehending vision, audition, touch, memory, and the vestibular sense. Humans have different strategies for finding their orientation and, generally, features in the environment or landmarks play an important role. In contrast to some animals, humans have no sense of magnetic fields that could provide orientation. A particularly difficult task is orientation in open waters as there are no landmarks or static features present. In this chapter, we explore how to create a digital sense of orientation.

Our use case and motivation is open-water swimming, which is gaining in popularity as an amateur sport. Excitement, health benefits and the possibility to connect with nature motivate more and more physically active individuals to swim in the open waters. The largest amateur event in the UK — the Great Swim – attracts 22,000 swimmers annually and the UK Outdoor Swimming Society has grown from 300 to 23,000 members over the last ten years [68].

Yet, despite its many advantages as a sport, swimming is cognitively complex, requiring precise coordination of a large number of muscles. Technique, proprioception, rhythm and stamina all have a key impact on performance. In this environment alien to our bodies, two main challenges arise: breathing and orientation. Breathing is a vital function and all swimming styles incorporate strategies to enable inhaling air effectively, but orientation in the water remains problematic.

Novak [179] showed that humans are incapable of swimming in a straight line over longer distances. The difference in strength between arms and between legs results in a slight deviation in direction when there is no frame of reference available. This effect is typically observed in people walking in circles in the desert or unfamiliar terrains caused by the lack of an external directional reference, thus making the directional recalibration impossible [219].

Consequently, competitive sport swimming pools are divided into lanes with colorful ropes and have guide lines on the bottom to aid the swimmers. This is not the case for lakes, rivers and the sea, where competitions mostly use sparsely positioned buoys. For recreational open-water swimming, reference points need to be identified by the swimmer. For open-water swimmers, their field of view is mostly obstructed by non-transparent water and, during events, by other swimmers. Knowing where to swim depends on raising their heads above the water and trying to find visual references, affecting their swimming rhythm and performance. Depending on the situation this may result in losing a competition, excessive exhaustion, or even drowning. Currently, open-water swimmers can use dedicated devices in the form of GPS-enabled wristwatches. This improves their sense of direction, but still disrupts their swimming. Alternatively, long-distance swimming competitions use boats or kayaks to guide athletes, but these are not feasible for recreational swimming, especially given the constant increase in numbers participating in the sport. As Human-Computer Interaction (HCI) is increasingly interested in understanding interfaces for physical activity [169], the problem of providing directional feedback while swimming provides a relevant challenge and an exemplary case for creating a digital solution.

In this work, we explore the means of enhancing swimmers' sense of orientation in water. We present *Clairbuoyance*, augmented swimming goggles designed to improve the directional perception of swimmers. Our design uses peripheral light feedback to convey information about the swimmer's current direction. We contribute (1) the design and implementation of a prototype capable of providing underwater directional information visually; (2) a proof of concept of the idea and an evaluation of the device that uses two feedback modalities: absolute and relative direction; and (3) insights for designing future systems that support directional guidance, in particular, but not limited to, swimming.

This chapter is based on the following publication:

- F. Kiss, P. W. Woźniak, F. Scheerer, J. Dominiak, A. Romanowski, and A. Schmidt. Clairbuoyance: Improving Directional Perception for Swimmers. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pages 237:1–237:12, New York, NY, USA, 2019. ACM. Honorable Mention Award

8.1 Related Work

A starting point of our inquiry was understanding the different types of feedback through which directional information could be provided to swimmers. We describe how past methods for providing directional feedback relied primarily on haptic output, then look at previously researched means of providing visual feedback in applications that convey space and direction. Finally, we present literature focused on applications for swimmers.

8.1.1 Feedback for navigation

Several researchers explored ways to convey navigation cues using alternatives to traditional displays. Tan and Pentland [226] investigated the use of tactile displays in wearable computing. They presented a wearable tactile directional display which used sensory saltation, a tactile illusion evoked by stimulating different regions of the skin in rapid succession. Their work pioneered the use of non-traditional approaches for conveying navigational information for implicit interaction.

Brewster and Brown [29] introduced the concept of *tactons*, or tactile icons. The authors proposed to use individually identifiable vibration patterns differentiated by their pulse intensity, duration and frequency to convey specific cues and signals to users. Later, Lin et al. [140] proposed *tactons* to provide navigation cues for pedestrians. The authors further reported on two experiments where they investigated the use of tactile feedback to present navigation information to pedestrians. Similarly, Pielot et al. [188, 189] investigated and designed a tactile compass and Kiss et al. [117] presented a wearable system to provide turn-by-turn navigational instructions for motorcycle riders. The wide use of tactile interfaces for navigational feedback shows that conveying direction through non-visual cues may be highly effective. These results inspire our work and prompt exploring the design space for directional feedback while swimming.

8.1.2 Visual feedback for directional cues

Another line of research explored how to design effective visual cues for conveying direction. Burke et al. [36] compared multimodal feedback in terms of error rates and reaction time in a meta-analysis of 43 studies. Their findings suggested that visual-auditory feedback resulted in better performance than visual-tactile for single tasks, while the opposite is true for participants performing multiple tasks in parallel. *Eye-q* [50] used a peripheral display embedded in glasses to provide subtle notifications, with an emphasis on the social acceptability aspect of the design.

A common application scenario for unobtrusive visual feedback for space and direction information is navigation. AmbiGlasses [190] was a system that consisted of a pair of glasses with 12 LEDs used to convey directional information, which was found to be effective. Tseng et al. [234] proposed to use peripheral light for navigation on scooters, finding a range of signals which participants could effectively recognize as commands. Similarly, Matviienko et al. investigated the use of light feedback for turn-by-turn navigation in cars [158]. Their findings suggested that ambient light-based cues are easy to use and understand. Our work builds on these past results by exploring specific considerations for feedback while swimming.

8.1.3 HCI for swimming

Human-Computer Interaction under water poses additional challenges for building interactive systems, both from a technical and a design point of view. Past work in HCI contributed several systems for swimmers. Davey et al. [54, 53] built a system based on a tri-axial accelerometer and developed an algorithm to measure performance for competitive swimmers. They showed that this type of system can provide results equal to or better than manually collected data. Callaway et al. [37] compared video- and sensor-based measuring of swimming performance, showing how electronic sensors enable a more accurate analysis of swimmers' performance and provide useful tools for coaches and trainers, compared to traditional means. We were inspired by these systems as they showed that technical interventions in the swimming sport can be effective and perceived as useful.

Several past works have explored how information may be effectively presented to swimmers while considering the special perceptual conditions in the water. Förster et al. [72] investigated different modalities to provide feedback to swimmers. Their findings suggest that audio feedback is not appropriate for interfaces for swimmers, whereas haptic and visual feedback are effective. The same authors later presented SwimMaster [13], a wearable assistant for swimmers. This system was able to calculate performance metrics of swimmers such as the time to swim a lane, the swimming velocity and the number of strokes per lane, and provide information about other important aspects regarding style-specific factors such as body-balance and rotation. Their findings confirmed the preference for visual and haptic feedback over audio in interfaces for swimmers.

Hagema et al. [81] used waterproof accelerometers to measure the rhythm of strokes from swimmers and provide LED feedback on swimming goggles. This system enabled swimmers to maintain a constant pace and train to a preprogrammed rhythm. A similar idea was presented by Marshal [155], making a smartphone waterproof and using its connectivity capabilities to enable a coach to monitor the swimmer's performance remotely. Mangin et al. [151] designed a wearable distributed system able to collect data about swimming kinematics and transmit it wirelessly to a personal computer. These works in swimming technology guided the design of *Clairbuoyance*, showing that the use of augmented swimming goggles may be effective and that swimmers are able to process a certain amount of additional information while swimming. Our work is interestingly different from these efforts as it focuses on the sense of direction and explores sensory augmentation.

More recent works explored swimming technology for social play and therapy. In SwimTrain, Choi et al. [48] explored the use of exergames for promoting fitness through group fitness activities. This work showed that additional feedback while swimming offered a playful experience. Parvis et al. [186] used waterproof inertial systems to measure the movements of swimmers and assess swimming symmetry during rehabilitation therapy. *Clairbuoyance* was inspired by the research above as it shows that technology can add additional meaningful elements to the swimming experience. In contrast, our work did not aim to create new swimming experiences. Instead, we aimed to design a device that would augment an existing activity.

8.2 Design

The design process of *Clairbuoyance* consisted of several steps and iterations, aimed to address different aspects and problems. Technical challenges such as making our device waterproof required extensive trial-and-error attempts, sometimes even resulting in the partial or complete destruction of a prototype. Choosing our feedback modality was, at first, based on existing literature, but fine-tuning required presenting the design to users, prompting discussion and collecting observations and insights both at the lab and at a local public swimming pool. Informal interviews provided useful feedback and ideas, mostly from other swimmers at the pool.

In this paper we summarize the most relevant aspects of our design process and omit failed attempts, flawed prototypes and annoying feedback modalities. In the following subsections, we present and explain our design decisions for creating *Clairbuoyance*, a system to improve directional perception for swimmers.

8.2.1 Choosing a feedback modality

We chose to use visual feedback over haptic based on Burke's findings since swimming is a single task [36]. We favored visual feedback over sound based on the findings of Förster, Bächlin and Tröster [13, 72].

Given the repeated success in providing visual feedback through augmented glasses, we chose a similar approach for our design [71, 190, 233, 234, 237]. Instead of glasses, we decided to augment swimming goggles, the most ubiquitous swimming gear besides the swimsuit. This decision also minimized the impact of our gadget on usability, portability, social acceptance and comfort. Further, light-based feedback was successfully used before in applications for interactions during physical activity [251].

8.2.2 Defining feedback modes

Based on the design used for ActiveBelt by Tsukada et al.[235], we identified two main types of orientation: absolute and relative. Absolute orientation describes a general awareness of directions, which is what we ideally experience in familiar environments. For example, even when we are not completely sure were the geographic North is, we can intuitively point towards particular places outside our immediate field of view.

In contrast, relative orientation depends on a given direction. This can easily be illustrated by a compass, which calculates directions respect the magnetic North. In this case, the general awareness of directions is not important, but the focus is on a given goal direction, and the important information is by how much are we diverging from it.

These two intrinsically different orientation concepts have different representations from an informational point of view. The absolute sense of orientation is a continuous signal, which can be represented as a single, uninterrupted stream of data. The relative sense of direction has a discrete nature since there are likely only three possible states for a given observer: the observer is facing the desired direction, the observer must turn to the right to face the desired direction or the observer must turn to the left to face the desired direction (we assume here that the likelihood of being completely opposite to the desired direction is close to zero).

Given these differences between the two orientation concepts, we proposed two different representations or *modes*, matching the numerical nature of each type of signal:

Absolute continuous feedback (ACF):

we mapped all directions to the RGB color spectrum, thus each direction was represented by a single RGB color (see Figure 8.1). We arbitrarily assigned red to North, and then pure green to 120° and blue to 240°. The color mapping was inspired by past work in HCI that effectively mapped hue to circular models [38]. We considered the use of a fixed color for the goal, but past work suggested that a fixed color pattern was easy to memorize [38]. Given that goals can be situated anywhere and North is not a preferred direction, a static color mapping can provide a consistent experience. Thus, a continuous spectrum without end was our choice.

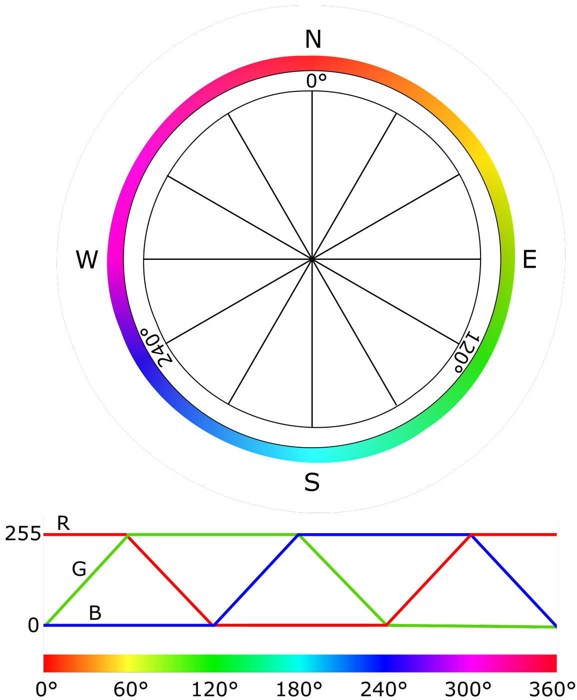


Figure 8.1: Mapping of cardinal directions to the RGB color spectrum.

This representation enabled an intuitive recognition of directions, as well as an observable variation in the signal that might indicate getting closer (or further) to a desired goal.

Relative discrete feedback (RDF):

since only feedback is needed when the swimmer needs to correct his/her direction, we decided to indicate which direction the swimmer must go to correct the course. When the heading is correct, no signal is displayed; when the swimmer deviates from the desired direction, a light indicates where to turn (see Figure 8.2).

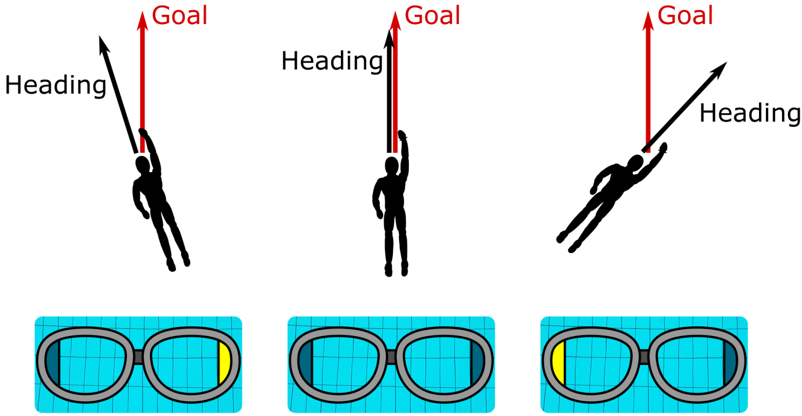


Figure 8.2: Relative discrete feedback mode: a light cue indicates towards which side to correct heading.

8.2.3 Design requirements

A system capable of providing the described above sensory augmentations needs to fulfill multiple requirements. From a functional point of view, the device must be able to calculate directions, both absolute and relative to a desired goal direction. The system must also be able to compensate rotations respective to the X and Y axis, providing consistent feedback while the user looks forward, downwards (while swimming) or rotating the head to the sides for breathing. Additionally, it is desired that the system ignores the small, rhythmic characteristic oscillations around the current direction caused by the swimming movements.

From an interaction perspective, the system must enable the user to switch between feedback modes, as well as select the desired goal direction in the *relative discrete mode*. Additionally, the system must be capable of presenting both feedback modes to the user in the visual periphery. That means that the device is required to provide light on both sides of the field of view and present the whole RGB color spectrum.

For the technical aspects, the system must be waterproof – or at least water-resistant for a few hours to a moderate depth. It also needs to be completely wireless, compact, robust and lightweight, since it would ideally be completely and exclusively attached to the goggles, without presenting any hindrance to the swimmer.

8.2.4 Final prototype

Based on the requirements, the definitive implementation of Clairbuoyance consists of standard swimming goggles augmented with custom made electronics. Two RGB LEDs are attached to the sides of the goggles and provide the visual cues (see Figure 8.3). The orientation is calculated using readings from a digital magnetometer, accelerometer, and gyroscope. All these electronics, plus a microcontroller and batteries, are encased in a waterproof container fixed to the strap of the goggles. A textile band can be added to provide extra stabilization (see Figure 8.4). A single push-button enables the user to switch feedback modes or select the target direction.

The device provides the two feedback modes: the *absolute continuous feedback* (ACF) is conveyed to the user by both LEDs, which display the RGB color mapped to the current heading. The *relative discrete feedback* (RDF) is provided by lighting a single LED on the side towards which the swimmer must turn. When the heading is within a predefined threshold, both LEDs are off. In this mode, the visual feedback is displayed with a yellow light, since this color is easy to notice in most underwater environments.

The user can switch between feedback modes by pressing the button for at least two seconds (long-press). The effect of pressing the button for less than two seconds (short-press) depends on the feedback mode. For the *relative discrete orientation*, a short-press sets the current heading as the goal direction. In this case, both LEDs blink and briefly change colors, to provide feedback about the user's action. From this point on, the feedback will provide information based upon this acquired direction. While providing *absolute continuous* feedback, the short-press has no effect.

8.3 Implementation

In this section, we describe our prototype to ensure the reproducibility of our study.

Figure 8.3: Final version of the prototype: the control button (right), the device working on the ACF mode (right) and a detail of the hardware (bottom).

8.3.1 Hardware

Taking the same approach as Parvil et al., we based our prototype on an Arduino-compatible microcontroller board, the Teensy 2.0[12] [186]. The directional information is collected with a three-axis gyroscope (L3DG20H) combined with a three-axis magnetometer and accelerometer combo (LSM303DLHC). All these sensors are included in the Adafruit 9-DOF IMU breakout board, which we used for our prototype.

The devices are powered with two CR2477 3V coin cell batteries in series, ensuring a constant 5V power supply with an LT1521CST-5 voltage regulator. The visual feedback is displayed using high-brightness common-cathode RGB-LEDs (7000/8000/4000 mcd) on each side of the goggles. To improve the

[12] https://www.pjrc.com/teensy/

Figure 8.4: Final version of the protoype stabilized with a textile strap.

visibility of the light feedback, we used laser-cut acrylic light diffusers. User input was enabled with a 6mm tactile button, attached to the right LED. The button was positioned to ensure ease of use and prevent unintentional actuation.

LEDs and a button were connected to the microcontroller with standard four pairs and six pairs cable. The LEDs and the button were waterproofed with transparent heat-shrink tubing and hot-glue to enable visibility and actuation while guaranteeing the integrity of the device. The controller, sensors and batteries were encased in an IP68 ingress-protected junction box, intended for outdoor installations. This method allows for easy access to the electronics for switching the device on, replacing batteries or reprogramming the board, while keeping it

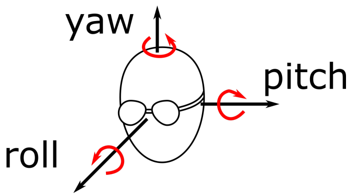


Figure 8.5: Aeronautical naming of the Tait-Bryan angles, in this case applied to the human head.

dry during its usage. To protect the electronics from condensation, we used silica bags.

8.3.2 Software

The software was written in the Arduino IDE using the Teensyduino plugin. It calculates the current heading as a unidimensional value, which is the projection of the measured magnetic vector on the horizontal plane. The horizontal plane is calculated using the accelerometer and gyroscope readings, under the assumption that the time-averaged acceleration exerted by the swimmer is negligible when compared with the gravitational pull.

Using the aeronautical angle nomenclature, where *yaw* is the heading, *pitch* the elevation and *roll* the lateral rotation (see Figure 8.5), the current orientation of the user with respect to the horizontal plane can be calculated with the following equations:

$$pitch = \arctan \frac{A_y}{\sqrt{A_x^2 + A_z^2}}, \quad roll = \arctan \frac{A_x}{\sqrt{A_y^2 + A_z^2}} \tag{8.1}$$

where A_x, A_y and A_z are the normalized values of acceleration for the three respective Cartesian axes of the sensors' local coordinate system. The calculated $pitch$ and $roll$ are angles in radians, with respect to the horizontal plane. The magnetic readings M_x, M_y and M_z are also normalized and projected on the calculated plane:

$$H_x = M_x \cdot \cos roll - M_z \cdot \sin roll \tag{8.2}$$

$$H_y = M_y \cdot \cos pitch + M_y \cdot \sin roll \cdot \sin pitch + M_z \cdot \cos roll \cdot \sin pitch \quad (8.3)$$

$$yaw = \arctan \frac{H_x}{H_y} \quad (8.4)$$

The obtained *yaw* is then the angular difference between the magnetic North and the current heading of the user. This value is converted to degrees and smoothed using an averaging filter with length 10.

The heading value are used either directly for the *absolute continuous orientation* feedback, or compared to the value stored by the user for the *relative discrete orientation* feedback.

8.4 Evaluation

In order to evaluate *Clairbuoyance*, we conducted a within-subject controlled experiment in an Olympic-sized swimming pool. We wanted to evaluate how both feedback modes performed in terms of performance and usability. Participants were asked to swim to a series of targets across the pool while using *Clairbuoyance* in one of its two modes and without additional aid.

8.4.1 Participants

We used social media and snowball sampling to recruit participants. We distributed the experiment call both on general channels (university mailing lists) and groups specific to swimming in order to get a spectrum of novice and advanced participants. Potential participants were asked to declare that they could swim a quarter-mile front crawl with interruptions as requested. Additionally, we required that participants should have had normal eyesight (or corrected to normal while swimming). We recruited 24 participants (16 male and 8 female), aged from 18 to 62 years old ($M = 26.54$, $SD = 10.93$). Participants stated that they swam an average of 7 times a month ($SD = 9.00$), with some swimming more than 5 times a week. We classified them into two groups: recreational and advanced swimmers. Advanced swimmers identified as one of the following categories: active lifeguard, active competitive swimmer, swimming coach, former competitive swimmer, club water polo player. There were 12 advanced and 12 recreational swimmers in our sample. Participants received USD 15 as remuneration for their time spent in the study. Additionally, isotonic drink and food was available for recovery after the study.

8.4.2 Apparatus

The participants swam towards a target on the opposite side of the pool. The target consisted of a yellow semi-circle, clearly visible on the other side of the pool (see Figure 8.7). To impede participants from using the lanes on the bottom of the pool as an orientation reference, we created paths for them to follow, as illustrated on Figure 8.6. A base path consisting of 6 straight segments, with a total length of 175 meters, was used to create three symmetric paths by inversion and rotation. This removes learning bias, since each path is perceived differently by the participant, and, still, the length and angle respective to the side of the pool of the individual segments remains equal. Counterbalancing using Latin squares was applied to the conditions and routes. To avoid confusion about the target, there was only one target on each side of the pool, which was relocated for each segment to predefined positions, marked with tape on the floor.

Through an iterative process, we produced a system robust enough to endure the study conditions. Waterproofing was the main challenge, but the effect of gyroscopic drift also required attention. Given that we expected that extensive exposure to water pressure under experimental conditions would produce enough stress on the sealing of the case to compromise the electronics, we produced three physical prototypes to be able to continue the experiment in case of damage. This proved useful since it allowed replacing batteries preventively without interrupting participation. To eliminate possible magnetic drift effects, the prototype was restarted between trials, which kept the effect negligible. Calibration was performed in situ, to account both for geographic magnetic deviation and environmental aberrations. To calibrate each individual prototype we used MotionCal[13], a software system specifically designed for this purpose.

8.4.3 Hypotheses

Using this controlled experiment design, we evaluated the following hypotheses:

[13] https://github.com/PaulStoffregen/MotionCal

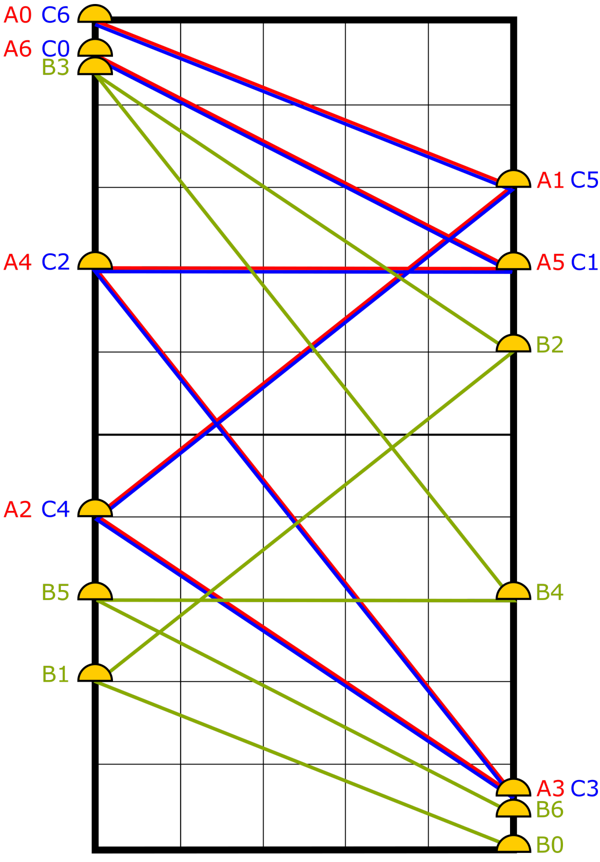


Figure 8.6: Path A (red), Path B (green) and Path C (blue) were assigned in a counter-balanced fashion to the three conditions (base, ACF and RDF). Path targets were ordered increasingly, starting at 0. The dimensions of an Olympic swimming pool are 25 meters by 50 meters, thus each path had a length of approximately 175 meters.

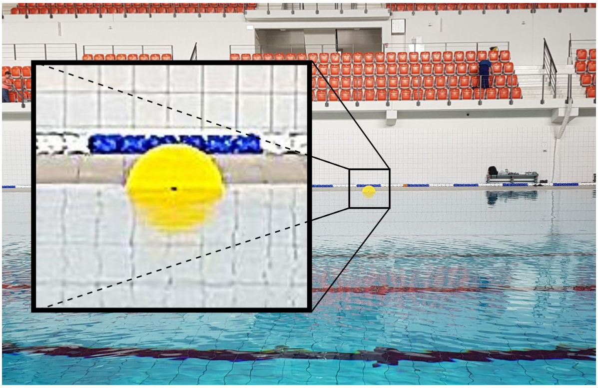


Figure 8.7: Target, consisting of a yellow semi-circle with a black center.

1. Peripheral visual feedback will reduce completion time
2. Peripheral visual feedback will reduce orientation errors

8.4.4 Conditions and measures

To evaluate *Clairbuoyance*, we asked participants to swim three times, once under each of the three conditions:

- BASE condition: no feedback
- RDF: Relative Discrete Feedback
- ACF: Absolute Continuous Feedback

We collected four different metrics: Task Completion Time (TCT), error distance, task load and a score describing the usability of the system.

TCT was measured with a chronometer, and consisted of the swimming time of each participant to complete each path. Thus, TCT is the sum of the measured time to reach each target and does not include the time the participants rested.

The error distance was measured as the separation between the point where the participants first touched the side of the pool (estimated goal) and the center of the yellow target (real goal). We use the sum of the absolute value of all error distances as a metric.

Task load was measured using the NASA Task-Load Index (TLX), an assessment tool to quantify and analyze the workload required to complete a task. For this, after each path was completed, participants were asked to fill a questionnaire, providing subjective feedback on six subscales [87, 88].

The usability of the system was measured using the System Usability Scale (SUS). For this purpose, participants were asked to fill the SUS questionnaires for each of the feedback methods [14].

Individual opinion on *Clairbuoyance* and the whole proposed interaction was collected in semi-structured interviews.

8.4.5 Procedure

We welcomed participants in a room adjacent to the swimming pool to brief them before the experiment. Participants were asked for their consent for participating in the experiment and processing the data gathered during the study. After explaining the experimental task in detail, we asked them to confirm in writing that they were fit to complete the task. They were specifically instructed to swim at a moderate pace and take breaks as needed in order to complete the required number of trials without perceived exhaustion. Next, we gathered demographic data and questioned them about their perceived swimming ability. We then gave them time to change to their swimming attire.

When ready to swim, each participant was assigned a starting condition and positioned at the starting point for the assigned path. An experimenter then positioned the target on the opposite side of the pool at the point specified by the target sequence (see Figure 8.6). We then reminded the participant to swim to the target at a moderate pace and not to look ahead, as well as use the same swimming technique for all paths. The participant was asked to raise an arm when ready, and an experimenter blew a whistle to start both the swimming and the chronometer at the target location. When the participant touched the border on the other side of the pool, the chronometer was stopped and the position respective to the target was noted. The participant was given time to rest as required, and then followed the same procedure towards the next target of the path. After completing all trials in a condition, we offered the participant an isotonic drink. After completing each path and while recovering, the participant completed a questionnaire containing the NASA TLX measures and SUS.

After completing all trials, we let the participant rest and change back to their regular clothes. Next, we debriefed the participant in a room adjacent to the pool, and conducted a semi-structured interview that focused on the experience of using *Clairbuoyance*. We asked which method the participant preferred and what strategies were employed to complete the task. We also explored usage scenarios beyond the swimming pool and possible use of the device in open waters.

8.5 Results

We collected data about two quantitative metrics, the total error respect to the individual targets (in meters), and the total time required to complete the task excluding the rest pauses (in seconds). Additionally, participants reported the perceived load of each task using TLX, and the usability of both feedback methods using the SUS.

The normalized error, normalized TCT and TLX were analyzed using ANOVAs, with the most relevant results summarized on Table 8.1. We decided to standardize the TCT and Error results per participant in order to account for the differences in swimming proficiency and physical fitness between participants. The values were standardized as normality assumptions (Shapiro-Wilk test) have been fulfilled.

Post-hoc Tukey HSD tests showed that participants missed the target by a significantly higher distance when using RDF than in the other two conditions, both times at a significance level of .001. We found no significant difference between the base condition and ACF. For the TCT, post-hoc Tukey HSD revealed that the base condition took significantly longer to complete than when using RDF, with significance level at $p = .05$, while the two other comparisons were not significant.

The mean score of the SUS was calculated as 70.72 for RDF and 71.57 for ACF. A Wilcoxon test showed that the difference in SUS scores was not significant.

8.5.1 Interviews

All pre- and post-study interviews were recorded and transcribed verbatim. We collected a total of 3hr 24mins of recordings. First, we coded the interviews to determine their preferred feedback type (each participant was explicitly asked to make a binary choice). We then printed the statements from the interviews on post-it notes and used affinity diagramming to classify them into common thematic

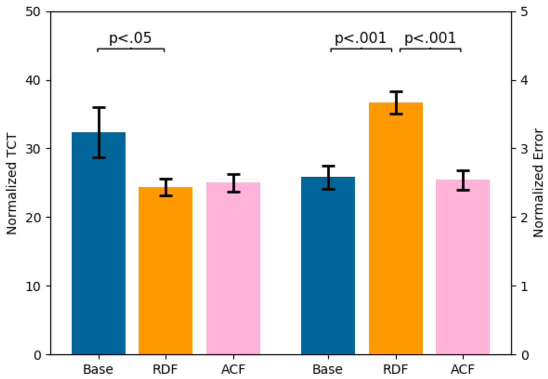


Figure 8.8: Mean and standard error of the normalized TCT and distance with respect to the real and estimated position of target reached by the participants, for each condition. Markings indicate post-hoc significance.

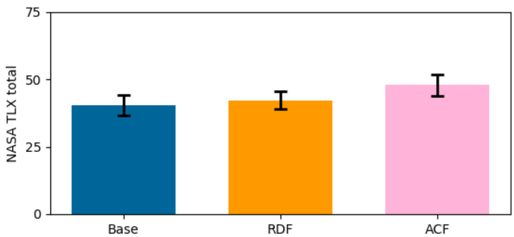


Figure 8.9: Mean value and standard error Total Load, according to NASA Task-Load Index, for each condition.

groups. We identified three themes in the data: WORKOUT INTEGRATION, TRUSTING THE DEVICE and USAGE SCENARIOS.

Preferred feedback type

Fifteen out of 24 participants preferred RDF. Interestingly, advanced users had a very high preference for RDF (11 out of 12 advanced swimmers), while recreational swimmers preferred ACF (8 out of 12 participants). Advanced swimmers were focused on minimizing distractions and appreciated the fact that

		Base	*RDF*	*ACF*	ANOVA
Error [m]	*M*	2.58	3.67	2.54	$F_{2,393} = 3.47$
	SE	0.17	0.16	0.14	$p < 0.05$
TCT [s]	*M*	32.36	24.37	25.01	$F_{2,393} = 16.25$
	SE	3.66	1.18	1.26	$p < 0.001$
NASA TLX	*M*	40.42	42.25	47.96	$F_{2,393} = 16.25$
	SE	3.87	3.27	3.95	$p = 0.32$

Table 8.1: Error rate, Task Competon Time (TCT) and Nasa TLX statistical analysis for the three evaluated conditions.

no feedback was produced in RDF when they were on the correct heading. One swimmer did not want to think about translating colors into directions:

> We swimmers focus on completing the distance fast and we don't want to think how far in we are or how much is left. (...) It's easy to mix up the colors and forget which color to follow. I saw pink most of the time and I'm not used to that. It was distracting.

In contrast, amateur swimmers valued ADF as it offered stronger feedback with constant confirmation of status. One participant was concerned that light on one side may not have been visible enough at all times:

> The colors were better and I reflected that I enjoyed that the colors carried a meaning. When the sun came out, I could still see them. So, I think the colors would be nice irrespective of what kind of water you're swimming in.

Workout integration

Participants were eager to share insights on how *Clairbuoyance* could be improved to better reflect their swimming practices and fit their workout routine. One participant reflected that they would choose a different goggle model for the task.

> I wanted to make sure the light was in my field of view and I wasn't always sure of that. I constantly thought there might have been more light that I didn't see. I guess I would need different goggles.

In the ADF condition, an often-repeated remark was the need to customize the sensitivity of the device. A dynamic directional range that would change as the user was approaching the target was suggested:

> You should be able to adjust the range of the device at the starting point. And then this range should become narrower as you get closer to the target. That would make sense.

Finally, some users remarked that the device could be integrated in their regular swimming routine and produce a positive experience by alerting them when a course correction was needed. This would amount for more time to relax and thus a more enjoyable swim:

> I liked it because the diodes told you when you were off course. It's intuitive and easily visible. You have to react to it when needed, but, most of the time, you can just enjoy the swim.

Trusting the device

Another aspect often addressed by the participants which emerged from the data was how users decided (or not) to trust the device and navigate based on *Clairbuoyance*'s indications. One participant was initially doubtful about the system, but they later decided that the feedback was useful:

> There was a moment when I didn't trust the system, I thought it broke down. Had I trusted it, I would have hit the target. So, then, I decided to trust it and ended up in the right spot. Then, it started making sense to me.

Many participants reported that they needed to think less about their direction of swimming. They saw *Clairbuoyance* as an opportunity to focus more on swimming technique or relax. One swimmer remarked that the device enabled him to relax while swimming:

> You can sort of stop thinking. You don't have to focus on where you are swimming, but you can freely move and the lights give you hints. You can just relax, just swim.

Usage scenarios

Finally, users explored possible uses of *Clairbuoyance* beyond the study extensively, considering the functionality of the device in open waters. Participants would propose usage scenarios for the system. One of the advanced swimmers suggested that *Clairbuoyance* would be useful even for complex tracks often seen in open-water swimming competitions.

> The longer the distance, the more useful it is. It also depends on the conditions... in open waters, there's the sun, the color of the water (...) when the sun blinds you, you could trust the device to navigate for you. Even for multiple buoys, this could work.

Another suggested scenario in open waters was not only making sure one was following the right heading, but also establishing the correct heading. An advanced participant commented that the system could eliminate the need to reassure oneself that one had not forgotten the swimming direction in a competition:

> I know that it will lock that point and stick to it. I would use it a lot. In open waters, the legs could be two kilometers or more, so you can't even see the target point when swimming. Sometimes, there are intermediate buoys, but it's usually not enough. It would help me keep reach the goal.

8.6 Discussion

Having explored augmented orientation for swimmers, we observed that *Clairbuoyance* offered benefits to the users that were observed both in quantitative data and qualitative feedback. Below, we summarize our findings and outline challenges and opportunities for future systems.

8.6.1 An augmented sense of direction reduced TCT

Swimmers using *Clairbuoyance* needed less time to complete the path than on the base condition, confirming the *Hypothesis 1*. Given that the normalized error for ACF and the base condition is approximately the same and that we found no correlation between the Error and TCT, the reduction in TCT can be interpreted as an increment in swimming speed for participants using *Clairbuoyance*.

Psychountaki and Zervas found a correlation between trust and performance, in particular speed [193]. This, combined with the feedback collected in the interviews with participants of the study, supports the idea that *Clairbuoyance* had a positive effect on the confidence that swimmers had on their sense of direction, resulting in an increment in their performance.

8.6.2 *Clairbuoyance* did not improve accuracy

The use of *Clairbuoyance* did not result in a more accurate estimation of the direction to the target. Swimmers missed the target by approximately the same distance both on the base condition and with the ACF, and a larger distance when using RDF. This suggests *Hypothesis 2* is false.

We observed that participants swimming on the base condition did not respect the request to avoid peeking and most of them raised their heads to locate the target visually. Additionally, the tiling of the bottom of the pool and the marking of the lanes provided some reference for orientation. However, because this was present in all conditions, it is possible to exclude its effect in the performance, although acknowledging that these issues weaken the comparison among the feedback methods and base condition. Despite this, it is logical to assume that the base condition error is the minimum error to be expected, and thus the ACF shows at least good performance. This, in addition to the participants reporting no issues with the direction of color transitions or hue choice, suggests that the chosen color mapping was not a hindrance.

The difference in performance between RDF and ACF can be explained as the first method giving feedback only beyond a threshold of 5°, an effect not present for the absolute feedback. It would be interesting to determine the relationship between the threshold and the error distance.

8.6.3 Advanced and novice users valued different features in the prototype

We observed differences between advanced and recreational swimmers in the qualitative feedback gathered. Differences in requirements between novice and professional users in sports are a known phenomenon in HCI for sports [123]. Our work showed that advanced users preferred RDF which offered a more holistic experience, while novices opted for ACF, which made users feel more in control. As we observed that advanced users wanted to control the dynamic range of RDF, our results resonate past work that suggested that professional users require more fine-tuned controls [229]. This shows that the design requirements for swimming applications for advanced swimmers are complex as they need to combine being unobtrusive (which is often achieved by minimizing input) with a large degree of control. On the other hand, our work suggests that constant visibility of system status is an important requirement for novice swimmers.

8.6.4 Limitations

Conducting a proof-of-concept study, we were forced to make a series of compromises to conduct a controlled experiment. The study was conducted in a swimming pool, which even given the experiment design, possesses visual characteristics and features that facilitate orientation. We expect that a real-life open waters scenario will completely lack visual cues. We observed that large structures of steel were present in the immediate surroundings of the pool. Ferromagnetic materials in large concentrations have a disruptive effect in the behavior of magnetometers, which might also suggest that the performance of the prototype in open waters would be better.

Our choice of the study design was primarily dictated by practical, ethical and liability considerations. While we recognize that evaluating the prototype in open water would have offered more ecological validity, there was no way to ensure the safety of the participants in a lake or sea. Conducting the experiment in a swimming pool enabled us to hire a dedicated lifeguard and fulfill the ethical standards required by our institutions. A future viable alternative would be evaluating *Clairbuoyance* in a competitive swim, among its participants. However, as we did not know if and how the device affected performance, we could not request that participants jeopardize their results.

8.7 Conclusion

In this article, we presented *Clairbuoyance*, a system that provides visual peripheral feedback about orientation, augmenting the non-classical sense of direction.

We described the design considerations behind the system, its implementation, and evaluation. Based on the collected data, we found that visual orientation feedback reduced TCT.

We also discovered a preference for *relative discrete feedback* among more proficient swimmers, while less experienced swimmers found *absolute continuous feedback* more useful and attractive.

In future work, we expect to refine the *relative discrete feedback* and evaluate *Clairbuoyance* in a controlled experiment in open waters.

On a more general level, this work assesses the potential for augmentation of non-classical senses using a classical sense as the information channel. Using the classification from the design space proposed in Chapter 5, it is possible to describe *Clairbuoyance* as a system that *enhances* the sense of direction through *vision*. This highlights the versatility of the design space and suggests its adequacy for exploring the augmentation of both classical and non-classical senses. Further, our research highlights the need for SA systems to be flexible and versatile to address the needs of different users.

Chapter 9

Research Probe III: VUM

> Where the telescope ends, the microscope begins. Which of the two has the grander view?
>
> Victor Hugo

The interaction with SA application presents particular challenges. Following the vision presented in Chapter 6 requires breaking the ground for new perspectives on interaction design. Without attempting to present new interaction elements, we conducted a research probe on this topic, aiming to assess how current interaction techniques compare in the context of SA. In this chapter, we report on the experimental evaluation of different techniques of interaction and display for a Virtual Ubiquitous Microscope (VUM), an augmentation system capable of magnification beyond natural vision.

9.1 Introduction

Humans have always used technology to enhance their sensory perception. Starting with simple monocles, today's modern technologies enable extraordinary abilities to enhance human vision, such as seeing beyond the naturally visible spectrum of light [1], or decelerating the speed of what we visually perceive [124].

We expect this trend to continue, offering ever-increasing levels of augmentation and empowering users to perceive much more than typically possible.

This chapter is based on the following publication:

- F. Kiss, P. W. Woźniak, V. Biener, P. Knierim, and A. Schmidt. VUM: Understanding Requirements for a Virtual Ubiquitous Microscope. In *19th International Conference on Mobile and Ubiquitous Multimedia*, New York, NY, USA, 2020. ACM

The magnification of vision is an established human need. Historical sources report on ancient attempts of overcoming the natural limits of vision through the use of lenses [217]. Since the 17th century, the use of microscopes, and for a few decades electronic microscopes, has allowed humans to perceive objects several orders of magnitude smaller than what the naked eye can see. Scientists, and to some limited extent the general population, have used these tools to great benefit. Ubiquitous access to different levels of visual magnification can be useful in many situations, both for technical and more trivial activities. Humans interact with small objects on a daily basis. Being able to take a closer look at small details can have a positive effect on a broad range of activities.

With the advent of ubiquitous and wearable computing, the incorporation of microscopy to Augmented Reality (AR) becomes foreseeable, or at least feasible [232]. See-through Head-Mounted Displays (HMD) offer the novel opportunity of presenting visual information ubiquitously, dynamically, and on-demand. They can enable users to magnify objects in their immediate surroundings at will. This kind of augmentation is aligned with Schmidt's vision of seamless integration of sensory amplification and enhancement into users' lives, with users perceiving the amplification not as a tool, but as a direct augmentation of their perception [208]. This kind of interaction, ubiquitous and seamless, poses a series of challenges quite different from those posed by traditional *ubicomp* interfaces. Screens with icons, text, or diagrams become an obstacle when the user only wishes to see better. Traditional interaction metaphors become rapidly cumbersome when they interrupt what we usually do intuitively in a highly automated way. Augmenting the human senses should make perception *easier* and *better*, so controlling augmented vision should not be more complicated than controlling natural vision, nor require more effort. From an HCI perspective, the study of a Virtual Ubiquitous Microscope (VUM) offers an exiting probe into sensory augmentation and an opportunity to gain a better understanding of how to design interfaces for this emerging field.

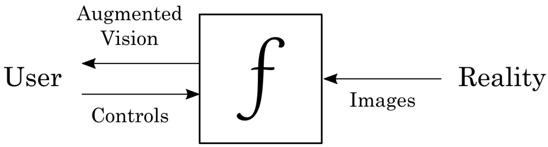


Figure 9.1: Functional model of the interaction with a Virtual Ubiquitous Microscope: the system receives input from the user in form of controls and from reality in form of images. The output of the system is the magnified image.

In this chapter, we conduct an exploratory probe in the design of interfaces for sensory augmentation using a VUM as a functional example. Recognizing the difficulty of an abrupt shift in interaction paradigms, we take an incremental approach, assessing which existing methods are perceived more favorably by users. For this purpose, we designed an AR system for the ubiquitous visualization of magnified imagery and conducted a user study comparing different input and output techniques. Admittedly, the development of a VUM presents an additional challenge that will not be addressed in this work, namely the technical implementations of sensors that enable real-time magnification. Our contribution is an interaction design for a VUM and the evaluation of four different interaction techniques and two different display techniques for such design. In the next section, we review past work and identify the critical aspects of the design for a VUM. Next, we present our design and describe the proposed interaction and display techniques. We follow by reporting on the experiment design and discuss its results. This chapter concludes with a description of the limitations of this work and a conclusion.

9.2 Related Work

Sensory augmentation can be modeled as a mediation of human perception [152, 199]. In the case of visual amplification, such mediation would consist of a filter affecting the amount of detail and field-of-view a user can see for a given object. For the particular case of a VUM, it is thus possible to model the interactive system as a function that receives as input the images to be magnified and the user controls, and returns the magnified image (see Figure 9.1).

To our knowledge, there is no previous work in the literature that investigates the challenges of interaction design for virtual microscopes. However, the visualization of magnified images has been subject of HCI research, and previous work in this area provides a basis for our work, both for the input and output characteristics of the system.

The dynamic display of magnified images has multiple aspects that require attention. The transition between normal to magnified vision, as well as between levels of magnification, can be performed in a continuous manner, or by incremental discrete intervals. Work by Bartram et al. suggests that continuous transitions benefits the understanding of information arranged hierarchically [15] while research by Chen et al. [46] and Pan et al. [182] provide support for its technical feasibility and highlight its relevance.

The usage of optical-see-through HMD offers a versatile platform for the augmentation of vision but prompts a series of design decisions. It is necessary to opt between displaying the images using the complete field-of-view supported by the device or to present the user with a delimited area displaying the image. Further, the images must be positioned respect a coordinate system that can either be the real world, or the centered in the user. This has significant implications for the behavior of the interaction when the user moves. There is no specific research for the case of virtual microscopes, but it is possible to gain some insights from past research on map representation and navigation. Previous work focuses on both the use of floating frames (lenses) to display delimited regions of magnified images [141], and use of full field-of-view for immersive zooming interfaces [170]. Both approaches present benefits and disadvantages and thus it is unclear which would result in better user experience for the particular case of a VUM.

Another aspect of the interaction is *control*. Controlling magnified views presents singular challenges, both due to the difference between the granularity of movements and different levels of magnification [11]. Past work proposes the control of zooming interfaces with tangible controllers [27, 163], and one-handed or two-handed gestures [75, 187].

Satriadi et al. explored the interaction space for AR and VR using gestures and handheld controllers [204]. Through user studies, the authors compared different control possibilities for maps in different formats, including flat maps, curved ones, and globes. The central contribution of this work resides in two techniques to investigate hybrid input for mid-air gestures. Although this work was conducted around a clearly different application, namely map navigation, the general approach used by the authors presents many inspiring elements. Dünser and Billinghurst recognized a number of challenges in the evaluation of AR systems and applications, and examined common evaluation techniques used in user studies [61]. In this work, we used the methodology proposed by Dünser and Billinghurst.

Past work highlights the necessity for a better understanding of input and output methods for sensory augmentation. Thus, our contribution to this area of research consists on a research probe aligned with the goals and methods of the literature and applied to the concrete case of a virtual ubiquitous microscope. Further, we provide general insights for interaction designs in line with the emerging new paradigm of sensory augmentation.

9.3 Virtual Ubiquitous Microscope

To assess the interaction challenges of ubiquitous augmentation, we designed, implemented, and evaluated a prototype of a VUM. We emphasize that the main contribution of our work is not the design itself, but the evaluation of different interaction methods, with the VUM serving only as a vehicle for our study.

A VUM has a defined functionality: when activated, it shows the user a magnified view of a particular object of interest. Further, the user should be able to control the level of magnification, or terminate the interaction and return to normal vision.

To simplify the interaction we defined two design constraints: First, the interaction must be started only intentionally by the user and thus avoiding accidental activation is necessary. Second, the operation of the system must in no way hinder the performance of the task in which the system aids the user. Consequently, the activation and termination, as well as the control of the level of magnification, need to be performed with simple, low effort actions.

Traditional microscopes do not offer the possibility of panning, even if such functionality would be desirable [11]. Given the level of magnification of a microscope, navigation is difficult to control, since even very small displacements result in the object of observation disappearing from the field of view of the device. Thus, the VUM avoids the problem of panning by presenting the magnification of the static image recorded at the beginning of the interaction.

These design choices allowed us to focus on the principal goals of our investigation: the INTERACTION and DISPLAY TECHNIQUES for VUM. Based on past work, we proposed four INTERACTION TECHNIQUES and two DISPLAY TECHNIQUES, which are exemplary of typical approaches in AR and VR applications. This comparison between different proposed methods was recommended by Dünser and Billinghurst for novel AR applications, given the lack of well-established frames of reference for such evaluations [61].

9.3.1 Interaction Techniques

We selected four INTERACTION TECHNIQUES which represent current approaches in interacting with AR: one- and two-handed gestures, a tangible input device, and voice commands. The first three are derived from previous work [27, 75, 187]. We also included voice commands since they efficiency over gestures for some particular cases led us to consider this technique a likely INTERACTION TECHNIQUE in future AR interaction designs [**?**].

We chose the specific gesture for each INTERACTION TECHNIQUE based on the approaches used in the discussed related work while prioritizing keeping the operation of the device efficient.

We defined the interaction with a VUM as having three stages: first, the VUM is activated, and the interaction starts. The next stage is viewing of the magnified object while controlling the zoom level. Finally, the interaction can be terminated by zooming out beyond the minimal level of magnification.

One-Handed Mid-Air Gestures

The one-handed mid-air gesture aims to resemble the movement a person would perform to move an object closer or further away from the eyes, to examine it in more detail or gain a more general overview. The user starts the interaction by raising the index finger and pressing down the finger towards the thumb. While keeping the fingers together, the user can move the hand along the z-axis, which means closer or further away from the body, to change the magnification level. This INTERACTION TECHNIQUE is illustrated in Figure 9.2. This particular hand gesture was selected due to the intuitive and controllable relationship between the performed distance and the level of zoom. Additionally, this gesture is commonly used in AR platforms, such as the Microsoft Hololens[14].

Two-Handed Mid-Air Gestures

Two-handed mid-air gestures offer a more extensive design space for interaction. However, performing two-handed gestures adds extra effort and the impossibility

[14] `https://docs.microsoft.com/en-us/windows/mixed-reality/gestures`

[t]

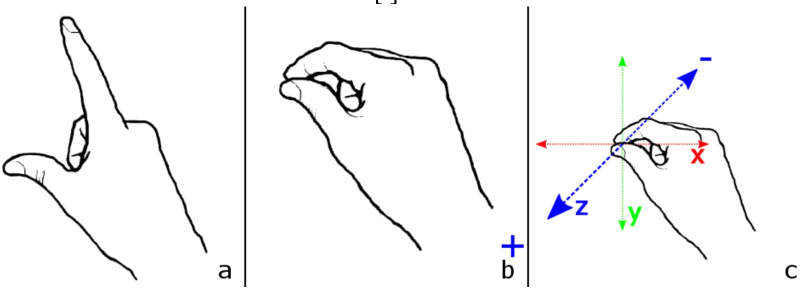


Figure 9.2: One-handed gesture

[t]

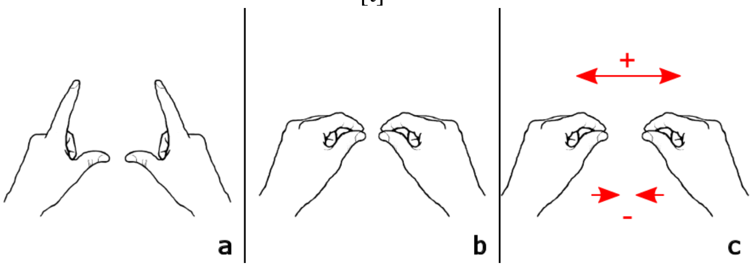


Figure 9.3: Two-handed gesture

[t]

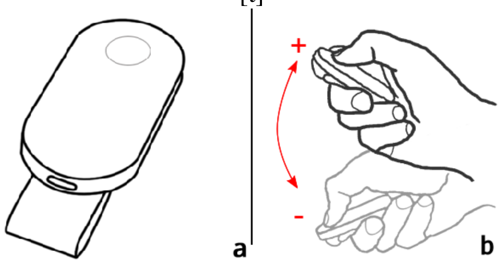


Figure 9.4: Physical controller

Figure 9.5: Four INTERACTION TECHNIQUES were evaluated to control the VUM: (a) one-handed gesture, (b) two-handed gesture, (c) a physical controller, and voice commands (not depicted).

of performing a parallel task with the other hand. Chaconas and Höllerer [42] compared different two-handed gesture sets for rotation and scaling and comparing them to a one-handed gesture set. Based on their findings, we choose the two-handed gesture to manipulate the magnification level of the AR microscope. Similarly to the one-handed gesture, users must raise both index fingers and press subsequently against the thumb to start the interaction. While holding the fingers together, the user can move the hands apart or together in order to modify the magnification level of the virtual microscope. This INTERACTION TECHNIQUE is illustrated in Figure 9.3. This gesture was inspired by previous work [75] and, similarly to the one-handed gesture, due to the relationship between the separation of the hands and the level of magnification. Further, it mimics a commonplace conversational gesture used to emphasize size.

Physical Controller

A physical controller represents the ability of the user to control a system and gives the user direct haptic feedback for command inputs. Further, this INTERACTION TECHNIQUE is more flexible respect its posture requirements and users can operate the system without keeping their arms mid-air, thus reducing physical fatigue. To change the magnification level, users press a button and rotate the tip of the controller up or down. The controller movement is illustrated in Figure 9.4. For this choice of gesture we consciously avoided using a traditional mouse to remain consistent with the other mid-air gestures. The particular choice of gesture considers typically tangible controllers, which commonly feature accelerometers and gyroscopes. Such sensors can readily measure rotational movements yet cannot quantify linear distances with precision.

Voice Command

As the last INTERACTION TECHNIQUE, we enable users to operate our prototype through voice commands. The user can command the system with speech to change the magnification level. By saying “smaller”, users can lower the magnification level, while “bigger” increases the magnification. The exact increment and decrement of magnification levels depend on the implementation and are discussed in the Apparatus section.

9.3.2 Display Techniques

We designed two different Display Techniques to show the magnified content to the user in line with previous work [141, 170]. On one hand, we developed the *head-lock* visualization. For this technique, the magnified content covers the entire available display space. On the other hand, the *tag-along* visualization displays the magnified content in a rectangular window floating in mid-space and positioned on top of the observed object. Although further alternative techniques would be possible, this selection reflects the contrasting approaches of augmenting the *observer* and augmenting the *observed*. This dichotomy is emphasized by the first technique completely immersing the user in the interaction, while the second technique limits the interaction interface in size and position. Further, these two Display Techniques are consistent with our exclusion of panning and navigation. Given the difficulty in finding a compromise between the two discussed Display Techniques, their comparison can yield useful insights for future design. In the following paragraphs, we explain each technique in detail.

Head-Lock Visualization

With the *head-lock* technique, the entire available display-space offered by the AR HMD is used to visualize the magnified content, frequently called *full-screen*. When the VUM is activated, the complete field of view of the user is filled by the magnified image. This visualization is locked to the user's head and follows the user's movements anywhere. Hence, we avoid forcing the user to remain completely still during the observations. The user can navigate through the different levels of magnification with the designated Interaction Technique and move his head around freely without altering the focus of magnification (see Figure 9.6, left). The advantage of this technique is the large display area, enabling a wider oversight of the magnified image. However, this mechanism hinders the user from diverging attention from the magnified image and impedes moving around safely. This translates to the inability to perform other tasks while zooming into things.

Tag-Along Visualization

In contrast to *head-lock*, the *tag-along* technique for visualization has a small virtual display area that is attached to the position of the observed object. When the user activates the ubiquitous microscope and zooms into an object, a postcard-sized virtual display appears in front of that object, showing its microscopic magnification (see Figure 9.6, right). The virtual display remains stationary in space and does not follow the users' movements. Therefore, the user can look in another direction, and the magnified image will stay where it was and save its current status. This allows the user to magnify multiple objects in parallel and switch between them, regaining control of a magnified view by focusing the gaze on it. In consequence, each instance of the magnification view should be terminated individually.

The main advantage of this display technique is the freedom of the user to control and compare multiple magnification views of objects, while also being able to interact with the real world. This comes at the price of a less immersive experience and a limited display area, making the observation of smaller details or the comparison of features within a particular level of magnification less effective.

9.4 Evaluation

We evaluated the performance and user preference of our design using a functional prototype. Further, we compared the different INTERACTION and DISPLAY TECHNIQUES in terms of these two aspects. In the following, we describe our methodology and experimental design, and present the collected data.

9.4.1 Methodology

We used a 4x2 within-subject experiment design with the independent variables INTERACTION TECHNIQUE and DISPLAY TECHNIQUE. For INTERACTION TECHNIQUES, we had four levels: *one-handed*, *two-handed*, *controller*, and *voice*. For DISPLAY TECHNIQUES we had two levels, *head-lock* and *tag-along*. The performance was measured while participants interacted with the prototype to solve a visual search task across different levels of magnification. We recorded the task completion time (TCT), training time, and error rate, as well as task workload [88], and personal preference once each task was finished.

9.4.2 Apparatus

For this experimental study, eight different versions of the prototype were implemented; one for each combination of INTERACTION and DISPLAY TECH-

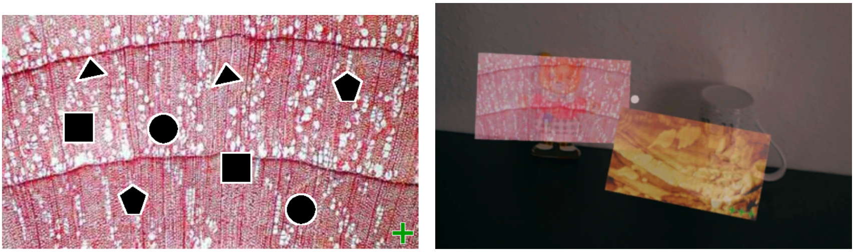


Figure 9.6: Right: The task is counting the squares across all levels of magnification (here for the *head-lock* Display Technique, showing the first level of magnification). Left:Floating frames of the *Tag-along* Display Technique.

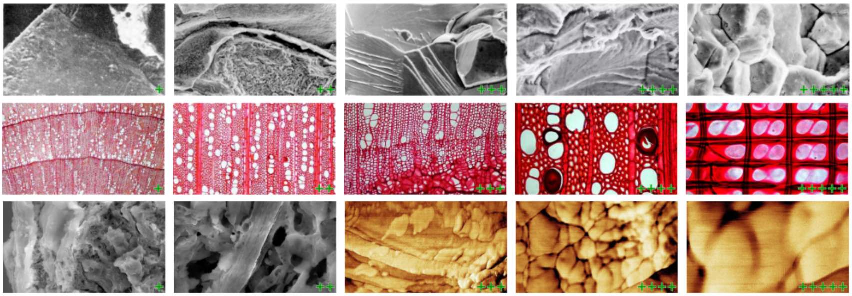


Figure 9.7: Images used for all levels of magnification for metal (top row), wood (middle row), and ceramic (bottom row). Image sources: [168, 211, 66]

niques. These were implemented in Unity[15] and presented to users on a Microsoft HoloLens[16]. The physical controller Interaction Technique was performed using the HoloLens clicker.

To simulate the magnification functionality, we implemented an object recognition algorithm using Vuforia[17]. This was trained for a small set of objects of different materials: a wooden toy, a metal candy box, and a ceramic cup. These objects were chosen for being distinctive yet familiar items that are typically made of the material they represent. The system can individually recognize these objects in the center of the user's field of view. To aid the user in the selection process, a green dot was displayed at the center of the field of view to indicate that the object was recognized and the magnification could be initiated for that object.

[15] `https://store.unity.com/de/products/unity-personal`

[16] `https://www.microsoft.com/en-us/hololens`

[17] `https://library.vuforia.com/articles/Training/Object-Recognition`

Due to the inherently discrete nature of the voice INTERACTION TECHNIQUE, allowing only step-wise increments and decrements, we opted to implement the magnification using fixed discrete levels. This also simplifies the prototype and ensures consistency across trials and conditions, since all users will experience the same levels of magnification, and thus reduce the distortions caused by *quid tertium*. We simulated the different levels of magnification using a set of static images [211, 168, 66] with different levels for each material (see Figure 9.7).

The discrete level of magnification was matched to the continuous distance of the movement performed by the user either with the hands or using the physical controller. The ratio between moved distance and level of magnification was determined empirically through a pilot study conducted specifically for this purpose. This way we determined comfortable ranges of movement for the three physical INTERACTION TECHNIQUES. For the *One-handed* gesture, the minimal level of zoom was the position where the interaction starts, which is the distance of a semi-extended arm, and the maximal level of zoom was reached for the hand positioned around 10 centimeters away from the face of the user. For the *Two-handed* gesture, the minimal level of zoom was fixed to the initial distance between the users' hands, 20 centimeters, and the maximal level of zoom was reached around shoulder-width, 50 centimeters. For the rotation of the *Controller*, the minimal level was fixed to the starting position and the maximal for a rotation of 90 degrees respect the initial angle. For all three methods, the intermediate levels of magnification with distributed at regular intervals between the minimum and maximum.

Using the corresponding interaction technique for each condition, the images were displayed when the interaction was initiated, starting with the lowest level of magnification. The user would control the level of magnification by switching between images using the different INTERACTION TECHNIQUES. To give the users a reference for the current level of magnification, it was visually represented with increasing + symbols in the bottom right corner of the image. Finally, when the user terminated the interaction by zooming out beyond the minimum level of magnification, the magnified image would disappear.

9.4.3 Task

To investigate the effects of the INTERACTION and DISPLAY TECHNIQUES, participants were asked to perform a task that required them to navigate through all levels of zoom and retrieve information from the magnified images. For this purpose, we designed a visual search task based on previous work, requiring the identification of shapes [18]. Each level of magnification for each material (5 levels per object, a total of 15 images) was marked with a varying amount of black geometric figures (squares, circles, triangles, and pentagons). Each image contained a total of eight symbols, with the quantity of each shape being different for each image (see Figure 9.6, left). The objective of the task was to count the total number of squares present across the different levels of magnification for each material. For each condition, one of eight sets of marked images was assigned in a counterbalanced fashion, with each set having a total number of squares ranging from 13 to 15. To minimize the chances of participants guessing the number, the range of possible solutions was unknown to the participants. With this design, each participant had to check the shape of a total of 120 objects and keep track of the squares, while controlling the level of magnification and switching objects.

9.4.4 Procedure

Each participant was first briefed and asked for consent and demographic information. During the experiment, the participant remained sitting on a chair. The three objects were placed in front of the participant on a small table. The participant was asked to wear the HoloLens and adjusted it to fit comfortably. The trial task was then explained to the participant. A prototype version with unmarked images was loaded and initiated on the HoloLens using a laptop. This initial set of images did not present geometric shapes and was intended to help the participant get familiarized with the HoloLens and the INTERACTION and DISPLAY TECHNIQUES for the current condition. The participant was encouraged to test the prototype until they felt comfortable interacting with it. The time needed for training was recorded and then the application for the actual task was started, this time with images presenting the geometric shapes. At this point, the stopwatch was restarted. The participant then proceeded to count the squares and upon finishing, reported their total number. This result and the TCT were noted and the participant was given a NASA TLX questionnaire to fill. These steps were repeated until all four interaction techniques for a given visualization were done. Next, this procedure was repeated for the other display technique. The order of the INTERACTION TECHNIQUES and DISPLAY TECHNIQUES were individually counterbalanced across participants, with the constraint of clustering conditions by display technique to avoid confusing the participant.

At the end of each trial, a short semi-structured interview was conducted, were participants indicated their preference for an interaction technique and explained

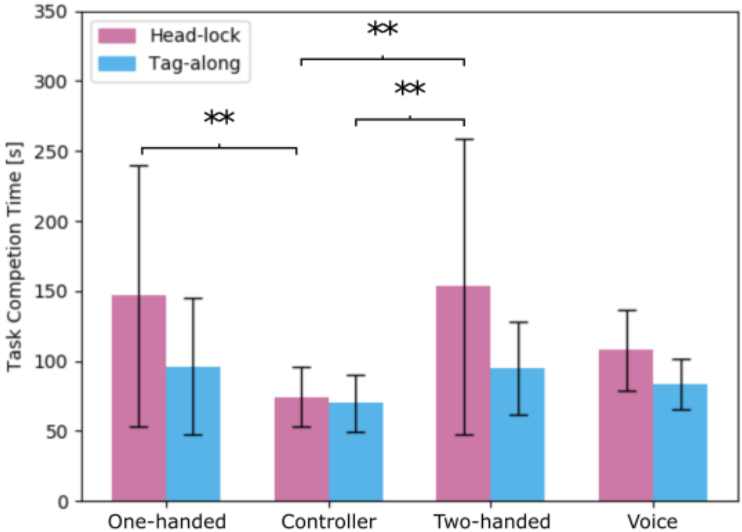


Figure 9.8: Mean value and standard error of Task Completion Time for the four INTERACTION TECHNIQUES and two DISPLAY TECHNIQUES. Post-hoc significance is indicated (p<0.01)**

the reasons for this preference. Additionally, participants were encouraged to suggest application scenarios for a VUM. The total participation time ranged from 30 to 60 minutes.

9.4.5 Participants

We recruited 16 participants. Five of them identified themselves as females and eleven as males, and their ages ranged between 18 and 27 years. All participants reported being in perfect health, presenting normal or corrected vision (with glasses or contact lenses). Eleven participants had no prior experience with the Microsoft HoloLens, four had little experience, attained through other studies in which they participated. One participant stated to be an experienced user.

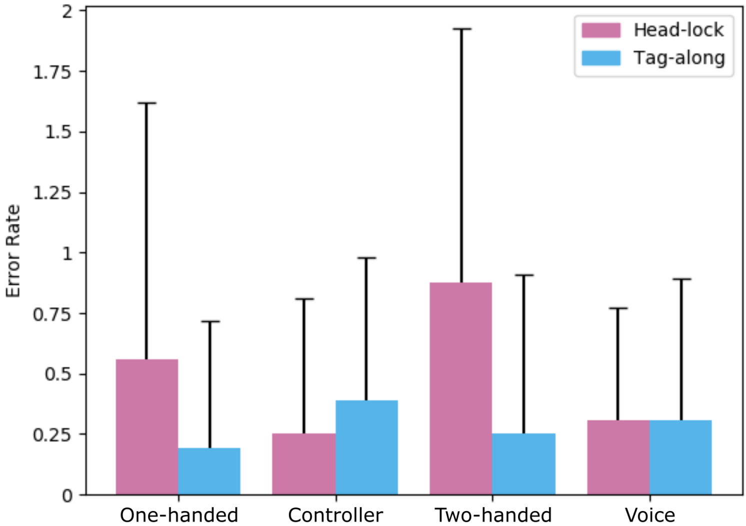


Figure 9.9: Mean value and standard error of Error Rate for the four INTERACTION TECHNIQUES and two DISPLAY TECHNIQUES. No post-hoc significance was found.

9.4.6 Quantitative Results

A two-way ANOVA was conducted to examine the effect of INTERACTION and DISPLAY TECHNIQUES on TCT. The test showed a significant effect of DISPLAY TECHNIQUE $F(3,119) = 6.30$, $p < .001$ and INTERACTION TECHNIQUE $F(1,119) = 12.41$, $p < 0.001$. There was no significant interaction effect, $F(3,119) = 1.61$, $p = .19$. Post-hoc Tukey HSD for the TCT showed significant differences between *one-handed* and *controller* ($p < .01$), and *controller* and *two-handed* ($p < .01$). Post-hoc comparisons of individual DISPLAY TECHNIQUE × INTERACTION TECHNIQUE pairs are shown in Figure 9.8. Group mean and standard deviation are shown in Table 9.1.

We investigated the effect of the two factors on error rate using a two-way ANOVA. We observed no significant effects to be caused by either DISPLAY TECHNIQUE, $F(3,119) = 0.80$, $p = .45$, or INTERACTION TECHNIQUE, $F(1,119) = 2.75$, $p = .10$. The results are illustrated on Figure 9.9

A two-way ANOVA which investigated the effect of DISPLAY TECHNIQUE and INTERACTION TECHNIQUE on the NASA TLX score revealed a significant effect

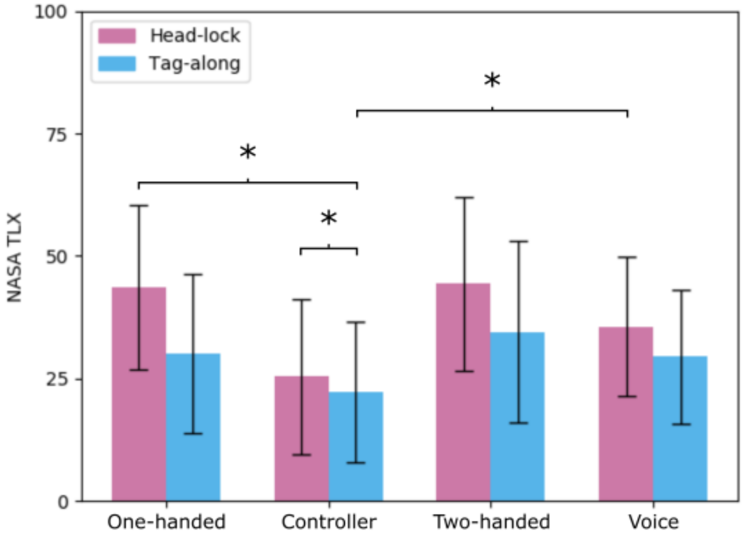


Figure 9.10: Mean value and standard error of the Task Load Index score for the four INTERACTION TECHNIQUES and two DISPLAY TECHNIQUES. Post-hoc significance is indicated (* p<0.05)

of INTERACTION TECHNIQUE: $F(1,119) = 24.47$, $p < .001$. There was no effect of DISPLAY TECHNIQUE, $F(3,119) = 0.75$, $p > .05$ nor an interaction effect, $F(3,119) = 0.76$, $p > .05$. Post-hoc Tukey HSD results are shown in Figure 9.10.

9.4.7 Qualitative Results

We interviewed participants about their preference for INTERACTION and DISPLAY TECHNIQUES, and prompted them to envision applications scenarios. The preferred INTERACTION TECHNIQUE was the *controller*, being the first choice for 12 from 16 of participants. Both *one-handed* and *voice* techniques were chosen by two participants, and *two-handed* was chosen by no participant. When asked about the least preferred method, *one-handed* and *two-handed* gestures were chosen by six participants each, and *voice* by four. The preference for the *controller* was mainly attributed to ease-of-use by nine of interviewees, to reliance by nine, and to intuitive control of the interaction by six.

Table 9.1: Mean value, standard deviation, and standard error for Task Completion Time and Error, calculated for each condition of the two variables.

	TCT		Error	
	Mean	**SD**	**Mean**	**SD**
ONE-HANDED	121.31	79.63	0.38	0.87
CONTROLLER	71.97	21.47	0.31	0.59
TWO-HANDED	124.06	84.91	0.56	0.95
VOICE	95.69	27.22	0.31	0.54
HEAD-LOCK	120.42	79.88	0.50	0.87
TAG-ALONG	86.09	34.53	0.28	0.60

The *tag-along* DISPLAY TECHNIQUE was preferred by nine participants. Participants explained this preference by noting the ability to see multiple images simultaneously. On the other hand, interviewees who preferred *head-lock* explained this to be caused by the image following the gaze of the user, allowing freedom of movement, and that improved the viewing in terms of the degree of detail and contrast.

The application scenarios envisioned by the interviewees were in the areas of clinical medicine, technical design, research in biology, materials science, and education.

9.5 Discussion

The collected results suggest that users appreciated both DISPLAY TECHNIQUES, but *tag-along* performed significantly better in terms of TCT. The INTERACTION TECHNIQUES also had a significant effect on the TCT, and show that using the *controller* is more effective other techniques. The rate of errors committed in the task provides little insight, only indicating a better performance of the *tag-along* visualization technique. The difference in task load was significant across INTERACTION TECHNIQUES, reinforcing the preference for the *controller*.

Based on these observations, we conclude that physical controllers are the preferred choice for controlling a VUM. This technique results in better performance and lower task load. While the study produced a clear result, we do not advise that future systems resign from developing alternative controls. Part of the results of the experiment can be explained by the fact that the remote may be the device which felt most familiar to the majority of the users. Perhaps, this is a case of legacy bias in both performance and preference. This might also explain, to some extent, the preference for the *head-lock* mode by some users despite better performance when using the *tag-along* display. Enabling users to switch between modes can bring further understanding on this topic, since it remains unclear if the effects in performance and preference are due to human physiological and perceptual characteristics. Arguably, design shapes perception as much as perception shapes design. Users trained and formed within a paradigm may develop preference and expertise within that paradigm. This suggests the need to introduce incremental steps in the transition to new interaction paradigms, such as augmented perception. Users performed better with physical controllers and voice commands, which is aligned with the common practices for controlling devices (e.g. TV remote) or asking people to perform an action.

Quantitative results indicate better performance for the *tag-along* technique. However, a large part of the participants preferred the *head-lock* technique. The results suggest that the visualization mode has a stronger impact on the interaction for the given approaches, while users easily adapt to different INTERACTION TECHNIQUES, obtaining similar performance and proficiency in controlling the system. This observation is supported by the error rate, TCT, and NASA TLX scores. These results are aligned with the findings by Lee et al. [136], confirming that the variation of typical INTERACTION TECHNIQUES has a limited effect in performance during the interaction. However, we must recognize that the selection of investigated INTERACTION and DISPLAY TECHNIQUES is by no means exhaustive, and thus the generalization of our findings needs to be applied with caution. On the other hand, this kind of limitation is typical for the evaluation of physical prototypes of AR applications, as discussed by Dünser and Billinghurst [61].

From an interaction design perspective, our findings highlight the challenges for interfaces in the age of ubiquitous computing. Design decisions such as visualization or control methods have a profound impact on the quality of the interaction, thus signaling the importance of *good design* in the successful incorporation of AR to human activities as suggested by Schmidt [208].

Consequently, future work on VUM controls should investigate combining both DISPLAY TECHNIQUES, enabling users to toggle between the views, and investigate the users' choices in different use cases. Many challenges remain to be addressed, such as the incorporation of navigation and panning in the interaction. Alternative display methods are still unexplored. Semi-transparent displays or enabling a transition between the two modes studied in this work offer further challenges to HCI.

9.6 Limitations

This research contributed an exploratory AR application, which particularly focused on sensory enhancement and augmentation. Thus, we investigated a limited subspace of the whole interaction design space for AR, namely only four input and two output methods. Further, the particular implementation of the methods is a research prototype, limited in terms of versatility and flexibility, such as forcing the users to start two-handed gestures with their hands apart. We designed the application as an isolated concept. If VUM were to be part of a generic AR interface, the selected techniques might be already assigned to control other functions. Finally, we cannot exclude the possibility of the discrete magnification levels affecting the interaction and this aspect needs to be addressed in future work.

9.7 Conclusion

In this chapter, we explored interaction and display techniques for a VUM. We proposed possible techniques that could help incorporate this system into human perception, following the vision of Sensory Amplification. We designed the interaction in two dimensions: interaction and display. We proposed four INTERACTION TECHNIQUES, one- and two-handed gestures, a physical controller, and voice commands, and two DISPLAY TECHNIQUES, one fixed to the user in a full-screen mode, and one fixed to the magnified object inside a floating frame. We implemented a functional prototype based on the Microsoft HoloLens that incorporates all four interaction modalities. We conducted a user study with 16 participants. In a search task, participants had to recognize and count squares among other shapes through different levels of magnification. Our results indicate an increased performance for *tag-along* visualizations and input based on tangible controllers. This advantage is also reflected by user preference in both cases. However, both DISPLAY TECHNIQUES was popular among users despite the differences in performance, suggesting that both modes are equally desirable. We hope that our work will inspire further research into how current developments in AR can be used to enable sensory augmentation.

On a more general level, this chapter sheds light on the complexities of interaction design for SA. Our findings pose an interesting challenge to our vision, since what we envision as *true* SA should not require tangible interaction elements, such as physical controllers. This phenomenon could be attributed to legacy bias (users are trained to use hardware) and technical limitations (interaction based on gesture detection may not be as robust as tactile buttons). However, these hypotheses require further research to be properly assessed. Additionally, we have learned that despite these limitations and the variety of user preferences, users are reliably capable of learning how to control SA to a high level of proficiency in a short time and for a wide variety of input methods. Finally, our work suggests that offering flexibility in both interaction and display techniques is a desirable feature for SA applications.

Chapter 10

Implications for Users of Sensory Augmentation

> Cognition attempts to make sense of the world: emotion assigns value.
>
> Donald A. Norman

In this chapter, we discuss the implications of SA for individual users in terms of attention, cognition, and stress.

This chapter is based on the following publication:

- F. Kiss and A. Schmidt. 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*, PervasiveHealth'19, pages 377–382, New York, NY, USA, 2019. ACM

10.1 Introduction

The incorporation of SA to human activities will likely benefit users in numerous ways but this may come at a cost. Additional information, even if useful, will require additional processing power from the users' brains, potentially taking it away from other tasks.

As we have witnessed during the last few decades, the adoption of computers, and particularly of mobile technologies, has empowered people in work and private life. But it has also exposed them to amounts of information and stimuli that surpass natural human capabilities. This overstimulation and informational overload result in a series of problems, that range from being unable to assess the quality of information due to its volume (e.g. fake news) to somatic anxiety disorders caused by the constant expectation of unpredictable stimuli (e.g. phantom notifications [131]). In particular, the overstimulation of the senses has been linked to stress as early as in 1969[261]. Even if this affliction may seem mild when compared with the ones mentioned above, it is its pervasiveness, and the deterioration of health it causes long term, which makes it arguably an urgent problem to deal with.

Current practices in interaction design, both on mobile devices and workstations, do not address this issue: they are the very cause of the problem. The way interaction with information is conceived does not scale up, and with an ever-growing amount of information and communication channels, the overexposure to information and stimuli can only get worse.

This phenomenon has grim implications for AR and, particularly, for SA. The augmentation of perception and sensation implies additional information being delivered to the user, which is the source of this problem. In the particular context of information overload and stress, bad design policies for the deployment of SA can have devastating effects on the human psyche. Nowadays, the physical source of information (and therefore the stress it causes) is limited mostly to mobile devices and stationary screens. Translating the current interaction space to SA means enabling systems to display information on every possible surface and about everything we normally perceive. Making additional streams of information omnipresent is likely to be very detrimental to human cognition and health.

This technological shift has the potential of increasing the baseline stress of interaction with machines and fellow humans; but also can be an opportunity to revise the interaction paradigm.

In this chapter, we review the literature that profiles information overload as a stress factor in HCI and then formalize a problem that has been so far recognized but yet not properly addressed at its source. Further, we propose a different approach to interaction design for upcoming interactive technologies, in particular for SA. Finally, we suggest a research agenda for the methodical investigation of concrete solutions to the problem.

10.2 Related work

The research on sensory overstimulation predates HCI and has since been an ongoing topic of scholarly interest. Within the existing literature about stress in relation to information and cognitive load, it is possible to differentiate more general psychology research from HCI-specific work. In the following sections, we first review the psychology research on stress caused by excessive information and sensory stimuli, and then the most relevant HCI publications on the topic of stress.

10.2.1 Psychology research on information stress

The physiological effect of sensory overload has been subject of research for a long time. Zuckerman et al. showed in 1969 that sensory overstimulation resulted in higher-than-normal levels of hostility and adrenocortical activation [261]. Along similar lines of work, Ludwig demonstrated how sensory overload can produce altered states of consciousness, similar to psychedelic experiences [142].

The incorporation of computers and digital means of communication into human activities encountered limitations of the human cognitive capabilities already at its early stages. In 1985, Hiltz and Turoff recognized the potential problem that CMC could pose for users [94]. Even within the limited context of office-automation and with the limited array of applications available at that time (namely e-mails, computerized conferencing and bulletin board systems), the authors foresaw how increments in the volume of information resulted in what they dubbed *information entropy*. The authors suggested the need for structures in charge of managing and curating information to avoid overwhelming users. Interestingly, the authors also recognized technological legacy bias as a relevant factor in this situation, calling for a change of interaction design paradigms. The authors also emphasized the importance of empowering users to manage and filter information themselves, already hinting at the benefits of taking a user-centered design approach instead of the predominant application-centered design. This research piece was published a decade before the world wide web was created. What we now understand as the Internet was hard to imagine. However, the problem of managing information flows and the resulting stress was already recognized by scholars and has been around since then.

In 1997, Hal Berghel wrote an article titled *Cyberspace 2000: Dealing with Information Overload*, in which he described some early attempts to protect the user from information excess in the context of the Internet [20]. On a more dystopian note, David Shenk's book *Data Smog: Surviving the Information Glut* enumerates the negative consequences of the Information Age in people's well-being [215]. The author described in this work the detrimental effects of informational excess on all spheres of life: mental and physical health, relationships, religion, education, politics and economy. Even if stress is directly included in the category of mental and physical health problems, it is possible to view all other areas were "data smog" has negative effects as indirect causes for additional stress.

In his work from 2003, Nielsen proposed the concept of *Information Pollution* to describe the negative impact of interruptions caused by CMC systems in productivity and performance of workers [175]. Even if his observations did not address health problems directly, Nielsen identified e-mail and instant messaging as problematic due to their ability to overload cognitive capacity. *Information Pollution* was further investigated by Bray in the context of modern Internet-powered technologies, in particular social media [28]. Bray observed that most available literature on information science examined the use of single systems while failing to consider the reality of users multi-tasking with other technological systems simultaneously. Additionally, Bray proposed three tenets to improve information management systems: (i) *replacing*[18] available information with better information, instead of adding information on top, (ii) avoiding the use of technologies that imply refocusing attention frequently, and (iii) treating human attention as a resource that needs to be conserved. All these derive from the idea that human attention and cognition power are scarce, exhaustible resources, proposing what we could call *cognition austerity*.

In 2009, Bawden and Robinson reviewed the negative effects of information technologies in health in an article titled "The dark side of information: overload, anxiety and other paradoxes and pathologies" [16]. The authors related the amount of information available and the pace at which it changes to multiple

[18] Emphasis of the original author

mental and physical health issues. Furthermore, they predicted new "pathologies of information" to emerge as information technologies would continue to evolve. It is remarkable that this work refers mostly to technologies of the so-called Web 2.0, and was published when smartphones were a brand new technology (the first iPhone was launched in 2007 and the first Android phone in 2008). Therefore, mobile devices had not yet gained such a prominent role as today.

In a study from 2011, Thomée et al. assessed the effects of mobile phone usage on over 4000 young adults [230]. Their findings suggested that overuse was associated with stress and sleep disturbances for female respondents, and *high accessibility* was associated with stress, sleep disturbances, and depression symptoms for both male and female participants.

Yin et al. extended previous research with investigations of the effects of mobile technologies in the context of what they call "technostress" in the workplace [253]. They proposed a model in which stress is caused by two main factors: techno-overload and techno-insecurity, both being caused by the excessive amount of information and choices provided by mobile technologies. Ledzińka and Postek revisited this topic in 2017, recapitulating the psychological foundations and empirical evidence on information overload and overflow [134]. The authors framed their concept of "infostress", or information stress, within the context of work and office environments. Nonetheless it is possible to extrapolate the effects of infostress to every other situation in which people are exposed to more information they can process. The work of Scheydt et al. reviews the concept of sensory overload from a clinical perspective, analyzing both conceptual definitions and clinical cases [206]. They suggested a difference between what they call "subjective overload" and "objective overload", proposing that people experience a combination of both, being the first more prominent in the context of mental disorder, while the second one seems more related to environmental factors.

10.2.2 HCI research on stress

In the last years, there has been growing interest within HCI research in understanding stress. Most of this work can be categorized in stress detection and stress reduction, with some cases exploring both aspects.

Detecting stress

Stress detection is a challenging research topic, in particular when trying to dispense with stationary medical-grade sensors and measure stress in the wild with wearable devices. Sanches et al. explored the design aspect of this problem, combining psychological foundations with bio-sensors and interface design [203]. Their findings highlight the difficulties in detecting and quantifying stress with wearable devices, as well as the importance of taking into account interaction aspects, such as *interactive history* or *fluent state transitions*. The authors proposed a set of tools for the users to enable them to reflect on their way of life and thus be able to address the causes of stress. Despite this, the authors proposed no design solution to the problem of stress itself.

In 2014, Sun et al. published a method to measure stress based on mouse usage [224]. Their work connected the mechanical model of the arm and muscle stiffness, which can be considered a stress indicator, with the way the user controls a mouse. The authors proposed to measure stress experienced by users based on the mouse input data and validated their model with a controlled experiment.

A different approach to measure stress was taken by Yoon et al. in 2014. They conducted a survey to assess how mobile messenger notifications caused stress in users, and how users perceived and managed this particular kind of stress [255]. The authors derived from their observations a set of design recommendations, proposing to reduce the information load by managing and categorizing incoming messages, and reducing the update frequency of notifications.

In 2015, Liapis et al. published a study of the *Valence-Arousal* scale using skin conductance. Their findings suggested the validity of using this scale for self-assessment of stress in HCI tasks, such as finding information in webpages with problematic information architecture or focusing on particular tasks despite web advertisements. Li et al. conducted an empirical study of perceived stress in everyday college activities [137]. The authors differentiated between eustress and distress, being the latter the form of stress that has a negative impact on health and cognitive performance. Their findings suggested that heart rate variability, combined with smartphone and computer usage, can be used for general stress classification. The authors emphasized the importance of *positive* eustress and its recognition, as an important design factor for HCI applications.

Managing stress with controlled breathing

Stress reduction or management has been also a focus of research in HCI, involving different strategies and approaches. A large part of work on this topic explored the used of controlled breathing as an intervention: Moraveji et al. developed a system that monitors and influences the pace of respiration of a person while working on a desktop computer [165]. The authors hypothesized that *peripheral paced respiration* influences user respiration and their findings support this idea, although the effect was not sustained for the duration of the task, and no additional metrics about stress besides the breathing rate were measured. Building upon this work, Wongsuphasawat, Gamburg and Moraveji evaluated a mobile version of the system [248], but their findings showed that matching breathing pace to external stimuli resulted in an increment of the cognitive load. Morajevi and Soesanto published later a list of ten heuristics bases on psychology to design user interfaces [166], aiming to reduce the stress caused by interactions, although they do not focus particularly in information overload.

Soyka et al. further explored the usage of breathing for managing stress in virtual reality [220]. The authors designed a virtual underwater environment, where rhythmically moving jellyfish would suggest the breathing pace to the user. Their findings supported their approach, suggesting the effectiveness of this kind of application as an intervention to help people cope with chronic stress.

The use of virtual reality to help users cope with stress was also explored by Bruggerman et al. [33]. Their designed proposed the use of *virtual healing spaces* to provide a space for meditation and stress reduction through controlled breathing.

Paredes et al. compared the effectiveness and impact of voice and haptic feedback at reducing breathing pace in drivers [184]. The authors evaluated their design in a controlled experiment using driving simulators. The results of this study suggested that both feedback modalities were effective in reducing the stress level of participants without detrimental effects on driving safety.

Managing stress with self-reflection

A frequently used alternative approach is encouraging self-reflection and thus enable users to recognize stress causes as a first step to reduce stress. Along these lines, MacLean et al. introduced Woodwing, a system that externalized manifestations of affective state to motivate reflection and self-regulation [149].

The proposed design consisted of an electro-mechanical butterfly attached to the user's wrist, and controlled by a computer. The device was actuated in response to the user's level of stress, which calculation was based on electrodermal-activity and electrocardiogram. The system was built into a driving simulator, and provided live feedback to the driver. This design was evaluated in a user study, with participants driving along predefined paths. The results indicated that participants drove in a safer way while wearing the butterfly, but also experienced higher levels of stress. The authors suggested that this was due to the unfamiliarity of users with the butterfly, and the volume and type of data it provided.

Also aiming to enable self-reflection, Wilson et al. proposed and application that collects data about sleep quality and duration, and physiological stress metrics [246]. These data was later translated into informative visualization to provide feedback to users, with the ultimate goal of raising awareness and encourage reflection. The authors suggested explicitly to use *Mindfulness* techniques as interventions.

Lee and Hong explored the causes of stress and the space for self-interventions in their publication *Mindnavigator* [135]. Their work was based on workshops in which college students explained which stressors and relief-strategies were present in their daily activities. Based on the collected data, the authors proposed an analytical framework for self-reflection.

Yu et al. proposed to us ambient light as an intervention to assist relaxation and reduce stress [256]. Their findings support the notion that light can improve awareness of stress and trigger deep breathing and validate the use of biofeedback-controlled ambient light as persuasive technology for stress reduction.

Other HCI explorations of stress management

Some publications address the problem of stress but cannot be assigned to the two general categories proposed above.

For example, Kettner et al. proposed a system to provide non-stressful tactile notifications [113]. They compared in a user study different kinds of feedback using wristbands that provided pressure-based and vibration-based feedback. The authors compared the effect in stress reduction of the design against vibrotactile feedback, showing the latter to have a greater impact, particularly in low-stress situations.

In the field of augmented reality, Jones and Dechmerowski proposed the use of computer-generated environments to simulate stress conditions that are impossible to explore in real life [107]. This would enable the usage of physiological measurements and classifiers in situations that safety risks or costs would render implausible to evaluate. In their contribution, the authors reviewed different stress measurement methods and outlined adaptive stress training models in augmented and virtual environments.

10.2.3 Summary

Despite the absence of SA as the focus of research in the context of stress and cognitive overload in HCI, it is possible to derive one general observation that applies both to general MR and SA: past efforts focus on tackling the negative effects of information excess and stressful interaction while the lack of work on addressing the cause of this problem is noticeable. The literature links stress to interaction design. Further, it presents methods to detect it and suggests different approaches for dealing with it, including mindful meditation and breathing techniques. Despite these techniques being useful for dealing with stress in general, we believe that they are not the optimal solution to address the stress and overstimulation problems associated with interaction and potentially extended to SA.

10.3 The risk of augmenting information stress

Research typically focuses on current problems since issues with technologies that are already deployed are easier to identify and understand. Also, the actual negative effects we experience are more urgent than hypothetical ones. However, solutions to these issues are mostly damage-control and mitigation, since the cost and effort of addressing the cause of the problem are too high. The replacement of current interaction paradigms implies replacing current systems. This does not only cost money and time, but also forces the users to learn new interaction methods. This highlights the importance of foreseeing problems and preventing them before technologies are deployed.

SA systems will likely replace some typical functions from mobile devices in the coming years. Interaction metaphors and concepts that were inherited by mobile devices from stationary computers are likely to be further inherited by augmented reality systems, including SA applications. Many of these symbolic elements proved to be of great utility for the interaction platforms of the time in which they were created, back then when personal computers started to proliferate. The use of files and folders, pop-up dialogues, icons, and windows, or even more abstract concepts, such as *selecting* and *pasting*, they all went from non-existence to be a transcultural part of common knowledge in a very short lapse of time. By the time people started to have access to mobile devices, the use of computers was already global, and a great part of humanity had some level of proficiency in basic computer concepts. The interaction elements were then inherited by mobile devices, mostly because people already knew them, but implicitly because the technological transition was incremental: every new iteration of consumer electronics profited from proven interaction elements, only adding new concepts or modifying existing ones when necessary.

This is likely the case for the next stage: AR and VR interfaces already rely to a great extent on interaction metaphors and elements that are easily traced to workstations from over 30 years ago. Of course, icons and windows are still useful (and popular) and are unlikely to go away. But along with the positive aspects of the current interaction design paradigm, all its negative aspects can also be transferred to new interaction platforms.

The adoption of current interaction metaphors and design in SA systems will likely reproduce the causes of stress present in mobile devices. Further, SA can potentially worsen the problem: mobile devices have limited space and capacity to present stimuli and information to a user, while in SA the augmentation is pervasive to everything humans can sense. Presenting users with multiple streams of information, mixed with pervasive advertising and ubiquitous notifications is not an appealing image of the future, but also not an unlikely one. The consequences of this future for the psyche of users can go from very annoying to completely devastating.

10.4 Strategies for stress-less interaction design

We propose to formulate design recommendations to minimize the negative effects of inheriting problematic design narratives. These recommendations build upon Bray's design recommendations [28], adapted to the particular case of SA but easily extensible to general AR.

1. Upgrade: Prioritize the replacement of available sensory information with better information, over the addition of more information.
2. Steadiness: Minimize shifts in the focus of attention, as well as interruptions.
3. Conservation: Consider human attention as a limited resource that needs to be conserved.
4. Control: Empower the user to manage the information flow at will.

We believe that these design principles can reduce information overload in HCIs and therefore stress. In the following sections, we explain each of these principles in more depth.

10.4.1 Information upgrade

SA typically has a strong emphasis in audiovisual information, and per definition implies an increment of the information we perceive. Still, the manner in which this information is presented can affect the cognitive load it takes from the user. Even in the context of SA, replacing information with better (or *richer*) information is possible.

Accordingly, we recommend the use of subtle methods of augmentation, which are only obvious under purposeful scrutiny, but remain unnoticed otherwise.

This translates to avoiding the use of text, highlights, or markers to present extra information. Instead, additional information can be subtly conveyed for instance by modifying the optical properties of individual objects, such as hue, brightness, or contrast. More abstract information can be transmitted with subtle but persistent sounds.

10.4.2 Information steadiness

Forcing the user to re-focus attention requires extra effort and is by itself a cognitive task. To minimize the toll this takes from the user's cognitive capacity, the presentation of additional information should take place overlapping with the augmented object, or in its immediate area.

Abstract information should be presented either as non-positional, for instance as ambient audio, or if it must be presented visually, it should be in the general direction that the user is looking at.

The use of sudden or dynamic stimuli, such as pop-ups, blinking text, sudden noises, or moving symbols, should be minimized, or if possible, completely avoided. Abrupt changes should be softened, although this should not result in excessively slow transitions since those might cause anxiety.

10.4.3 Attention conservation

Attention is a limited, scarce resource and interaction designs should attempt to preserve it.

An additional aspect to take into account is SA's multi-modal nature. Even in the cases in which the provided additional feedback is purely visual, this will be overlaid on top of reality, which we constantly perceive through all our senses. Since attention is limited across senses, it follows that it should be considered as a pooled resource.

Based on these assumptions, it is possible to reduce information load, and therefore stress, by minimizing simultaneous multi-modal feedback.

When possible, the sense over which information will be conveyed to the user should also be chosen based on the actual load of each sense.

Additionally, systems could detect when the user acknowledges the presented information and cease the stimuli automatically.

10.4.4 Information flow control

Since the constant exposure of additional information is detrimental to the user, it is important to decide what and when to display. User-centered design should empower the user, since the user is the agent of decision and choice. Thus, the information should be presented on demand by the user, or only when it is extremely important that the user becomes aware of it.

Information with low levels of priority should only be presented to the user when the level of attention of the user allows it, without compromising other activities.

In this way, the user decides about what is presented, thus perceived. This principle could also include all of perception, giving the user power of choice over both augmented sources of information and environmental cues.

10.4.5 Some considerations

The above-described principles are recommendations aiming to reduce the informational overload of users during computer-mediated interactions with machines or people. The reduction of informational overload should result in a reduction of the stress experienced by the user during the interaction. The greater the extent to which these principles are implemented, then greater will be the reduction of stress.

Of course, the stress reduction will only affect the portion of stress caused by the interaction. Other sources of stress will likely remain unaffected by our design recommendations.

10.5 Future Work

The vision described in Chapter 6 lays ahead in the future. Despite it not being an imminent reality, we believe it is due to happen eventually. Thus, it makes sense to design SA with this long term scenario in mind, since once in place, interaction paradigms are difficult to modify. To reach the envisioned scenario with an interaction paradigm that does not tax users' attention and health, we propose a series of research topics that we deem necessary, starting with a systematic literature review about information stress, as defined by Ledzińka [134], with an emphasis in the augmented senses. This can set the foundations for establishing a design space for stress-less interactions in AR and SA applications. Next, the validity of this design space must be assessed to determine to which extent it can reduce interaction stress. Finally, an exploration of techniques to empower user control over information flows with intent can be conducted independently from the aforementioned research. This can be performed using a wizard of Oz approach, thus escaping current technical limitations.

10.6 Conclusion

In this chapter, we have identified the main challenges for SA in the context of user interfaces, namely interaction stress and information overload. The reviewed literature highlights the negative cognitive effects of exposure to excessive information and links them to stress. This phenomenon has been recognized by psychologists and Information Science researchers, and a corpus of work in HCI proposed different methodologies to detect and reduce stress, mostly by helping the users control their breathing, or encouraging self-reflection and self-awareness. Still, there is so far no clear attempt to address the causes of stress as a strategy for interaction design. We see in this problem a clear opportunity to rethink how interactions are designed and propose design recommendations to lower stress generation. This is particularly important in the context of SA since the novelty of the field allows for significant freedom when designing interfaces.

Chapter 11

Social Implications of Sensory Augmentation

> Cognition attempts to make sense of the world: emotion assigns value.
>
> — Gabriel García Márquez

In Chapter 6, we identified aspects of SA that may result in negative consequences from a social perspective, e.g. privacy threats or discrimination. The social consequences of empowering human perception beyond the natural limits are diverse and their study encompasses many disciplines. In the field of competence of HCI and within the scope of this thesis, the social acceptability of SA seems the clearest challenge and critical indicator for the successful incorporation of these technologies into human societies. In this chapter, we investigate the topic of social acceptability of SA in terms of its most likely source of conflict: the asymmetry of sensory capacity between users and third parties. We approach this task through a series of interviews and a survey, from which we derive a set of social implications for SA.

This chapter is currently ***under submission*** *and planned to be published as follows:*

- F. Kiss, C. Eghtebas, N. Goyal, and P. W. Woźniak. Understanding Acceptability for Ubiquitous Augmented Reality. ACM Transactions on Computer-Human Interaction (TOCHI).

11.1 Introduction

Privacy is not only a right, but it is an extremely important mechanism for humans to function in society [231, 194]. Data we decide to disclose or not is what gives us the agency to choose the way we interact with the society around us. Sensitive data about ourselves can be misused in detriment to our interests, enabling scenarios as extreme as identity theft or social engineering. Even seemingly innocuous information can be used to infer data of a more serious nature, affecting life circumstances like employment or safety [145]. Ultimately, privacy is based on an individual's capacity to manage access of their information to others [212] where information pertains to identifiers regarding a person; from the name, to savings account, personal preferences or activities to medical records.

In consequence, augmentation technologies, capable of peeking from further away, eavesdropping remotely, or generally allowing third parties to invade privacy, can be received by society with mistrust and rejection. Given the imminence of this issue becoming a problem for the deployment of novel interfaces, the timeliness of understanding public reactions towards SA is of critical importance, since a failed first introduction of these technologies to societies can result in a lasting rejection.

The research reported in this chapter consists of a review of related work, a series of interviews, and a survey. To study the reactions towards the hypothetical scenario of people using SA in social environments, we conceived a fictional technology that enables ubiquitous visual zoom, inspired by the *Bionic Scope* [64]. Since our fictional technology augments sight, it relies on a HMD to present magnified vision to users. We first conducted a series of interviews to understand the general concerns of the public about privacy in the context of SA. Building on the gained knowledge, we designed and conducted a survey to evaluate public perception on a larger scale. Based on our findings, we derived a set of recommendations to minimize the threat to privacy and maximize social acceptance of SA applications in with respect to privacy.

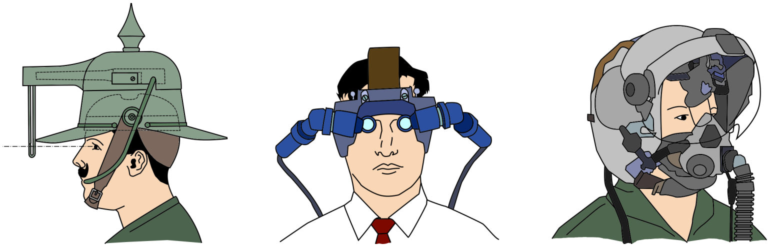


Figure 11.1: Left: Pratt's patent from 1916 can be considered an early precursor of HMDs [191]. Center: Sutherland's design was a revolutionary optical see-through display [225]. Right: A display embedded into a fighter pilot's helmet [52]

11.2 Related Work

In this section, we review the prior literature on privacy concerns that is relevant in the context of SA. Given the lack of a formal definition in previous work for SA, we focused on work related to AR and ubiquitous computing. We identified two main aspects of how technology is perceived: first, to which extent this technology can be actually perceived, i.e. its *visibility*, and second, the factors that make a technology socially acceptable (or not).

11.2.1 Visibility

Past work studied the visibility of wearable technologies and, particularly relevant to our fictional augmentation device, HMDs. The capabilities and form factor of HMDs changed dramatically since its earliest known instance from 1916 [154] and continued to do so with Sutherland's optical see-through display in 1968 [225] (see Figure 11.1).

During the 1970s, military researchers investigated HMD applications, including helmet displays for pilots [52] (see Figure 11.1). Aiming to increment wearability and taking advantage of miniaturization of electronic components, prototypes for AR HMDs started trending towards a more minimalist design of glasses [153].

By the 1990s, AR was a well-established field, with the HMDs' limitations clearly outlined and pursued to be overcome [236]. The leading paradigm in the field was for a wearable HMD [223], pioneered by the Google Glass. Currently, HMDs for AR are starting to reach the consumer marker based on the lessons learned by developed-oriented products such as the Microsoft HoloLens, Meta, or Magic

Leap. Research in its early phase also demonstrates a prototype of AR contact lenses [138]. This evolution of HMDs, especially in the context of AR, suggests a trend towards miniaturization. Even if arguably motivated by wearability and furthered by technical advances, this trend also results in a progressive decrement in the visibility of the device. It is reasonable to assume this trend to continue, resulting in technologies capable of providing AR interfaces that are invisible to third parties.

11.2.2 Social acceptability

As discussed in the Introduction, the main social concern for technologies that augment the sensory ability of users results from the potential to misuse SA and breach the privacy of third parties. Addressing this problem is a complex task since “attitudes to what is and what is not private data vary between people in different contexts and roles” [86]. In the case of HMDs, and in particular for our fictional zooming augmentation, the presence of a camera has a negative impact in how the public perceives the interaction [96]. The nature of the device has been shown to influence its social acceptability, as well as stereotypes arising from the level of computer literacy or proficiency [58]. The difference in technical knowledge has also been shown to have an effect in the acceptability of technologies, indicating that *experts* see emergent technologies as greater threats to society than *non-experts* [218]. In general, the audience involved in the interaction, as well as the location, can be seen as determinant factors for the social acceptability of a technology [144]. Further, the role of a given individual in interaction and its goal can result in more socially acceptable interactions, as shown by Profita et al. for the particular case of the Google Glass [192]. However, despite the many perks offered by the Google Glass in medicine, education, and entertainment [205, 45], it failed to reach ubiquity due to privacy being among one of the core concerns [130, 257].

The reviewed work suggest the importance of the role of interaction’s participants in how this is perceived. Further, it suggests that the visibility of the augmentation affects its acceptability: the negative reaction towards sensing technology (as was already the case for early cameras [96]) is directly caused by the third parties noticing the technology. Paradoxically, invisible augmentations would not trigger this response, despite arguably enabling greater potential for misuse, by “avoiding negative attention” as highlighted by Koelle et al. in their survey of design strategies for socially acceptable HCI [127].

11.3 Research Questions

The potential threat to privacy by SA may have an effect in the social acceptability of these technologies. This is aligned with Schmidt's suggestion about both expected benefits and perceived threats playing a role in the social acceptability of individual technologies [210].

Further, the asymmetry of sensory power between the two or more participants of interaction also plays a role in acceptability, as well as the context in which the interaction takes place, e.g. a person using binoculars to watch a close up of a sports event versus a close up of children playing in a park.

These factors gain additional prominence in the particular context of SA since the goal of these technologies is the increment of sensory capacities, possibly resulting in a stark asymmetry of perceptual capabilities respect non-augmented individuals. To investigate the role of these factors in the social acceptability of a SA, we formulated the Research Questions enunciated in the Table 11.1. These questions can be considered as sub-questions of this thesis' RQ6 and are numbered accordingly.

Table 11.1: Research Questions about social acceptability of SA. These questions can be regarded as sub-questions of RQ6 in the general context of this thesis.

	Research Question (RQ)
RQ6.1	Is SA socially acceptable?
RQ6.2	Is social acceptability of SA dependent of the level of asymmetry between users and non-users?
RQ6.3	Is the visibility of a SA related to its acceptability?
RQ6.4	Does the relationship between perceived usefulness and potential misuse of SA play a role in its social acceptability?

11.4 Methodology

We user a mixed-methods study design to assess public perception of a hypothetical SA that enables users to magnify vision, thus allowing close-ups to objects or people from a long distance. This augmentation was chosen due to its high feasibility, potential benefits, and also the potential threat to the privacy of third parties.

Our study consisted of two stages. In the first stage, we conducted a series of interviews to obtain a general understanding of expectations and concerns that users and third parties might have regards SA. The second stage consisted of a survey, based on the insights gained in the first stage, to obtain more generalizable findings regarding the social acceptability of SA and its implications for interaction design. These two stages are described in detail in the following sections.

11.5 Interviews

To gain an initial understanding of the public's expectations and concerns for SA, we conducted a series of interviews with the aid of a functional prototype. In this section, we describe the used prototype, the procedure of the interviews and their analysis, and our findings.

11.5.1 Visual Zooming Technology Prototype

To illustrate the potential and the effects of SA, interviewees were presented with the prototype of a Visual Zoom Technology (VZT) using a Microsoft HoloLens, so they could explore the interaction freely. We proposed a VZT that enables users to magnify vision and thus grant close-ups of objects or people within direct line-of-sight. Our choice is based on the documented predisposition against visual augmentations due to privacy concerns [96, 130]. Current trends suggest visual augmentations becoming accessible to the consumer market in the coming years, leading us to consider our proposed scenario highly likely.

Given the limited resolution of the camera in the Microsoft Hololens (2 megapixels), we opted for printing physical copies in different sizes of high-resolution images. A copy of one image was hanged on the furthest wall of the room where the interviews were conducted, ten meters away from the participant, while a copy of a different image was placed on a table next to the interviewee. When the user of the prototype faced a copy, the original digital image was displayed on top, thus allowing a controlled high-resolution magnification and clear display of details by moving this virtual image closer to the user. This enabled a tangible demonstration of the fictional VZT, which we could present to interviewees and convey to them the characteristics of this augmentation consistently.

Further, participants were told this was an early prototype and to envision this technology allowing unlimited magnification and a much smaller form factor.

11.5.2 Procedure

The interviews followed a semi-structured approach and were conducted individually in a meeting room. Participants were asked to remain sitting on an office chair for the whole duration of the interview. They were briefed about the goals of the study and asked to provide consent and demographic information.

First, the participant was introduced to the Microsoft Hololens and the gestures to control it by means of a calibrating program. Once familiarized with the device, the participant was invited to try out a first demonstration, which consisted on zooming into a large poster on the furthest wall of the room. A second demonstration followed, allowing the participant to zoom into a small flier on top the table next to the participant.

After the two demonstrations, we began with the interview by asking about which scenarios could the participant imagine for the VZT. The next question prompted participants to imagine situations in which the user would be in a hurry to gather information, and to which extent could this technology have an effect in them. This was followed by asking the participants to suggest gestures to control the magnification, as well as different methodologies, e.g. continuous versus discrete magnification levels.

At this point, we showed a third demonstration, that allowed participants to magnify into the previous printed copies with three different speeds (slow, medium, and fast), and asked the interviewees about which speed they preferred and why.

Next, we asked the participants about zooming into other people, followed by their thoughts on being zoomed into by others. We followed these questions by asking about notifying others about zooming in, as well as being notified themselves.

We recorded audio and video during the interviews, with additional remarks written down as notes. The interviews had durations of 30 to 45 minutes.

11.5.3 Participants

The twelve interviewees (six males and six females with an average age of 33.5 years, $SD = 12.7$) had professional backgrounds in diverse Engineering disciplines, Computer Science, Administration, and Public Relations. They were recruited in the University's campus and received no monetary remuneration for their participation.

11.5.4 Analysis

The recordings of the interviews were transcribed and revised using Atlas.ti by two researchers.

The transcriptions were then divided into individual statements, which were printed separately, coded and classified using affinity diagrams by two researchers. This procedure resulted in multiple groups, which were organized in three main themes: Usability Expectations, Personal Implications, and Social Implications.

Usability Expectations

A recurrent theme in the interviews was what participants expected of the technology in terms of usability and functionality.

Despite the main focus being on technical aspects, such as the speed of zoom and control methods, some acceptability aspects transpire from the recorded feedback. Control suggestions resonated with social acceptability and visibility, as the participants mentioned usability within a social context:

> It is okay as long as the gesture is semantically undefined …in combination with wearing this would communicate that I am not crazy. (P3)

In addition to the interaction methods, the appearance of the device and its acceptability were also mentioned in the context of the technical discussion:

> It is something I would use, I mean, if it was like everyone had it or it was really popular. If I used this right now people might think I was Robocop or something like that. You know what I mean? (P12)

Personal Implications

Multiple recorded statements can be interpreted as implications for the user in a social context. Most frequent were the implications of staring at other people, or being stared by other people using the SA, as a form of social advantage. Further, participants discussed ways to mitigate the effect of these asymmetrical sensing capabilities through notifications.

The augmentation technology being worn by only one of the parts of the social interaction was perceived as an unbalance. This resulted in implications beyond the enhanced perception enabled by the hardware. Interviewees suggested use cases where users of SA could leverage this enhanced perception to gain social advantages over third parties.

Zooming into other people was generally recognized as detrimental to the privacy of third parties. Most participants stated reluctance to do so, with only one suggesting this to be the primary purpose of this technology. Other participants' responses pondered about the level of certainty of whether it was fundamentally benign or not to use this technology in that way.

> I am being honest here. I think most of us would do it at some point because I think it is kinda natural that you have some kind of interest. You can do it perverted style or you could do it normal style . . . or accidentally. But I think yes I would zoom in on other people. (P7)

Participants' responses to being stared at by others ranged from emphatic rejection to *I don't expect people to stare at me*. No participant specifically wanted to be stared at but some showed acceptance that it was inevitable to be stared at and ultimately accepted it. Additionally, P7 mentioned the existence of ubiquitous security cameras and our current obliviousness to being constantly watched.

> It is okay for them to look at me why shouldn't they be able to zoom on me. If you live on this planet you are seen. You cannot do anything against it. (P4)

> If I don't recognize it I don't have a problem with it. I think that is the whole point. You walk around and nowadays you have security cameras.. and you don't recognize it or have the feeling that you are being watched. (P7)

Discussions around our current obliviousness to ubiquitous vigilance and the need to be informed if you are being watched emerged during the interviews, which led to the idea of notifications. One participant compared staring and being stared at with VZT to WhatsApp chat feature that consists of checkmarks that indicate read messages. Surprisingly, participants mostly preferred not to be notified of being stared at by users of SA and some found that it would be "offensive" to inform others that they stared at them:

> I think that is a bad idea because, you don't want to offend someone, you don't want to look at them. (P4)

Social Implications

The theme Social Implications identifies socially acceptable contexts for VZT in public and private spaces and introduces permissions to stare at others based on a social media model. Participants who at first rejected the idea of staring at others with VZT reassessed their views when finding use cases where this would be socially acceptable, such as artistic performances or monitoring their own children.

Similarly, for interviewees who didn't want to be stared at, they recognized that limiting the use of VZT would come at the cost of some use cases perceived as beneficial. Such was also the case for people with specially assigned roles in society, for instance police or people in charge of ensuring the safety of others.

> But then you don't have the use case for police or looking for people in a crowded place. Then you just have zoom-in into information displays, public displays. (P5)

Participants also mentioned that malicious acts using VZT would still be possible even if zooming-in on people wasn't allowed, as for instance stealing sensitive information like including bank credentials. Ultimately, the main distinction made by participants was between public and private spaces, where privacy concerns were viewed unanimously. Participants discussed that the technology could have built-in restrictions on viewing a private space, as for instance looking inside a building through a window. A public space, what is normally regarded as a conventional public setting, e.g. outdoor locations such as train station or city streets, had varying levels of social implications. The VZT was also compared to existing methods of zooming using binoculars and already existing means of intruding on someone's privacy.

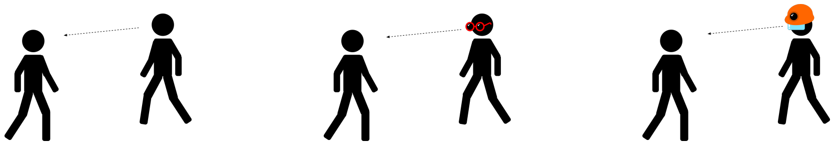


Figure 11.2: Surveyed scenarios: lenses (left), glasses (center), and a helmet (right)

> If you are in public there is no difference between someone standing close by or 100 meters away... there is no problem. But if you are in your own house or your own apartment; then I think that this privacy issue is not that big of an issue because you can already do it with binoculars. (P2)

Interviewees suggested to control the misuse of the SA by only permitting to zoom-in onto people who have previously granted permission to do so. One participant compared such permissions to existing equivalents in social media, for instance the metaphor of *friending*, and the management of access to public and private digital spaces. Participants recognized that this approach would only be viable once the technology was widely deployed and has a considerable user base.

The even more far-reaching issue is of heterogeneous SA types which would have different capabilities that would need to interact in a plausible reality. Lastly, a participant discussed the particular case of a user having permissions to observe anyone, but given none permission to others.

> I think the idea of invisibility would be a bad idea, except for spies or people in the witness protection program. (P10)

11.6 Survey

The interviews showed diverse reactions to VZT, in some cases depending on the role of the interviewee in the interaction. Further, they suggested that being aware of the augmentation played an important role in how they perceived the interaction. To gain a better understanding of how these aspects affect the social acceptability of SA, we conducted an online survey. We used a repeated-measures within-subject design to measure the perceptions on three variations of SA hardware with different levels of visibility. The SA was presented in the context of a social encounter and participants were asked to assess the interaction from three different social roles. In this section, we report on the survey design, collected data, and derived findings.

11.6.1 Survey Design

The survey consisted of six parts: Demographics, hardware, three scenarios, and the end section. The Hardware section presents three variations of the form factor of an VZT, identified from previous work and with different levels of visibility: a clearly visible helmet [52], inconspicuous sunglasses [153], and almost invisible contact lenses [138]. In this section, the survey asks about USABILITY EXPECTATIONS surrounding the three variations of this technology through questions such as "Someone will misuses this technology" or "This technology should be available to everyone" requesting answers with a 5-value Likert scale ranging from complete disagreement to complete agreement.

The three scenarios probed the perceived advantages from the PERSONAL IMPLICATIONS and SOCIAL IMPLICATIONS of VZT in different social situations. This topic was suggested in the interviews, with the interviewees emphasizing the relevance of the role in the interaction when determining the social acceptance of the technology. Thus, in the survey, we investigated different scenarios for SA and how users perceived advantage in those scenarios. The three scenarios consisted of two people walking towards each other, were one of them was wearing the SA hardware as shown in Figure 11.2. The scenarios varied in the role the participant had to play, either wearing (W) the SA hardware, not wearing the hardware but being involved (I), and not involved (NI) but observing the scenario as an external spectator. These alternatives encompass all possible roles of the participant in the scenario and aim to assess the effect of such roles in how the VZT is perceived. Following what we learned from the SOCIAL IMPLICATIONS theme in the interviews, we reflected on the need of permissions to use VZT and asked if zoom should be allowed in each scenario. The scenarios were kept abstract to allow every participant to relate to it and afford more interpretations from survey participants. For each scenario, the participants had to answer six questions: three with Likert Scale items (with 5 possible values), two *yes* or *no* questions, and one requesting text input for the participant about what should participants not be allowed to zoom into. In this section some questions requires a *yes* or *no* answer, other a Likert Scale core (again from 1 to 5), and a few of them asked the participant to input text.

Lastly, the end survey investigates SOCIAL IMPLICATIONS of VZT as compared to existing technology, and the need of permissions to use the technology. Again, some questions required text input while others Likert Scale scores. The survey was made using lime survey and published on social media for four weeks. The within-participant repeated measures order of scenarios were randomized per participant.

11.6.2 Participants

We recruited 100 participants through social media to take the survey. The 100 survey participants had an average age of 34.85 years old ($SD = 13.96$) who came from different professional backgrounds (Engineer 35, Academic 36, Manager 10, Medicine 9, Mix 8, Other 2). Thirty-seven participants wore glasses, six contact lenses, and fourteen both.

11.6.3 Results

Perceived Potential for Misuse

We conducted a one-way repeated-measures ANOVA of Aligned Rank Transformed (ART) Data [247] to investigate the effect of the DEVICE used for zooming on the perceived potential for misuse. Cell frequencies were verified to ensure the reliability of ART. CONTACT LENSES were perceived as most likely to be misused ($M = 4.22$, $SD = 1.13$) with GLASSES ($M = 2.29$, $SD = 1.19$) and HELMET ($M = 3.32$, $SD = 1.27$) scoring lower (see Figure 11.3). There was a significant effect of the DEVICE on misuse, $F(2,198) = 9.05$, $p < .001$. Post-hoc tests revealed that contact lenses scored significantly lower than glasses ($p < .01$) and the helmet ($p < .001$, Bonferroni-Holm corrected p-values).

Perceived Advantage

Next, we investigated the differences in the perceived advantage offered by the three DEVICEs in the three SCENARIOs. To that end, we conducted a two-way repeated-measures ANOVA of Aligned Rank Transformed Data with DEVICE and SCENARIO as the factors and perceived advantage as the response variable.

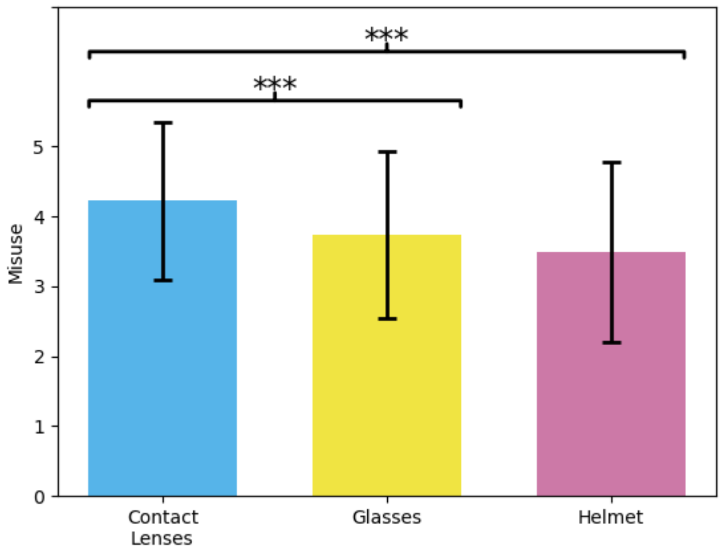


Figure 11.3: Mean value and standard deviation for the perceived misuse scores across the three DEVICEs. Significant differences ($p < .001$) are indicated with $***$.

The results are presented in Figure 11.4. A significant effect, $F(2,792) = 63.60$, $p < .001$, of DEVICE was present. No effect of SCENARIO, $F(2,792) = 0.61$, $p = .54$, or interaction effect, $F(4,792) = 0.54$, $p = .70$, was present. In post-hoc analysis, we found that all condition pairs were significantly different with CONTACT LENSES scoring significantly higher than both GLASSES and HELMET (both at $p < .001$). The HELMET was also ranked as producing significantly less advantage than GLASSES, $p < .05$. All p-values were Bonferroni-Holm corrected. The scores for the three DEVICEs are shown in Figure 11.5.

Perceived Permissions and Useful Contexts for SA

With one being low and five being high, respondents rated the augmentation with above-average acceptability if everyone carried around binoculars ($M = 3.38$, $SD = 1.24$) and if binoculars were already integrated into other devices ($M = 3.59$, $SD = 1.10$) . Participants rated the expected easiness of learning to use a VZT on an everyday basis as intermediate ($M = 2.53$, $SD = 1.07$). Participants were

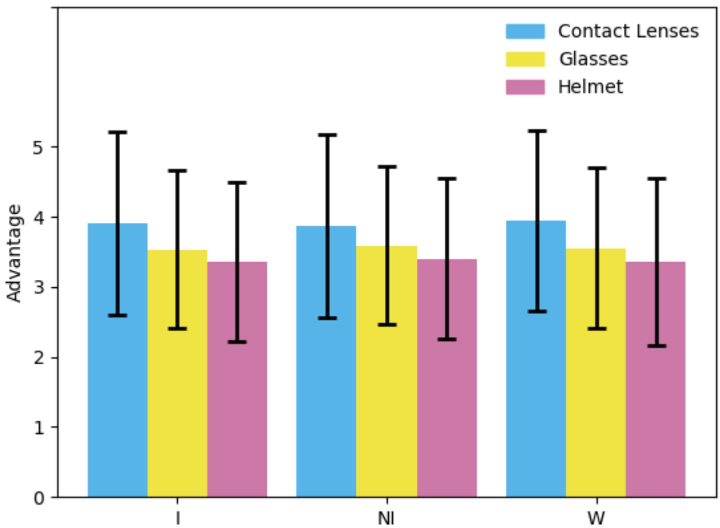


Figure 11.4: Perceived advantage scores across the SCENARIOs and DEVICEs. No significant difference was found across SCENARIOs.

split, where roughly half of them ($n = 51$) thought special permission would be required to use augmented zoom technology, and the remainder ($n = 49$), thought no special permission is needed. The general agreement across participants was that VZT should be available to law enforcement was ($M = 3.44$, $SD = 1.25$).

11.7 Discussion

The results indicate that contact lenses are perceived to offer the highest advantage and highest potential for misuse. Across three devices, the contact lenses offer the most inconspicuous appearance, followed by the glasses, with the helmet being a clearly salient feature in most contexts. This gradient of visibility is reflected consistently in both perceived advantage and potential for misuse, suggesting a proportional relation between them, and an inverse proportionality to visibility. This behavior is consistent across scenarios, suggesting that perceived advantage

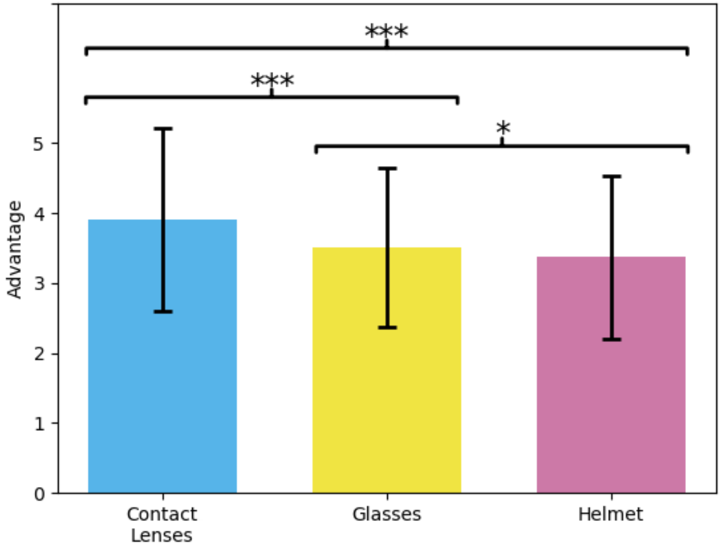


Figure 11.5: Perceived advantage scores across DEVICEs. Significance differences of $p < .05$ and $p < .001$ are indicated respectively with $*$ and $***$.

and misuse are independent of the role in the interaction of the person assessing these traits.

These findings suggest that the visibility of a zooming SA has a direct effect in its perceived potential to be misused, thus affecting its social acceptability. A possible reason for this behavior is that third parties becoming aware of a potential threat to their privacy are better capable to take proactive or palliative measures against the threat. Additionally, the exposure of the user to public scrutiny can be perceived as a deterrent against misuse. We consider the underlying factors of this phenomenon to be general features of SA and thus cautiously suggest that our finds can be broadly extended to all applications extending the human senses. Hence, designing interactions and producing wearable devices intentionally noticeable could increase their social acceptability, which is aligned with the *candidness* and *transparency* highlighted by Koelle et al. as a determining factor in social acceptability [127].

However, the aesthetic value of technological devices has also an effect on its acceptance [108]. Finding an optimal compromise would possibly foster the incorporation of wearable technologies in general. This resonates with the concept of *seamful* interaction, in opposition to seamless [99], which proposes a series of valuable gains for the user experience at the expense of minimalist simplicity and comfort.

The first research question guiding this investigation inquired about the social acceptability of SA. Our findings suggest that the social acceptability of SA is a continuum and the product of multiple factors. In our research, we identified some of these factors and provided a first assessment on how they influence acceptability. The visibility of a SA was observed to be inversely proportional to the potential to misuse assessed by participants, thus answering RQ6.3 positively. Given the social implications of the potential for misuse of technology, it follows that acceptability is affected by visibility. The nature of this relationship is still unclear and further investigation is necessary to form conclusions. However, our work suggests that it is possible to increase the social acceptability of SA technologies by making them more visible.

Another factor that influences social views on SA is the asymmetry in the interaction between users and third parties. All participants from the interviews and surveys saw a clear potential for misuse of the technology, intrinsic to the asymmetry of perceptual capabilities. This answers our second research question positively. Interestingly, the assessment of risk was generally independent from the role of the participant in the interaction.

Finally, the observed data suggests the existence of a strong link between perceived advantage and potential misuse, thus answering RQ6.4. Given the logical correlation between advantage and usefulness, our results indicate that this relationship does play a role in the social acceptability of SA.

11.8 Limitations and Future Work

Our work presents limitations inherent to the user methods: larger numbers of interviews and surveys can possibly produce richer insights about the investigated methods. Further, our findings are limited to the proposed levels of visibility and interaction scenarios. Future work should improve this assessment with a larger variety of devices and scenarios, including additional levels of visibility.

Additionally, future work should investigate the relationship between visibility and social acceptance in more detail, by taking into account the aesthetic and subjective expectations users have for wearable devices, and thus aiming to find an optimal compromise.

Further, the reported work is limited to a particular application, namely visual magnification. We believe that our findings can be generalized to other technologies but ultimately, this should be investigated separately, assessing this phenomenon with other SA technologies. Some possible alternatives are the analogous version of magnification for the other senses, thus enabling to hear what others say from long distances, as well as enabling to smell, taste, or touch other individuals. We envision these three last augmentations as potentially highly controversial and an interesting study topic for HCI and other disciplines, such as Psychology and Sociology.

11.9 Conclusion

In this chapter, we present data obtained through qualitative interviews and a survey about aspects of social acceptability of SA technologies. We investigated the perceived potential of misuse, and thus risk, of wearable technologies that provide a perceptual advantage to the user. Our findings suggest a strong link between the perceived adversarial advantage that technology provides and its potential to be misused. Further, the perceived advantage and misuse potential of a technology decrease for increasing visibility of the wearable technology.

These findings have implications for the design of AR technologies and particularly for SA. The trend recognized in past work suggests a continuous reduction in the visibility of AR technology, which improves its wearability and aesthetic value. However, our findings suggest this might have a negative impact in the social acceptability of these technologies. Thus, transparency should be considered as a highly relevant factor when designing interactions that involve SA.

We believe that SA will face challenges in terms of social acceptability in the coming years. Our work identified critical factors influencing acceptability and suggests that ensuring transparency of the interaction by means of visibility is a promising strategy to address these challenges.

Chapter 12

Lessons Learned: Designing Sensory Augmentation

> If knowledge can create problems, it is not through ignorance that we can solve them.
>
> Isaac Asimov

User-Centered Design (UCD) is a fundamental tool in HCI research and especially in the development of SA. The UCD process standard [100] describes different usability principles, planning, and activities with a high level of abstraction [106]. This abstraction is necessary to enable the application of this process to a wide range of design problems. However, within a certain area of application, the refinement and adaption of the process to the specific characteristics of the area can provide a clearer framework to address individual design problems. We propose a Design Procedure for Sensory Augmentation (DPSA) based on the knowledge gained through the work conducted in this thesis. In this chapter we present the design procedure, provide practical recommendations, and discuss its differences and commonalities respect the standardized UCD process.

12.1 Design Procedure for Sensory Augmentation

As described by Edelson, "the design procedure specifies the processes and the people involved in the development of a design" [63]. Design is a highly complex activity that involves multiple areas of expertise and a combination of systematic and creative processes. To manage the complexity of designing SA, we define a design procedure that specifies the requirements, methods, and steps to successfully design, implement, and evaluate SA applications.

The DPSA relies on the assumption of a predefined goal in terms of what additional sensory information should be presented to the user. It consists of five stages, with the last two stages forming an iterative cycle to be repeated as often as necessary (see Figure 12.1).

Our design procedure resembles the UCD process since it is partially based on this approach. However, the proposed procedure is complementary to the UCD process and aims to serve as a practical development tool whereas the UCD process provides a theoretical design structure and context. A detailed comparison highlighting differences and commonalities between the two approaches can be found in Section 12.2.

12.1.1 Select the supporting sense

The first step of the workflow consists in the selection of a supporting sense. This choice may be trivial in many cases, e.g. if the goal of the augmentation is to improve vision in low visibility conditions, the supporting sense of choice will be the vision in most scenarios.

However, some augmentation applications offer more flexibility. For instance, conveying the temperature of a given object could be achieved using vision, sound, vibrations, or thermal stimuli. In these cases, the choice will depend on many factors: the nature of the information to be conveyed, its urgency, the frequency with which it will be conveyed, the environmental context in which this information will be conveyed, etc.

Table 12.1 provides a list of general recommendations for the selection of sensory domains. A detailed reasoning for these recommendations, as well as further insights on the selection of supporting senses, can be found in Chapter 10.

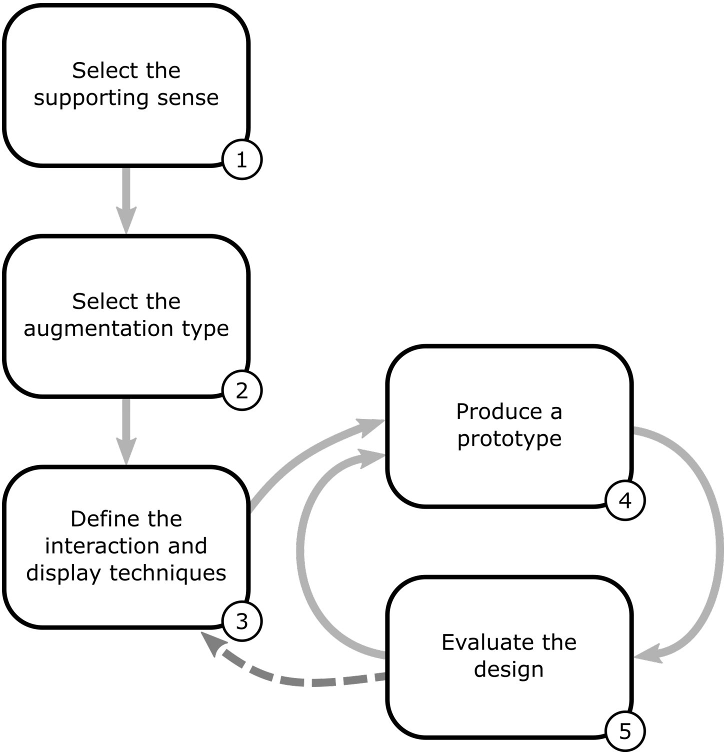


Figure 12.1: Design Procedure for Sensory Augmentation. The first three stages provide a preliminary design, while the last two refine iteratively.

12.1.2 Select the augmentation type

As discussed in Chapter 5, the choice of augmentation type is closely related to each particular application. In many cases, the choice is trivial, since the goal of the application will dictate the type of augmentation. For instance, enabling users to visually magnify objects will be achieved optimally through *amplification*.

In other cases, the designer must select an augmentation type between possible alternatives. In this case, it is recommended to choose the method that ensures the closest resemblance to natural perception. This is aligned with the recommendations for reducing cognitive load and stress from Chapter 10. Thus, *annotation* augmentations should only be used when every other type is incapable

Table 12.1: General recommendations for the selection of supporting senses.

Balance cognitive load	When the augmentation aids a visual task, avoiding the unnecessary use of visual output can reduce the distraction caused by the feedback. The same can apply for auditory tasks and visual output.
Vision for focus	Visual feedback allows the delivery of highly complex information. Additionally, it allows the user to navigate the information at will, permitting skimming and quick focus switch.
Hearing for multitasking	Auditory output allows the user to focus visual attention in other activities, but complex information such as speech can only be presented linearly.
Touch for harsh conditions	Touch feedback, such as vibration or temperature stimuli, can be used to convey simple cues and is particularly useful when auditory feedback is inadequate.
Avoid smell and taste	Smell and taste outputs are generally not recommended due to the technical difficulty of their implementation and signal latency.

Table 12.2: Summary of Sensory Augemntation types and some examples. For a complete overview, see Chapter 5.

Type	Definition	Examples
Amplification	Perceive circunstacially unavailable phenomena.	binoculars, directional microphone
Extension	Perceive beyond natural sensitivity.	night vision, microscope
Enhancement	Synesthesia, perceive outside the domain of a sense.	metal detector, radar
Annotation	Symbolic or abstract representation of reality.	navigation arrow, text
Addition	Purely new sense.	transcraneal brain stimulation

of delivering the desired information, and *enhancement* only when *extention* and *amplification* cannot provide the target functionality.

Table 12.2 provides an overview and explanation of the augmentation types discussed in the design space.

12.1.3 Produce a prototype

A prototype is necessary for the evaluation of SA applications. Admittedly, approaches such as *design fictions* [26] can foster critique and conceptual development, but physical prototypes often a tangible dimension that allows a more thorough investigation. In the early stages of the design process, prototypes generally serve as demonstrators with the main purpose of illustrating an augmentation design to users. Given the novel nature of SA, most users are not familiar with these technologies. Therefore it is critical to provide tangible examples when involving them in, for instance, the determination of requirements, expectations, or concerns about these technologies.

In later stages of the design process, the prototype should gain fidelity and functionality, becoming the vehicle of more sophisticated evaluations to assess performance, user experience, usability, and other success criteria.

The implementation of physical prototypes carries technical challenges that are beyond the scope of this work. However, it is worth to review some *good prototyping practices* that are generally useful, and in our experience, also apply to the particular case of SA. In the following list, we summarize the ten most relevant practices.

Good Prototyping Practices

1. Identify the target functionality to be implemented and clearly define the goals of the prototype.
2. Consider that some features desired in a final design may not be relevant for the purpose of the prototype, while some others that are not necessary for the final design might be relevant for its development and evaluation.
3. Make an early assessment to decide if a physical prototype is the most adequate approach, or if alternative methods (e.g. Wizard of Oz) can provide the desired functionality with fewer resources and less time.
4. If the purpose of the prototype is to be used in the context of a user study, ensure that the implemented system cannot harm or endanger participants or experimenters in any way.
5. Start implementing simple, low-fidelity prototypes to investigate individual functionalities.
6. Take advantage of the numerous and versatile prototyping platforms, such as Arduino or Processing.
7. Involve potential users in the iterative prototyping process through pilot studies to refine the user interface of the prototypes.
8. Gradually improve the fidelity of prototypes and combine the individual functionalities, once each of them has been successfully implemented.
9. When possible, avoid reusing parts of early prototypes in the following iterations. Sometimes, it is helpful to go back to earlier versions. Further, debugging systems with repurposed components tends to be more complex and taxing.
10. Test prototypes thoroughly in pilot studies to detect failures or inconsistent behavior, before launching large-scale user studies where malfunction will result in larger losses of time, effort, and resources.

12.1.4 Evaluate the design

The evaluation of the design for the particular case of SA overlaps with the Activity 4 of the UCD process. We have found that developing SA is hardly a linear process and repetitions of the prototype implementation step, or even revising the interaction and input techniques, are necessary multiple times.

Especially in the early stages, the design stage involves the active evaluation of intermediate solutions through interviews, focus groups, or pilot studies, to determine design aspects of the solution. Also, given the novel nature of SA, determining user requirements may require a tangible prototype since many aspects of the interaction can be hard for participants to imagine without concrete references. However, for this initial development phase, a full-blown user study may be an unnecessary expenditure of resources.

Evaluating single functionalities of the design, as seen in Chapter 7, results in simpler yet more robust experiment designs and contributes to the reliability of the results and the findings derived from them. Evaluating multiple functionalities simultaneously can reduce the costs of the experiment, but it might result in the addition of undesired confounding variables. Similarly, to the recommendations for the implementation of prototypes, we suggest a progressive and incremental evaluation of the design, assessing separately effectiveness, functionality, usability, and user experience.

As discussed in Chapter 8, the evaluation of SA applications should be conducted in realistic use case scenarios, particularly for the later stages of evaluation. This may pose particular challenges, as exemplified in the evaluation of Clairbuoyance, a design intended to aid swimmers in open waters. However, it is critical to assess the user experience and the usability of the augmentation in an environment and with a task that closely resemble the target use case. For the particular case of Clairbuoyance, we closely emulated the critical factors at play in a real-life scenario, in a safe and manageable experimental setup. However, the challenge of creating realistic use-case scenarios is not limited to safety issues but may reside in ensuring consistency across trials of an experiment. For instance, the study of directional hearing reported in Chapter 7 required the simulation of a noisy environment to provide consistent cues, noise, and SNR to participants.

Evaluating how effective is an augmentation in delivering additional information to the user is usually achieved by having experiment participants perform predefined tasks and collecting information about their performance. Despite the inspiration provided by previous work, we have encountered that the creation of tasks has also a remarkable creative component, since the object of evaluation is novel by definition.

Through the work conducted in this thesis, we have used multiple HCI research and design methods, with different degrees of success. Traditional quantitative experiments, in which participants perform a task and metrics such as task completion time and error rate are collected, have proved to be a reliable method of evaluation. They provided insights on the performance of the system and to what extent the augmentation improves human skills. They also provide indirect feedback about the quality of the experiment design, indicating, for instance, the need for better calibration of the difficulty of a task (see Study II in Chapter 7). Along these lines, pilot studies are highly recommended, especially for designs involving physical prototypes, as exemplified in Chapters 7 and 8.

Workshops, interviews, and focus groups have also played an important role in the design and evaluation of SA during our work. These tools helped us to understand the expectations, concerns, and sometimes prejudices that people may have about SA. These methods produce valuable data but present the particular challenge of consistently conveying a novel concept to participants. The use of prototypes and props helped to reduce this effect. For instance, during the Future Technology Workshop (FTW) reported in Chapter 7, participants were shown a video illustrating the SA, and in the interviews reported in Chapter 11, interviewees interacted with a physical prototype. However, experiments must place special care in managing this problem and consider that despite all possible efforts, some limitations cannot be avoided. For instance in Chapter 7, participants of the FTW provided valuable feedback on many aspects of the design for HA but their proposed interaction method, once implemented, was disliked by participants of the user study, who clearly stated to prefer something different.

Similarly to interviews, the use of surveys can produce valuable data to understand how users perceive a SA technology, even more so in terms of data volume. However, the challenge of conveying novel concepts is also increased by this medium, since surveyees cannot ask questions or interact with a prototype. We recommend to use a combination of both, as described in Chapter 11, first collecting general feedback through face-to-face interviews, and then use it to design a survey.

The assessment of usability and user experience is supported by well-established methods, such as the System Usability Scale [14] and the NASA Task Load Index [88]. In our experience, the use of post-trial semi-structured interviews with participants provided valuable insights that helped to articulate the results obtained through the above-mentioned forms.

Our work did not require the use of ethnographic studies. However, we envision that this method will be useful in the coming years, once SA gains ubiquity, and will help understand how the augmented sense affect both users and society. On a more technical perspective, the study of SA can benefit by the development of metrics and methods to determine aspects of the augmentation such as sensory throughput or quantified attention. We expect this research topic will gain traction in the coming years.

12.2 Commonalities and differences with the ISO User-Centered Design process

Both the DPSA and UCD process provide a structured approach to design interactions with a clear emphasis on the involvement of users in the process. Further, both approaches share a similar structure and phases, differing mostly in their focus and specificity. We see the DPSA as a specialization of the UCD process applied to SA. In this section, we explain the differences between these two approaches. But first, we provide a summary of the ISO UCD process [100].

12.2.1 The ISO User-Centered Design process

The UCD process has six phases, or *activities* as defined in the standard (see Figure 12.2). Further, the standard also defines the rationale, principles, and planning of UCD.

The rationale for UCD can be translated to benefits for the user in terms of usability. The principles of UCD identify four aspects: involvement of the user, mediation between users and technology, iteration of design solutions, and multi-disciplinary design. The planning of UCD provides recommendations for a structure approach to the process and emphasizes the importance of reserving time for iterations and user feedback.

The activities of the UCD process have a defined order (see Figure 12.2), starting with (0) the identification of the need for UCD. The model then moves to (1) understanding and specifying the context of use of the design, followed by (2) specifying the requirements of the target user and organisation. The next activity is (3) the production of design solutions, followed by (4) the evaluation of the solution against the specified requirements. At this point, the results of the evaluation determine if a new iteration is needed and steps 1 to 4 are repeated, or if (x) the system satisfies the requirements, the design process is concluded.

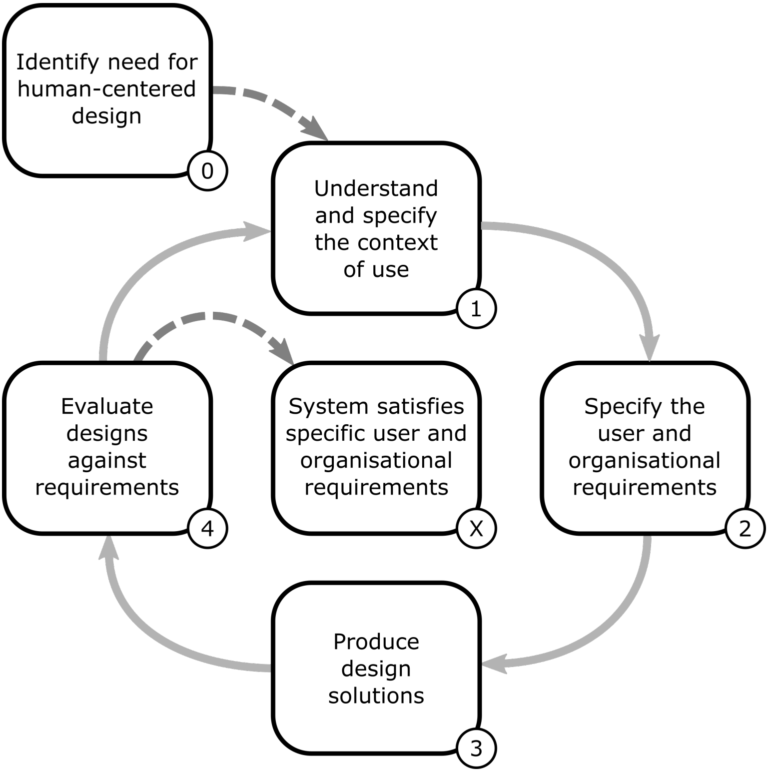


Figure 12.2: Activities of the User-Centered Design, as described in the ISO standard [100].

12.2.2 Key differences

The main difference between the two approaches is their level of abstraction: while UCD is a general method that can be applied to many fields in HCI, DPSA is particularly defined for SA. This distinction pervades all aspects of the design process.

The rationale for DPSA are the same as those of UCD, namely the benefits for the user in terms of usability and user experience, but in this case specific, for SA.

The principles remain the same, but redefined: involvement of the users is intrinsically necessary for the development of SA, as exemplified by the design and evaluation of ESH in Chapter 7. Mediation between users and technology gains a new meaning in the context of DPSA, since SA mediates between users and reality. What in UCD are functions between users and technology, in DPSA are tools to manage human perception, and thus the appropriate allocation of these functions is directly linked to the measure of success of SA design. The iteration of design solutions is also an integral part of the DPSA, although different from the UCD process. The design of SA requires a more flexible strategy, resorting to shorter cycles in the initial phase and complete cycles for the final stages of the process. This reflects the need of an exploratory phase at the beginning of the process, which enables the rapid creation of an initial sketch of the design, jointly with a crude prototype, to enable early user feedback to help define the concept. This is necessary to circumvent the difficulty of involving users in the design of a concept usually foreign to their experience and worldview, as was the case for ESH in Chapter 7. Finally, multidisciplinary design does not differ conceptually between the two methods. In our particular experience, DPSA involved collaborations with experts, as well as the use of resources, from diverse disciplines, such as HCI, Computer Science, Psychology, Design, Sports, Electronics, Ergonomics, and Philosophy.

We recognize a general conceptual overlapping between the *activities* from the UCD process and the stages in the DPSA, particularly in the design-implementation-evaluation path. This is to be expected, since the UCD process aims to be a general template and these archetypical steps are arguably intrinsic to any design process that involves the evaluation of the design. However, DPSA has a set of fundamental requirements that apply to users and usage universally, and while the particular design goals may differ, the success criteria are the same. The DPSA places particular emphasis in the critical aspects of the process and provides an efficient workflow for the conception, realization, and evaluation of SA applications.

Admittedly, the UCD process can be applied to SA problems, and the DPSA can be used for applications that do not involve SA. However, the key benefit of the DPSA resides in the reduction of overhead design activities

Chapter 13

Conclusion

> All our knowledge is the offspring of our perceptions.
>
> Leonardo da Vinci

This thesis explored the concept of SA following a user-centered approach. This goal was achieved through research probes involving participatory design techniques, physical prototypes, and user studies. The following sections summarize the contributions of our research, offer answers to the research questions that guided this thesis, discuss the limitations of the conducted research, and provide perspectives for SA.

13.1 Summary of Research Contributions

In this thesis, we offered a definition for SA within the context of HCI. Further, we identified instances of SA in the literature, which allowed us to formulate a design space. Inspired by a vision for SA, we conducted three research probes, which resulted in a system for improving auditory attention, a system for aiding swimmers in open waters, and a ubiquitous virtual microscope. Further, we proposed a design space for HA. Additionally, we investigated the implications of SA for users and society as a whole. Finally, we provided a set of design recommendations to aid future research and development of SA applications. In the following, these contributions are discussed individually.

Definition of Sensory Augmentation

In Chapter 3, we recognized the need for a formal definition of a research topic that has already been an object of research but was yet never formalized. We proposed a definition for SA, framing within the theoretical context from classic work in HCI. This definition does not only provide a structural fundament to the work conducted during this thesis, but also serves as emphasis for the identity of this field, setting the basis for future research efforts.

Design Space for Sensory Augmentation

In Chapter 5, we proposed a design space for SA based on the literature. This served two purposes: first, to classify existing work, thus highlighting gaps in the systematic investigation of SA, and second, to aid in the ideation of novel SA concepts.

Enhanced Selective Hearing

In Chapter 7, we presented the design and evaluation process of a system to augment the classical sense of hearing. Through an iterative, user-centered design process, we developed a system capable of increasing the human ability of selective auditory attention.

Improved Orientation for Swimmers

In Chapter 8, we presented a wearable device to aid swimmers in open waters. The system was implemented in a physical prototype and evaluated in a controlled experiment, providing insights on the performance and usability of the application, along valuable insights for general SA applications.

Virtual Ubiquitous Microscope

In Chapter 9, we reported on the design and evaluation of a Virtual Ubiquitous Microscope. This project allowed us to gain a deeper understanding of input and output methods for visual augmentations.

Design Space for Hearing Augmentation

Based on the knowledge gained in Chapter 7, we proposed a design space for Hearing Augmentation. This space offers a higher level of detail and specificity than the design space discussed in Chapter 5, both aiding in the classification and ideation of novel applications, and highlighting the possibilities for expanding and refining the design space for SA.

Implications of Sensory Augmentation

In Chapters 10 and 11, we discussed the implications of SA for users and society. Through interviews, surveys, and literature review, we identified challenges and opportunities for the deployment of SA applications, and consequently produced design recommendations.

Design Recommendations

In Chapter 12, we presented a set of design recommendations based on the knowledge gained throughout the work conducted for this thesis. We hope these recommendations can aid designers and researchers in future efforts within the novel field of SA.

13.2 Research Questions

In Chapter 1.6, we formulated a set of Research Questions to guide the investigation on SA. In this section, we address these questions with knowledge gained through the efforts that resulted in this thesis:

RQ1: How to structure a design space for SA?

In Chapter 4, we conducted a literature review to identify work involving the augmentation of the human senses. Based on the resulting corpus of work, we articulated a design space for SA on two main axes: the type of augmentation and the supporting sense. We propose this design space as a first general approach to this novel topic that allows to easily classify existing work and identify yet unexplored areas of research. However, as shown in chapter 7, it is possible to propose more specific design spaces for each sense and application scenario. This suggests the feasibility of taking a similar approach for each sense and application scenario, resonating with the work of MacLean et al. [148].

RQ2: Can SA improve perception through classical senses?

We showed empirically that the augmentation of classical senses is possible, and further, it improves perception. This was achieved explicitly in Chapter 7 and implicitly in Chapter 9. In the first case, we showed how enabling users to increase their control over auditory attention selectivity resulted in a better performance for information retrieval. In the second case, we did not evaluate the benefits of augmentation (in this case, ubiquitous microscopic vision) but it is possible to assume that such a system would allow for a richer observation of reality.

RQ3: Can SA improve perception through non-classical senses?

We showed in Chapter 8 how SA can help swimmers orient themselves in environments without sufficient visual cues. Our results suggest that augmenting non-classical senses is possible, although the nature of non-classical senses diverges among them. We believe that our work shows that other non-classical senses can be augmented to benefit of the user, albeit further research is necessary before drawing a more conclusive answer to this question.

RQ4: What type of interfaces are suitable for SA?

In Chapter 9, we presented an evaluation of different input and output methods based on a user study. Our findings suggest that gestures are not recommendable over voice or tangible controllers. Additionally, using small virtual displays associated to real objects showed advantages over covering the whole field-of-view of users in terms of performance, time, and task load. However, participants of the experiment preferred this visualization type, stating they would like to have both types available and easily switch between.

RQ5: What implications presents SA for users?

Making generalizations about technologies that are currently not available is challenging. Thus, to minimize speculative arguments, we can only answer from the narrow perspectives offered by our probes and the existing literature. In Chapter 6 of this thesis, we have shown is that augmenting the senses can benefit users in terms of control over the amount and quality of information they obtain from reality. However, increments in the amount of information that humans perceive result also in some negative effects, among which we count stress and additional effort to process information, as discussed in Chapter 10. We formulated recommendations for minimizing these effects, but further research is required to understand the extent to which SA can negatively affect human perception and general well-being.

RQ6: What are the social implications of SA?

We identified privacy as a social concern for SA technologies. Through data collected in interviews and surveys presented in Chapter 11, we confirmed that people perceived a threat in adversarial usage of SA. This threat arises from the asymmetry of sensory power created by augmentations, posing as the main reason for opposition against social acceptance.

Our research suggests that the visibility of the physical system providing the augmentation is perceived as an indicator of its potential for misuse, with less visible systems posing a stronger threat. This hints that the transparency of the interaction plays a strong role in the social acceptability of interactive technologies. This phenomenon could be leveraged to improve social acceptance of SA applications by making devices and interactions obvious.

13.3 Limitations

We have answered our Research Questions to a significant degree but we recognize the limitations of the conducted research. Although the reviewed literature is arguably the most relevant for the scope of this thesis, we have hardly exhausted the potential to find further examples and instances of unlabeled SA applications in previous work. We are convinced these likely additions will not modify the proposed design space or definition in a significant way but they will surely enrich and reinforce our understanding of SA.

We have conducted an exploration of a research topic that had so far no formal definition, thus lacking strong foundations to build upon. We have based our theoretical framework on classical HCI work but recognize the possibility of additional perspective contributing to a much stronger and deeper theory for SA.

Our assessment of the design space and possibilities was guided by traditional values and goals of HCI. However, we recognize that future work should conduct a systematic investigation of many aspects left unexplored by our work, such as the ranges and throughputs for each sense.

13.4 Perspectives for Sensory Augmentation

In this thesis, we have begun the exploration of SA. We have acknowledged its importance as a separate research topic by merit of its potential impact on how humans interact with machines and reality, of its existence in the literature, even without being named as such, and for its recognizable fit within the existing theoretical framework. Our probe-based exploration is far from exhaustive: we have only started exploring a research field with the potential to revolutionize HCI, setting the basis for a new interaction paradigm. We hope our work will inspire other researchers and developers, who will chart this mostly unexplored space. We also hope that our work will serve as a guide, even if crude, to approach this very promising subject.

In our work, we have learned that the development and deployment of SA will likely face several challenges. Beyond the technical requirements, which we expect will be solved in the short term, there is still much work needed in terms of usability, user experience, and also acceptability. Further, SA is still seen by many as merely accessibility, a very important field of research, but drastically different in its goals and methodology. Additionally, the augmentation of the human senses can cause complex ethical problems in terms of privacy and social order, which fall well beyond the scope of our work and the expertise of those involved in it.

We conclude by emphasizing again the potential of SA to help people in all kinds of activities, from trivial tasks to life-saving missions. Even more, these technologies have the potential of producing a much richer experience of reality.

Chapter 14

Future Work

> Perception is naturally surpassed toward action; better yet, it can be revealed only in and through projects of action. The world is revealed as an "always future hollow", for we are always future to ourselves.
>
> Jean-Paul Sartre

We hope that our work has opened the doors for new ideas and designs based on the augmentation of the human senses. Some suggestions for further paths of research are discussed individually in the respective *future work* sections of the chapters reporting on our research probes. However, we envision a larger corpus of work that the research reported in this thesis can inspire. In this chapter, we suggest possible future work topics as a direct continuation of our work, as well as research possibilities for a larger scope of SA.

Parts of this chapter are based on the following publication:

- F. Kiss, S. Mayer, and V. Schwind. Audio VR: Did Video Kill the Radio Star? *Interactions*, 27(3):46–51, Apr. 2020

The work conducted in the context of this thesis provided insight in SA to address the research questions satisfactorily. However, the topics highlighted by the questions remain vastly unexplored. First, the design space for SA provides a structure for analysis and classification, but extensions and refinements are certainly possible, as demonstrated in Chapter 7. We envision further design spaces specific to sensory domains and maybe even augmentation types. The literature review served as a basis to develop our design space, yet a larger review, including further venues of publications and also other scientific disciplines, will likely yield a richer corpus of SA applications. The exclusion of *accessible technologies* from our review may have also resulted in the exclusion of systems and designs that could potentially be used for SA, despite this not being the original intent of their creators. Fields of research with highly specialized application areas, for instance, Human-Habitat Multimodal Interaction, have their own publication venues and may contain isolated instances of SA systems.

Our research also suggests further emerging themes that unveil research opportunities, some of them extensive enough to be considered for further Ph.D. thesis topics. For instance, understanding the effects of SA on cognitive load, overstimulation, and stress require a significant amount of work and the investigation of social implications provides an excellent dissertation theme. We envision that parts of these topics can also motivate multiple theses for Bachelor's and Master's courses of studies in Computer Science, Psychology, Sociology, or even Philosophy, and generate various opportunities for multi-disciplinary collaborations.

Our work recognizes the need for a shift in paradigms, moving away from the old interaction metaphors of workstations and past ubiquitous computing. We believe that SA will trigger a change in HCI that will propagate across its disciplines and beyond, reshaping how we interact with computers and reality. We believe that technology will allow users to have complete control over what they perceive, enabling to customize reality to their needs and preferences. In the short term, we see particularly promising opportunities for auditory augmentations, a modality so far overlooked back research when compared to vision. Enabling users to completely reshape the soundscapes they perceive can have a strong impact on the quality of life for people living in crowded cities and being constantly exposed to loud environments. In the long term, we expect a variety of SA applications empowering users to see, hear, taste, smell, and touch much more of reality, maybe even additional senses that we cannot yet imagine. These applications will need to be developed and studied, as well the dramatic implications they might bring to human societies. We also foresee a need for a timely stance on regulations for the development and commercialization of SA applications. Private interests and undemocratic governments may weaponize these technologies in detriment to the public interest and against the ideals that motivate our work. Neglecting to ensure an ethical development and deployment of SA can have terrible and unforeseen consequences, from threats to privacy, discrimination, persecution, or maybe even the direct manipulation of how others perceive reality. HCI researchers must find ways to minimize and prevent these problems through design and research, but also through proactive and extensive discussion of the topic.

Therefore, we dearly encourage the community to pursue this area of research. The opportunities and potential rewards are many, and the necessity for this research is clear. We hope that our work will serve as a strong basis for the suggested research, hinting towards the paths to follow and providing the first tools to design and investigate Sensory Augmentation.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] Y. Abdelrahman, P. Wozniak, P. Knierim, N. Henze, and A. Schmidt. Exploration of alternative vision modes using depth and thermal cameras. In *Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia*, MUM 2018, pages 245–252, New York, NY, USA, 2018. ACM.

[2] G. D. Abowd. Beyond Weiser: From ubiquitous to collective computing. *Computer*, 49(1):17–23, 2016.

[3] M. Ackerman, T. Darrell, and D. Weitzner. Privacy in context. *Human-Computer Interaction*, 16(2):167–176, 2001.

[4] M. G. Akritas. The rank transform method in some two-factor designs. *Journal of the American Statistical Association*, 85(409):73–78, 1990.

[5] R. Albrecht and T. Lokki. Auditory Distance Presentation in an Urban Augmented Reality Environment. *ACM Trans. Appl. Percept.*, 12(2):5:1–5:19.

[6] R. Albrecht, T. Lokki, and L. Savioja. A Mobile Augmented Reality Audio System with Binaural Microphones. In *Proceedings of Interacting with Sound Workshop: Exploring Context-Aware, Local and Social Audio Applications*, IwS '11, pages 7–11, New York, NY, USA, 2011. ACM.

[7] S. American National Standards Institute. Committee on Bioacoustics and M. Sonn. *American national standard psychoacoustical terminology*. American National Standards Institute, 1973.

[8] J. Amores and P. Maes. Essence: Olfactory interfaces for unconscious influence of mood and cognitive performance. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 28–34, New York, NY, USA. ACM.

[9] E. R. Anderson. The five senses. *Folk-taxonomies in early English*, 2003.

[10] P. M. Aoki, M. Romaine, M. H. Szymanski, J. D. Thornton, D. Wilson, and A. Woodruff. The Mad Hatter's Cocktail Party: A Social Mobile Audio Space Supporting Multiple Simultaneous Conversations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '03, pages 425–432, New York, NY, USA, 2003. ACM.

[11] C. Appert, O. Chapuis, and E. Pietriga. High-precision magnification lenses. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 273–282, New York, NY, USA, 2010. ACM.

[12] M. Aytekin, C. F. Moss, and J. Z. Simon. A sensorimotor approach to sound localization. *Neural Computation*, 20(3):603–635, 2008.

[13] M. Bächlin, K. Förster, and G. Tröster. Swimmaster: A wearable assistant for swimmer. In *Proceedings of the 11th International Conference on Ubiquitous Computing*, UbiComp '09, pages 215–224, New York, NY, USA. ACM.

[14] A. Bangor, P. T. Kortum, and J. T. Miller. An empirical evaluation of the system usability scale. *Intl. Journal of Human–Computer Interaction*, 24(6):574–594, 2008.

[15] L. Bartram, A. Ho, J. Dill, and F. Henigman. The continuous zoom: A constrained fisheye technique for viewing and navigating large information spaces. In *Proceedings of the 8th annual ACM symposium on User interface and software technology*, pages 207–215, 1995.

[16] D. Bawden and L. Robinson. The dark side of information: overload, anxiety and other paradoxes and pathologies. *Journal of information science*, 35(2):180–191, 2009.

[17] J. Beare and G. Ross. De sensu et sensibilibus, 1991.

[18] S. I. Becker, A. M. Harris, D. Venini, and J. D. Retell. Visual search for color and shape: When is the gaze guided by feature relationships, when by feature values? *Journal of Experimental Psychology: Human Perception and Performance*, 40(1):264, 2014.

[19] G. Bell and P. Dourish. Yesterday's tomorrows: notes on ubiquitous computing's dominant vision. *Personal and ubiquitous computing*, 11(2):133–143, 2007.

[20] H. Berghel. Cyberspace 2000: Dealing with Information Overload. *Commun. ACM*, 40(2):19–24, Feb. 1997.

[21] M. Berning, F. Braun, T. Riedel, and M. Beigl. ProximityHat: a head-worn system for subtle sensory augmentation with tactile stimulation. In *Proceedings of the 2015 ACM International Symposium on Wearable Computers*, pages 31–38. ACM, Sept. 2015.

[22] V. Best, E. J. Ozmeral, N. Kopčo, and B. G. Shinn-Cunningham. Object continuity enhances selective auditory attention. *Proceedings of the National Academy of Sciences*, 105(35):13174–13178, 2008.

[23] A. Blandford, D. Furniss, and S. Makri. *Qualitative HCI research: Going behind the scenes*, volume 9. Morgan & Claypool Publishers, 2016.

[24] J. Blauert. *Spatial hearing: the psychophysics of human sound localization*. MIT press, 1997.

[25] J. R. Blum and J. R. Cooperstock. Expressing human state via parameterized haptic feedback for mobile remote implicit communication. In *Proceedings of the 7th Augmented Human International Conference 2016*, AH '16, pages 41:1–41:2, New York, NY, USA. ACM.

[26] M. Blythe. Research through design fiction: narrative in real and imaginary abstracts. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 703–712, 2014.

[27] F. Bourgeois and Y. Guiard. Multiscale pointing: Facilitating pan-zoom coordination. In *CHI '02 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '02, pages 758–759, New York, NY, USA, 2002. ACM.

[28] D. A. Bray. Information pollution, knowledge overload, limited attention spans, and our responsibilities as IS professionals. In *Global Information Technology Management Association (GITMA) World Conference-June*, 2008.

[29] S. Brewster and L. M. Brown. Tactons: Structured tactile messages for non-visual information display. In *Proceedings of the Fifth Conference on Australasian User Interface - Volume 28*, AUIC '04, pages 15–23, Darlinghurst, Australia, Australia. Australian Computer Society, Inc.

[30] S. Brewster, D. McGookin, and C. Miller. Olfoto: Designing a smell-based interaction. In *Proceedings of the SIGCHI Conference on Human Factors*

in Computing Systems, CHI '06, pages 653–662, New York, NY, USA. ACM.

[31] D. Broadbent. Perception and communication. Elmsford, NY, US, 1958.

[32] D. E. Broadbent. Effect of noise on an "intellectual" task. *The Journal of the Acoustical Society of America*, 30(9):824–827, 1958.

[33] K. J. Bruggeman and S. W. Wurster. The hiatus system: Virtual healing spaces: Low dose mindfulness based stress reduction virtual reality application. In *ACM SIGGRAPH 2018 Appy Hour*, SIGGRAPH '18, pages 8:1–8:2, New York, NY, USA, 2018. ACM.

[34] Z. Bujas and A. Chweitzer. Goût électrique par courants alternatifs chez l'homme. *CR Soc. Biol.(Paris)*, 126:1106–1108.

[35] A. Bulling and T. O. Zander. Cognition-aware computing. *IEEE Pervasive Computing*, 13(3):80–83, 2014.

[36] J. L. Burke, M. S. Prewett, A. A. Gray, L. Yang, F. R. B. Stilson, M. D. Coovert, L. R. Elliot, and E. Redden. Comparing the effects of visual-auditory and visual-tactile feedback on user performance: A meta-analysis. In *Proceedings of the 8th International Conference on Multimodal Interfaces*, ICMI '06, pages 108–117, New York, NY, USA. ACM.

[37] A. J. Callaway, J. E. Cobb, and I. Jones. A comparison of video and accelerometer based approaches applied to performance monitoring in swimming. *International Journal of Sports Science & Coaching*, 4(1):139–153.

[38] M. G. Carcedo, S. H. Chua, S. Perrault, P. Wozniak, R. Joshi, M. Obaid, M. Fjeld, and S. Zhao. Hapticolor: Interpolating color information as haptic feedback to assist the colorblind. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 3572–3583, New York, NY, USA, 2016. ACM.

[39] P. Cariani and C. Micheyl. Toward a theory of information processing in auditory cortex. In *The human auditory cortex*, pages 351–390. Springer, 2012.

[40] S. Carlile, P. Leong, and S. Hyams. The nature and distribution of errors in sound localization by human listeners. *Hearing Research*, 114(1-2):179–196.

[41] A. Carton and L. E. Dunne. Tactile Distance Feedback for Firefighters: Design and Preliminary Evaluation of a Sensory Augmentation Glove. In *Proceedings of the 4th Augmented Human International Conference*, AH '13, pages 58–64, New York, NY, USA, 2013. ACM.

[42] N. Chaconas and T. Höllerer. An evaluation of bimanual gestures on the microsoft hololens. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 1–8, March 2018.

[43] L. Chaieb, E. C. Wilpert, T. P. Reber, and J. Fell. Auditory beat stimulation and its effects on cognition and mood states. *Frontiers in psychiatry*, 6:70.

[44] A. Chamberlain, A. Crabtree, T. Rodden, M. Jones, and Y. Rogers. Research in the wild: understanding 'in the wild' approaches to design and development. In *Proceedings of the Designing Interactive Systems Conference*, pages 795–796. ACM, 2012.

[45] J. S.-K. Chang, M. J. Henry, R. Burtner, O. Love, and C. Corley. The heroes' problems: Exploring the potentials of google glass for biohazard handling professionals. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, pages 1531–1536. ACM, 2015.

[46] H.-I. Chen, C. Chen, and Y.-Y. Chuang. A deep convolutional neural network for continuous zoom with dual cameras. In *SIGGRAPH Asia 2017 Posters*, SA '17, pages 13:1–13:2, New York, NY, USA, 2017. ACM.

[47] E. C. Cherry. Some experiments on the recognition of speech, with one and with two ears. *The Journal of the acoustical society of America*, 25(5):975–979, 1953.

[48] W. Choi, J. Oh, D. Edge, J. Kim, and U. Lee. Swimtrain: Exploring exergame design for group fitness swimming. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 1692–1704, New York, NY, USA. ACM.

[49] A. R. Conway, N. Cowan, and M. F. Bunting. The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic bulletin & review*, 8(2):331–335, 2001.

[50] E. Costanza, S. A. Inverso, E. Pavlov, R. Allen, and P. Maes. Eye-q: Eyeglass peripheral display for subtle intimate notifications. In *Proceedings of the 8th Conference on Human-computer Interaction with Mobile Devices and Services*, MobileHCI '06, pages 211–218, New York, NY, USA. ACM.

[51] N. Dahlbäck, A. Jönsson, and L. Ahrenberg. Wizard of oz studies—why and how. *Knowledge-based systems*, 6(4):258–266, 1993.

[52] M. R. Dansereau, J. M. Colombi, M. E. Miller, and M. J. Robbins. A design evaluation framework for helmet mounted displays in fighter aircraft. In *IIE Annual Conference. Proceedings*, page 1961. Institute of Industrial and Systems Engineers (IISE), 2015.

[53] N. P. Davey, M. E. Anderson, and D. A. James. Validation trial of an accelerometer-based sensor platform for swimming. *Sports Technology*, 1(4-5):202–207.

[54] N. P. Davey, M. E. Anderson, and D. A. James. An accelerometer-based system for elite athlete swimming performance analysis, 2005.

[55] K. E. DeLeeuw and R. E. Mayer. A comparison of three measures of cognitive load: Evidence for separable measures of intrinsic, extraneous, and germane load. *Journal of educational psychology*, 100(1):223, 2008.

[56] D. Dobler, M. Haller, and P. Stampfl. ASR: Augmented Sound Reality. In *ACM SIGGRAPH 2002 Conference Abstracts and Applications*, SIGGRAPH '02, pages 148–148. ACM.

[57] M. S. Dobreva, W. E. O'Neill, and G. D. Paige. Influence of aging on human sound localization. *Journal of neurophysiology*, 105(5):2471–2486, 2011.

[58] D. C. Dryer, C. Eisbach, and W. S. Ark. At what cost pervasive? a social computing view of mobile computing systems. *IBM Systems Journal*, 38(4):652–676, 1999.

[59] L. Duarte, T. Antunes, and L. Carriço. Can You Feel It?: Sharing Heart Beats with Augmento. In *Proceedings of the 3rd Augmented Human International Conference*, AH '12, pages 4:1–4:5, New York, NY, USA, 2012. ACM.

[60] S. Dubnov, N. Tishby, and D. Cohen. Polyspectra as measures of sound texture and timbre. *Journal of New Music Research*, 26(4):277–314, 1997.

[61] A. Dünser and M. Billinghurst. Evaluating augmented reality systems. In *Handbook of augmented reality*, pages 289–307. Springer, 2011.

[62] M. Ebata, T. Sone, and T. Nimura. Improvement of hearing ability by directional information. *The Journal of the Acoustical Society of America*, 43(2):289–297, 1968.

[63] D. C. Edelson. Design research: What we learn when we engage in design. *The Journal of the Learning sciences*, 11(1):105–121, 2002.

[64] S. Ekuni, K. Murata, Y. Asakura, and A. Uehara. Bionic scope: wearable system for visual extension triggered by bioelectrical signal. In *ACM SIGGRAPH 2016 Posters*, page 41. ACM, 2016.

[65] D. S. Elvitigala, R. Peiris, E. Wilhelm, S. Foong, and S. Nanayakkara. GrabAmps: Grab a Wire to Sense the Current Flow. In *Proceedings of the 8th Augmented Human International Conference*, AH '17, pages 30:1–30:4, New York, NY, USA, 2017. ACM.

[66] M. Emami, Y. Sakali, C. Pritzel, and R. Trettin. Deep inside the ceramic texture: A microscopic-chemical approach to the phase transition via partial-sintering processes in ancient ceramic matrices. *Journal of Microscopy and Ultrastructure*, 4(1):11 – 19, 2016.

[67] D. C. Engelbart. Augmenting human intellect: A conceptual framework. *Menlo Park, CA*, 1962.

[68] J. Etherington. Why open water swimming is the trend of 2015. *The Telegraph*, 2015.

[69] K. Fan, J. Huber, S. Nanayakkara, and M. Inami. SpiderVision: Extending the Human Field of View for Augmented Awareness. In *Proceedings of the 5th Augmented Human International Conference*, AH '14, pages 49:1–49:8, New York, NY, USA, 2014. ACM.

[70] S. Firestein. How the olfactory system makes sense of scents. *Nature*, 413(6852):211–218, 2001.

[71] A. Firouzian, Y. Kashimoto, Z. Asghar, N. Keranen, G. Yamamoto, and P. Pulli. Twinkle megane: Near-eye led indicators on glasses for simple and smart navigation in daily life. In K. Giokas, L. Bokor, and F. Hopfgartner, editors, *eHealth 360*, pages 17–22. Springer International Publishing, 2017.

[72] K. Förster, M. Bächlin, and G. Tröster. Non-interrupting user interfaces for electronic body-worn swim devices. In *Proceedings of the 2Nd International Conference on PErvasive Technologies Related to Assistive Environments*, PETRA '09, pages 38:1–38:4, New York, NY, USA. ACM.

[73] C. Freigang, N. Richter, R. Rübsamen, and A. A. Ludwig. Age-related changes in sound localisation ability. *Cell and tissue research*, 361(1):371–386, 2015.

[74] J. Frich, L. MacDonald Vermeulen, C. Remy, M. M. Biskjaer, and P. Dalsgaard. Mapping the landscape of creativity support tools in hci. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pages 389:1–389:18, New York, NY, USA, 2019. ACM.

[75] V. Gentile, D. Fundarò, and S. Sorce. Elicitation and evaluation of zoom gestures for touchless interaction with desktop displays. In *Proceedings of the 8th ACM International Symposium on Pervasive Displays*, PerDis '19, pages 16:1–16:7, New York, NY, USA, 2019. ACM.

[76] E. B. Goldstein. *Sensation and perception. 8th edition.* Thomson Wadsworth, 2010.

[77] B. M. Gorman. VisAural:: A Wearable Sound-localisation Device for People with Impaired Hearing. In *Proceedings of the 16th International ACM SIGACCESS Conference on Computers & Accessibility*, ASSETS '14, pages 337–338. ACM.

[78] T. D. Griffiths and J. D. Warren. What is an auditory object? *Nature Reviews Neuroscience*, 5(11):887, 2004.

[79] E. Grönvall, J. Fritsch, and A. Vallgårda. FeltRadio: Sensing and Making Sense of Wireless Traffic. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*, pages 829–840. ACM, June 2016.

[80] G. Haas, E. Stemasov, and E. Rukzio. Can't you hear me?: Investigating personal soundscape curation. In *Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia*, MUM 2018, pages 59–69, New York, NY, USA, 2018. ACM.

[81] R. M. Hagema, T. Haelsig, S. G. O'Keefe, A. Stamm, T. Fickenscher, and D. V. Thiel. Second generation swimming feedback device using a wearable data processing system based on underwater visible light communication. *Procedia Engineering*, 60:34 – 39, 2013. 6th Asia-Pacific Congress on Sports Technology (APCST).

[82] M. Hagood. Quiet comfort: Noise, otherness, and the mobile production of personal space. *American Quarterly*, 63(3):573–589, 2011.

[83] G. Hamilton-Fletcher, M. Obrist, P. Watten, M. Mengucci, and J. Ward. "i always wanted to see the night sky": Blind user preferences for sensory substitution devices. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 2162–2174, New York, NY, USA. ACM.

[84] J. P. Hansen, A. W. Andersen, and P. Roed. Eye-gaze control of multimedia systems. In *Advances in Human Factors/Ergonomics*, volume 20, pages 37–42. Elsevier, 1995.

[85] E. R. Harper, T. Rodden, Y. Rogers, A. Sellen, B. Human, et al. Human-Computer Interaction in the year 2020. 7 J J Thomson Avenue, Cambridge, CB3 0FB, England, 2008. Microsoft Research Ltd.

[86] R. H. R. Harper, M. G. Lamming, and W. M. Newman. Locating systems at work: implications for the development of active badge applications. *Interacting with Computers*, 4(3):343–363, 1992.

[87] J. Hart, D. Onceanu, C. Sohn, D. Wightman, and R. Vertegaal. The attentive hearing aid: Eye selection of auditory sources for hearing impaired users. In *IFIP Conference on Human-Computer Interaction*, pages 19–35. Springer, 2009.

[88] S. G. Hart. Nasa-Task Load Index (NASA-TLX); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(9):904–908, 2006.

[89] S. Hasegawa, S. Ishijima, F. Kato, H. Mitake, and M. Sato. Realtime Sonification of the Center of Gravity for Skiing. In *Proceedings of the 3rd Augmented Human International Conference*, AH '12, pages 11:1–11:4, New York, NY, USA, 2012. ACM.

[90] F. Heller and J. Borchers. AudioScope: Smartphones As Directional Microphones in Mobile Audio Augmented Reality Systems. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 949–952. ACM.

[91] F. Heller and J. Borchers. AudioTorch: Using a Smartphone As Directional Microphone in Virtual Audio Spaces. In *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services*, MobileHCI '14, pages 483–488. ACM.

[92] M. Hilbert. How much information is there in the "information society"? *Significance*, 9(4):8–12, 2012.

[93] M. Hilbert. How much of the global information and communication explosion is driven by more, and how much by better technology? *Journal of the Association for Information Science and Technology*, 65(4):856–861, 2014.

[94] S. R. Hiltz and M. Turoff. Structuring Computer-mediated Communication Systems to Avoid Information Overload. *Commun. ACM*, 28(7):680–689, July 1985.

[95] T. Hiraki, S. Fukushima, and T. Naemura. Sensible shadow: tactile feedback from your own shadow. In *Proceedings of the 7th Augmented Human International Conference 2016*, pages 1–4, 2016.

[96] J. Hong. Considering privacy issues in the context of google glass. *Commun. ACM*, 56(11):10–11, 2013.

[97] N. Houri, H. Arita, and Y. Sakaguchi. Audiolizing body movement: its concept and application to motor skill learning. page 13. ACM, Mar. 2011.

[98] O. Hugues, P. Fuchs, and O. Nannipieri. New augmented reality taxonomy: Technologies and features of augmented environment. In *Handbook of augmented reality*, pages 47–63. Springer, 2011.

[99] S. Inman and D. Ribes. Beautiful seams: Strategic revelations and concealments. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pages 278:1–278:14, New York, NY, USA, 2019. ACM.

[100] I. ISO. 13407: Human-centred design processes for interactive systems. *Geneva: ISO*, 1999.

[101] Y. Itoh and G. Klinker. Vision Enhancement: Defocus Correction via Optical See-through Head-mounted Displays. In *Proceedings of the 6th Augmented Human International Conference*, AH '15, pages 1–8, New York, NY, USA, 2015. ACM.

[102] Y. Itoh, J. Orlosky, K. Kiyokawa, and G. Klinker. Laplacian Vision: Augmenting Motion Prediction via Optical See-Through Head-Mounted Displays. In *Proceedings of the 7th Augmented Human International Conference 2016*, page 16. ACM, Feb. 2016.

[103] J. Jacoby. Perspectives on information overload. *Journal of consumer research*, 10(4):432–435, 1984.

[104] D. Jain, L. Findlater, J. Gilkeson, B. Holland, R. Duraiswami, D. Zotkin, C. Vogler, and J. E. Froehlich. Head-Mounted Display Visualizations to Support Sound Awareness for the Deaf and Hard of Hearing. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 241–250. ACM.

[105] N. Jirakittayakorn and Y. Wongsawat. Brain responses to a 6-hz binaural beat: Effects on general theta rhythm and frontal midline theta activity. *Frontiers in neuroscience*, 11:365.

[106] T. Jokela, N. Iivari, J. Matero, and M. Karukka. The standard of user-centered design and the standard definition of usability: analyzing iso 13407 against iso 9241-11. In *Proceedings of the Latin American conference on Human-computer interaction*, pages 53–60. ACM, 2003.

[107] D. Jones and S. Dechmerowski. Measuring Stress in an Augmented Training Environment: Approaches and Applications. In *Proceedings, Part II, of the 10th International Conference on Foundations of Augmented Cognition: Neuroergonomics and Operational Neuroscience - Volume 9744*, pages 23–33, New York, NY, USA, 2016. Springer-Verlag New York, Inc.

[108] M. Kalantari and P. Rauschnabel. Exploring the early adopters of augmented reality smart glasses: The case of microsoft hololens. In *Augmented reality and virtual reality*, pages 229–245. Springer, 2018.

[109] R. Kamien and A. Kamien. *Music: an appreciation*. McGraw-Hill New York, NY, 1988.

[110] S. Kasahara, M. Ando, K. Suganuma, and J. Rekimoto. Parallel Eyes: Exploring Human Capability and Behaviors with Paralleled First Person View Sharing. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 1561–1572, New York, NY, USA, 2016. ACM.

[111] R. S. Keast and A. Costanzo. Is fat the sixth taste primary? evidence and implications. *Flavour*, 4(1):5.

[112] S. Keates. Pragmatic research issues confronting hci practitioners when designing for universal access. *Universal Access in the Information Society*, 5(3):269–278, 2006.

[113] R. Kettner, P. Bader, T. Kosch, S. Schneegass, and A. Schmidt. Towards Pressure-based Feedback for Non-stressful Tactile Notifications. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '17, pages 89:1–89:8, New York, NY, USA, 2017. ACM.

[114] G. Kidd. Enhancing auditory selective attention using a visually guided hearing aid. *Journal of Speech, Language, and Hearing Research*, 60(10):3027–3038, 2017.

[115] G. Kidd Jr, S. Favrot, J. G. Desloge, T. M. Streeter, and C. R. Mason. Design and preliminary testing of a visually guided hearing aid. *The Journal of the Acoustical Society of America*, 133(3):EL202–EL207, 2013.

[116] S. Kim and A. K. Dey. Augmenting human senses to improve the user experience in cars: applying augmented reality and haptics approaches to reduce cognitive distances. *Multimedia Tools and Applications*, 75(16):9587–9607, 2016.

[117] F. Kiss, R. Boldt, B. Pfleging, and S. Schneegass. Navigation systems for motorcyclists: Exploring wearable tactile feedback for route guidance in the real world. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pages 617:1–617:7, New York, NY, USA. ACM.

[118] F. Kiss, S. Mayer, and V. Schwind. Audio VR: Did Video Kill the Radio Star? *Interactions*, 27(3):46–51, Apr. 2020.

[119] F. Kiss and A. Schmidt. 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*, PervasiveHealth'19, pages 377–382, New York, NY, USA, 2019. ACM.

[120] F. Kiss, P. W. Woźniak, V. Biener, P. Knierim, and A. Schmidt. VUM: Understanding Requirements for a Virtual Ubiquitous Microscope. In *19th International Conference on Mobile and Ubiquitous Multimedia*, New York, NY, USA, 2020. ACM.

[121] F. Kiss, P. W. Woźniak, F. Scheerer, J. Dominiak, A. Romanowski, and A. Schmidt. Clairbuoyance: Improving Directional Perception for Swimmers. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pages 237:1–237:12, New York, NY, USA, 2019. ACM. Honorable Mention Award.

[122] J. Kjeldskov and J. Paay. A longitudinal review of mobile hci research methods. In *Proceedings of the 14th International Conference on Human-computer Interaction with Mobile Devices and Services*, MobileHCI '12, pages 69–78, New York, NY, USA, 2012. ACM.

[123] K. Knaving, P. Woźniak, M. Fjeld, and S. Björk. Flow is not enough: Understanding the needs of advanced amateur runners to design motivation technology. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 2013–2022, New York, NY, USA. ACM.

[124] P. Knierim, T. Kosch, G. LaBorwit, and A. Schmidt. Altering the speed of reality? exploring visual slow-motion to amplify human perception using augmented reality. In *Proceedings of the 1st Augmented Humans International Conference*, AHs'20, New York, NY, USA, 2020. ACM.

[125] M. Kobayashi and C. Schmandt. Dynamic Soundscape: Mapping Time to Space for Audio Browsing. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, CHI '97, pages 194–201. ACM.

[126] W. E. Kock. Binaural localization and masking. *The Journal of the Acoustical Society of America*, 22(6):801–804.

[127] M. Koelle, S. Ananthanarayan, and S. Boll. Social acceptability in HCI: A survey of methods, measures, and design strategies. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, pages 1–19, New York, NY, USA, 2020. Association for Computing Machinery.

[128] M. Koelle, S. Boll, T. Olsson, J. Williamson, H. Profita, S. Kane, and R. Mitchell. (un) acceptable!?!: Re-thinking the social acceptability of emerging technologies. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, page W03. ACM, 2018.

[129] N. Koizumi, M. Sugimoto, N. Nagaya, M. Inami, and M. Furukawa. Stop Motion Goggle: Augmented Visual Perception by Subtraction Method Using High Speed Liquid Crystal. In *Proceedings of the 3rd Augmented Human International Conference*, AH '12, pages 14:1–14:7, New York, NY, USA, 2012. ACM.

[130] K. Krombholz, A. Dabrowski, M. Smith, and E. Weippl. Ok glass, leave me alone: towards a systematization of privacy enhancing technologies for wearable computing. In *International Conference on Financial Cryptography and Data Security*, pages 274–280. Springer, 2015.

[131] D. J. Kruger and J. M. Djerf. High ringxiety: attachment anxiety predicts experiences of phantom cell phone ringing. *Cyberpsychology, Behavior, and Social Networking*, 19(1):56–59, 2016.

[132] T. J. Lapis, M. H. Penner, and J. Lim. Humans can taste glucose oligomers independent of the hT1R2/hT1R3 sweet taste receptor. *Chemical Senses*, 41(9):755–762.

[133] J. Lazar, J. H. Feng, and H. Hochheiser. *Research methods in human-computer interaction*. Morgan Kaufmann, 2017.

[134] M. Ledzińska and S. Postek. From metaphorical information overflow and overload to real stress: Theoretical background, empirical findings, and applications. *European Management Journal*, 35(6):785–793, Dec. 2017.

[135] K. Lee and H. Hong. MindNavigator: Exploring the Stress and Self-Interventions for Mental Wellness. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pages 572:1–572:14, New York, NY, USA, 2018. ACM.

[136] M. Lee, M. Billinghurst, W. Baek, R. Green, and W. Woo. A usability study of multimodal input in an augmented reality environment. *Virtual Real.*, 17(4):293–305, Nov. 2013.

[137] C.-T. Li, J. Cao, and T. M. H. Li. Eustress or Distress: An Empirical Study of Perceived Stress in Everyday College Life. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, UbiComp '16, pages 1209–1217, New York, NY, USA, 2016. ACM.

[138] J. Li, Y. Wang, L. Liu, S. Xu, Y. Liu, J. Leng, and S. Cai. A biomimetic soft lens controlled by electrooculographic signal. *Advanced Functional Materials*, 7 2019.

[139] K. Lilija, H. Pohl, S. Boring, and K. Hornbæk. Augmented Reality Views for Occluded Interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pages 446:1–446:12, New York, NY, USA, 2019. ACM.

[140] M.-W. Lin, Y.-M. Cheng, W. Yu, and F. E. Sandnes. Investigation into the feasibility of using tactons to provide navigation cues in pedestrian situations. In *Proceedings of the 20th Australasian Conference on Computer-Human Interaction: Designing for Habitus and Habitat*, OZCHI '08, pages 299–302, New York, NY, USA. ACM.

[141] J. Looser, M. Billinghurst, and A. Cockburn. Through the looking glass: The use of lenses as an interface tool for augmented reality interfaces. In *Proceedings of the 2Nd International Conference on Computer*

Graphics and Interactive Techniques in Australasia and South East Asia, GRAPHITE ’04, pages 204–211, New York, NY, USA, 2004. ACM.

[142] A. M. Ludwig. Self-Regulation of the Sensory Environment: Conceptual Model and Potential Clinical Applications. *Archives of General Psychiatry*, 25(5):413–418, Nov. 1971.

[143] S. S. Lundgaard and P. A. Nielsen. Personalised soundscapes in homes. In *Proceedings of the 2019 on Designing Interactive Systems Conference*, DIS ’19, pages 813–822, New York, NY, USA, 2019. ACM.

[144] Z. Lv, A. Halawani, S. Feng, S. Ur Réhman, and H. Li. Touch-less interactive augmented reality game on vision-based wearable device. *Personal and Ubiquitous Computing*, 19(3-4):551–567, 2015.

[145] D. Lyon. *Surveillance as social sorting: Privacy, risk, and digital discrimination*. Psychology Press, 2003.

[146] L. Lyons, M. Tissenbaum, M. Berland, R. Eydt, L. Wielgus, and A. Mechtley. Designing visible engineering: supporting tinkering performances in museums. In *Proceedings of the 14th International Conference on Interaction Design and Children*, pages 49–58. ACM, 2015.

[147] A. Macfarlane. Olfaction in the development of social preferences in the human neonate. *Parent-infant interaction*, 33:103–113, 1975.

[148] A. MacLean, R. M. Young, V. M. Bellotti, and T. P. Moran. Questions, options, and criteria: Elements of design space analysis. *Human–computer interaction*, 6(3-4):201–250, 1991.

[149] D. MacLean, A. Roseway, and M. Czerwinski. MoodWings: A Wearable Biofeedback Device for Real-time Stress Intervention. In *Proceedings of the 6th International Conference on PErvasive Technologies Related to Assistive Environments*, PETRA ’13, pages 66:1–66:8, New York, NY, USA, 2013. ACM.

[150] F. Macpherson. Taxonomising the senses. *Philosophical Studies*, 153(1):123–142, Mar 2011.

[151] M. Mangin, A. Valade, A. Costes, A. Bouillod, P. Acco, and G. Soto-Romero. An instrumented glove for swimming performance monitoring. In *icSPORTS*, pages 53–58, 2015.

[152] S. Mann. Mediated reality with implementations for everyday life. *Presence Connect*, 1, 2002.

[153] S. Mann. Continuous lifelong capture of personal experience with eyetap. In *Proceedings of the the 1st ACM workshop on Continuous archival and retrieval of personal experiences*, pages 1–21. ACM, 2004.

[154] G. F. Marshall. Back from the past: the helmet integrated system of albert bacon pratt (1916). *Optical Engineering*, 28(11):281247, 1989.

[155] J. Marshall. Smartphone sensing for distributed swim stroke coaching and research. In *Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication*, UbiComp '13 Adjunct, pages 1413–1416, New York, NY, USA. ACM.

[156] K. Martin, B. Dalton, and M. Jones. Crafting urban camouflage. In *Proceedings of the Designing Interactive Systems Conference*, DIS '12, pages 797–798, New York, NY, USA, 2012. ACM.

[157] V. Mateevitsi, B. Haggadone, J. Leigh, B. Kunzer, and R. V. Kenyon. Sensing the Environment Through SpiderSense. In *Proceedings of the 4th Augmented Human International Conference*, AH '13, pages 51–57, New York, NY, USA, 2013. ACM.

[158] A. Matviienko, A. Löcken, A. El Ali, W. Heuten, and S. Boll. Navilight: Investigating ambient light displays for turn-by-turn navigation in cars. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '16, pages 283–294, New York, NY, USA. ACM.

[159] D. J. McFarland, W. A. Sarnacki, and J. R. Wolpaw. Electroencephalographic (eeg) control of three-dimensional movement. *Journal of neural engineering*, 7(3):036007, 2010.

[160] L. A. Meyerovich and A. S. Rabkin. Empirical analysis of programming language adoption. In *Proceedings of the 2013 ACM SIGPLAN International Conference on Object Oriented Programming Systems Languages & Applications*, OOPSLA '13, pages 1–18, New York, NY, USA, 2013. ACM.

[161] P. Milgram, H. Takemura, A. Utsumi, and F. Kishino. Augmented reality: A class of displays on the reality-virtuality continuum. In *Telemanipulator and telepresence technologies*, volume 2351, pages 282–292. International Society for Optics and Photonics, 1995.

[162] D. Moher, A. Liberati, J. Tetzlaff, and D. G. Altman. Preferred reporting items for systematic reviews and meta-analyses: the prisma statement. *Annals of internal medicine*, 151(4):264–269, 2009.

[163] P. Mohr, M. Tatzgern, T. Langlotz, A. Lang, D. Schmalstieg, and D. Kalkofen. Trackcap: Enabling smartphones for 3d interaction on mobile head-mounted displays. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pages 585:1–585:11, New York, NY, USA, 2019. ACM.

[164] G. Mone. Feeling sounds, hearing sights. *Commun. ACM*, 61(1):15–17, 2017.

[165] N. Moraveji, B. Olson, T. Nguyen, M. Saadat, Y. Khalighi, R. Pea, and J. Heer. Peripheral Paced Respiration: Influencing User Physiology During Information Work. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*, UIST '11, pages 423–428, New York, NY, USA, 2011. ACM.

[166] N. Moraveji and C. Soesanto. Towards Stress-less User Interfaces: 10 Design Heuristics Based on the Psychophysiology of Stress. In *CHI '12 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '12, pages 1643–1648, New York, NY, USA, 2012. ACM.

[167] N. Moray. Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly journal of experimental psychology*, 11(1):56–60, 1959.

[168] M. Möser. Elektronenmikroskopische Fraktographie. in Elektronenmikroskopie in der Festkörperphysik. VEB Deutscher Verlag der Wissenschaften, Berlin, 1982.

[169] F. Mueller, R. A. Khot, A. D. Chatham, S. Pijnappel, C. C. Toprak, and J. Marshall. HCI with sports. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '13, pages 2509–2512, New York, NY, USA. ACM.

[170] A. Mulloni, A. Dünser, and D. Schmalstieg. Zooming interfaces for augmented reality browsers. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services*, MobileHCI '10, pages 161–170, New York, NY, USA, 2010. ACM.

[171] M. Murer, I. Aslan, and M. Tscheligi. Lollio: Exploring taste as playful modality. In *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction*, TEI '13, pages 299–302, New York, NY, USA. ACM.

[172] H. Nakamura and H. Miyashita. Development and Evaluation of Interactive System for Synchronizing Electric Taste and Visual Content. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 517–520, New York, NY, USA, 2012. ACM.

[173] S. Naseh Hussaini. Mobile SoundAR: Your Phone on Your Head. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '11, pages 1777–1782. ACM.

[174] J. W. Nichols. Using handhelds as controls for everyday appliances: A paper prototype study. In *CHI '01 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '01, pages 443–444, New York, NY, USA, 2001. ACM.

[175] J. Nielsen. IM, Not IP (Information Pollution). *Queue*, 1(8):76–75, Nov. 2003.

[176] A. Niijima and T. Ogawa. Virtual Food Texture by Electrical Muscle Stimulation. In *Proceedings of the 2016 ACM International Symposium on Wearable Computers*, ISWC '16, pages 48–49, New York, NY, USA, 2016. ACM.

[177] K. Nojiri, S. Low, K. Toda, Y. Sugiura, Y. Kamiyama, and M. Inami. Present Information Through Afterimage with Eyes Closed. In *Proceedings of the 5th Augmented Human International Conference*, AH '14, pages 3:1–3:4, New York, NY, USA, 2014. ACM.

[178] L. Norooz, M. L. Mauriello, A. Jorgensen, B. McNally, and J. E. Froehlich. BodyVis: A New Approach to Body Learning Through Wearable Sensing and Visualization. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 1025–1034. ACM, Apr. 2015.

[179] J. Novak. Swimming direction and visual control. *Biomechanics and medicine in swimming IV (International Series on Sport Sciences)*, pages 345–349.

[180] S. R. Oldfield and S. P. Parker. Acuity of sound localisation: a topography of auditory space. i. normal hearing conditions. *Perception*, 13(5):581–600, 1984.

[181] A. E. O'Sullivan, C. Y. Lim, and E. C. Lalor. Look at me when i'm talking to you: Selective attention at a multisensory cocktail party can be decoded using stimulus reconstruction and alpha power modulations. *European Journal of Neuroscience*, 2019.

[182] K.-Y. Pan and Y.-Y. Chuang. Continuous zoom with two fixed-focal-length lens. In *SIGGRAPH ASIA 2016 Posters*, SA '16, pages 14:1–14:2, New York, NY, USA, 2016. ACM.

[183] R. Pan, S. Singhal, B. E. Riecke, E. Cramer, and C. Neustaedter. "myeyes": The design and evaluation of first person view video streaming for long-distance couples. In *Proceedings of the 2017 Conference on Designing Interactive Systems*, DIS '17, pages 135–146, New York, NY, USA, 2017. ACM.

[184] P. E. Paredes, Y. Zhou, N. A.-H. Hamdan, S. Balters, E. Murnane, W. Ju, and J. A. Landay. Just Breathe: In-Car Interventions for Guided Slow Breathing. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 2(1):28:1–28:23, Mar. 2018.

[185] H. K. Park and W. Lee. Motion Echo Snowboard: Enhancing Body Movement Perception in Sport via Visually Augmented Feedback. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*, DIS '16, pages 192–203, New York, NY, USA, 2016. ACM.

[186] M. Parvis, S. Grassini, E. Angelini, and P. Scattareggia. Swimming symmetry assessment via multiple inertial measurements. In *2016 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, pages 1–6.

[187] K. Pfeuffer, J. Alexander, and H. Gellersen. Partially-indirect bimanual input with gaze, pen, and touch for pan, zoom, and ink interaction. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 2845–2856, New York, NY, USA, 2016. ACM.

[188] M. Pielot, B. Poppinga, and S. Boll. Pocketnavigator: Vibro-tactile waypoint navigation for everyday mobile devices. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services*, MobileHCI '10, pages 423–426, New York, NY, USA. ACM.

[189] M. Pielot, B. Poppinga, W. Heuten, and S. Boll. A tactile compass for eyes-free pedestrian navigation. In P. Campos, N. Graham, J. Jorge, N. Nunes, P. Palanque, and M. Winckler, editors, *Human-Computer Interaction – INTERACT 2011*, pages 640–656. Springer Berlin Heidelberg.

[190] B. Poppinga, N. Henze, J. Fortmann, W. Heuten, and S. Boll. Ambiglasses - Information in the periphery of the visual field. In *Mensch und Computer 2012: interaktiv informiert–allgegenwärtig und allumfassend!?*, 2012.

[191] A. B. Pratt. Weapon., May 16 1916. US Patent 1,183,492.

[192] H. Profita, R. Albaghli, L. Findlater, P. Jaeger, and S. K. Kane. The at effect: how disability affects the perceived social acceptability of head-mounted display use. In *proceedings of the 2016 CHI conference on human factors in computing systems*, pages 4884–4895. ACM, 2016.

[193] M. Psychountaki and Y. Zervas. Competitive worries, sport confidence, and performance ratings for young swimmers. *Perceptual and Motor Skills*, 91(1):87–94, 2000. PMID: 11011876.

[194] J. Rachels. Why privacy is important. *Philosophy and Public Affairs*, 4(4):323–333, 1975.

[195] R. Raisamo, I. Rakkolainen, P. Majaranta, K. Salminen, J. Rantala, and A. Farooq. Human augmentation: Past, present and future. *International Journal of Human-Computer Studies*, 2019.

[196] N. Ranasinghe, A. D. Cheok, O. N. N. Fernando, H. Nii, and G. Ponnampalam. Electronic taste stimulation. In *Proceedings of the 13th International Conference on Ubiquitous Computing*, UbiComp '11, pages 561–562, New York, NY, USA. ACM.

[197] L. Rayleigh. On our perception of the direction of a source of sound. *Proceedings of the Musical Association*, 2:75–84, 1875.

[198] L. Rayleigh. Xii. on our perception of sound direction. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 13(74):214–232, 1907.

[199] J. Rekimoto and K. Nagao. The world through the computer: Computer augmented interaction with real world environments. In *Proceedings of User Interface Software and Technology'95*, 1999.

[200] C. A. Running, B. A. Craig, and R. D. Mattes. Oleogustus: The unique taste of fat. *Chemical Senses*, 40(7):507–516, 2015.

[201] D. Saha and A. Mukherjee. Pervasive computing: a paradigm for the 21st century. *Computer*, 36(3):25–31, 2003.

[202] N. A. Samshir, N. Johari, K. Karunanayaka, and A. David Cheok. Thermal sweet taste machine for multisensory internet. In *Proceedings of the Fourth International Conference on Human Agent Interaction*, HAI '16, pages 325–328, New York, NY, USA. ACM.

[203] P. Sanches, K. Höök, E. Vaara, C. Weymann, M. Bylund, P. Ferreira, N. Peira, and M. Sjölinder. Mind the Body!: Designing a Mobile Stress Management Application Encouraging Personal Reflection. In *Proceedings of the 8th ACM Conference on Designing Interactive Systems*, DIS '10, pages 47–56, New York, NY, USA, 2010. ACM.

[204] K. A. Satriadi, B. Ens, M. Cordeil, B. Jenny, T. Czauderna, and W. Willett. Augmented reality map navigation with freehand gestures. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 593–603. IEEE, 2019.

[205] J. Scheffel and G. Kockesen. Wearable web technology: Google glass and the mirror api. *MSc Media Technology, Leiden University*, 2013.

[206] S. Scheydt, M. M. Staub, F. Frauenfelder, G. H. Nielsen, J. Behrens, and I. Needham. Sensory overload: A concept analysis. *International Journal of Mental Health Nursing*, 26(2):110–120, 2017.

[207] C. Schmandt and A. Mullins. AudioStreamer: Exploiting Simultaneity for Listening. In *Conference Companion on Human Factors in Computing Systems*, CHI '95, pages 218–219, New York, NY, USA, 1995. ACM.

[208] A. Schmidt. Augmenting human intellect and amplifying perception and cognition. *IEEE Pervasive Computing*, 16(1):6–10, 2017.

[209] A. Schmidt. Technologies to amplify the mind. *Computer*, 50(10):102–106, oct 2017.

[210] A. Schmidt, B. Pfleging, F. Alt, A. Sahami, and G. Fitzpatrick. Interacting with 21st-century computers. *IEEE Pervasive Computing*, 11(1):22–31, 2012.

[211] W. Schoch, I. Heller, F. Schweingruber, and F. Kienast. Wood anatomy of central european species. 2004.

[212] B. Schwartz. The social psychology of privacy. *American Journal of Sociology*, 73(6):741–752, 1968.

[213] L. Shen, M. Y. Saraji, K. Kunze, and K. Minamizawa. Unconstrained Neck: Omnidirectional Observation from an Extra Robotic Neck. In *Proceedings of the 9th Augmented Human International Conference*, AH '18, pages 38:1–38:2, New York, NY, USA, 2018. ACM.

[214] R. Shen, T. Terada, and M. Tsukamoto. A System for Visualizing Sound Source Using Augmented Reality. In *Proceedings of the 10th International Conference on Advances in Mobile Computing & Multimedia*, MoMM '12, pages 97–102. ACM.

[215] D. Shenk. *Data Smog: Surviving the Information Glut*. HarperCollins Publishers, New York, NY, USA, 1997.

[216] L. E. Sherman, A. A. Payton, L. M. Hernandez, P. M. Greenfield, and M. Dapretto. The power of the like in adolescence: effects of peer influence on neural and behavioral responses to social media. *Psychological science*, 27(7):1027–1035, 2016.

[217] G. Sines and Y. A. Sakellarakis. Lenses in antiquity. *American Journal of Archaeology*, 91(2):191–196, 1987.

[218] M. W. Skirpan, T. Yeh, and C. Fiesler. What's at stake: Characterizing risk perceptions of emerging technologies. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pages 70:1–70:12, New York, NY, USA, 2018. ACM.

[219] J. L. Souman, I. Frissen, M. N. Sreenivasa, and M. O. Ernst. Walking straight into circles. *Current biology*, 19(18):1538–1542.

[220] F. Soyka, M. Leyrer, J. Smallwood, C. Ferguson, B. E. Riecke, and B. J. Mohler. Enhancing Stress Management Techniques Using Virtual Reality. In *Proceedings of the ACM Symposium on Applied Perception*, SAP '16, pages 85–88, New York, NY, USA, 2016. ACM.

[221] M. Speicher, B. D. Hall, and M. Nebeling. What is mixed reality? In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pages 537:1–537:15, New York, NY, USA, 2019. ACM.

[222] C. Spence, M. Obrist, C. Velasco, and N. Ranasinghe. Digitizing the chemical senses: Possibilities & pitfalls. *International Journal of Human-Computer Studies*, 107:62 – 74, 2017. Multisensory Human-Computer Interaction.

[223] T. Starner, S. Mann, B. Rhodes, J. Levine, J. Healey, D. Kirsch, R. W. Picard, and A. Pentland. Augmented reality through wearable computing. *Presence: Teleoperators & Virtual Environments*, 6(4):386–398, 1997.

[224] D. Sun, P. Paredes, and J. Canny. MouStress: Detecting Stress from Mouse Motion. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pages 61–70, New York, NY, USA, 2014. ACM.

[225] I. E. Sutherland. A head-mounted three dimensional display. In *Proceedings of the December 9-11, 1968, fall joint computer conference, part I*, pages 757–764. ACM, 1968.

[226] H. Z. Tan and A. Pentland. Tactual displays for wearable computing. *Personal Technologies*, 1(4):225–230.

[227] T. Tao, P. Ratsamee, Y. Uranishi, K. Kiyokawa, T. Mashita, and H. Takemura. An Interactive 4d Vision Augmentation of Rapid Motion. In *Proceedings of the 9th Augmented Human International Conference*, AH '18, pages 21:1–21:4, New York, NY, USA, 2018. ACM.

[228] N. Terzimehić, R. Häuslschmid, H. Hussmann, and m. schraefel. A review & analysis of mindfulness research in hci: Framing current lines of research and future opportunities. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pages 457:1–457:13, New York, NY, USA, 2019. ACM.

[229] J. Tholander and S. Nylander. Snot, sweat, pain, mud, and snow: Performance and experience in the use of sports watches. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 2913–2922. ACM.

[230] S. Thomée, A. Härenstam, and M. Hagberg. Mobile phone use and stress, sleep disturbances, and symptoms of depression among young adults - a prospective cohort study. *BMC Public Health*, 11(1):66, Jan. 2011.

[231] J. J. Thomson. The right to privacy. *Philosophy & Public Affairs*, pages 295–314, 1975.

[232] E. Tremblay, I. Stamenov, R. Beer, A. Arianpour, and J. Ford. Switchable telescopic contact lens. *Optics express*, 21:15980–15986, 07 2013.

[233] H.-Y. Tseng, R.-H. Liang, L. Chan, and B.-Y. Chen. Lead: Utilizing light movement as peripheral visual guidance for scooter navigation. In

Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services, MobileHCI '15, pages 323–326, New York, NY, USA. ACM.

[234] H.-Y. Tseng, R.-H. Liang, L. Chan, and B.-Y. Chen. Using point-light movement as peripheral visual guidance for scooter navigation. In *Proceedings of the 6th Augmented Human International Conference*, AH '15, pages 177–178, New York, NY, USA. ACM.

[235] K. Tsukada and M. Yasumura. Activebelt: Belt-type wearable tactile display for directional navigation. In *International Conference on Ubiquitous Computing*, pages 384–399. Springer.

[236] D. Van Krevelen and R. Poelman. Augmented reality: Technologies, applications, and limitations. *Vrije Univ. Amsterdam, Dep. Comput. Sci*, 2007.

[237] T. van Veen, J. Karjanto, and J. Terken. Situation awareness in automated vehicles through proximal peripheral light signals. In *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, AutomotiveUI '17, pages 287–292, New York, NY, USA. ACM.

[238] G. N. Vavoula and M. Sharples. Future technology workshop: A collaborative method for the design of new learning technologies and activities. *International Journal of Computer-Supported Collaborative Learning*, 2(4):393–419, 2007.

[239] M. Vinnikov, R. S. Allison, and S. Fernandes. Gaze-contingent auditory displays for improved spatial attention in virtual reality. *ACM Trans. Comput.-Hum. Interact.*, 24(3):19:1–19:38, Apr. 2017.

[240] C. von Hardenberg and F. Bérard. Bare-hand human-computer interaction. In *Proceedings of the 2001 Workshop on Perceptive User Interfaces*, PUI '01, pages 1–8, New York, NY, USA, 2001. ACM.

[241] M. Wang, S. L. Lyckvi, C. Chen, P. Dahlstedt, and F. Chen. Using Advisory 3d Sound Cues to Improve Drivers' Performance and Situation Awareness. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 2814–2825, New York, NY, USA, 2017. ACM.

[242] A. M. Warren, A. Sulaiman, and N. I. Jaafar. Social media effects on fostering online civic engagement and building citizen trust and trust in institutions. *Government Information Quarterly*, 31(2):291–301, 2014.

[243] M. Weiser. The computer for the 21 st century. *Scientific american*, 265(3):94–105, 1991.

[244] F. L. Wightman and D. J. Kistler. Resolution of front–back ambiguity in spatial hearing by listener and source movement. *The Journal of the Acoustical Society of America*, 105(5):2841–2853, 1999.

[245] G. Wilson, G. Davidson, and S. A. Brewster. In the Heat of the Moment: Subjective Interpretations of Thermal Feedback During Interaction. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 2063–2072, New York, NY, USA, 2015. ACM.

[246] M. Wilson, J. Doyle, and G. McTaggart. Designing Stress Management Interventions for Older Adults to Improve Wellbeing. In *Proceedings of the 2015 British HCI Conference*, British HCI '15, pages 307–308, New York, NY, USA, 2015. ACM.

[247] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 143–146, New York, NY, USA, 2011. ACM.

[248] K. Wongsuphasawat, A. Gamburg, and N. Moraveji. You Can'T Force Calm: Designing and Evaluating Respiratory Regulating Interfaces for Calming Technology. In *Adjunct Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology*, UIST Adjunct Proceedings '12, pages 69–70, New York, NY, USA, 2012. ACM.

[249] N. Wood and N. Cowan. The cocktail party phenomenon revisited: how frequent are attention shifts to one's name in an irrelevant auditory channel? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(1):255, 1995.

[250] K. J. P. Woods and J. H. McDermott. Schema learning for the cocktail party problem. *Proceedings of the National Academy of Sciences*, 115(14):E3313–E3322, 2018.

[251] P. Woźniak, K. Knaving, S. Björk, and M. Fjeld. Rufus: Remote supporter feedback for long-distance runners. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '15, pages 115–124, New York, NY, USA, 2015. ACM.

[252] S. Yamano, T. Hamajo, S. Takahashi, and K. Higuchi. Eyesound: Single-modal mobile navigation using directionally annotated music. In *Proceedings of the 3rd Augmented Human International Conference*, AH '12, pages 22:1–22:4, New York, NY, USA. ACM.

[253] P. Yin, R. Davison, Y. Bian, J. Wu, and L. Liang. The sources and consquences of mobile technostress in the workplace. *PACIS 2014 Proceedings*, Jan. 2014.

[254] M. Yokoyama, Y. Okano, M. Sato, S. Fukushima, M. Furukawa, and H. Kajimoto. Looming Silhouette: An Approaching Visual Stimulus Device for Pedestrians to Avoid Collisions. In *Proceedings of the 3rd Augmented Human International Conference*, AH '12, pages 23:1–23:4, New York, NY, USA, 2012. ACM.

[255] S. Yoon, S.-s. Lee, J.-m. Lee, and K. Lee. Understanding Notification Stress of Smartphone Messenger App. In *Proceedings of the Extended Abstracts of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, CHI EA '14, pages 1735–1740, New York, NY, USA, 2014. ACM.

[256] B. Yu, J. Hu, M. Funk, and L. Feijs. DeLight: Biofeedback Through Ambient Light for Stress Intervention and Relaxation Assistance. *Personal Ubiquitous Comput.*, 22(4):787–805, Aug. 2018.

[257] Q. Yue, Z. Ling, X. Fu, B. Liu, W. Yu, and W. Zhao. My google glass sees your passwords. *Proceedings of the Black Hat USA*, 2014.

[258] P. Zahorik and F. L. Wightman. Loudness constancy with varying sound source distance. *Nature neuroscience*, 4(1):78, 2001.

[259] H. Zhao, C. Plaisant, and B. Shneiderman. "i hear the pattern": Interactive sonification of geographical data patterns. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '05, pages 1905–1908, New York, NY, USA. ACM.

[260] D. N. Zotkin, R. Duraiswami, and L. S. Davis. Rendering localized spatial audio in a virtual auditory space. *IEEE Transactions on multimedia*, 6(4):553–564, 2004.

[261] M. Zuckerman, H. Persky, L. Miller, and B. Levin. Contrasting effects of understimulation (sensory deprivation) and overstimulation (high stimulus variety). *Proceedings of the Annual Convention of the American Psychological Association*, 4(Pt. 1):319–320, 1969.

APPENDICES

Appendix A

Auxiliary Material

For the work reported in this thesis, we have used material such that was not published in this volume for the sake of maintaining a reasonable length. This includes questionnaires and forms, as well as the letters of approval of the Ethics Committee of the University of Konstanz. This material has been made available online by means of a data management system for academic institutions. The material can be accessed at the following URL:

https://figshare.com/projects/Thesis_Auxiliary_Material/82295

Eidesstattliche Versicherung

(Siehe Promotionsordnung vom 12. Juli 2011, § 8, Abs. 2 Pkt. 5.)

Hiermit erkläre ich an Eides statt, dass die Dissertation von mir selbstständig und ohne unerlaubte Beihilfe angefertigt wurde.

Stuttgart, den 8. Mai 2020

Francisco Kiss

Francisco Kiss

Reshaping Ubiquitous Interaction Through Sensory Augmentation

Current technological advances enable unprecedented approaches for aiding human perception through digital technologies. Powerful Sensory Augmentation becomes feasible. Our concept aims to improve the natural sensory capabilities of users while transcending the traditional interaction concept of tools. This approach promises to facilitate a more natural interaction with ubiquitous computers and support the enhancement of interaction with reality through computing machines. The goal of this research is to investigate the possibilities of this emerging paradigm and to provide a formal structure for future efforts in this field.

In this thesis, we propose a definition for Sensory Augmentation and a design space to structure its study. We present a vision for its application to human activities and identify opportunities and challenges for the incorporation of Sensory Augmentation to human activities. Further, we report on three research probes that help investigate these opportunities and challenges, as well as users' experiences and preferences for this technology. Each of these probes explores a specific area of augmentation, both assessing specific theoretical aspects of Sensory Augmentation and providing practical insights in the technical challenges posed by the development of applications.

Additionally, we investigate and discuss the implications of augmenting the senses from the perspectives of users and society, focusing on benefits but also addressing the main foreseeable negative effects and possible strategies to minimize them. Finally, we present a design process for Sensory Augmentation applications alongside with practical recommendations and discuss future research directions of this new field.

Universität Stuttgart