Exploring Cooperative Musical Interaction through a Mobile Augmented Reality Application

Master's Thesis in Computer Science

Adrian Jensby Sandaker

June 3, 2019
Halden, Norway

Østfold University College www.hiof.no

Abstract

As mobile augmented reality technology slowly becomes a part of our everyday lives, the door is opened for designers and developers to explore new and creative ways of interacting together through this medium. While the technology has been available for some time, few have explored how we might cooperatively interact and engage with sound and music in augmented reality space. In this thesis I explore how we can utilize mobile augmented reality technology to design cooperative musical interactions. Adopting the methodology of Research through Design, and building on ideas from existing research and work in the field, I constructed a design concept with the aim of exploring the problem space, and shedding light on my research questions. This led to a highly iterative prototyping process, which culminated in the creation of a high-fidelity prototype - the mobile application Petals. The prototype application was subsequently evaluated through three separate field deployments, where it was assessed by six different users. Through these field deployments I was able to collect rich and detailed qualitative data in the form of observations and interviews, which was subsequently transcribed and analyzed through a process of open coding. The resulting analysis indicated that the prototype was successful in creating a highly immersive and ludic experience, and findings show that users are positive to engaging cooperatively with music mediated by mobile augmented reality. Furthermore, I found that use of binaural audio can be significant in strengthening users perceived immersion in the experience, and might also be effective in provoking movement and active participation within the augmented reality space. Finally, the results also shed some light on which awareness mechanisms are needed to better support cooperative interaction in augmented reality.

Keywords: Augmented Reality, Research through Design, Computer-Supported Cooperative Work, Sound and Music Computing

Acknowledgments

I would firstly like to thank all participants who helped test and evaluate the prototype. Without your assistance I would not have been able to finish this project. An enormous thanks to my supervisor Susanne Koch Stigberg for all your support and assistance. Your guidance has been essential in shaping my project, on both a methodological and conceptual level. I would also like to thank my family for their unconditional support and encouragement, and my friends for keeping my spirits high during this entire process. Furthermore, I would like to express my gratitude towards the open source community behind libpd and pd-for-android for providing the tools I needed to realize this project. Lastly, to Jeanette, thank you for your incredible support, patience and advice during this entire process. You mean the world to me.

Contents

List of Figures

List of Tables

Listings

Chapter 1

Introduction

> "*I'm always interested in what you can do with technology that people haven't thought of doing yet.*"
>
> \- Brian Eno

1.1 Background and Motivation

The fields of music and computer science have always been deeply intertwined. Ever since Alan Turing's Manchester Mark II [46, p. 197], computers have been used to play music, no matter how simple and primitive it might have sounded. Nowadays, almost all of the music we consume is at some point touched or mediated by computers in some way, and the technological evolution keeps pushing the envelope of what is possible. As new computing platforms emerge and mature, they present new opportunities for designers to probe and explore, to once again change our perspective on how we might engage with and create music. In the last decade or so, the introduction of the internet connected smartphone has forever changed how we view mobile computing, and has pushed the limits for how we play, create and communicate with each other every day. Today, it is possible to compose and produce complex and sophisticated musical pieces, simply through the use of a mobile device and a creative mind. What used to require bulky hardware synthesizers and complex studio setups can now be done in the palm of your hand, using widely available equipment and applications that can be wielded by anyone. Looking beyond traditional mobile applications, with mobile augmented reality now rapidly emerging as an important arena for new interactions to take place, it is up to developers and designers to shape the future, and imagine how we might creatively express ourselves in new ways by dreaming up and bringing to life new modes of interaction on this platform. I see this project as an opportunity to explore this space, to probe and examine what might be the next step in computer mediated musical expression.

During the first year of the master's programme I enrolled in the course *Interaction Design*, wherein I was introduced to the fields of Human Computer Interaction and Interaction Design through both practical exercises and project work, as well as through hallmark papers from the field presented throughout the semester. This way of thinking about computer science resonated with me, and this class is where the first seeds of inspiration for this thesis were sown. My personal motivation for this project stems from my lifelong love of music. For as long as I can remember, music has been an essential part of my life. From my earliest piano lessons as a child, to discovering the world of computer music as a young adolescent, all the way to the work done for this thesis - music has been the one common denominator. Consequently, I saw this project as an opportunity to consolidate my interests in music and computer science, while also contributing knowledge to the fields of Human Computer Interaction (HCI) and Interaction Design. I see this combination as a good fit, as the broader concepts relating to music-making and creative expression can be said to mesh well with the general ideas and conceptions found within the third-wave of the HCI continuum [78, p. 149]. Approaching the combination of augmented reality and music interaction with this perspective opens the door to exploring new and fresh ways to enable creative and ludic expression, treading new ground and playing with novel and unfamiliar ways of co-interacting with music.

1.2 Research Question

The field of augmented reality is growing rapidly, and is well on its way to becoming an important platform for new and innovative interactions between humans and computers to transpire. The widespread availability of augmented reality technology, in combination with the introduction of high speed mobile internet, has opened up a whole new space for shared immersive experiences to take place, right in the palms of our hands. This also introduces the possibility of exploring existing fields of research in new and creative ways, utilizing the new possibilities offered by recent technological advancements to think differently about how we interact together in virtual spaces. These spaces are instrumental in supporting not only utilitarian and work related functions, but also to encourage ludic and playful interactions that remind us to stay curious about what new and creative possibilities technology can offer.

Rooted in both the personal and professional motivation described in section 1.1, this project aims to explore several facets of cooperative work and augmented reality. My main research question is as follows:

RQ 1 *How can mobile augmented reality technology be used to design cooperative musical interactions?*

Here I define interaction with music as the creation or manipulation of sounds - of either vocal or instrumental origin - in a harmonious way. Furthermore, I limit the cooperative interaction to those taking place in the context of *same time - same space* when seen through the CSCW-matrix as described in section 2.2.1. In addition the project aims to shed light on the following sub-questions.

RQ 2 *How does the inclusion of binaural audio affect the immersion of users in a cooperative mobile augmented reality space?*

RQ 3 *What awareness mechanisms are needed for ludic mobile real time cooperative interaction in augmented reality?*

1.3 Report Outline

This thesis has been arranged according to the following structure:

Chapter 1 introduces the background and motivation behind the work for this thesis, giving insight into my reasoning for choosing to explore the area of study on both an academic and personal level. Moving on, the project research questions are presented, and some concepts are defined and clarified.

Chapter 2 provides a review of related work in the fields of augmented reality, computer-supported cooperative work and sound and music computing, placing the work in an academic context and providing the theoretical framing for the project work. Furthermore, some important exemplars are presented, which have been key sources of inspiration in the formation of the initial design concept.

Chapter 3 gives insight into the methodological approach of the work, and describes the various methods employed throughout the project. The methodology Research through Design is presented and discussed, with the focus directed on how it is applied to the project in practice. Furthermore, I present a framework employed to classify the various prototypes developed, before outlining how the field research was to be performed. This chapter also provides a brief discussion of what contributions to knowledge might be made, before assessing some ethical considerations of the work.

Chapter 4 describes the entire process of designing and developing a high fidelity prototype, from concept definition to prototype finalization. The chapter begins by constructing a design concept, drawing on influences from key exemplars of work in the field. Moving on, I describe my process of using video to create a low-fidelity augmented reality prototype, before presenting the various technology choices made to compose the technological makeup of the high fidelity prototype. This is followed by a brief section describing the process of designing the synthesizer patch used in the prototype. Finally, I describe the entire process of designing the high fidelity prototype, documenting each step along the way.

Chapter 5 briefly assesses how I performed the open coding of data, before presenting the most significant findings from the analysis. These findings have been grouped into three main categories and are presented as concepts created through the open coding process.

Chapter 6 provides a discussion on the findings presented in the prior chapter, focusing on how they have enabled me to explore the project research questions, while also framing them in the context of related work in the field. In addition, this chapter also contains some reflections on my experience with using Research through Design as a methodology for the project, assessing how the documentation of the design work might have contributed knowledge as Research through Design work.

Chapter 7 presents the conclusion, and briefly assesses some limitations of the study before presenting some suggestions for future work.

Chapter 2

Related Work

"I have to follow my instinct and intuition and curiosity."

- Ryuichi Sakamoto

This chapter provides insight into the domains of augmented reality, computer-supported cooperative work and sound and music computing. My intention is to present a broad overview of each separate domain, while highlighting the work most relevant to the themes and concepts explored within this thesis. I begin by introducing and defining augmented reality from a historic perspective, before giving some insight into the technical aspects of developing augmented reality experiences. Moving on, I introduce the field of computer-supported cooperative work, focusing mainly on the concepts of context and awareness. This is followed by a brief look at collaboration within augmented reality space. Next, I give some insight into the domain of sound and music computing. The focal point of this section is on the design and creation of controllers and interfaces for musical expression and audio in the context of augmented reality experiences. I also shed some light on collaborative musical expression in the context of augmented reality. Following this, I present some selected exemplars of work in the field that have been particularly relevant in the development of my design concept. Lastly, I provide a short summary which briefly revisits the major themes discussed throughout the chapter.

2.1 Augmented Reality

The introduction of internet connected smart devices such as mobile phones, watches and glasses to the general public, has forever changed how we interact with information in our daily lives. With the technology commonly available today, we are no longer restricted to displaying information as text on a traditional computer monitor or handheld tablet screen. Ordinary smartphones are now capable of delivering experiences that seamlessly intertwine high quality digital graphics and information with a live feed of the world around us, overlaying both textual information and high-fidelity 3D objects onto our surroundings. The technology which enables us to superimpose digital graphics and information onto a live view of the real world is what is known as *augmented reality*.

2.1.1 Definition and History

According to Azuma [1], an augmented reality system should include the following three key characteristics:

1. Combines real and virtual
2. Is interactive in real time
3. Is registered in three dimensions [1, p. 356]

While they were initially defined over 20 years ago, these points still outline the general requirements of any AR experience worth its salt. Furthermore, the first point (1) is essential in separating AR from it's at times more commonly seen relative - *virtual reality*. The term augmented reality itself was first coined by Thomas Caudell in 1990 during his work at The Boeing Company [14], although the broader idea and concept of augmenting reality with information has been around for decades. The first documented description of what would today be considered an augmented reality device dates back to 1901, when writer Frank L. Baum portrayed an apparatus he dubbed the *character marker* in his science fiction novel "The Master Key" [3]. In the novel, Baum describes a device shaped like a pair of glasses, which gives the user an augmented view of the world, much like that made possible through the use of a head-mounted display such as Google Glass[1] or MagicLeap One[2]. When worn, this character marker device would display a single character on the forehead of anyone you met, revealing their true nature and intentions. Good people would be marked with the letter G, while evil people would be marked with the letter E. This way, the device allowed the wearer to determine the true intentions of anyone, simply by casting a glance at them [3]. While some elements of this device are still firmly located in science fiction, the technology needed to realize the augmented reality functionality is now commonly available to anyone. What was once a far-flung dream deep in the realm of science fiction, is nowadays commonplace technology accessible to most people. Augmented reality is still to some extent a nascent technology with regards to its maturity in the broader technological landscape, but a wide range of different applications have already been proposed and implemented successfully, both in commerce [37], engineering[11, p. 4] and entertainment [34].

An example of early work in the field is the system developed by L. B. Rosenberg in the early 90s, wherein he achieves the effect of augmented reality using what he calls *virtual fixtures* [69]. His aim was to improve the performance of operators working with telepresence systems, i.e. systems developed to give the user the sense of being present in a remote environment, often with the ability to manipulate it [69]. His findings indicated that the use of augmented reality was highly beneficial to support processes performed by the teleoperator [69], stating that the use of virtual fixtures improves performance by altering how the operator conceptualizes the task, giving a simplified perception of the workspace and utilizing alternate sensory pathways to supply information [69, p. 81]. This system is regarded by some as the first fully functional augmented reality system [67]. Another forerunner in the practical use of AR is the *Digital Desk* system [83] developed around the same time, which allowed users to perform computer based interactions

[1] `https://x.company/glass/`
[2] `https://www.magicleap.com/magic-leap-one`

using paper documents on their physical desktop. This was made possible using a system of cameras and projectors in order to establish the augmented reality environment for the user to interact with. Several different prototypes were developed for the system, among these a language translation application [83, p. 20] and a collaborative drawing environment [83, p. 24]. While several applications were developed for this environment, they were all anchored to the fixed desk system. In stark opposition to this very stationary desktop based arrangement, the TransVision system developed in 1995 at the Sony Computer Science Laboratory [66] is more reminiscent of the mobile solutions commonly seen today. The TransVision offered two separate configurations - one using a palmtop computer, the other using a head-mounted display. Both configurations utilized a camera, and enabled the user to perceive an augmented view of the world, akin to how one might today experience augmented reality using either a smartphone or a pair of smart glasses [66]. While many of these early AR systems were widely differing in approach and technical realization, they were all important stepping stones towards developing the cutting edge AR systems seen today.

2.1.2 The Reality-Virtuality Continuum

Within the spectrum of immersive computing in general, it is possible to create experiences and interactions with considerably different levels of immersion. In general, augmented reality and virtual reality can both be said to exist within the same continuum, with an increasing level of immersion taking place as we gradually move from the digital overlays in a real world (AR) to a fully digital and completely enveloping virtual environment (VR). In order to define and place an experience within this spectrum of immersive computing, one can utilize the Reality-Virtuality (RV) Continuum scale developed by Paul Milgram in 1995 [57] as seen in Figure 2.1.

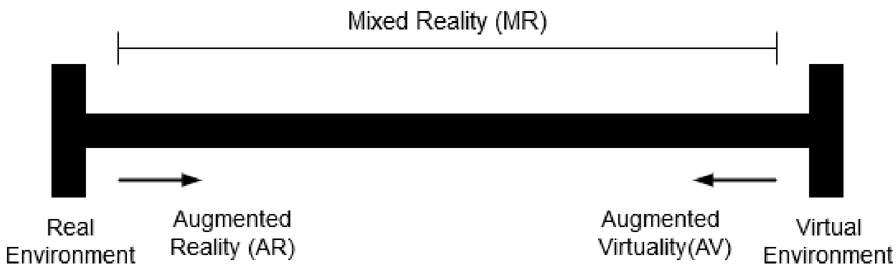


Figure 2.1: Model of the Reality-Virtuality (RV) Continuum scale [57]

Through use of this scale, we can place any augmented- or virtual reality experience within a spectrum, enabling us to classify the *extent of augmentation* taking place. On the leftmost side of the scale in figure 2.1 we would place reality as is, without any virtual enhancement taking place at all. On the far right we would place an entirely virtual environment, much like that experienced by the use of virtual reality headsets like Oculus Rift[3] or HTC Vive[4]. Finally, the area in between these outer points, what is described as *mixed reality*, represents any experience falling in between these, e.g. those made possible with the mobile augmented reality technology of today [57]. Furthermore, Milgram separates this middle spectrum into the concepts of *augmented reality* (AR) and *augmented*

[3]`https://www.oculus.com/rift/`
[4]`https://www.vive.com`

virtuality (AV). In the associated definition, Milgram defines AR as the process of adding computer generated enhancements to the real world environment being currently observed [57]. On the other hand, AV describes the addition of 'reality' to an otherwise virtual environment [57]. Though this distinction might initially seem less important in comparison to the separation of AR and VR, it arguably increases the granularity of the scale, making it possible to define or classify a particular system with far greater precision. This ability to define the level of immersion could also be considered important when developing experiences taking place within the spectrum of reality and virtuality.

2.1.3 Developing for Augmented Reality

In the early years of augmented reality development, those looking to get started had relatively few options available with regards to tools and software development kits. A few different alternatives were available, such as the Virtual Reality Distributed Environment and Constructive Kit (VR-DECK) developed at IBM [19], or the Minimal Reality Toolkit [29] used in the TransVision system [66]. However, a definitive industry standard had yet to be established. This changed in 1999, when the open source library ARToolkit[5] was introduced to the general public. The library was developed by Hirokazu Kato [48], and was released by the University of Washington under the GNU GPL license. Suddenly, there was a viable base for AR development available to anyone wanting to try their hand at it. While originally developed mainly for use with a certain type of head mounted display [48, p. 2], ARToolkit was quickly ported and spread to other platforms post release, and was successfully being used to create AR experiences on mobile phones by 2003 [33]. Furthermore, there has been considerable growth in the AR field in general which has introduced a whole sleeve of different libraries and frameworks for developers to choose from, supporting most of the major platforms of today [82, p. 52].

2.2 Computer Supported Cooperative Work

The field of computer-supported cooperative work (CSCW) initially got its start at a workshop held in 1984 by Irene Greif and Paul Cashman [72]. At the time of its inception, computer systems had already been in use within office environments for decades, supporting both routine activities and business critical tasks from day to day. In addition, the term *groupware* had already been coined at the time, defined as "intentional group processes plus software to support them" [43]. Though the use of computers for collaborative work was already an established fact, the research into it was still scattered and spread out among many different fields and disciplines [72]. Therefore, the establishment of CSCW as a field was at the time critical in uniting the interests of both developers and researchers from different areas of expertise within a collective community [31]. As CSCW is a highly varied and interdisciplinary field, it can still today often be difficult to narrow down and define the collective scope and focus of work within its domain. The term *Computer Supported Cooperative Work* itself can be defined as "*...the study of how people use technology, with relation to hardware and software, to work together in shared time and space.*"[65, p. 1]. After emerging as a separate field, CSCW quickly gained momentum and is today a large field dealing with a variety of different topics, even spouting

[5]`https://github.com/artoolkit`

its own annual conference[6] hosted by the Association for Computing Machinery (ACM).

2.2.1 The Context of Cooperative Work

Within definition of CSCW stated above, the inclusion of *time* and *space* as concepts points toward an important notion within the field, namely the *context* in which a particular computer system is being used to support cooperative work. In any given context, two separate dimensions can be said to exist: time and space [65]. Within the dimension of time, cooperative work can take place either at the *same time*, or at *different times*[65]. Furthermore, within the dimension of space, cooperative work can be happening within the same space, or in a distributed way across different spaces [65]. This concept of context will typically be depicted as a matrix, first introduced by Robert Johansen in 1988 [42], seen below in table 2.1.

	Same Space	**Different Space**
Same Time	Face to face interactions	Remote interactions
Different Time	Continuous task	Communication + coordination

Table 2.1: The CSCW Matrix.

As visible from the matrix, face to face interactions are a typical example of cooperative work taking place at the same time in the same space. Within the domain of CSCW, this can be represented through e.g. the use of groupware, wall displays or a shared table where work is occurring. Through the use of technology, cooperation can also take place at the same time in different spaces as remote interactions, ex. by way of video conferencing software such as Skype or FaceTime, or something even simpler such as an ordinary telephone call. The common denominator between all these interactions is the fact that they are happening synchronously, regardless of space [65]. However, as people take part in cooperative work, it is not always possible to make this a synchronous process. Sometimes it might also be desirable to contribute at different times depending on the nature of the work, requiring the tools to support this course of action too.

In times where asynchronous work processes are taking place, it might be beneficial to anchor the work to a specific space through the use of ex. a design room - a physical space dedicated to a specific project or continuous task taking place. This enables visiting participants to share information and partake in work at different times as long as the physical space is accessible [65]. On the other hand, sometimes it is not feasible to attach work to a specific physical space due to various constraints or limitations. Perhaps some participants are working remotely abroad, possibly even in a different time zone altogether. Such a scenario is an example of a situation requiring support for both asynchronous and

[6] `https://cscw.acm.org`

distributed work. This is a typical use case for internet enabled tools such as email, version control software or collaborative software such as wikis [65], all enabling participants to perform work regardless of their current position in time and space. Therefore, by considering and classifying cooperative work according to these dimensions, it is possible to better understand and tailor to the requirements of any situation when developing systems for support.

2.2.2 Awareness in CSCW

Irrespective of how we classify work within space and time, a fundamental goal of CSCW systems is to achieve a high degree of coordination [22, p. 40], i.e. harmonious collaboration between work participants [30, p.426]. Tied to this is a fundamental idea in the CSCW field: the concept of *awareness*. Early work by Dourish and Bellotti in the nineties presented a definition of awareness as "...an understanding of the activities of others, which provides a context for your own activity" [21, p. 107]. However, the idea of awareness within the field is often regarded as rather ambiguous, having been interpreted and discussed at length throughout the years [71] [30]. This has made it difficult to maintain one all-encompassing and consistent definition of the term.

A more recent definition by Gross presents awareness as "*...a user's internal knowing and understanding of a situation including other users and the environment that is gained through subtle practices of capturing and interpreting information*"[30, p. 432]. Gross argues that this information is the product of a duality as it partly exists in the environment, and is partly provided by awareness technology [30]. Furthermore, projects such as the *ambientRoom* [40] project by Ishii et al. has explored how awareness technology can be integrated into the environment itself through use of technology. In this project, they explored how one might communicate ex. the number of unread emails using sound. The room was outfitted with a soundtrack of a soothing natural soundscape, which increased and decreased in volume and density according to ex. the number of unread email messages of a user [40, p. 2]. While the ambientRoom project uses a somewhat creative approach to communicating awareness, it shows how we might creatively approach the challenge of awareness support through inventive use of technology. In order to further define and clarify the concepts relating to awareness, Gross [30] separates the concepts of *coexistence awareness* and *cooperation awareness*.

Coexistence awareness

In CSCW, *coexistence awareness* is defined as "users' mutual person-oriented information on each other" [30, p. 434]. A notable amount of research into awareness within CSCW has explored the use of media spaces and collaborative virtual environments (CVEs). Generally, the goal of media space systems has been to enable cooperation between two or more physically separate locations through permanent video and/or audio links, creating a shared environment accessible through this connection [30]. On the other hand, CVEs are distributed virtual reality environments designed to promote and support collaborative activities and information sharing between participants within this virtual space [18, p. 4]. In simpler terms, media spaces are a type of virtual environment created by connecting real world spaces through video and audio links, while CVEs are environments where the space itself is fully virtual. Both types of systems, if constructed and implemented correctly, should be capable of fulfilling the goal of providing information on participants presence and availability, thereby fostering what is known as coexistence awareness.

Cooperation awareness

Gross defines *cooperation awareness* as "...users' mutual information on their activities — either as background information in a collaborative working environment, or as foreground information in a cooperative application" [30, p. 438]. In the current software landscape of today, with software solutions increasingly implementing internet connectivity as a core part of the functionality, it is not unusual for systems to provide information about other users' activities within a shared workspace. As an example, the web-based word processor Google Docs[7] implements several mechanics that provide information on the activities of co-workers [30, p. 440]. When used collaboratively, it will automatically email a notification to the owner of a shared document whenever another participating user creates a new comment within the document. In addition, it also presents awareness information by displaying the time of the last edit within the document, along with icons representing every user currently active and editing the document. These are clear examples of both background and foreground information being presented to users as they participate in the shared collaborative space.

2.2.3 Design Tensions in Awareness Research

While this separation helps us better understand the nature of awareness, it is still a very abstract and often difficult concept to comprehend. In further clarifying its processes, Gross underlines that awareness is a "*...dynamic construct — that is, a process rather than a point in time*" [30, p. 452], and points out how work itself is inherently social, which requires CSCW systems to communicate a vast amount of different information to provide awareness over a distance [30]. Through his research, Gross has identified four design tensions that must be considered when designing for awareness, namely *availability*, *privacy*, *conventions* and *tailoring* [30]. Here I would like to shed some further light on two of these that I have considered in my project, namely the tensions of *privacy* and *conventions*.

Privacy

The tension of privacy describes the difficult balance between sharing sufficient awareness information, while also maintaining a sufficient level of privacy [30, p. 455]. This can be challenging to achieve, and Gross proclaims how social interaction over remote channels might introduce new challenges, as the traditional social protocols used in face-to-face interactions might not apply [30, p. 457]. Furthermore, he underlines the need to better understand how we might design in a way where we provide enough information on each others activities to achieve coordination, while still having the possibility to keep some information private [30, p. 457]

[7] `https://docs.google.com`

Conventions

The design tension of conventions refers to the agreements in teams on how its members should behave [30, p. 457]. In the establishment of conventions, we might achieve better communication efficiency and reduce the effort needed to communicate [30]. Furthermore, he points to how conventions largely manifest as a dynamic construct, in that "*...group members create and maintain a growing mutual understanding*" [30, p. 458].

While these tensions describe challenges that exist in many different scenarios where work processes are supported by technology, the novelty and nascence of cooperative augmented reality might introduce additional challenges as we develop our understanding of these.

2.2.4 Collaborating within Augmented Reality

The general use AR technology has seen a surge in popularity after the adaptation of smart phones by the general populace, but the idea of using AR to support collaborative processes is hardly new. Pioneering work has been utilizing augmented reality to support collaborative processes since the nineties [47][10].

An early example of collaborative augmented reality is the *Studierstube* system [80] developed in the late nineties. With applications aimed mainly at scientific visualization, the system lets multiple users interact with a computer generated 3D model through use of their bespoke control interface dubbed the *Personal Interaction Panel* (PIP). The interface consists of two physical components; a hand-held panel for displaying objects in the augmented reality space, and a pen for interacting with these objects [80]. In addition to enabling manipulation of any displayed model, the panel can be used to virtually display both a traditional 2D computer display, as well as to display the control interface for the Studierstube system itself, removing the need for users to exit the augmented reality environment in order to reconfigure the system. Several of the earlier uses of augmented reality for collaboration were centered around improving productivity and collaboration in a professional workplace environment [80][10][83]. However, there were also those exploring this through more entertainment-centered means. One such example is the MagicBook system [6], which explored the use of physical objects to transition users between the real world and immersive AR/VR environments. Using a children's book as a tangible metaphor, the system provides access to immersive 3D experiences through use of a customized interface. Users can choose to read and experience the "magic book" without any technology, but through the system interface its contents can be enhanced and explored either though augmented reality overlaying the real world, or in full virtual reality [6].

A recent study looked at the use of projected augmented reality to enable fully immersive interactions between remote participants [61] in a co-present way. In the context of virtual- and mixed reality, co-presence is achieved when a user is actively perceiving the presence others, while simultaneously sensing that others are actively perceiving them [60]. The study employs a prototype system composed of several projectors and Kinect[8] units, built using the RoomAlive Toolkit[9]developed at Microsoft Research in 2014 [45]. This system, dubbed Room2Room, explores how one can enable co-presence by projecting

[8] `https://developer.microsoft.com/en-us/windows/kinect`
[9] `https://www.microsoft.com/en-us/research/project/roomalive-toolkit/`

a life-size image of a person onto nearby seating furniture in the room. While the system is somewhat limited in scope, supporting only one-on-one interactions in compatible spaces, findings indicate that this AR-based system provides a feeling of presence significantly higher than that achieved using traditional video-conferencing solutions such as Skype [61]. When assessed in the light of coexistence awareness research, the use of projected augmented reality such as this could have a significant effect on how we develop and design systems for remote cooperation and communication in the future. Nonetheless, the various systems presented above show the breadth and variety found within the field of collaborative augmented reality.

2.3 Sound and Music Computing

Using a computer for the purposes of creating music is today a very common occurrence, as both the recording, production and performance of music usually relies heavily on computer usage in many or all parts of the process. Similarly, the consumption of music is today highly dependent on technology now that digital music streaming is the prevailing way of listening for many people. This long standing relationship between computers and music dates all the way back to the dawn of the earliest computers themselves, as the first primitive piece of computer programmed music was executed and played in 1951 on Alan Turing's Manchester Mark II machine [46, p. 197]. Nowadays, the study of computer music has it's own dedicated field within the umbrella term of *Sound and Music Computing* (SMC). Bernadini [5] proposes the following definition: "Sound and Music Computing (SMC) research approaches the whole sound and music communication chain from a multidisciplinary point of view. By combining scientific, technological and artistic methodologies it aims at understanding, modelling and generating sound and music through computational approaches." [5, p. 144]. Furthermore, a roadmap has been defined by the Sound and Music Computing Network[10] in an attempt to identify challenges related to the field at large. Among these, challenges there are two points addressing issues pertaining to research within the field:

- **Design better sound objects and environments:** "The growing abundance of electronically generated sounds in our environment, coupled with the rapid advances in information and sensor technology, present SMC with unprecedented research challenges, but also opportunities to contribute to improving our audible world."[68]
- **Understand, model, and improve human interaction with sound and music:** "The human relation with sound and music is not just a perceptual and cognitive phenomenon: it is also a personal, bodily, emotional, and social experience. The better understanding of this relation from all these perspectives will bring truly useful and rewarding machine-mediated sonic environments and services." [68]

While the above points are both rather open ended and broad in terms of scope, they also invite opportunity for designers interested in contributing to the field. As the continuing technological innovation introduces the possibility of new types of interactions through use of technologies such as AR, there is also a potential to utilize these for musical expression and performance. Mechanics like those used to drive an AR experience could potentially unlock new ways of creating, performing and thinking about music.

[10] `http://www.smcnetwork.org/index.html#roadmap`

2.3.1 Creating New Interfaces for Musical Expression

The tightly knit connection between music and computers opens the gate for new and aspiring technologies to innovate this shared domain, and at thereby provide new modes of interaction with music. A central hub for this kind work is The International Conference on New Interfaces for Musical Expression[11] (NIME). This annual conference, born from a workshop at CHI[12] in 2001 [63], brings together a multidisciplinary crowd of researchers and musicians every year. At this intersection of art and technology, there is also a space to explore and challenge the accepted ideas and notions of what is proper music interface design. While creativity is a vital part of this, some principles for the design of computer music controllers have been suggested and put forth in order to avoid common pitfalls, providing a theoretical base for the creation of new interfaces to interact with music [20]. The outlined principles concern both human and technological factors, and while they are rather informal in their presentation, they provide a general guide for anyone attempting to create a new musical interface. As an example, one principle reads simply - "Make a piece, not an instrument or controller" [20, p. 1] - referring to the fact that the creation of a tool or instrument without a specific musical or compositional idea to drive the direction or goal is not necessarily productive. While it might sometimes produce new and interesting research questions, projects with no rooting in a specific musical idea risk being dead ends due to not actually having a significant product or future direction to drive the work[20, p. 2]. Another principle states that "programmability is a curse" [20, p. 1], pointing to the need to establish certain limitations of use within the instrument interface. In other words, if the instrument has no conceivable limit to its configurability or customization, there is a chance that users will spend time doing experimentation with the instrument and its configurations itself, rather than using the instrument to actually create pieces of music [20]. While there is in theory nothing inherently wrong with this, it is likely to be considered counterproductive wherever the end goal is the creation of a piece of art.

2.3.2 Audio Augmented Reality

In the wider discussion on augmented reality, the focus has typically been centered on mainly visual experiences, often by way of overlaying digital graphics and 3D models onto the users' view of the world. This view can be further supported by the characteristics required for an AR system as defined by Azuma [1], presented in section 2.1.1, wherein one characteristic (3) is the system presenting content in three dimensions. However, several projects have explored how the combination of audio and augmented reality can produce new and innovative experiences, without the need of a visual component. Some of the earliest work was done with the goal of using an audio augmented reality device as an automated museum tour guide[4]. Users wearing a prototype would get automatic and location based access to audio based descriptions of museum pieces as they moved around the physical museum space. In other words, when the wearer approached a specific piece of art, the prototype would register their position and start playing the recorded description automatically [4], thus reducing the need to follow specific paths or directions as one might have when using a manual cassette-tape player at the time. Another example of pioneering work can be seen in the *Audio Aura* project developed at Xerox PARC in 1997 [59], which explored the use of audio augmented reality for the purpose of conveying work

[11] `https://www.nime.org`
[12] `https://sigchi.org`

related information in the context of an office environment. Some suggested scenarios included the use of both physical artifacts and locations to trigger auditory cues, such as playing a pre-recorded greeting message when approaching the office of a colleague who currently is away for the day [59]. The project was dependent on custom hardware for location tracking at the time [58], but today one would likely be able to utilize commonly available features of smart-phones for these purposes. Looking towards the domain of mobile augmented reality, a modern example of audio AR is the iOS application Fields[13], which uses both visual and auditory augmented reality in order to explore how physical spaces can be transformed into virtual places of music. The app allows users to place either existing musical pieces available within the app, or their own recorded audio clips, in a given spot where they are currently standing. After its initial placement, the audio will persist in the specific position as you move around the room, allowing one to create what is in essence a three-dimensional sound installation enabled by AR technology[85]. Furthermore, recording a piece of audio while in motion will make the recorded audio move along the same path within the space, providing a immersive and dynamic experience of the sound in three dimensional space [85]. Interestingly, while the application itself is firmly rooted in musical ideas, by introducing a collaborative element to the system it could potentially be used to achieve what the Audio Aura project set out to do in 1997, by placing audio in a spatial context and making it available for others.

2.3.3 Spatial Audio through Binaural Filtering

Existing work [86] [70] has suggested that there is a significant importance in correctly representing the relationship between spatial positioning of visual and aural content in an augmented reality space. The study by Zhou et al. [86] indicates that using 3D audio in an augmented reality both improves task performance, while also contributing to the overall feeling of immersion and presence in augmented reality space [86]. Furthermore, this study also revealed that 3D sound also helped improve the feeling of collaboration between participants in augmented reality space when performing a task together [86]. A later study by Sodnik et al. further strengthened these results [75]. I see this as very relevant when trying to develop immersive musical and auditory experiences in augmented reality space, such as in this project. The effect of spatial audio can be achieved through various ways. As described in the previous section, projects such as Audio Aura [4] and the Fields application [85] have succeeded in virtually placing audio in a physical space, making it accessible to users through technological means. However, another way of representing audio spatially is making it appear as if the audio source is placed somewhere in the environment, without actually placing it there. In other words, this makes it possible to create the illusion that sound is being emitted from a specific direction in space, only through clever use of audio processing. While this can be achieved in practice through the use of several different techniques, a common way of doing this is through *binaural filtering*[14].

Put simply, the process of binaural filtering requires an input audio signal, and a direction from which it is supposed to be emitted. It then manipulates the audio signal to sound like it is being emitted from the given direction [54]. This directional value is generally referred to as the *azimuth*, and is provided in degrees either between -180 and

[13] `https://fields.planeta.cc/`
[14] An example of binaural filtering: `https://www.youtube.com/watch?v=4jv8QFTmwqU`

180, or between 0 and 360. No matter which model we employ, an azimuth value of 0 degrees generally means that the sound source is directly in front of the listener. Then, if using the former model a value of -90 means that the sound is being emitted from the left, and a value of 90 means that the sound is being emitted from the right. Using the latter model, these values would instead be 90 and 270. Furthermore, while azimuth places the sound in a direction, by changing the *elevation* we can place it higher or lower in vertical space. A common way of calculating the spatial position of a sound source is by using a Head Related Transfer Function (HRTF). In practice, a HRTF consists of measurements recorded using the ears of either mannequins or humans from several different directions [16]. One commonly used source of measurements is the KEMAR dataset[15]. These values make it possible to describe the sounds propagation from a source in space, and all the way to the users ear [16]. While binaural filtering often incorporates both azimuth and elevation, previous work [75] has shown that humans are generally much better at locating sounds according to azimuth than either elevation or distance [75, p. 117].

2.3.4 Collaborative Musical Expression in an AR Space

Over the years several projects have explored the use of augmented reality for the creation of new and innovative musical instrument interfaces. One notable historical example is the *Augmented Groove* system presented in 2001, which enabled users to play and perform electronic music collaboratively [62]. This was made possible by manipulating a set of vinyl records customized with fiducial markers, enabling an overhead mounted camera to recognize and track their current position [62]. In addition, the system overlaid 3D virtual controllers onto the controllers, which enabled instant and direct communication to the user of the current state of the composition [62, p. 3]. A more recent study explored how augmented reality can be used to enable collective musical expression [49]. This was examined through an art installation, where participants explored an augmented reality experience through use of a smart-phone application in a controlled environment, constructed and set up for this specific purpose. The participants, wearing headphones, were subjected to an altered version of any sounds occurring in their surrounding environment such as their own footsteps or vocals. This was done by capturing and processing the audio input from the microphone in each phone, before feeding it back through the headphones, now altered by the processing. Using audio from the environment like this in order to control or manipulate a different piece of audio or music is commonly known as *reactive music*[8, p. 236]. In addition to this, two acoustic reflectors were constructed and installed, with the intention of creating a deeper sense of listening for the participants [49]. The results of the study indicate that the experiment was successful in creating an immersive space for the playful expression of music in a collective setting. Users reported feeling safe and comfortable when exploring the AR experience, as well as feeling brave enough to make sounds and vocalize in the environment, despite other people also being present in the room [49, p. 27]. In addition, the findings from this project showed how an environment could function as an instrument, mediated only by manipulating an audio signal [49, p. 28]. Considering the limited sources of sound available to participants, this can potentially be an indicator of how effective the construction of such an audio space really is, as simple vocalizations and ambient sounds became compelling within this virtual environment.

[15] `https://sound.media.mit.edu/resources/KEMAR.html`

2.3.5 Selected Exemplars of Work

Also operating within the realm of collaborative music in augmented reality is the art installation *Bloom: Open Space*[16], by Brian Eno and Peter Chilvers. However, contrary to the academic perspective found in the works mentioned above [49] [62], Bloom: Open Space has a more commercial and/or industry driven origin, being rooted in a project by Microsoft[17] to explore the crossover between technology and music. However, as I consider it as a highly significant exemplar of related work in this project, I will present it in detail within this section.

The installation itself was created by Brian Eno and Peter Chilvers and used the Microsoft HoloLens headset to achieve its augmented reality functionality [81]. This interaction enabled users to collaboratively create music through augmented reality by using simple pinch gestures performed in mid-air [81]. This pinch gesture, when recognized by the HoloLens system, would result in a virtual bubble being created and rendered graphically within the mixed reality space. Each one of these bubbles would then emit a single musical tone, which would subsequently be mixed with the bubbles created by other users, dissolving into thin air and creating a shared and collaborative musical AR experience between the users [81].

Furthermore, the system was deliberately designed in a way that made it difficult to precisely play specific notes, reducing the level of control somewhat and creating a more intuition-driven experience [81]. Then, to avoid this turning into a dissonant and noisy soundscape due to the lack of precision, the system would only give participants access to a limited range of notes when playing, which helped ensure a harmonic experience as people played in the shared space [81].

While it was not possible for me to attend the installation myself, it is built on concepts found within the existing mobile application "Bloom" [18], also developed by Eno and Chilvers. As part of the research progress for this project, I explored the Android version[19] of the application myself to gain insight and gather inspiration.

The application can likely be said to exist somewhere in the space between a musical instrument and a piece of interactive art. One might classify it as a *musical experience*, as it provides only a limited amount of control to the user, with no actual interaction required to create music. During use, you interact by way of touch input. Each tap on the screen results in a shape being slowly formed on the area of the screen that was touched. This shape is accompanied by a sustained tone, which gradually fades out over time along with the visual shape. Each new touch produces another circle, accompanied by another tone. This becomes part of a repeating pattern, which slowly loops over and over again, creating a melodic rhythm[20]. If the user at any point stops interacting with the interface, the application will slowly begin modifying the existing pattern by itself, essentially improvising and creating a unique piece of music without needing any additional input. A screenshot of the application in use can be seen in figure 2.2.

[16] http://bloomopenspace.com/
[17] https://www.microsoft.com/inculture/musicxtech/
[18] http://www.generativemusic.com/bloom.html
[19] https://play.google.com/store/apps/details?id=com.opallimited.bloom10worlds&hl=en
[20] https://www.youtube.com/watch?v=DdODb0t-Jn8

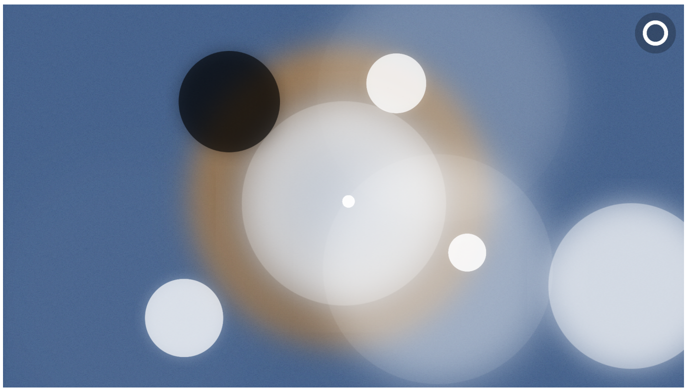


Figure 2.2: A Screenshot of the Android application *Bloom: 10 Worlds*. Each circle represents a musical note, and is the result of a tap on the screen by the user.

Conceptually, the application heavily incorporates what is known as algorithmic composition, which can be described as "*...the application of a rigid, well-defined algorithm to the process of composing music.*" [41, p. 157]. In the application, all music is created as a combination of input from users and the rules defined by the developers, a system which could be said to provide both constraints and opportunities, depending on which perspective one has. While the mobile application Bloom lacks the cooperative and augmented reality aspects of my project theme, it is useful as an exemplar of how one might design a mobile musical experience in practice. In addition, it also provides tangible insight into the conceptual backdrop which lies behind the creation of the augmented reality experience seen in Bloom: Open Space.

2.4 Summary

In this chapter I have provided an overview of augmented reality, presenting both the history and background of the field, as well as providing exemplars of more recent developments. Furthermore, I have given a brief introduction into the field of Computer Supported Cooperative Work (CSCW), focusing mainly on the concepts of *context* and *awareness* within the field. In addition, I have provided insight into the field of Sound and Music Computing (SMC), focusing on the design and development of interfaces for musical expression. These three fields have been selected to form the theoretical backbone of this thesis. Along with this, I have also presented some key exemplars of work from both field and industry that exemplify the executive ideas and concepts found within this project.

Chapter 3

Method

"And when you realize that everything is staged, then nothing is staged. There's a kind of liberation to that."

- Nils Frahm

In this chapter I present the underlying methodological approach I have applied in the project work, as well as the various methods used in every stage of the process. The core methodology used in this project is *Research through Design*. Consequently, the central thoughts and conceptions found within this methodology have guided each individual process and activity throughout the project, from the beginning phases of prototype design and development, through field deployment and evaluation, to finally the analysis of collected data and subsequent presentation of results. I have made an effort to deliberately select and employ tools and methods in my work that I see as connecting well with the ideas and practices of Research through Design as a methodology. Furthermore, I recognize that the resulting composition of methods within this work diverges somewhat from the typical configuration found in a master's thesis research project. I therefore consider it important to present, describe and contextualize each individual method in relation to the overarching methodology throughout the chapter. I do this both to provide detailed insight to understand the reasoning behind the method selection, as well as to encourage transparency and allow for critical reflections on the choices made.

The following section begins with a short introduction to Research through Design from a general perspective. I present the core conceptions of the methodology and discuss some epistemological perspectives, before giving insight into how one might employ the method in practice. Here i also briefly present a paper that has inspired and guided my executive process throughout. Following this, I give an introduction to a framework used to guide the design and prototyping process. This is followed by a description of the methods used to conduct a field study, wherein qualitative data is gathered through observations and group interviews. The next section outlines methods employed to analyze this data material. Here I have adopted the process of open coding as my approach. Following this, I present a brief overview of how I intend to contribute knowledge as a result of the thesis project. Lastly, I discuss some of the ethical considerations when doing field work, and provide insight into how I intend to ensure ethical practice in the research within this project.

3.1 Research through Design

As inspiration for my project comes from both the fields of art [73] and academia [49], i wanted to employ a methodology that could support an explorative and concept-driven approach, while still being firmly grounded within the confines of academic research. In this work, I wished to not only study the problem space through the existing literature and theory, but also to seek further understanding and contribute knowledge by constructing artifacts and exploring the realm of possibilities provided by the problem space. Rooted in these criteria, and driven by my research questions, I arrived at Research through Design as a suitable methodology for this work. *Research through Design* (RtD) is a particular approach to research that makes use of methods and processes traditionally associated with design practice [87, p. 167]. In RtD, research is performed not only by using traditional scientific methods of inquiry, but also by making and evaluation of artifacts that propose solutions to a given research problem. Historically, the foundations can be traced to a 1993 paper by Christopher Frayling [26], wherein he presents RtD as a union of practices from both *research* and *design*. According to some, these practices can at times be considered as divergent and heading in opposite directions, resulting in a tension between them [87, p. 167]. Others such as Bardzell et al. observe several similarities between them, going on to ask "*Are design and research not merely different variations of inquiry?*" [2, p. 99]. In general, *design* is concerned with solving specific and individual problem instances in the present moment, by creating specialized and precise solutions to improve or solve a given challenge [76, sec. 43.1.1]. The resulting knowledge produced by design processes is therefore often less abstract and conceptual than that produced by traditional scientific research and might therefore be difficult to apply in a generalized sense outside of its specific case or problem area of origin [76]. However, while strong adherence to the traditional scientific research principles is essential in tackling some problems, others might be better resolved through a design-oriented approach. In my opinion, the problem space explored in this project currently belongs to the latter category.

Using methods from design we are able to ask a different set of questions, and might therefore arrive at different solutions, resulting in the creation of knowledge that would perhaps not emerge through traditional research methods. The utility of design practice in knowledge creation can be further underlined by paraphrasing Zimmerman & Forlizzi [87], who state that sometimes artifacts and systems must be invented before they can be critically studied and henceforth proven to be the a suitable solution to a given problem [87, p. 168]. In their paper they point to the computer mouse as an example of an experimental and innovative artifact that had to be designed and developed before it could be studied, and only then consequently proven to be an exemplary solution through use of traditional research methods [87, p. 168]. In a similar vein, Stappers et al. [77] argue that design functions as a way to show that something is *possible* - regardless of necessity - where it might not have been apparent until now, thereby contributing knowledge in the form of existence proof [77, p. 172]. However, RtD has also received criticism from those who claim its original proposal by Frayling [26] supplied insufficient theoretical guidance to instruct practice [50, p. 5]. While this initial work was critical in establishing the fundamental connecting links between research and design, it provided little in the way of directly applicable theory to guide those wanting to adopt these ideas. On the other hand, this argument no longer carries much weight, as there now exists a growing body of

RtD work to learn from [50, p. 6]. Within this existing work, there are some key pieces that have had a significant influence on my process. Among these, the framework for RtD proposed by Koskinen et al. [50] has been vital in providing a sense of direction in my work and setting the stage for the research performed as part of this project.

3.1.1 The Field Approach to Research through Design

My approach in this project has been guided by a framework for RtD proposed by Koskinen et al. [50]. I have adhered to this framework with the intent of maintaining an explicit direction in the executive research work, as well as to help support the creation of reliable and valid knowledge through research. In essence, the framework separates RtD work into three different categories, namely *lab*, *field* and *showroom* [50]. Each practice presents a certain set of conventions to follow, which in turn affects how the research is carried out. Put simply, the lab approach favors strict controlled experiments, the field approach aims to study systems in their natural context, and showroom practitioners seek to generate reactions rather than research [50]. In this project, my work follows the practice of *field*. As mentioned, research within the field practice is generally aimed at examining how a system or artifact works in its natural surroundings, i.e. in the context or environment in which a system or artifact is supposed to be used. The research is then often performed using methods borrowed from interpretive social science. This stands in contrast to work within the lab practice, which is characterized by strict and controlled studies using experimental research methods [50, p. 69]. This difference is underlined by Koskinen et al. who state that "*The lab decontextualizes; the field contextualizes*" [50, p. 69]. When people interact with a piece of technology, they make sense of it based on their individual perspective both of the subject itself, and its *context* of use. This in turn informs and dictates how they act, meaning that two different people can have contrasting attitudes towards the same system or artifact in a given context, which leads to them using it in two different ways [50, p. 69]. This is the kind of information that we aspire to find through a field study, and the kind of information I consider relevant in exploring my research problem. Consequently, this also means that any findings will likely not have the same level of replicability as those attained through a lab study. However, this is not my intention either, and I will nonetheless aspire to contribute valid knowledge by conducting the research with rigor, and documenting all my choices and moves within this report. While this framework places the project within a greater frame of reference and outlines certain high-level principles on how to conduct research within the selected practice, the framework [50] does not give many practical instructions as to how one should proceed when applying RtD to a project. In light of the criticism towards RtD raised by Stappers et al. [77] regarding what they consider a lack of structure in RtD work, I have chosen to arrange my process roughly according to the five-step blueprint for RtD project proposed by Zimmermann & Forlizzi [87, p. 184]. Through this, I want to not only contribute research resulting from the field study, but also to contribute by showing how RtD work can be performed in practice in a project like this.

3.1.2 Applied Research through Design

To assist in bringing structure to my work, I have adopted the five-step plan for RtD suggested by Forlizzi & Zimmerman [87, p. 184] to guide my process. This plan carries a RtD project through the sequential phases of (1) selecting a problem; (2) conduct a literature review, finding a RtD paper to guide the process and beginning to iteratively design and develop ideas; (3) evaluating the resulting artifact according to the concerns of the RtD practice selected; (4) reflecting on and disseminating the findings; and (5) returning to explore the problem once more [87, p. 185]. Throughout my work I will carry out each step accordingly, with the exception of step five, as this arguably goes beyond the scope of what is feasible within a single master's thesis project. However, I will attempt to provide a starting point for subsequent work by presenting some possible future avenues of research in section 7.3 of this report. Although these five steps provide a stable structural framing for the work, there are otherwise few concrete restrictions given by the authors regarding how to perform each step, and how to carry the work out in practice. However, Forlizzi & Zimmermann do emphasize the importance of meticulously documenting as much of the design process as possible (including any missteps or failures) [87, p. 185], echoing the points made by Bardzell et al. [2, p. 105] and Koskinen et al. [50, p. 94] regarding the importance of documentation in RtD.

As suggested by the five-step model [87, p. 185], I have selected a paper within the field practice to function as scaffolding for my process. This 2007 paper by Sara Ljungblad [53] was recommended as an example of field-research by Zimmermann & Forlizzi [87, p. 186], and can be seen as having a certain thematic similarity to my project as it deals with enabling creative expression through interactive technology. In the project, Ljungblad explores what is referred to as *context photography*, meaning the use of real-time context data to affect digital pictures taken using a mobile device. Here it is worth noting that the notion of context differs from that discussed in sec 2.2.1, and that Ljungblad operates with a definition of context similar to that commonly used in everyday speech. In addition, the project drew inspiration from an alternative type of photography practice called *lomography* to inspire the design of the prototype [53]. The core concept in lomography is using substandard cameras to take pictures that are on one hand flawed when measured by traditional photography standards, but on the other hand interesting when seen through the lens of lomography enthusiasts [53, p. 360]. This concept guided the design process, which resulted in a prototype in the form of a camera application for mobile devices, which was then evaluated by amateur photographers in a exploratory user study. The participants were unfamiliar with this alternative practice and were not aware of the project before taking part. Through its incorporation of concepts from lomography, combined with contextual information from the users environment, the project shows how "*...alternative practice could contribute to the design of novel digital photography*" [53, p. 372]. In addition to providing a detailed account of how one might accomplish such a project, the project produced a prototype which in some ways foreshadowed photography practices employed by current applications such as Instagram [87, p. 186]. It is worth noting that due to its age (2007), the paper does not incorporate the vocabulary and tools currently used in the contemporary RtD paradigm. However, there is a clear lineage to be traced in the process employed in the work, and its suitability as a methodological inspiration can be supported by it being endorsed by Zimmerman & Forlizzi [87, p. 186].

However, while Ljungblad approaches the design in a participatory way by including people, I have approached this differently, focusing on developing a prototype informed by theory and personal choices. To help guide my work in the design process, I have made use of a dedicated framework to evaluate the purpose of each prototype.

3.2 Prototyping

Adopters of the five-step model [87] for RtD research are given few restrictions with regards to how design activities should be carried out. In my project, the design and prototyping is performed by the construction of various prototypes that are designed to explore new ideas and probe different aspects of the problem space. In order to guide this process, and to clarify the purpose of each prototype within the context of the project, I will make use of a framework developed by Stephanie Houde & Charles Hill during their time at Apple Computers [36]. This framework employs a three-dimensional model (seen in figure 3.1) which separates the aspects of any artifact (i.e. interactive computer system) into three distinct dimensions: *role*, *look and feel* and *Implementation.* This expands somewhat on the typical vocabulary used with regards to prototypes, where they are classified according to fidelity only. Prototypes designed with a high degree of finish are generally considered as *high-fidelity* prototypes. On the other end of the spectrum are the *low-fidelity* prototypes, which are simpler and less polished prototypes designed to quickly test an idea [36, p. 2]. However, by adapting the three-dimensional model we expand our vocabulary beyond just low- and high-fidelity, and might therefore more clearly communicate the purpose of a prototype.

Firstly, the dimension of *role* is concerned with what function an artifact will provide for the user, and how it can be considered useful to them [36]. Prototypes within this space can be anything from a static storyboard to an interactive click-through prototype. The key purpose it must fulfill is to enable the asking of design questions relating to the usefulness and function of an artifact. Next, the dimension of *look and feel* involves questions relating to the direct sensory experience of using an artifact. How does it look when in use? How does it sound? How should the artifact feel when handled by users? These factors are explored within this space [36]. Finally, the dimension of *implementation* deals with the concrete technical and architectural aspects of an artifact [36]. Sometimes it is not enough to simulate the experience of using an artifact with props, and an actual technical implementation is required in order to answer whatever question being explored. In this project, the various prototypes developed inevitably touch upon each dimension in some way or another. Furthermore, while my use of this model is useful in communicating the purpose of a prototype to readers of this report, it also allows me to audit the design process myself, and ensure that I probe each dimension sufficiently before developing the main prototype.

However, this kind of prototype might explore all three dimensions at once, thereby also representing a state of balance between them. Prototypes within this space (the center of the model) are meant to closely resemble the final artifact with regards to user experience, and can help balance the constraints of each of the other three dimensions. Within the framework, these are known as *integration prototypes.* With a integration prototype, it is possible to verify whether a given design is consistent and complete across all three dimensions at once [36].

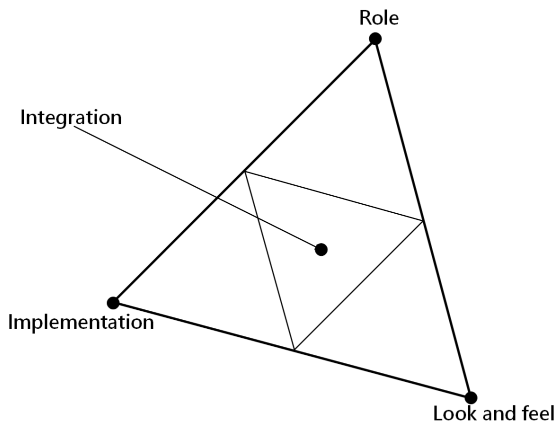


Figure 3.1: The three dimensional model for prototypes created by Houde and Hill [36]. Model has been recreated for use in this report.

In addition to applying the above model to my prototypes, I choose to adopt the perspective on prototypes held by Houde & Hill, who state that "*We define prototype as any representation of a design idea, regardless of medium. This includes a preexisting object when used to answer a design question.*" [36, p. 3]. As even the creation of simple augmented reality experiences requires a certain level of sophistication in the technological implementation, it might be reasonable to instead explore design questions through prototypes in the form of pictures or videos where this is applicable.

3.2.1 Prototyping in Practice

Throughout the project I have created several different prototypes to explore various dimensions, as well as one high fidelity integration prototype. As mentioned, I have used the prototyping framework [36] as a way to perform some evaluation activities regarding the purpose of each minor prototype during the prototyping phase, as this work was performed without involving external participants. I chose to not evaluate the low fidelity prototypes with users as most of these has been constructed mainly to verify the fitness of a technology or software library before incorporation and use with the integration prototype. However, testing and evaluation with users was performed with the finished integration prototype in order to produce qualitative data for analysis. As this project follows the *field* approach, this largely dictates which methods to apply in the phases of gathering and analyzing data.

3.3 Evaluation and Data Collection

This section describes the research methods that I have used in the process of testing and evaluating the integration prototype. As suggested in the five-step model [87, p. 185], evaluation of a prototype is performed according to the RtD-practice [50] to which the project belongs. In the case of this project, the RtD work follows the field-approach. In addition to building on the framework by Koskinen et al. [50], the operative elements of my field study have been informed by Siek et al. [74] and Blandford [7], while drawing on inspiration from Ljungblad [53]. In practice, evaluation was performed through prototype field deployment, wherein a a semi-structured qualitative study was performed to gather qualitative data. The main sources of data in this study were observation and group-interviews. I begin this section by giving a brief introduction to semi-structured qualitative studies, before moving on to describing the practical details of my approach.

3.3.1 Semi-Structured Qualitative Studies

The field study in this project was carried out by performing a semi-structured qualitative study (SSQS). The reasoning behind using qualitative methods lies in my desire to examine the research questions by collecting rich data from exploitative hands-on prototype evaluation with participants. Furthermore, the research questions themselves are of a character that makes them arguably more suited for this approach, as they aim to study phenomena related to concepts such as creativity, musical expression and cooperative interaction between people. Typically, a SSQS is focused on addressing research questions or problems to develop an understanding through *exploration*, instead of testing pre-defined hypotheses [7]. This was also my intention with the field deployment, and a key part of my reasoning to employ this method. According to Blandford, using an SSQS-approach within HCI research is a great way to both *evaluate the effects of using new technologies* and *understand current existing practices and needs*, framed in a real-world perspective [7, sec. 52.1.2]. These are both perspectives that I see as harmonizing well with the theme of my research. Furthermore, the method generally outlines some structural aspects by default, e.g. how to organize the study and how data collected should be analyzed, but at the same time it allows for change and evolution to take place throughout the study [7, sec. 52.2], thereby being *semi*-structured. As I had no way of predicting how users would react to the prototype once deployed, I considered this element of freedom as very beneficial to allow for a more adaptive and reactive approach.

3.3.2 Data Gathering

In the field study, data was gathered from participants by performing **non-participatory observation** during the testing of the prototype, followed by **semi-structured group interviews**.

Observation

In the observation part of my study, I adopted a non-participatory and overt approach [7, sec. 52.6.1]. Put simply, this means that I did not participate in using the prototype during the session, and the participating users were aware of my presence and observation. This choice to abstain from participation is motivated by my bias as the designer of the prototype, as I could risk having a significant effect on the participants use of the interaction. Furthermore, the test session itself was carried out mainly to allow the participants to explore the prototype in preparation for the group interview, and my observation thereof was simply an attempt to maximize the information gathered from the situation. During the observation, the notes taken were recorded using pen and paper. This was a conscious decision, as it allowed me to more quickly and fluently take notes in the form of both text and drawings if needed. It also promoted a certain transparency between myself and the participants, as they could clearly see whenever I noted anything down. Contrary to this, if I had taken these notes using a laptop, it might have been difficult to distinguish whether I was taking notes, or perhaps using the computer for other activities.

Group interview

Each practical testing session was followed by a group interview, which involved myself and the participating users. The group interview had a semi-structured design, and was conducted on the basis of an interview guide. My choice to perform the interviews in group was inspired by the work of Kiefer & Chevalier [49, p. 27], wherein a similar approach was used. In designing the questions for the interview guide, my approach was inspired by Charmaz [15] who states that "*By creating open-ended, non-judgmental questions, you encourage unanticipated statements and stories to emerge*" [15, p. 65]. Therefore, it was important for me that some of the questions were designed in a way that allowed me to extract raw and unexpected responses from participants. At the same time, I also created some more specific, less open questions designed to probe for insight into particular aspects of the prototype i saw as important, in case they did not organically emerge during the conversation. Furthermore, In the process of crafting the interview guide, I did not intend for all questions to be asked during the interview. They were designed with the intent of providing a starting point for discussion, and to help stimulate discussion if a dialogue did not emerge naturally. A copy of the interview guide used can be found in appendix D.

3.3.3 Participant Sampling

According to Kvale & Brinkmann, the purpose of the study should be the deciding factor when determining the number of participants to recruit [51, p. 148]. Depending on the research questions explored, this number can vary from just a few people (e.g. by performing in depth interviews), to tens of thousands (e.g. using standardized questionnaires) [51, p. 148]. I argue that the research questions explored in this project lend themselves well to employing a qualitative approach. Comparative studies by Kiefer & Chevalier [49, p. 27] and Ljungblad [53] involved a total of 9 and 7 participants respectively. The collaborative aspects of my prototype are designed for use by two people simultaneously, and as a result the study therefore requires an even number of participants. In total, 6 participants were recruited for testing and evaluation. Of these 6 participants, 2 were unfamiliar to me at the time of recruitment, while 4 were friends or acquaintances with varying degrees of closeness.

3.3.4 The Purpose of the Prototype

In situations where design questions are explored through participants' interactions with prototypes, it can be crucial to clarify the purpose of the prototype to participants [36, p. 15]. Prototypes can be ambiguous and complex, and therefore by defining the explicit purpose of the prototype and the study being carried out, one can avoid misunderstandings and correct any erroneous perceptions the users might have before evaluation has begun. As an example, participants exploring the project prototype might direct their focus only towards the mobile application user interface (ex. the placement of buttons within the app), despite the goal of the experiment being to study the interactions happening within augmented reality space. This could result in participants only reporting their experiences from a perspective of less interest to the study, or as put by Koskinen et al, " *The last thing* *any designer wants is feedback focusing on surface features of the expression rather than the thinking behind it.*" [50, p. 81]. I was therefore careful to clearly brief participants on this before carrying out the experiment. However, any feedback provided by participants should naturally be recorded nonetheless, regardless of which dimension it concerns.

3.3.5 Participant Recruitment

The choices made in the process of participant recruitment for the field deployment have been somewhat inspired by Ljungblad [53]. Firstly, the study by Ljungblad involves people with "*...a general interest in photography*" [53, p. 368]. Similarly, in recruiting participants I explicitly set out to reach users who had at least somewhat of an interest in musical and/or creative expression in a general sense. Much like how users with no interest in photography might be less than ideal for testing a creative camera application, users with no interest in music might be less suited as testers of an application aimed at musical expression and improvisation. Furthermore, in her study Ljungblad consciously involved only people who were not already familiar with the concept [32, p. 265]. This was also a deliberate choice in participant recruitment for me, as I wanted to avoid contaminating the ecological validity of the study by involving participants already familiar with the project. To get in touch with potential participants, I first attempted to recruit people through a "call for participation" posted to Facebook, and spread through information flyers posted at bulletin boards at the Halden campus of Østfold University College. These can be seen in figures 3.2 and 3.3 respectively. However, recruitment through these channels was largely unsuccessful, as I was only able to recruit one participant this way. Therefore, additional participants were recruited through alternative measures. Of these, two participants were recruited through direct instant messaging via Facebook Messenger, one participant was recruited face to face, and the final two were recruited by two of the participants upon my request. All participants were rewarded for their participation with a gift voucher to a cinema of their choice with a value of 150 NOK.

3.3.6 Consent Form

Prior to taking part in the project, participants were required to sign a consent form. This form verbosely described the nature of the research and the specific data to be collected, in addition to informing participants of their rights to withdraw their consent at any time during the project period. During the field deployment, the form was presented and signed on paper. The design of the form was roughly based on a template[1] provided by the Norwegian Centre for Research Data (NSD), with the contents of the form tailored to fit this project specifically. Furthermore, two versions of the form were created, one in English and the other in Norwegian. A copy of each form can be found in appendix C.

3.3.7 Test Session Duration

The duration of each testing session must be such that participants gain a certain amount of experience and understanding of the prototype. However, running a session for too long might leave a participant frustrated and impatient, which in turn could negatively affect the subsequent interview. Inspired by the study performed in [49, p. 27], I aimed for a session duration of between 10 and 30 minutes.

[1] `https://nsd.no/personvernombud/hjelp/samtykke.html`

Figure 3.2: Facebook post with a call for participation.

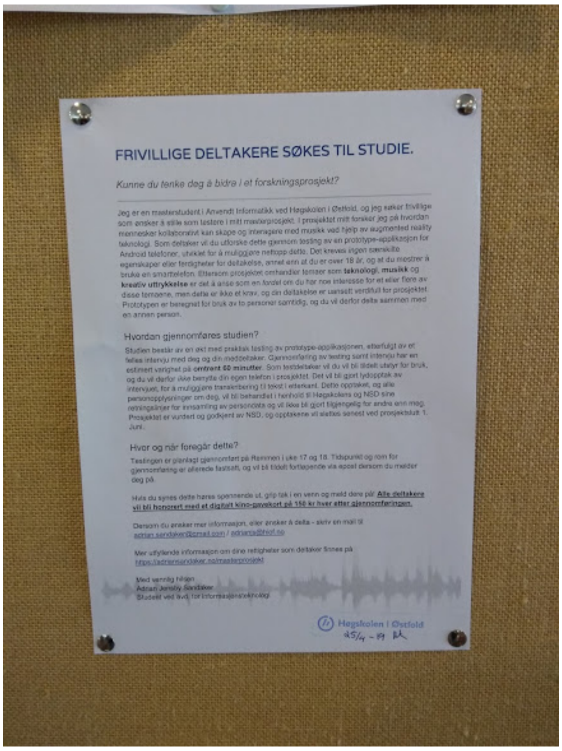


Figure 3.3: Call for participation flyer posted to notice board.

3.3.8 Pilot Test

Prior to executing the experiment, a pilot test was performed. This pilot test was intended to mirror the actual study, and consisted of a prototype test session, as well as an interview. As I participated myself, I was not able to take the role of an observer in this test. My co-participant in this pilot test was my domestic partner. The session consisted of a practical application test, followed by a test interview. The practical session lasted for 23 minutes, and the interview lasted for 26 minutes.

3.4 Analysis of Data

Following the field deployment, and the associated data gathering activities, the next step was to analyze the data. It is during the analysis phase that one might actually begin to make sense of the collected data, and any findings might begin to appear. The activities in the analysis should be carried out in a way that is appropriate both for the research questions, as well as for the qualitative nature of the collected data. Based on this, I decided to adopt an open coding approach as my main method of analysis.

3.4.1 Transcription of Data

Before any coding can take place, the data needs to be transcribed from audio to text. In their work, Kvale & Brinkmann argue that there exists no "one true way" of translating information from a spoken format to a written one [51, p. 212]. However, they state that there is one basic rule that we must conform to in any report: "...write explicitly in the report how the transcriptions have been performed" [51, p. 207]. In this project, I chose to transcribe the interviews to a digital format manually by hand. While this is a time consuming process, some have argued that "*...the very act of transcribing, and maybe making notes at the same time, is a useful step in becoming familiar with the data and getting immersed in it.*" [7, sec. 52.7]. In the act of transcribing, I felt that this held true for me as well. During transcription, I chose to ignore most instances of filler words such as "uhh", "umm" and other similar utterances that might occur during normal conversation. In addition, I have omitted detail about any minor pauses or involuntary vocal disruptions such as stuttering. This was done with the intent of increasing the fluency and readability of the text. Furthermore, I anonymized any information such as the names of people or places in order to protect the privacy of the participants. In the transcripts, participants are in the transcripts referred to simply by a letter and a number. This can be seen in the excerpt of one of the transcribed interviews, which is included in E. This excerpt is intended to give a tangible example of the style and form of the transcribed text.

3.4.2 Open Coding

Due to the qualitative nature of the collected data, I considered open coding as a fitting process for data analysis. Open coding can be described as "*...the analytic process by which concepts are identified and developed in terms of their properties and dimensions.*"[24, p. 310]. This process is performed by asking questions about the data and comparing the various aspects of the recorded phenomenon to find similarities or differences. In this process, data describing similar events is labeled and grouped, which over time produces a list of codes which can be grouped to form categories. While my work does not adopt the comprehensive Grounded Theory method, I was inspired by this process of open coding, and utilized this in the process of categorizing and labeling my transcribed data. Furthermore, in my analysis I took a combined *deductive* and *inductive* approach to the coding, meaning that some concepts were developed by me based on theory, while others were developed based on the transcribed text [24, p. 311].

3.4.3 Qualitative Data Analysis Tools

After having transcribed the data, I performed the coding using the CAQDAS software ATLAS.ti 7.

3.5 Contributions to Knowledge

In light of the methodological approach used in the project, theory [2] [87] dictates that knowledge might be produced not only the traditional way (e.g. as a result of the field research and subsequent analysis of data), but also as a result of the design processes themselves. Therefore, I have included this section to discuss how this might be done, and what kind of knowledge we might hope to produce from the RtD work. I begin by giving a quick insight into how I intend to ensure quality in the qualitative research. Moving on, I then briefly assess the epistemological potential of Research through Design work.

3.5.1 Ensuring Quality in Research

It can sometimes be difficult to agree on what constitutes quality in qualitative research [7]. How someone might rate this quality depend on factors such as their philosophical stance, and which research paradigm they subscribe to. In this research, I will attempt to ensure the quality of my research by adhering to the "*...four essential characteristics of good qualitative research*" as stated by Yardley [84, p. 219]:

- Sensitivity to context
- Commitment and rigour
- Transparency and coherence
- Impact and importance

The characteristic of *sensitivity to context* relates to matters such as being familiar with existing research, as well as taking measures to ensure ethically sound behavior [84, p. 219]. In this project, I have performed a detailed literature review, and have acted according to ethical guidelines of external origin. *Commitment and rigor* can be achieved by performing a thorough data collection as well as providing an analysis of sufficient depth [84, p. 219]. The scope of my data collection and analysis has been outlined in the previous chapters, allowing the reader to assess this from their perspective. In order to ensure *transparency and coherence* one should be transparent in the application of methods and presentation of data, as well as have a good fit between theory and method [84, p. 219]. Whether or not this has been achieved can be judged based on the contents of this chapter. And finally, the *impact and importance* can be achieved by clearly presenting the theoretical and practical relevance of any findings [84, p. 219]. A typical way of doing this in HCI work is presenting some implications for design [7, p. 52.10.1]. This has been done to the best of my abilities. These are principles I have strived to adhere by throughout the project.

3.5.2 Epistemology in Research through Design

In this project I use Research through Design as the methodology to frame my work. A central part of this project is the field study, in which I use tried and tested qualitative methods with the intent of securing a certain degree of validity in the generated knowledge. However, a key conception in RtD is the idea that knowledge is generated not only as a result of applying these traditional research methods (e.g. qualitative and/or quantitative research), but also as a result of the design processes themselves. Therefore, I would like to briefly assess the epistemological position of this by quoting Höök et al. and asking "*...how do we articulate, validate and constitute the knowledge gained through design research?*" [35, p. 1]. If the activity of design is to provide inherent value "by itself" in the generation of knowledge, there needs to be a clear way of extracting and sharing this knowledge gained from the design actions (and the associated reflections) themselves [77, p. 172]. This has been an ongoing debate for some time [23] [35], and the jury is still out on how exactly this should be done. Some have argued that RtD work can be vital in that it leads to the production of *design knowledge* [35, p. 3], knowledge which is mainly intended to be useful for designers. In this case this is considered as separate from knowledge created through HCI research, which is empirical in nature and generally exists on either the level of theory or instance [35, p. 1]. Therefore, in order to assess the potential knowledge contributions of my RtD work in this project, I will review and discuss this process as a part of chapter 6.

3.6 Ethical Considerations

As with any form of research, there are certain ethical considerations that must be addressed as part of the work. In the general context of my work I have consciously worked to act in accordance with the guidelines for research ethics defined by The Norwegian National Research Ethics Committees [25]. These guidelines build upon the four principles of *Respect*, *Good consequences*, *Fairness* and *Integrity*. As a researcher, it is my responsibility that the work I do is carried out in an ethically sound way.

3.6.1 Ensuring Ethical Accountability in Research

In the qualitative study I adopt the mindset of Kvale & Brinkmann [51], who state that there are four key areas of ethical considerations to keep in mind when performing a qualitative interview, namely *informed consent*, *confidentiality*, *consequences of the research* and *the role of the researcher* [51, p. 102]. In this section I will briefly assess each dimension as it relates to the research-aspects of this project.

Informed consent

In order to obtain informed consent, all research participants must be informed about the goals of the research, the potential risks and benefits of participating, as well as the general structure and design of the study they are taking part in [51, p. 104]. In this project I aim to achieve this by providing detailed information about the project both in writing through the consent form signed before participation, as well as verbally during introduction and debriefing phases of the interview. The consent forms have been reviewed by both my thesis advisor, as well as the Norwegian Centre for Research Data (NSD) before use. Furthermore, all quotes from the interview used throughout the thesis have been reviewed and cleared by their originator.

Confidentiality

The concept of confidentiality involves coming to terms with how any data gathered from participants should be handled by the researchers in a given study [51, p.106]. This is especially important in research dealing with themes of a sensitive nature (ex. a project wherein participants discuss their political beliefs). While no sensitive information is collected or discussed within this project, I am collecting personal data in the form of a participants' voice. This was something I deemed necessary in order to obtain a sufficient amount of qualitative research data for the analysis, which in turn required me to notify the Norwegian Centre for Research Data[2] of my project before I could proceed with data collection. I filed a request which was subsequently evaluated, and finally approved. A copy of the application used can be found in appendix A. I was notified of their approval on the 11th of April 2019. A copy of their assessment can be found in appendix B. In submitting a notification form to NSD, I have agreed to process the collected information according to the practices described therein. In practice, this means that the raw interview recordings are stored securely on an external server run by University of Oslo through their Nettskjema[3] service. This information is only to be accessed by me, my advisor or the party responsible for the processing and storage of data (i.e. UiO). All interviews were recorded using the application *Nettskjema-Diktafon* published by University of Oslo[4].

Consequence

The participation in any form of qualitative research could result in both the injury and/or gain for a participant [51, p. 107]. According to Kvale & Brinkmann, one should conduct research in such a way that the potential gain for the participant, in addition to the resulting contribution of knowledge, is greater than any potential risk of harm to a participant [51, p. 107]. Within my project there is little risk of any injury for a participant from the outset. However, the length and intensity of the interviews could potentially lead to participants divulging information of a sensitive nature. I am aware of this risk, and I will aspire to act in a respectful and dignified manner towards participants. If such a situation should arise, the participant is free to withdraw his or her information from the study as described in the consent form.

The Role of the Researcher

In a qualitative research project, the quality of the research relies heavily on the integrity and moral fibre of the researcher [51, p. 108]. There are certain ethical requirements both with regards to how one should act in process of data gathering, as well as with regards to

[2] `https://www.nsd.no`
[3] `https://www.uio.no/tjenester/it/applikasjoner/nettskjema/`
[4] `https://play.google.com/store/apps/details?id=no.uio.mobileapps.diktafon&hl=no`

validity of the knowledge presented. One should strive for transparency in all published results, while making sure to conduct the research in a respectful and honest manner [51, p. 108]. In this project I will do my utmost to act in an ethical manner. Through this report I aim to present my work in a fully open and transparent way, describing in detail every single step I take to arrive at my conclusions.

3.7 Summary

This chapter has provided insight into the methodology of Research through Design, as well as the various other methods used in each phase throughout this project. In practice, my methodological approach has been informed by both theoretical frameworks [50] [87] as well as a case example [53]. These have been essential in providing both a structural blueprint as well as an academic grounding to the work. Furthermore, the design and prototyping process has made use of a dedicated framework [36] to help evaluate and contextualize each single prototype developed throughout the project. The design phase culminates with the completion of a high-fidelity prototype developed to explore the project research questions through field deployment. This field research takes the shape of a semi-structured qualitative study, and uses group interviews in combination with observation to provide rich qualitative data. This data is then transcribed, coded and subsequently analyzed. In addition to constructing knowledge through analysis of data, Research through Design aims to produce knowledge from the design processes themselves. In order to extract academically sound knowledge contributions from these processes, I introduce the ideas of intermediate level knowledge and strong concepts. Lastly, this chapter also provides a brief discussion regarding some ethical considerations relevant to research project such as this.

Chapter 4

Prototype Design

"I don't have a theoretical language for music. I have this abstract dream language. I'm really inspired by sculpture, so I like to play this trick on myself and say, 'You're not making music, you're creating a space. You're building a room, putting some objects in it, and seeing what happens to the objects over time.' From then on, I'm totally free."

- Daniel Lopatin

This chapter describes the design and implementation of a prototype in the form of a mobile application, developed to enable exploration of the research questions defined in section 1.2. The presentation of the design process is strongly rooted in the project methodology of Research through Design, as described in section 3.1.2. As many have emphasized [87] [2] [50], an essential part of any RtD project is documenting the design process meticulously from the initial stages, until the very end. This also includes presenting any failures or unsuccessful ideas that result in dead ends or less fruitful results [87, p. 185]. This line of thought is further echoed by Gaver [27], who in speaking of RtD work explains how practitioners can be seen as taking part in a form of implicit conceptual work by discussing the various influences and rationales for their design decisions throughout a project [27, p. 938]. By doing this, and at the same time assessing what they have made, they are said to be:

> "*...highlighting important issues, dimensions of similarity, and criteria for choices and success.*" [27, p. 938].

Rooted in this way of thinking, the contents of this chapter describe and present the central activities, thoughts and ideas that have contributed to the design process throughout. By doing this, and reflecting on the process along the way, I aim to shed light on how each decision along the way helped continuously shape and evolve the framing of the problem space, which ultimately led to the creation of the final prototype artifact in the form of the application *Petals*.

I begin by defining the initial design concept behind the prototype. In this section, I clarify the goal of the prototyping process, and describe how the core ideas behind the prototype were formed. Moving on, I present video prototyping as a way of creating low fidelity prototypes for augmented reality, before describing how I performed this process myself as a part of the prototyping. This is followed by a section describing the various technologies selected for use in the high-fidelity prototype, and the reasons for their selection. Following this, I briefly describe the process of designing the sound used within the prototype. Moving on, the subsequent section describes the complete implementation of the final high-fidelity integration prototype. This section is broken down into several sub sections each describing how I designed and implemented a specific part of the prototype. Lastly, the chapter concludes with a brief summary which reiterates the most important moments of prototyping process, highlighting the major points of significance in each previous section.

Recounting this process in a coherent way has been somewhat of a challenge, as the design work has more than likely been influenced not only by the related work described in chapter 2 and the methods described in chapter 3, but also by tacit knowledge, existing skills and subconscious processes within myself. I have tried to remain conscious of this along the way, providing additional insight and information where needed. My presentation in this chapter draws inspiration from the tone and style seen in the work of Carcani, Hansen & Maartman-Moe [12], as well as Ljungblad [53]. Similar to Carcani et al. [12, p. 271] I began my project by defining a design concept.

4.1 Initial Concept Definition

As stated by Zimmerman & Forlizzi, an essential part of the initial design activities in RtD projects is performing a literature review [87]. This is necessary both to gain an understanding of the current state of the art within the field, as well as to learn which problems other researchers are currently working on in a similar space [87, p. 185]. Based on my adoption of the five-step model, this is where I too began this project.

The work done in the initial research process resulted in the literature review found within chapter 2, which provides insight into key pieces of research in the fields of *augmented reality*, *computer-supported cooperative work* and *sound and music computing.* Furthermore, in section 2.3.4 of this chapter I present a few very concrete exemplars of collaborative interaction with music and sound in augmented reality, namely the work done by Kiefer & Chevalier around audio augmented reality [49], as well as the art installation *Bloom: Open Space* [73].

The research performed during this process was fundamental in shaping my thinking early on in the process, and the above-mentioned exemplars had a significant influence on the definition of my initial design concept. However, before I started outlining the main conceptual design of my prototype, I began by clarifying the goal of the prototyping process and defining the requirements for a prototype.

4.1.1 Defining the Goal of Prototyping

The primary purpose of the integration prototype was to enable exploration of the research questions defined in section 1.2. Therefore, the system developed as a result of the prototyping should provide the necessary functionality required to gather data relevant to the problem space. As described in chapter 3, this data was to be gathered by performing a field deployment. In more practical terms, this means that the prototype should allow for the cooperative interaction with music in an augmented reality space, done in a real-world environment outside of the strictly controlled conditions of laboratory experiments. In addition, the prototype should make use of binaural audio, in order to enable exploration of the second research question.

In the context of this chapter, the prototyping process should be seen in light of it being an executive part of Research through Design work. This means that the documentation generated through the prototyping process could provide value in that it narrates and describes a highly generative design process. As discussed in section 3.5.2, this process could be considered to have epistemic value in itself [2].

Having outlined the conceptual requirements of the prototype, I moved on to establish the initial framing of the prototyping process, and defining a design concept rooted in the theory presented in chapter 2.

4.1.2 Constructing a Design Concept

During the initial concept definition phase, my primary goal was to expand my understanding of the problem space, and from that work towards proposing a problem framing for my research question. Seen through the five step RtD model by Zimmerman & Forlizzi, I was at this point "*..searching to understand what the state of the world is...*", and how my work might "*...offer a new perspective, a new problem framing, which provides a path to a preferred future.*" [87, p. 185]. With my understanding of the problem space rooted in the existing work and exemplars presented in chapter 2, I began exploring possible avenues of research which could introduce a new perspective on the ideas and concepts found therein.

During this process, I considered the art installation *Bloom: Open Space* and the *Listening Mirrors* project [49] as particularly interesting and relevant exemplars of work, as I saw these as encapsulating the major concepts found within my primary research question, namely the use of augmented reality as a medium for cooperative interaction with sound and music. Furthermore, both projects revolved around interactions that took place in the context I had envisioned for my prototype, namely the *same time-same space* dimension (as seen in the CSCW matrix described in section 2.2.1). In addition, I was inspired by the distinctly cooperative nature of musical expression taking place between users in Bloom: Open Space. Within the augmented reality space, everything you do becomes part of the shared soundscape, blending together and creating a collective expression created and experienced by everyone [73]. This line of thought also echoes through the work by Kiefer & Chevalier, wherein they state that

> "*...When a sound installation environment itself becomes more instrument-like then it can become a collective instrument and consequently a shared channel of communication between audience members in the AR world, despite the personal or even intimate nature of their experience.*" [49, p. 28]

This quote resonated with me, and I began asking myself how I might aspire to create a similar experience, where the collective instrument itself exists in a shared augmented reality space. Additionally, i found the mobility provided in the work by Kiefer & Chevalier [49] very appealing, as it only required the use of mobile phones, a format I saw as ideal for creating augmented reality spaces out in the field. On the basis of these exemplars, I began considering how I might draw on each of the above-mentioned concepts in my design. I wanted to incorporate the concurrent musical cooperation found in Bloom: Open Space, and present it in a mobile augmented reality format, suitable for execution in a field environment. Furthermore, while Bloom enables cooperation between up to 12 people, I wanted to reduce the scope of the interaction to involve only 2 people at a time, similar to Kiefer & Chevalier [49].

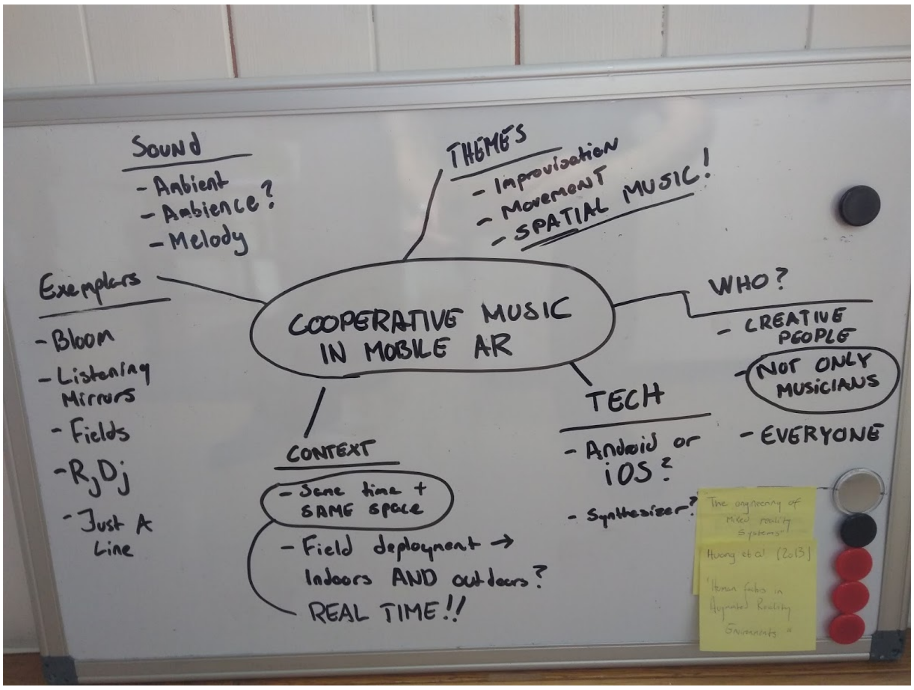


Figure 4.1: Brainstorming ideas for the design concept

4.1.3 Final Concept Definition

The above mentioned ideas ultimately became the foundation for my resulting design concept, which drew on inspiration from the art installation *Bloom: Open Space* and the work by Kiefer & Chevalier [49], both discussed in section 2.3.4, as well as the *Context Camera* project [53] presented in section 3.1.2. These concepts - along with theory presented in chapter 2 - were combined together to form the design concept, which can be summarized the following way:

> *Exploring synchronized and ludic cooperative interaction with music in mobile augmented reality space.*

Connection to Research Questions

The design concept was developed with the primary goal of exploring the main research question within the project, e.g. **how we can design for cooperative creation and experience of music in mobile augmented reality**. This was to be done through a prototype, as described in section 4.1.1. The creation and deployment of the prototype is itself a way of probing this question, and exploring how we might design and model ludic interactions with music and sound in augmented reality space. Furthermore, the prototype was to be utilized for exploring how binaural audio might affect immersion in mobile augmented reality. The motivation behind this primarily came from the related work on spatial audio discussed in section 2.3.3 and its effects on immersion in augmented reality. This was also motivated by the goals of sound and music computing addressed in section 2.3. The connection to the third research question is perhaps most clearly seen when considering the context of the prototype interaction, namely its use in the "same time, same space" dimension of cooperative work. As discussed in section 2.2.1, a central goal in CSCW research is to achieve a high degree of harmonious collaboration between participants. I therefore saw the prototype as an opportunity to shed light on which mechanisms of awareness might be required to enable a high degree of coordination when cooperating in a ludic fashion within an augmented reality space, using the prototype as a probe to examine this.

As discussed in chapter 3, evaluation was to be done in the field context, and the resulting artifact or system needed to be capable of - and suited for - deployment in the field. Furthermore, as the concept deals with mechanics such as cooperative augmented reality and creation of music, I considered it likely that a prototype capable of everything required would need to have some degree of complexity, potentially requiring significant development time. Before moving to the implementation phase, I therefore began by exploring the concept in a simpler and more immediate manner through low-fidelity prototyping.

4.2 Video Prototyping for Augmented Reality

The value of exploring ideas through iterative prototyping can also be clearly seen in the work by Ljungblad [53], wherein several early prototypes are constructed before the main high-fidelity prototype is realized. Each of the earlier prototypes contributed in its own way to the final design and helped iteratively develop both the functionality and the conceptual framing of the project along the way. Furthermore, I consider this as representative of the executive RtD process in general, which largely relies on carrying out an iterative design process that evolves and refines an idea from its nascent beginnings, to the final completed form [2, p. 185].

According to Zimmerman & Forlizzi, the design activities in an RtD project can be performed in a variety of ways [87], and might be performed by ex. exploring a new material or investigating ideas in a studio environment [87, p. 185]. At this point I was still in an explorative phase in the project. While I had defined the broad strokes of my design concept it still only existed mostly as a loose assemblage of ideas and thoughts, and had yet to be actually materialized in a prototype. As the first step in this process, I began exploring the concept through a low fidelity prototype. This way, I intended to evolve the concept into something more tangible, while also generating additional ideas before beginning implementation of the actual prototype. In a way, this work functioned as a sort of tangible brainstorming, allowing me to delve deeper into the concept by realizing parts of it in a simpler form.

4.2.1 Low Fidelity Prototyping through Video

Low-fidelity prototyping for mobile applications often involves drawing sketches, creating index cards or developing wireframe representations of an idea[1]. This way, one might quickly explore key concepts such as the navigation or user interface in a given application. However, when first beginning the prototyping process I felt like these mediums were less appropriate for my concept, and I was unsure of how I should approach this process. The medium of AR differs significantly from traditional mobile or desktop applications, and therefore might benefit from a different approach in the prototyping too.

Through researching how to prototype for augmented reality, I considered two alternative approaches to quickly generate solutions for augmented reality

1. Create a mock-up using AR prototyping tools.
2. Using video to simulate the experience.

The first option was to mimic the experience using a dedicated AR prototyping tool [2][3]. These provide a platform for quickly representing 3D models and graphical content in AR space, bypassing the manual implementation in code. However, research also led me to a talk given at the Apple Worldwide Developer Conference event in 2018 [39], in which video-prototyping is presented as a way of performing low-fidelity prototyping for augmented reality. Using a camera-enabled phone and some commonly available software, several key aspects of an augmented reality experience can be explored almost instantly in a cheap and simple way, without having to write a single line of code. At this point, I was feeling drawn to video prototyping, as it felt like a very immediate and tangible way to explore the design space. In addition, I felt like the problem space was still very ambiguous and complex, and I was unsure what I would gain from using an AR prototyping tool at this stage. I therefore elected to create a video prototype and began thinking of how I would do this in practice. The following section describes how I performed this process in my project, and presents some findings which have been carried over to the integration prototype.

4.2.2 Designing a Musical Video Prototype

As a starting point for my video prototype, I asked myself how sound and music could be represented within augmented reality. As I began developing the video prototype, I was looking for a way to embody or portray the creation of sound or music that might be realized in a video without requiring either special props or objects (or advanced video editing skills). To guide my process, I looked towards the sources of inspiration described in section 2.3.4. Here I considered how sound was represented in the central exemplars [85] [49] [73].

[1] https://www.interaction-design.org/literature/article/don-t-build-it-fake-it-first%2Dprototyping-for-mobile-apps

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

[3] https://www.torch.app

The Fields application [85] visually represents audio in augmented reality as something resembling floating clouds[4]. While I found this manifestation to be somewhat novel and aesthetically pleasing as an experience, it did not seem feasible to recreate in a physical real-world setting for the video prototype. Moving on, the work by Kiefer & Chevalier [49] bypassed the visual dimension entirely, and relied only on providing an experience using only audio as the medium of delivery. While seemingly effective (and also possible to do using video), I was interested in incorporating also exploring the visual dimension in the video prototype, and therefore did not pursue this any further.

I then moved on to looking at the Bloom: Open Space installation[5], which represents the musical notes as spheres floating in mid-air [73]. While this would still be difficult to recreate, it had a clear visual dimension and aesthetic, and the sound had been given a distinct physical form. I saw it as likely that something similar could be approximated either using physical props while recording, or by editing the video and using simple digital effects to simulate the experience. After some brainstorming, I came up with the idea to use rubber bouncing balls as props to represent an mock-up of the floating orbs seen in Bloom: Open Space. Although it would not be possible to make the rubber balls actually float in mid-air like the nodes in Bloom, I would at least be able to make them airborne for a short time by bouncing them on the floor and sending them flying. Furthermore, their shape could be seen as reminiscent of the spheres of Bloom: Open Space. While playing around with the idea of musical rubber balls, I also collected some additional exemplars for inspiration[6][7] wherein music was created from spherical objects. Inspired by Bloom and these exemplars I set about implementing the video prototype, starting by purchasing a few rubber balls from a local toy store. These can be seen in figure 4.2. As the environment of execution was less important, I quickly began recording the video in my own living room, as all I needed to get started was a floor and a mobile phone.

4.2.3 Executing the Idea in Practice

I began the session by opening the camera application on my phone. Then, while aiming the camera forwards in a slight downward angle, I threw the two rubber balls in front of me in a sequential fashion, one after another. While these bounced nicely off of the floor, they simply made a slight thump when hitting the ground. During this process, I played around with the idea of the balls instead emitting a pleasant and harmonious tone on impact, inspired by the exemplars presented in the previous chapter. I repeated this process a few times over, while simultaneously recording videos of each consecutive session. After a few minutes I had a few different videos of sufficient quality to work with. Then, as the next step I transferred the videos to my computer to continue the work.

With the creation of music being a key idea in my design concept, It was natural to incorporate this as a part of the video prototype as well. Using the digital audio workstation software FL Studio[8], I recorded a few short audio tracks for some of the clips wherein musical notes play in time with the video, each tone being played as the rubber

[4] `https://vimeo.com/269919192`
[5] `https://www.youtube.com/watch?v=-vQ_DYWh734`
[6] `https://youtu.be/hyCIpKAIFyo`
[7] `https://www.youtube.com/watch?v=IvUU8joBb1Q`
[8] `https://www.image-line.com/flstudio/`

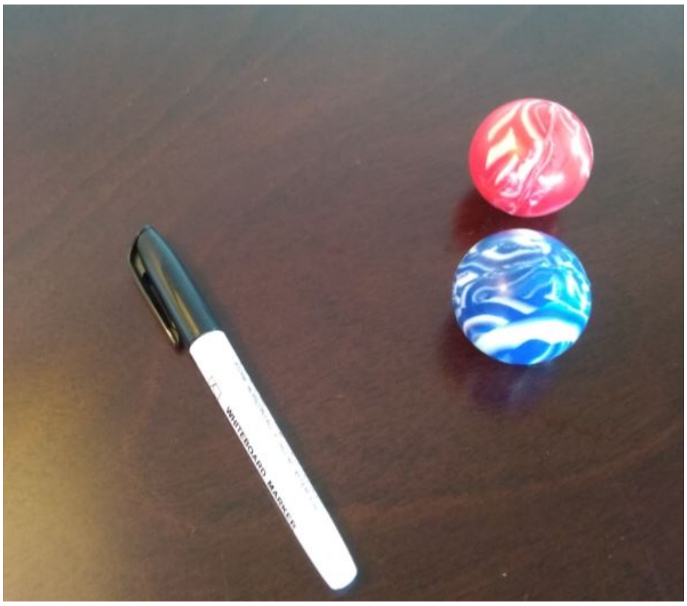


Figure 4.2: A photography of the rubber balls used for the video prototyping. A whiteboard marker is included for size comparison.

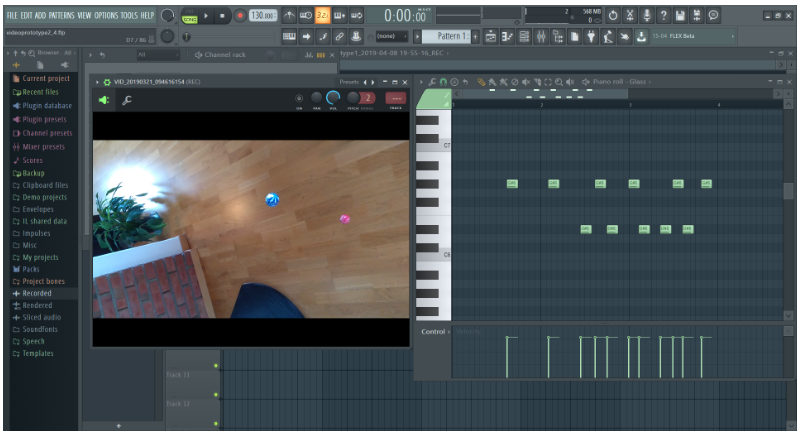


Figure 4.3: Preparing the audio track for the prototype in FL Studio.

balls hit the floor. I then used the basic video editing features of the built in Windows-software Microsoft Photos to edit each clip, muting the original sound and overlaying the audio to each respective video. By doing this, I was creating the illusion that each rubber ball produced musical notes.

Finally, I transferred the videos back to the phone and returned to the space where I had recorded the video before playing the video back on my phone, suddenly allowing me to experience the novelty of musical rubber balls playing out in front of me. Even though the experience was short, and the rubber balls came to a halt rather quickly, the video prototype had successfully taken the concept from an abstract idea and towards something more tangible. In addition, I was able to discover a few key takeaways from the video prototype that would later be of use in the development of the integration prototype. I wrapped up the work on the video prototype, and moved on to reflecting on what I had learned.

4.2.4 Reflecting on the Low-Fidelity Prototype

In retrospect, i consider the video prototype to have been successful. First of all, It had value in that it allowed me to quickly experience and test a version of the design concept without having to spend time implementing it as an actual augmented reality application. Seen in the context of the RtD model, it also offered a "*...different framing through its embodiment of a solution*" [2, p. 185], and helped expand my cognitive model and vision of what might constitute a suitable execution of my design concept. While I had initially been captured by the Bloom installation and its floating nodes, the video prototype had provided a different interpretation of the concept, by seemingly producing sound from contact between the rubber ball and the floor. I was intrigued by this idea, and decided to bring this with me to the next iteration of prototype development.

In addition, the video prototype provided some hands-on knowledge with regards to the look and feel of the experience. The first (1) important takeaway relates to the movement speed of a node. In several of the attempts, the rubber ball rapidly bounced in the direction of the throw. My efforts to then follow the ball with the camera were rather fruitless and resulted in a chaotic and unsatisfactory experience. Therefore, the movement of any graphical elements in an augmented reality prototype should likely not be of a high velocity such that users have to move their device a great deal. Another takeaway (2) relates to the range of movement, as during two attempts the rubber ball bounced with such height that it moved outside the vertical frame of the camera, effectively disappearing from my view for a short time. It would therefore likely be beneficial for users if any graphical components are given limited vertical movement range within the application. In addition (3), the size of any virtual objects should be assessed with care during development. My experience during video prototyping was that the rubber balls were too small, and when this was combined with the high velocity it provided a sub-par experience.

While the lack of a cooperative component in the video prototype somewhat reduced the overall usefulness in the context of the main research question, it did provide value in the form of the above-mentioned findings. I carried this with me to the next stage of prototyping. The prototype mainly explored facets relating to the look and feel of the prototype and bypassed the dimension of implementation entirely. The latter was intentional and can be considered a benefit of the format. The model in figure 4.4 shows the video prototype as seen through the prototyping model [36]. Two examples of videos resulting from the video prototyping process can be found in appendix G. After having carried out the low fidelity prototyping, I was more confident in the validity of my concept, and I therefore decided to move on to creating an augmented reality implementation of my concept. The first step in this process was to assemble the technology required to realize the concept.

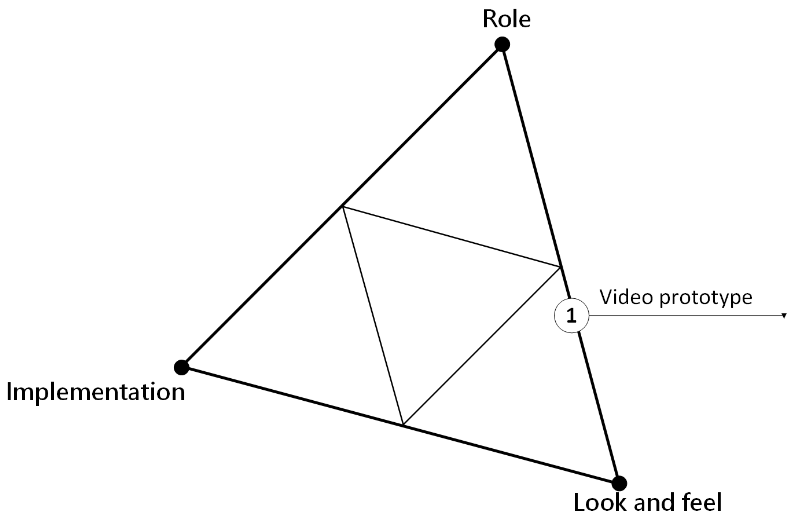


Figure 4.4: The video prototype as seen through the Houde & Hill model

4.3 Technology Choices

In this section I describe the requirements and the thought process behind the technology choices made during prototyping. The choices made in this phase had a significant impact on the final design and of the prototype itself. Furthermore, some theoretical concepts presented in this section are vital in providing background knowledge for understanding later sections. By assembling this configuration of components I intended to provide a technical framing around which I could creatively explore ideas, and enable the iterative generation of solutions to move the prototype development forwards. I began this work by selecting a mobile platform for the prototype.

Moving on from the low-fidelity prototyping, I felt like the design concept had become slightly less abstract. Through the video prototype I had created a (albeit simple) manifestation of the concept, and while it was still far from capable of exploring the research questions, I had gotten some new creative input through the experience. I felt a little more confident that I was moving in the right direction, and I began considering how to implement the concept as an augmented reality application. However, before I could do this I needed to select which technical tools I would use to realize this.

4.3.1 Deciding on a Mobile Platform

The use of mobile devices as the platform for this prototype application was rooted both in my primary research question, as well as in inspiration gathered from existing work in the field [49] [53]. Before I could assemble the technological stack on which to develop the prototype, the first step along the way was to decide on which mobile platform to use. Today, most of the mobile devices available to consumers generally use either iOS[9] or Android[10] as the operating system, narrowing the choice down to these two alternatives. Looking towards the methodology, Zimmerman & Forlizzi describe how it is essential to consider the skills possessed by the research team when beginning work on a RtD project [87, p. 185]. In the case of this project, the research team consisted of only myself. With regards to my own skill-set, I have a reasonable amount of experience developing for the Android operating system, both as a hobby developer and from working as a teachers assistant for two semesters at Østfold University College in the course "Android Programming". Furthermore, I already own a mobile phone capable of running augmented reality applications on the Android platform. I therefore quickly decided to develop the prototype as a mobile application for the Android operating system, and began exploring the available software development kits for augmented reality on the Android platform.

4.3.2 Selecting an Augmented Reality Development Kit

While researching which AR kit to use I quickly became aware of the many alternatives out there at the time, each with their own strengths and weaknesses. To guide my selection process, I looked towards the most central themes found within both my research question and my design concept as defined in section 4.1. As previously mentioned, an important factor when selecting a development kit was the ability to support *cooperative interaction* in a "same time, same space" context, i.e. the ability for users to engage in a real-time shared augmented reality experience within the same space. As I had little experience within this realm, I began researching whether any of the alternatives would support this functionality. During this research, I quickly discovered the ARCore framework developed by Google which provides all the necessary capabilities to enable cooperative AR experiences[11]. Furthermore, the framework is free and open source, and could be assumed to have a high degree of compatibility with the Android platform due to being developed by Google. On the basis of these factors, I decided to use the ARCore platform to realize the augmented reality functionality in the prototype. After having decided on this, I immediately began exploring the framework through examples provided through the official documentation[12]. In doing this, I intended to both familiarize myself with the technical aspects of ARCore, while also verifying that my device was fully supported and capable of running the example apps. I experienced it as easy to get started using the examples, and felt confident in my choice of platform for the augmented reality elements of the app. To provide some insight into how the framework itself works, I will over the next few sections present the framework and give a brief introduction to the main underlying concepts within ARCore.

[9]https://www.apple.com/ios/ios-12/
[10]https://www.android.com/
[11]https://tcrn.ch/2KMzqa3
[12]https://developers.google.com/ar/develop/java/quickstart

4.3.3 The ARCore Augmented Reality Toolkit

From a technical point of view, ARCore is dependent on three fundamental capabilities to achieve its augmented reality functionality, namely *motion tracking*, *environmental understanding* and *light estimation*. The motion tracking capabilities are required for the software to know the position of the phone relative to the physical environment around it. This is done through a technique called *Concurrent Odeometry and Mapping* (COR), which uses visually distinct features in the environment captured from the camera feed (so called *feature points*) to calculate any changes in location. This information is then combined with data from the built in gyroscope of a device to estimate its position and orientation over time within the real world. Within the domain of ARCore, this orientation is generally referred to as the *pose*. This is what makes it possible to present the three-dimensional content from the correct perspective at any given time, even when the user moves around it in the physical space. Furthermore, ARCore achieves its environmental understanding by identifying clusters of these feature points on horizontal or vertical surfaces, typically a table or floor. This is how the framework is able to get an understanding of the room. When registered by ARCore, the surfaces are then made available through the application as *planes*, and can then be used to place any digitally rendered content on. As the system relies on the discovery of these feature point clusters in order to generate planes, it is generally recommended to avoid targeting surfaces with little to no distinct features or texture[13]. As an example, an empty white wall in a dimly lit room might not register properly when scanning, as there are to few feature points to register it as a plane. Finally, in order for the virtual content to blend in with the environment, ARCore detects information about the lighting of the surrounding environment from the camera image and uses this information to light the virtual objects accordingly. This concept of **light estimation** helps make the rendered object look more natural, and helps increase immersion within the experience. When placing a rendered object, ARCore requires it to be attached to an *anchor*. An anchor in ARCore represents a permanent location and orientation within the real world as seen through the camera, and is what helps make sure that any placed objects stay put. These are typically created by tapping the screen over an area of a scanned plane. In ARCore, an anchor can have multiple virtual objects attached simultaneously. As the objects are attached to the same anchor, their relative positions will be preserved even when the anchor adjusts its pose, making them resilient to changes in pose. Anchors are a central concept in ARCore, and is essential to realizing cooperative AR experiences through the use of *cloud anchors*.

Cloud Anchors

In addition to providing the fundamental augmented reality capabilities such as environmental understanding and motion tracking, ARCore provides the necessary building blocks required to create collaborative augmented reality experiences. This is made possible using the *Cloud Anchors*[14] API, which is included in the ARCore platform by default. As previously discussed in section 4.3.2, the existence and availability of this service was crucial in my choice to use ARCore as the framework in this project, as it was an essential component in providing the prototype with its required functionality of cooperative real time interaction in augmented reality. Using the Cloud Anchors API within ARCore, it

[13] `https://developers.google.com/ar/discover/concepts`
[14] `https://developers.google.com/ar/develop/java/cloud-anchors/overview-android`

is possible to provide several devices with a shared point of reference in the environment, thus enabling the construction of a shared augmented reality experience. Similar to how a regular ARCore anchor functions, a cloud anchor also represents a reference to a point in physical space. However, instead of only existing on a local device like the ordinary anchors, cloud anchors can be shared between devices and can therefore be used to provide several devices with the same reference point in physical space. After creating and hosting a cloud anchor on one device, it can then be resolved by other devices using what is known as a *cloud anchor id*. This takes the form of an encoded string, and is generated and returned from the Cloud Anchors API upon the successful hosting of a cloud anchor.

As long as both users have scanned and detected the same feature points in the physical space, meaning that their phones have a similar independent understanding of the same environment, this Cloud Anchor id can be used to resolve the same ARCore anchor on both devices, thereby creating a shared reference point in augmented reality space. In practice this makes it possible to place objects in the augmented reality space on one device, and have them appear at the same location on other devices as long as they share this anchor. Reflecting back on the concept again, this is what would enable me to construct a shared AR space similar to seen in Bloom: Open Space, where the same content was visible for multiple users simultaneously. However, in order to actually render and present 3D content within the augmented reality space, a library or framework capable of 3D rendering would also be required.

4.3.4 Using the Sceneform Framework for 3D Rendering

In ARCore, there are three supported alternatives for 3D rendering: Unreal[15], Unity[16] and Sceneform[17]. The two former alternatives are extensive and sophisticated game engines, with capabilities far beyond what I would likely require for this project. I am also unfamiliar with both, and they both use languages with which I am not very familiar (namely C++ and C#). However, in exploring the examples from the ARCore documentation, I had briefly been exposed to the Sceneform library through its use in the examples. While significantly less advanced than both Unity and Unreal, it would allow me to use the Java language when implementing the graphical elements of the application. I therefore decided to adopt Sceneform as my 3D rendering framework.

The framework is currently being developed and maintained by Google, and is intended to provide 3D a simple way of creating and rendering 3D content on Android. It has been designed to work seamlessly together with ARCore, and is optimized for delivering 3D to augmented reality experience on mobile platforms. A key concept behind Sceneform is to enable Java developers to create 3D scenes without needing much knowledge of computer graphics. This was beneficial to me in that it enabled me to leverage my existing knowledge of Java to get started quickly, without having to spend much time familiarizing myself with 3D graphics programming in OpenGL. The Sceneform API enables the importing and use of existing 3D models in supported formats (*.obj, *.fbx, or *.gltf), but also natively supports the creation of simple shapes such as spheres, cubes and cylinders through the API. My familiarity with the Java language and the Android platform were key factors in my decision to use Sceneform, in addition to the library itself being compatible for use with the ARCore framework. As I have no practical experience with OpenGL and

[15]`https://www.unrealengine.com`
[16]`https://unity.com/`
[17]`https://developers.google.com/ar/develop/java/sceneform/`

graphics programming in general, the simplicity and ease of use were attractive features in my choice to use Sceneform. The library also contains several classes that provide a basic scaffolding for setting up a basic ARCore application, making it relatively uncomplicated to start development of an application. Similar to my approach with ARCore, I again explored the samples provided on the official GitHub page[18], to familiarize myself with the framework and API. In addition, I completed a short training course intended to introduce the most important concepts of the framework [19]. With ARCore and Sceneform providing the basic capabilities required to create the visual part of the prototype, I felt like I was moving in the right direction. As the next step, I began looking at how I should implement the auditory aspects of the prototype.

4.3.5 Selecting an Audio Engine

In addition to providing an augmented reality experience, the prototype had to in some way enable users to create and manipulate sound. Furthermore, to support the second research question of the project, the prototype had to be capable of providing a binaural audio experience. To guide this process, I once again turned towards the methodology and its focus on considering the skills of the research team, as well as my role as a designer.

Through many years of practicing music production as a hobbyist, I have become very familiar with the practice of synthesizer programming and sound design. Therefore, using a synthesizer module for the generation of sound would likely allow me to leverage this existing knowledge. Based on this, I decided to use a synthesizer as the main sound source in my prototype. However, while I had experience programming virtual synthesizers in the context of music production software, I had no experience working with synthesis on a mobile platform. In beginning the process of researching alternatives, I consulted my main sources of inspiration to see how they achieved their audio functionality.

In the *Musical Moves* system by Carcani et al. [12] they employed the Max/MSP[20] synthesizer system to generate audio in real time based on user input [12, p. 274]. The system appeared capable of doing most of what I was looking to achieve, and I saw this as a very attractive alternative for my own prototype. Similarly, the project by Kiefer & Chevalier [49] employed the Pure Data[21] synthesizer system to manipulate and process audio in real time on a mobile device [49, p. 26]. At a glance, Pure Data appeared to provide a lot of the same functionality as Max/MSP, while also being open source and free to use. I also quickly discovered that Pure Data was relatively easy to integrate into an Android app, and I therefore elected to adopt Pure Data as the synthesizer module within the prototype.

4.3.6 Embedded Pure Data on Android with libpd

As a language, Pure Data is written with sound synthesis in mind, and many common synthesizer components are provided as existing objects within the language. Pure Data itself is a full-fledged programming language and software system originally written in the C programming language, and it is therefore not immediately compatible with the

[18]`https://github.com/google-ar/sceneform-android-sdk`
[19]`https://codelabs.developers.google.com/codelabs/sceneform-intro/`
[20]`https://cycling74.com/`
[21]`https://puredata.info/`

Android platform by default. However, by using a bridging library, it can be implemented as part of an application. In the prototype, the use of Pure Data within the app was made possible by the external Java-libraries *libpd*[22] and *pd-for-android*[23]. Firstly, the libpd-library is an embeddable wrapper for Pure Data, which is built on top of the base distribution, i.e. Pd Vanilla. The libpd library enables Pure Data to be run in an even wider range of different environments, and reduces the need to write platform specific audio code for a given device (as long as it is supported by the library itself). In this case, libpd functions as a layer between the Android application and the Pure Data patch which it controls, making it possible to communicate with the Pure Data patch through an API. In this case, the prototype makes use of libpd through the library *pd-for-android*, which contains the necessary classes and interfaces to allow for painless integration into an Android application.

A model of the architecture as presented in the book *Making Musical Apps* [9] can be seen in figure 4.5. Much like Pure Data itself, both libpd and pd-for-android are open source and freely available online. Through following the official documentation[24], I was able to construct a relatively basic synthesizer without needing intimate knowledge of digital signal processing and audio programming. While the format and language was new, most of the concepts seemed similar to what I was familiar with through my experience with synthesizer programming. Feeling somewhat confident in my choice of Pure Data as a sound source, I moved on to identifying and selecting the required back end components needed to support the prototype functionality.

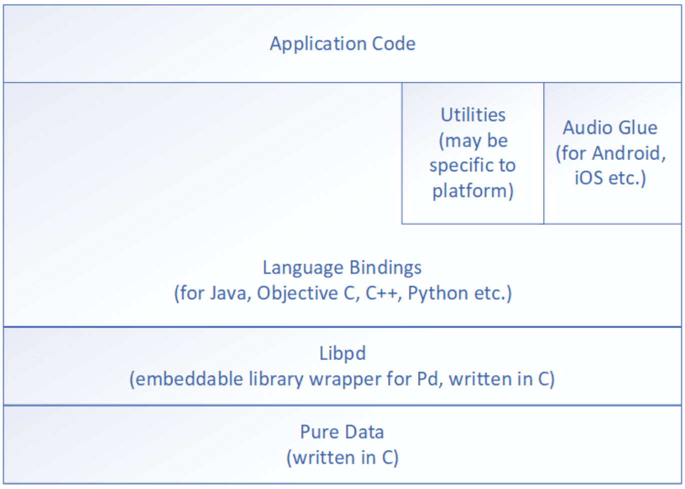


Figure 4.5: Layered model depicting the architecture of the libpd-library. The model has been reproduced based on a model in the book ”Making Musical Apps” [9, p. 44]. Within the context of this model, the prototype application is represented by the upper block labeled Application Code

[22] `http://libpd.cc`
[23] `https://github.com/libpd/pd-for-android`
[24] `https://puredata.info/docs`

4.3.7 Firebase Platform as a Backend

While exploring the cloud anchors system through the ARCore documentation[25], It became clear that I would need a way of sharing data between devices to achieve a cooperative augmented reality experience. To do this in practice I considered two different solutions, either I would (1) synchronize the data between users through an online database system, or alternatively (2) use a dedicated service aimed at providing real time multiplayer experiences[26].

While I had no experience using the latter, I had some experience using the online database provided through the platform Firebase. I also learned that Firebase was used within the Cloud Anchors tutorials provided by Google[27], as well as being recommended for the creation of shared augmented reality experiences within the book *Augmented Reality Game Development* [52].

The choice therefore seemed relatively easy, and I decided to make use of the Firebase Realtime Database[28] service. Firebase is a platform managed by Google that provides various back-end services and analytics targeted for mobile and web applications. Among these services is the *Realtime Database*, which is a cloud hosted NoSQL database that stores data in a JSON-like schema structure. This was the service which would provide me with the functionality needed to synchronize data across devices, and the final piece of the puzzle needed to achieve a shared augmented reality.

By using a service such as Firebase, I also removed the need to construct a dedicated back end server architecture for the prototype myself, thereby potentially reducing the required development time. There were a few drawbacks to this however. One such drawback was the fact that Firebase is developed and operated by Google. This meant that I would in practice give up some of my control over the data, in exchange for the convenience of not having to create and manage my own back-end server. I consider this a fair compromise, as none of the data recorded and used by the prototype is of sensitive nature, and would only be connected to myself and my devices. In addition, as previously mentioned the Cloud Anchors API which is central to the prototype functionality is also developed by Google, and I was therefore already subject to Google's data collection, privacy policy and terms of use anyway.

4.3.8 Version Control System

In order to track changes in the prototype code, the version control system Git was used during development. In addition, an online repository was created on the development platform GitHub[29]. Using GitHub has the advantage of enabling development across multiple machines, as well as functioning as a cloud solution for backup of code. During the project work, I utilized GitHub for version control of both the application code and the report source files in the raw latex-format.

[25] `https://developers.google.com/ar/develop/java/cloud-anchors/overview-android`
[26] `https://www.gamesparks.com/`
[27] `https://developers.google.com/ar/develop/java/cloud-anchors/quickstart-android`
[28] `https://firebase.google.com/docs/database`
[29] `https://www.github.com`

4.3.9 Development Environment

During development of the prototype application, the following development environment was utilized.

Software

- Windows 10 w/ Windows Subsystem for Linux
- Android Studio version 3.2.1
- Pure Data version 0.49.0
- GitHub Student Developer Pack

Hardware

- (2x) Motorola Moto G5s Plus 32GB
- Sennheiser Momentum Headphone
- Sennheiser HD 25-1 II Headphone

As the prototype relies heavily on use of augmented reality, I opted to continuously deploy the application to physical devices during the prototyping stage, mostly bypassing the use of an emulator. The devices used during development were also utilized for the testing and evaluation with users, which reduced the chance of bugs and errors occurring due to variations in hardware. Furthermore, while the work of Kiefer & Chevalier [49] employs bone-conducting headphones as a part of the experience, I have instead chosen to use ordinary over-the-ear headphones in my study. The setup used for both testing and the field deployment can be seen in figure 4.6.

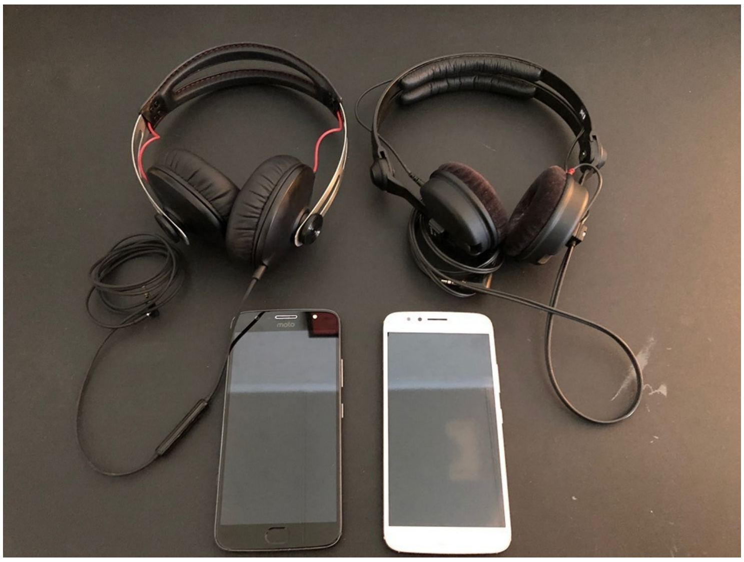


Figure 4.6: The equipment used for both testing and deployment.

4.4 Designing the Sound

In addition to the visual aspect of the prototype, another essential component of the user experience was the auditory aspect. Just as a 3D model must be designed and sculpted before it can be put into an environment, so must the sound.

As described in section 4.3.5, the prototype would use the Pure Data system to generate synthesized sound. In practice, this would enable me to construct and load a synthesizer patch within an Android application, which could then be controlled programmatically through the libpd-library. In addition to providing a high degree of customizability, using Pure Data as the synthesizer engine for the prototype had the beneficial side effect of enabling me to take a somewhat modular approach to the prototyping, meaning I could work on creating the synthesizer patch separate from the development of the prototype application. In other words, I could design and program the sound before actually having integrated libpd into the application. As long as i made sure to prepare the required input and output channels within the patch, the patch itself could be designed without knowing the exact details of the application implementation. Therefore, I began this process by considering *how* the patch should sound. As a starting point I considered the context of use for the prototype within this project, and how the sound would be experienced during use. This led to some reflections on what would be desirable characteristics of the sound.

Firstly, the collaborative nature of the experience meant that sounds would likely be played in an unpredictable and impromptu manner, as each participant would have the freedom to play a tone at any time during use. As a result, using a sound with a harsh and jarring character could at worst lead to users experiencing discomfort and anxiety during use. In addition, users would likely be playing notes simultaneously within the augmented reality space, potentially with a high level of activity, leading to sounds being played rapidly and with little space in between. In this case, using sounds with a long duration could have led to a feeling of clutter and messiness, as the overlapping sounds would bleed into each other for a sustained time period (much like when playing a piano while simultaneously holding down the sustain pedal). Furthermore, I considered the sound I had used in the video prototype, which can be said to contain some of the same characteristics as detailed above.

During this process I also turned once again towards the initial sources of inspiration, and decided to once again explore the *Bloom: 10 Worlds* application for Android as an exemplar of sound design. Within this application, one particular sound caught my interest, and further guided my vision of what I wanted the prototype to sound like. This sound[30] has a characteristic reminiscent of a bell, with a round and glass-like texture. In general, this type of sound is relatively simple to create using a synthesizer, and can be achieved by multiplying sine waves [13, p. 106] using a synthesizer. Having performed this assessment of how the instrument should sound, I now had an idea of the desired characteristics of the patch, and could begin the actual sound design using the Pure Data system.

As the Pure Data language might be unfamiliar to many, I will briefly present an overview of the language itself and its mechanics. This is done with the intention of producing knowledge from the sound design process as well, making it easier to understand it from both a technical and an aesthetic point of view.

[30] `https://youtu.be/_R7eY_xX8rk`

4.4.1 An Introduction to the Pure Data Programming Language

Pure Data (often abbreviated as Pd) is an open source visual programming language developed in the nineties by Miller Smith Puckette [64]. It is designed for multimedia processing, and offers tools to create and edit sound, video and digital graphics. The software and language has a lot in common with the previously mentioned software Max/MSP[31] used in the Musical Moves project [12], likely due to both Pure Data and Max/MSP being initially developed by Puckette himself. The base distribution of Pure Data runs on a wide variety of hardware configurations and operating systems, including Windows, Linux, FreeBSD and macOS.

Furthermore, through its active online community additional support has been developed for several other hardware configurations such as the Raspberry Pi single-board computer, and smartphones running iOS or Android. In addition to the base Pure Data distribution, several other community developed distributions are available for download on the internet. These are modified version of the original software, and introduce features such as improved GUI or extended functionality through the incorporation of external user-made modules not present in the original software. Popular examples of alternative distributions are *Pd-extended*[32] and *Pd-L2Ork*[33].

In order to avoid confusion when dealing with Pure Data, the unaltered and original base distribution is generally referred to as *Pd-vanilla*. This is also the version used in this project. In addition to its above mentioned use in the *Listening Mirrors* installation [49], Pure Data has been used in several notable large scale projects such as the first edition of the experimental tabletop musical instrument Reactable [17], and the major video game title Spore [44]. While these exemplars are somewhat different from my project in theme, they nonetheless function as key indicators of what can be achieved through the language.

Programs created with Pd are generally referred to as *patches*, alluding to modular hardware synthesizers where sound is created and manipulated by connecting physical patch cables on a hardware unit. In a similar manner, a Pure Data patch is made by instantiating objects within the programming environment and connecting these together via their ports (inlets and outlets) using virtual patch cables. The objects themselves are instantiated using keywords, which form the vocabulary of the language. The different keywords[34] represent different objects with various properties, and these are the basic building blocks of any Pure Data patch. As an example, typing the keyword **osc** when creating an object will produce a sine wave oscillator, while the keyword **dac** will create an audio output object converter. The base distribution of Pd contains the most essential objects, and users can add more through the use of *externals*. The image in figure 4.7 shows a very simple Pd patch. In this patch, a sine wave oscillator is generating a tone with a frequency of 440 Hz. This oscillator object is connected to an audio output, which renders the sound through the speakers of the computer.

[31] `https://cycling74.com/`
[32] `https://puredata.info/downloads/pd-extended/`
[33] `http://puredata.info/downloads/Pd-L2Ork`
[34] `https://puredata.info/docs/tutorials/pd-refcard`

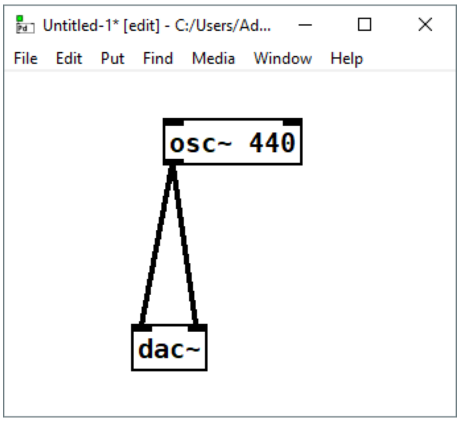


Figure 4.7: A simple Pure Data patch which produces a constant 440Hz tone.

4.4.2 Developing the Sound in Pure Data

Having established some guidelines with regards to the aesthetic qualities of the patch, I started out by considering if there were any functional requirements for the patch, and how these might be developed. The main requirement I identified was that the patch would likely need to be polyphonic[35], meaning it should support the simultaneous playing of several individual notes at the same time. As the application was intended to be experienced in a collaborative fashion, it was likely that at some point both participants would attempt to trigger the playing of a note within the app at the same time, either intentionally or just by chance. If the patch did not support polyphony (i.e. if it was monophonic), only one note would be rendered at a time, resulting in one participant not getting their note played. I therefore made sure that the patch was created with support for polyphony.

The sound prototyping process resulted in a patch which produces a short and smooth "bell"-like sound. The patch is structured to receive messages in the form of MIDI-notes from the application. Whenever a message is received by the patch through the "notein" object, it is processed and routed to one of ten voices within the patch. These ten voices are separate but identical instances of the same sub-patch. This voice-patch is what actually generates the bell-sound. This voice consists of one oscillator, which produces a sine wave tone according to the input received. This signal is then controlled using an envelope generator, which is what controls the changes in gain over time. The envelope is programmed to have a short attack time, and a short decay and sustain, meaning it decreases in volume quite quickly. In addition, the voice patch is set to receive the *release* parameter as an input. The release time of a note defines how long it takes for the sound to fade out to silence after having been played. If a note is played with a release time of 3000 ms, it will take three seconds before the note fades out to silence.

As mentioned, the main synthesizer patch is programmed to use up to 10 voices. Using multiple voices like this ensures polyphony, and makes it possible for multiple notes to be played simultaneously. In the patch, the voice object receives a message containing the MIDI note and velocity, and renders the resulting signal through its outlet. This signal is then routed through an external object which applies a reverb[36] effect on the signal,

[35] https://www.soundonsound.com/techniques/introducing-polyphony
[36] https://musicterms.artopium.com/r/Reverberation.htm

before it is finally sent to the main output. The reverb effect was provided by an external object [37], and was added to accentuate the feeling of space in the interaction.

Furthermore, when interacting with the prototype the users would ideally be wearing headphones most of the time in order to get the full sonic experience. This would in turn likely make them less aware of their surroundings, with both their visual and auditory focus directed towards the application. I therefore saw it as useful to enable users to incorporate input from the device microphone into the playspace. By simply creating a line-in object in the patch, and connecting it to the line out, I had enabled audio input within the patch. To enable some control, I implemented a simple volume control mechanism which I would use in the code to turn the microphone on or off. Following this, I routed the microphone input through the reverb module, to make it flow better with the rest of the interaction. The decision to apply effects on the input signal was inspired by the work of Kiefer & Chevalier [49], in which a core part of the concept was altering the real time sounds experienced by users [49]. Finally, I implemented the binaural audio capabilities by incorporating the **+binaural~** external as described in section 4.5.8. To make sure that the patch would be capable of positioning single nodes in space, I placed the binaural external within the voice-patch, and connected a receiver object. This way, it would be possible to rapidly update the position of a note according to its received azimuth value, before triggering it to render the filtered audio. With the incorporation of binaural audio the patch was functionally complete, and contained all the properties required to produce the intended audio experience. The Pure Data patch itself, along with a visual representation of it can be found in appendix G.

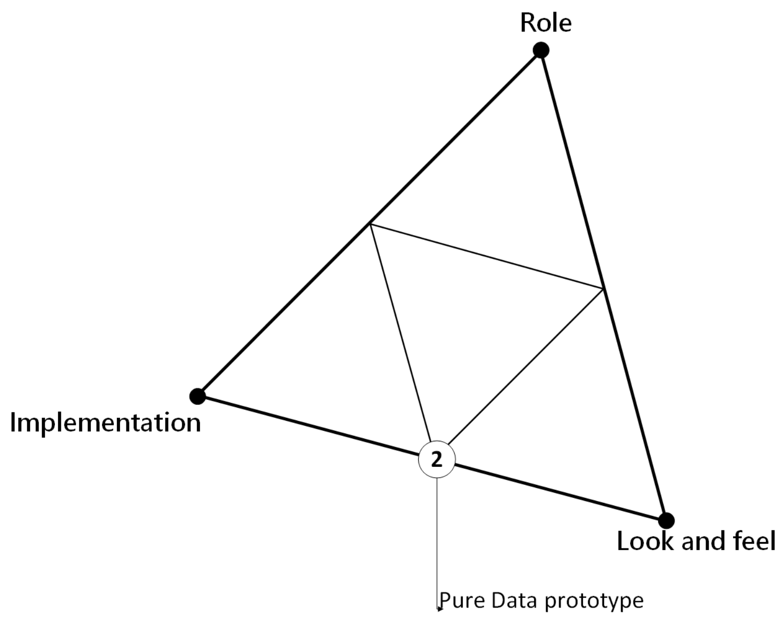


Figure 4.8: The Pure Data patch as seen through the Houde & Hill model

[37] `http://tre.ucsd.edu/wordpress/wp-content/uploads/2016/05/pdreverb.zip`

4.5 Integration Prototype Development: The Design of Petals

This section describes the process of designing and developing the final integration prototype, which became the Android application *Petals*. This name was chosen as a homage to the Bloom installation by Brian Eno & Peter Chilvers described in section 2.3.4, as my design concept draws inspiration from both Bloom: Open Space and Bloom: 10 Worlds. The name was chosen as the prototype represents only parts of the experience provided by these exemplars, much like how a flower petal represents only part of a flower in bloom.

The focus of this chapter is mainly directed towards communicating the high level concepts and ideas behind the many design choices made in the prototype. Some technical details will be given where they are deemed relevant or otherwise interesting as seen in light of the RtD methodology, but the focus will be more on *why* a design choice was made, instead of *how* it was implemented in code. Any details about the latter can be revealed by looking at the complete prototype source code which is accessible through appendix G, giving full insight into every detail of the development and implementation. The chapter begins with a section describing the transition from low-fidelity video prototype to the integration prototype.

Reflecting back to the points discussed in section 4.1.1, it was important to ensure in every step of the design activities that I correctly facilitated for exploration of the research questions within the project, and that I worked towards creating a prototype that would be suitable for use in a field deployment. At the same time, I wanted the prototype to deliver an experience that encouraged users to follow their ludic impulses, and that enabled spontaneous and improvisational musical interaction. In other words, in addition to supporting the research I wanted the experience to also be engaging and fun. A certain tension can be said to exist between the conceptions of usability central to most interaction design work, and the concepts generally driving the design of musical interactions. Where interaction design is often focused on making systems *easy to use*, musical instrument interfaces will often sacrifice some usability to allow for *long-term engagement* and *virtuosity* [56]. This intersection is described by McDermott et al. who state that

> "*A musician who learns and plays for love of music is in a very different mind-set from that of a software user, impatient to carry out a task.*"
> [56, p. 44].

This quote resonated with me, as it described what I saw as a possible pitfall in my design process. As the very heart of this project revolved around the combination of elements from both music and interaction design, I considered it important to assess this tension between usability and difficulty when designing the musical interface.

4.5.1 Translating the Video Prototype to Augmented Reality

At this point, I had assembled a technological footing for the prototype. Having played around with both ARCore and Sceneform during the technology selection process, I had already created a few small test apps to validate various parts of the functionality, and I therefore felt ready to begin working towards implementing my design concept as a proper augmented reality application. As a starting point, I looked to the video prototype for inspiration on how to begin. In the video prototype, described in section 4.2, simple bouncing rubber balls functioned as a metaphor for what was imagined to be a musical element in augmented reality. While this concept was perhaps quite simple, it had a very tangible and genuine feel, and there was a clear lineage to be traced between the execution of the video prototype and the ideas brought over from the sources of inspiration discussed in section 2.3.4. Furthermore, I felt the process of converting the design concept described in 4.1.3 to the video prototype had helped strengthen the idea, while also setting it apart from the main sources of inspiration [73] [49].

As discussed in the previous section, there is a tension between the principles of interaction design, and the principles of musical interface design. However, in thinking of this I became more aware of how the prototype I was aiming to create was not necessarily a comprehensive instrument, but rather a musical experience where people could play with and explore the space through music. This reminded me of one of the principles for designing computer music controllers, which reads

> "*Make a piece, not an instrument or controller*" [20, p. 3]

This quote resonated with me, and by grounding my design in the video prototype, I guided my work towards more strongly creating a piece. Revisiting the video prototype provided a starting point for implementing the simulated interaction from the video prototype, as an actual augmented reality experience. With this in mind, I began developing something which resembled a virtual counterpart of the video prototype.

I set up a new project in Android Studio with all the required dependencies, and using the helper class [38] class provided by the Sceneform framework, I immediately had a very simple app capable of rendering content in AR space. Then, using the ARCore and Sceneform frameworks, I was able to quickly implement the basic functionality required to generate a simple 3D sphere and render it in augmented reality space. As described in section 4.3.4, the Sceneform API provides the basic building blocks required for the generation of simple geometric shapes, removing the need to create a 3D model using external software. Then, by animating the movement of these spheres along the Y-axis, I was suddenly quite close to reproducing the look of a bouncing rubber ball similar to what I had used in my video prototype. This connection was further strengthened by assigning the color red to the sphere. I then set the sphere to automatically animate upon creation, so that each sphere I placed in the environment would instantly begin moving up and down upon creation. With the current prototype, I could simply open the app, scan a part of the environment, then tap on a registered surface, and I would be presented with a red sphere bouncing up and down like a rubber ball. The image in figure 4.9 shows a screenshot of this in action.

However, while the rubber ball used in the video prototype came to a halt rather quickly after being dropped on the floor, the virtual bouncing ball rendered in AR space continued bouncing endlessly until I stopped the application. While It was starting to look somewhat similar to the prototype, the virtual ball was still missing the element of sound that I had used in the prototype. In the video, each rubber ball produced a harmonious note upon contact with the floor. To then recreate this illusion within the AR counterpart, I made it so that a randomly generated tone was sent as a MIDI message to the Pure Data patch running within libpd whenever the virtual sphere appeared to

[38] `https://developers.google.com/ar/reference/java/sceneform/reference/com/google/ar/sceneform/ux/ArFragment`

Figure 4.9: Screenshot from the first iteration of the integration prototype. Four red spheres are animated to move up and down, similar to the bouncing ball in the video prototype.

touch the ground (when the animation restarted). This resulted in each sphere emitting a harmonious and pleasant tone whenever it bounced on the floor in augmented reality space. Suddenly, I had translated the idea and concept behind the video prototype to augmented reality. At this point I spent some time playing around and exploring the experience of this early version to generate ideas, which would hopefully guide me towards my next design move. While in this brainstorming state, I quickly noticed that since the animation in the prototype was set to perpetually repeat, this resulted in the formation of a musical loop within the augmented reality space. Furthermore, since the playing of each musical note was directly tied to the animation of the relevant node in augmented reality space, the structure of the looping musical arrangement was in practice represented physically in augmented reality by the 3D spheres. Therefore, manipulation of the 3D spheres could effectively be used as a ways of manipulating the musical arrangement itself. In a way, the musical arrangement had become spatial, and was represented visually through the bouncing nodes. This became the foundation for my next design move, as I decided to further explore this idea in the next iteration.

4.5.2 Obtaining Structure Through a Looping Arrangement

The concept of looping animations seemed like a very suitable way of connecting the dimension of *space* in augmented reality to the representation of *music* and *sound*. However, this way of representing music came with one potential limitation, namely the short repetition time of the animation and therefore by extension the musical arrangement. As the animation was set to repeat roughly every two seconds, the musical loop would then repeat too, and it was therefore not possible make the musical arrangement last longer than the animation without also losing the glue which tethered the *spatial* and *musical* dimensions together.

One alternative would be to simply extend the duration of the animation, thereby also extending the duration of the loop. This would in practice mean either (1) slowing down the movement of the balls to allow for a longer duration, or (2) increasing the vertical range of the animation by making them bounce higher up in the air. The former (1) would potentially risk breaking the immersion, as well as the connection to the design concept, as the spheres would behave less like their real world counterpart from the video prototype. However, implementing the latter change (2) would potentially mean allowing the 3D nodes to move outside the users field of view, thereby going against the findings discovered in the low-fidelity prototyping described in section 4.2.4, wherein the rubber balls bounced outside the camera field of view. During this process, I asked if it was actually necessary to extend the loop duration. While it could arguably be seen as a significant limitation, the existence of such a constraint might also be a good thing, functioning as a way of provoking creative processes and improvisation within the provided boundaries. Guided by these thoughts, I decided to keep the loop duration as it was, repeating every two seconds and forming a short but sweet loop.

In contemplating this further, I also saw how it could represent an opportunity to provoke action, as the quick repetition time made it so that users could likely quickly grow tired of hearing the same loop repeat, and might therefore also feel motivated to change it. Therefore, I was hoping that by representing the arrangement this way, and keeping the short loop length, it could be seen by users as something dynamic and malleable during play, provoking users to intervene and change the structure. In addition, the use of repetition as a technique within music is nothing new. In music theory, the term *ostinato* describes a short pattern of either musical or rhythmic nature, that repeats throughout a composition [39]. This technique has been used extensively in music for centuries, and can be found everywhere in both classical as well as modern music.

While the looping structure provided a frame for the arranging of music in the augmented reality space, it did not dictate the actual musical makeup of the arrangement. In other words, while the arrangement defined *when* tones should be played, it did not define *which* tones to actually play. Up until now, whenever a node was placed in AR space it was randomly assigned a tone selected from a range slightly larger than an octave (from A4 to C6). This was implemented mainly to replicate a mechanic found in the early version of Bloom: Open Space, wherein the system played an entirely random note whenever you interacted with it [81]. I was curious to how this would affect my experience of the interaction. In addition, it allowed me to test the Pure Data patch with a range of different notes, without having to immediately decide on how to map the parameter controls to augmented reality. However,

as a result, there was no actual control to be had over which notes were played, and each session produced a largely random composition. Furthermore, as the notes were randomly selected, they did not always fit together. This lead to the generated music often being inharmonious and less pleasant. I recognized this problem as similar to that faced by Carcani et al. in their work on the *Musical Moves* project [12, p. 274], and as my next move I set about implementing a solution to this.

[39] `https://www.britannica.com/art/ostinato`

4.5.3 Ensuring Harmony for Novice Musicians

Inspired by the *Musical Moves* project by Carcani et al. [12], I decided to implement a mechanism that mediates the selection of notes during play. By controlling which notes were available, this would lead to users always being presented with a harmonious outcome. I did this a conscious effort to both improve the feeling of flow during the interaction, as well as to reduce the perceived difficulty of use while still granting the user a fair amount of expressivity and control. While I saw it as likely that this would limit the potential for long-term engagement and virtuosity, it also likely lowered the bar and allowed anyone to take part in the interaction and create music, regardless of their previous level of experience. In other words, mediating the selection of notes helped shift the interaction towards a more ludic method of operation, focusing on providing an experience rather than a programmable instrument. I also saw this as corresponding somewhat with the idea of making a piece rather than an instrument [20], at least on the surface. According to McDermott et al. this form of musical interaction can be described as *transient and frivolous*, which is a form of interaction where

> "*In the ideal case, the user will get something good, but something different, no matter what he or she does*" [56, p. 41].

During the research I had experienced this in practice when playing around with the *Bloom: 10 Worlds* application as described in section 2.3.4. In this application, the available range of notes is selected from a diatonic scale[40], and the music produced as a result of input is generally pleasant [56, p. 41]. However, there is also the question of whether it becomes too effortless to achieve good results when such a mechanism is implemented, potentially leading to the users quickly feeling unsatisfied and bored with the interaction. In her paper on the context camera [53], Ljungblad describes how users testing an early version of the prototype described it as being "*...too easy to succeed with the pictures taken with the prototype*"[53, p. 366], as the prototype had provided a direct shortcut to producing novel and interesting results. However, these users were themselves active practitioners of lomography, and their criticism must therefore be seen in light of their existing expertise, and their expectations for what a lomography camera should be. While my prototype drew inspiration from concepts such as algorithmic composition and loop-based music, it has less of a direct connection to a specific existing practice, and also presents a more ludic experience in comparison with the context camera. At the same time, I saw it as likely that professional musicians might quickly lose interest in the prototype application, as it likely provides insufficient complexity to meet their needs. However, looking back once again towards the research questions explored, I consider it more important to facilitate fluency between the participants in the interaction, rather than promoting complexity and virtuosity for experienced professionals. Also, guided once more by the principles of designing computer music controllers, I saw this as somewhat representative of the principle "*Instant music, subtlety later*" [20, p. 1].
In practice, the mechanism was implemented by simply limiting the range of notes available for use at any time in the application. Instead of the assigning a MIDI note randomly when a user touched the ground, it was now selected from a pre-defined array of values where each value represents a note in a set musical scale. This mapping was inspired by the mapping in Bloom: 10 Worlds, and was set to use the E minor scale. To find the actual

[40] `https://www.pianoscales.org/diatonic.html`

notes used, I used a music theory website to look up the notes in the scale [41]. After having identified which notes should be used, it was a simple process of finding the corresponding MIDI note values to each note. The image in figure 4.10 shows a visual representation of the scale used. In addition, my initial mapping with the MIDI note values can be seen in the listing below.

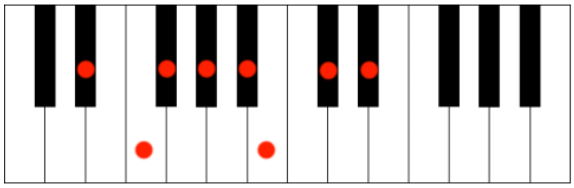


Figure 4.10: The notes in the E Flat Minor scale.

```
int[] E_MINOR_MIDI_NOTES = {39, 41, 42, 44, 46, 47, 49, 51, 53, 54, 56, 58,
59, 61, 63, 65, 66, 68, 70, 71, 73, 75, 77, 78, 80, 82, 83, 85, 87};
```

Listing 4.1: Integer array containing a range of values for MIDI notes in the E minor scale starting at D-sharp 3 and ending at D-sharp 7.

4.5.4 Mapping Synthesizer Parameters to Augmented Reality

While the note selection was now mediated as described in the previous session, each note was still randomly selected when placed. After having defined a basic framing for the musical arrangement through looping, and establishing some rules for harmony, the next step was deciding how to perform the mapping between the synthesizer and properties of the augmented reality interface. Inspired to how Carcani et al. explored mapping of dance moves to produce harmonious output via a Kinect [12, p. 274], and how Ljungblad incorporated contextual information from the environment into the photos generated by the context camera [53, p. 368], I saw the potential of doing something analogous to this by using the properties of the augmented reality space to shape the creation of music in the prototype.

In theory, any parameter within the Pure Data patch capable of receiving input could be controlled and manipulated through the libpd API, simply by sending a message containing whatever values to be used. However, controlling every minute detail of the sound might not necessarily be desirable in this case, as this could quickly lead to a high degree of complexity, making the interface more difficult to handle. Considering the research questions again, a fundamental aspect of exploring these is to enable users to cooperatively create and interact with music together. Furthermore, as described in section 4.5.3 I wanted the participants to experience a certain level of fluency when using the interaction, focusing on providing a ludic experience rather than a utilitarian instrument. Therefore, creating a highly complex mapping would likely have made it difficult to facilitate for immediate harmonious cooperation, without users first having to spend time to learn and master it before use. Instead, I wanted to create a mapping of parameters that would

[41] https://piano-music-theory.com/2016/06/01/d-sharp-minor-and-e-flat-minor-scales/

encourage almost immediate synergy between participants without requiring any training, aiming to foster the feeling described by Swift [79] as *flow*. This notion of flow is characterized as

> "*...the state in which an individual's skill level is commensurate to the difficulty of the complex task being performed.*" [79, p. 87]

This is often seen as a key component of what makes collective musical expression pleasurable and satisfying to participants. I therefore decided to be wary of increasing the complexity in the mapping of the interface, hoping to instead inspire immediate and harmonious cooperation by keeping it simple and accommodating for effortless use by novice users. During this process I explored various sources of existing work to gain inspiration for my mapping. Key influences in this process were the *context camera* project by Ljungblad [53], the Bloom application and art installation, as well as the *Musical Moves* project by Carcani et al. [12]. In the final mapping I decided to focus on the musical properties of *pitch*, *velocity* and *key* (e.g. minor or major).

Pitch

In music theory, *pitch*[42] refers to how a sound is perceived by the human ear, and is what allows us to perceive the difference between notes. Within the prototype, the pitch is mapped along the X-axis of the AR playspace when seen from the users point of view (i.e. in the forward direction when viewing the environment through the camera). This is illustrated in figure 4.11. Nodes placed closer to the camera emit a lower tone, while nodes placed further from the camera emit a higher tone. This holds true regardless of how much of the playspace has been mapped. My intention with this mapping was to enable users to access the full range of notes even on very small surfaces. When a node is placed, it gets assigned a tone based on the distance from the device, with 1 meter being the lower threshold and 5 meters being the upper threshold. The lower threshold was chosen as placing nodes closer than 1 meter from the device makes for a confusing experience. The upper threshold was chosen as placing nodes too far away leads to the velocity being too low. This note value is then kept until the node is either deleted or manipulated. By changing the position of the node, the pitch also changes, again relative to the user position, producing a different tone than before. This mapping was inspired in part by "Bloom: 10 Worlds" application, where several of the different pieces map the pitch along either the X or Y axis of the screen. Furthermore, this mapping was chosen with the intention of maximizing the range of tones available to users at any time, the thought being that by mapping pitch to distance, you would be able to play a large range of tones even if the scanned surface area was small.

Velocity

In the context of MIDI, the *velocity* of a note refers to the force with which it is played [43], and is described with values between 0 and 127. A MIDI note with a low velocity value will be rendered with a low volume of sound, and consequently a note with a high velocity will be rendered at a high volume, as if a pianist was striking the keys with force. In the prototype system, the velocity of a given note is determined based on its proximity relative

[42] `https://en.wikibooks.org/wiki/Music_Theory/Fundamentals_of_Common_Practice_Music/Notation#Pitch`

[43] `https://ask.audio/articles/midi-velocity-what-it-is-how-it-works`

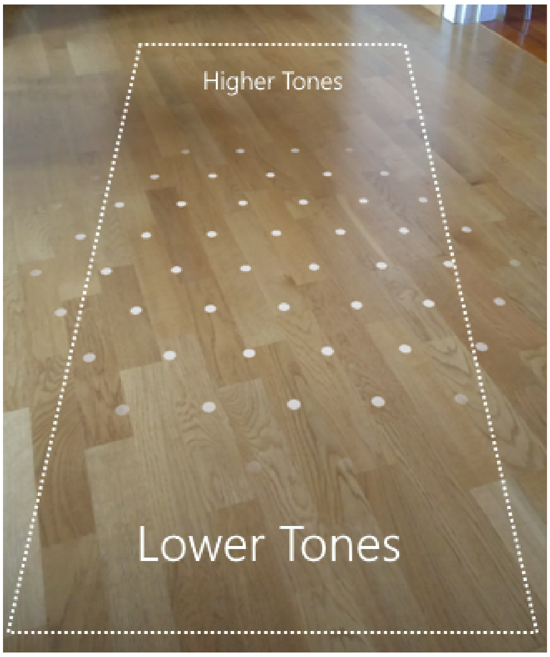


Figure 4.11: Illustrating the mapping between pitch and distance in the prototype.

Figure 4.12: Screenshot of Bloom: 10 Worlds showing the pitch mapping.

to the user. This value is calculated each time a message is sent from the application code to the Pd patch. The implementation is rather simple, and uses the ARCore API to find the distance between the user and a given node by calculating how far away it is from the device. This mapping was done with the goal of making the augmented reality space feel more dynamic and alive to users, and to encourage users to move around during play to experience the soundscape changing as they move. The image in figure 4.13 illustrates how this mechanic works in practice.

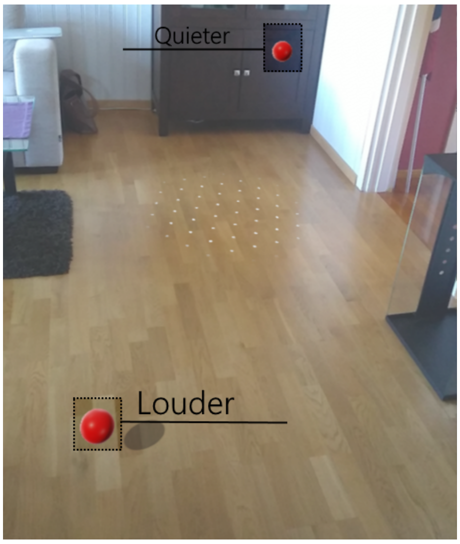


Figure 4.13: Illustrating the mapping between velocity and proximity.

Key / Scale

When speaking of *key* with regards to music, we refer to a group of notes that all fit a certain scale[44]. In other words, if one knows the key of a piece of music, finding the compatible notes is a simple task. Within the prototype, the range of notes is limited to certain keys in order to ensure harmonic output as described in section 4.5.3. Furthermore, the key (minor or major) of available notes is determined by the ambient light level of the surrounding environment. This mapping was mainly inspired by the use of contextual information in Ljungblads *context camera* [53]. Using the APIs provided by the augmented reality frameworks, I was able to access the current estimated level of light in the environment at any time during a session[45]. This estimate is returned as a value in the range between 0.0 and 1.0, which is then used to determine the scale. Within the prototype, there are four possible scales available to the user. Which one the application uses is determined by the value returned by the getLightEstimate() method. If the value is below 0.5, the available range of tones are from a minor scale. If the value is closer to one (i.e. above 0.5), the available tones are selected from a major scale. The full mapping can be seen in table 4.5.4. The light estimation is fetched in the beginning of a session, and the key is set before any nodes are put into play. In addition, the value is stored in the database and communicated to the other participant as they begin a session, to ensure that both devices have the same range of notes available during play despite possibly having slightly different light estimations.

LightEstimation Value	**Scale Used**
Below 0.25	A Minor
Above 0.25 && Below 0.50	E Minor
Above 0.50 && Below 0.75	C Major
Above 0.75	G Major

Though the above mapping is overall quite simple, informal testing performed by myself indicated that it was effective in providing variation with regard to the range of musical expression available during play, while still being simple enough to use without too much practice. In the current version of the prototype, any node put into the playspace immediately became a part of the looping arrangement. While the looping mechanism had become a core element of the prototype design, I wanted to further broaden the interaction by expanding upon this. To guide my design I looked towards music theory, seeking how to expand upon the concept of ostinato. During my research I came across a resource[46] which described how it was common for ostinato patterns to be accompanied by a more free form melody line over the top. This was also something I had been exposed to during my own history with playing and creating music. This idea therefore resonated with me, and I henceforth began exploring the idea of introducing a way of playing free notes in the playspace, without them becoming part of the looping arrangement.

[44] `https://www.dummies.com/art-center/music/piano/playing-the-piano-understanding-musical-keys/`

[45] `https://developers.google.com/ar/reference/java/arcore/reference/com/google/ar/core/LightEstimate`

[46] `https://www.8notes.com/glossary/ostinato.asp`

4.5.5 Expanding the Interaction

At this point, I was now looking to expand the interaction by making it possible to play notes outside of the looping arrangement. I therefore began to ask how I should implement this in practice within the prototype. In the context of the prototype, I was hoping that the introduction of a non-looping node type would help expand the interaction by letting users arrange looping nodes to create a repeating musical sequence, then using the other nodes to improvise on top of the looping arrangement - similar to how a musician would perform a *solo*[47]. As discussed in section 4.5.2 the use of looping in the prototype had its roots in the concept developed through the video prototype, which itself was created on the basis of ideas I gathered from *Bloom: Open Space*. To reiterate, the video prototype itself was carried out using rubber bouncing balls as a metaphor for musical notes. These were selected for use mainly to function as a real world approximation of the virtual floating orbs found within Bloom: Open Space, seeing as it was not possible to make physical objects actually float in mid air. However, as I had began the augmented reality implementation, I was no longer bound by the same physical limitations as in the video prototype with regards to how nodes should look and behave. Therefore, it was now fully possible to actually create something akin to the floating nodes fond in Bloom: Open Space[48].

By adapting the code used for the bouncing nodes, I was able to rather quickly implement a new node type in the application code. In practice, it was only a matter of replacing the animation of the node, making it slowly float through the air, instead of bouncing up and down. Thus, the *floating node* was born. To further separate the new node from the existing bouncing node, the floating nodes were assigned a significantly longer release time[49] when being sent to libpd. In practice, this meant that the sound of the node would linger in the soundscape for much longer, thereby making it easy to discern which tones are produced by which kind of notes. The looping node produced a tone with a release time of 2000 ms, while the new floating node produced a tone with a release time of 10000 ms. This would also add a slight sonic diversity to the soundscape, while still maintaining consistency with the sound of the looping nodes.

In order to switch between the bouncing node, and the now implemented floating node, I added a toggle-button to the user interface which would let users easily switch between node types while playing. During the initial testing, I found the new floating node to be a beneficial addition, providing more variety to the interaction. However, when reflecting on the behavior of the bouncing nodes during these testing sessions, they felt a little difficult to control. The repeating animation provided a framing for creation of rhythmic structures, but felt at times too random and undirected, and it was difficult to feel totally in control of the interaction. A similar kind of challenge was faced by Ljungblad and her team during testing of the context camera [53], wherein users felt that they had insufficient control over the picture taking process, leading to a high degree of similarity in the output [53, p. 367].

To attempt mitigation of this sense of lacking control in my prototype, I altered the behaviour of the node animation. Previous to this, the bouncing nodes would begin animating on creation, and would not cease their movement until they were completely

[47]`https://doi.org/10.1093/gmo/9781561592630.article.26159`
[48]`https://www.youtube.com/watch?v=I5xf9J9oyRE`
[49]`https://www.wikiaudio.org/adsr-envelope/#Release`

removed from the scene. However, to increase the level of control I implemented a functionality to stop the animation whenever a user touches the node. The node would then stay still as long as a user has selected or is transforming a node (e.g. manipulating its position in space). In other words, when a user is interacting with the node, the animation is stopped and the node is reset to the initial position on the ground. Then, when the user finishes moving the node and releases it back into the playspace, the animation starts again, and continues in the now re-calibrated loop.

In essence, this means that two or more nodes could now be synchronized by selecting them and releasing them at the right time. This not only introduced a greater sense of control to the interaction, but also enabled the creation of musical *chords* within the app. While I had considered looking at the use of chords earlier in the design process, seeing as it is a prominent focus in Carcani et al. [12, p. 275]), I had initially dismissed the idea of chords due to the limitations of the loop-based arrangement. My reasoning was that using single nodes provided for a greater amount of freedom, and a larger amount possible combinations during the two second interval. In this scenario I had envisioned implementing a dedicated chord-making functionality, which would allow the user to select a chord before putting it into the playspace. However, this way the chords could appear more organically, while still providing the user with the freedom to combine any notes together. While testing this I noticed that creating chords by simply stopping the animation required a bit of finesse to achieve in practice, but I decided to not further simplify this process in order to possibly foster a slight feeling of long-term engagement, and to possibility introduce a greater sense of achievement in participants who were able to master this act.

However, as I envisioned this new functionality in play during a cooperative session, I became aware of another potential point of contention. I had not yet assessed whether both users should be assigned the same sound in the playspace, or whether I should introduce a new sound. As described in section 4.4, I had only developed one synthesizer patch for the sound. This patch was was both anchored in a specific source of inspiration, and programmed according to my vision of what the experience should sound like in use. However, I had neglected to consider the effect of extending the available range of sounds. Introducing another sound would potentially be beneficial in that it would expand the variety in the interaction, while also giving each user a stronger identity in the playspace.

By having a separate sound, a user would be able to make out their individual contributions in the musical space. On the other hand, I saw it as interesting to keep the prototype as it currently was, assigning both users the same sound. From what I had been able to discern, this was how the Bloom: Open Space installation operated. Each participant interacted with the shared AR space using the same sound, blending together to create a shared musical expression [81]. I was intrigued by this, as I felt this more strongly encouraged a fully collective expression, with the lack of sonic separation diluting the boundaries between identities in the playspace, and directing the focus towards creating a shared musical output instead. I therefore decided to not extend the sound palette, keeping the same sound for both users. This also further echoes the ideas described in section 4.5.3, where I aim to facilitate a high level of fluency in the interaction.

Up until this point, the mobile user interface (i.e. the visual elements overlaying the augmented reality view) had not been my primary area of focus. It had currently only been outfitted with the toggle-button described in the previous paragraph. However, as I began to feel satisfied with how the augmented reality components functioned, I turned my attention towards the user interface and began looking at how it should be designed to further extend the functionality provided by the prototype.

4.5.6 Designing the User Interface

When discussing the *user interface* in this section, I am referring to the visual components (e.g. buttons, icons and dialog-windows) overlaying the augmented reality view within the application. While the prototype interaction itself takes place within augmented reality space, the user interface framing it is arguably vital in providing access to the critical functionality of the application. Furthermore, it was important to me that I provide a user interface in the app which actively supported the immersiveness of the experience, i.e. providing the required controls without disrupting the users flow within the augmented reality experience. In designing the user interface, I mainly drew inspiration from two sources: (1) the guidelines for augmented reality design proposed by Google [28], as well as (2) the application Just a Line[50]. The main user interface of Just a Line can be seen in figure 4.14. As seen in this screenshot, any buttons or widgets rendered to the screen are given transparent backgrounds, thereby providing the AR view with more screen real-estate. The user interface of Just a Line provided vital inspiration for the prototype design, and by reproducing several of the key design choices I hoped to both support a high degree of immersiveness, while also reinforcing these design patterns and thereby creating a sense of uniformity for users familiar with existing augmented reality applications such as these. This latter decision was anchored in the design guidelines, which argue the benefit of using familiar UI patterns[51].

Figure 4.14: Screenshot from the Android application *Just a Line* showing the main user interface

[50] https://play.google.com/store/apps/details?id=com.arexperiments.justaline&hl=en_US

[51] https://designguidelines.withgoogle.com/ar-design/interaction/ui.html#ui-onboarding-instructions

Extending the Functionality Through Building the User Interface

As previously mentioned, the user interface had at this point only been outfitted with one simple button which allowed users to toggle between bouncing and floating nodes as described in section 4.5.5. Apart from this, the interface was a blank canvas. As i began the process of designing the main user interface, I considered which actions I would likely require to support, as this would dictate the makeup of the interface. After looking at the currently implemented functionality, and considering how the user might use the application during the field deployment, I decided on implementing the following buttons in the interface:

1. Audio on/off
2. Microphone on/off
3. Clear all nodes
4. Connect with other user.
5. App information

The icons used for each of these buttons can be seen in figure 4.15.

1. Audio on/off button

The button for volume control was implemented as a simple on/off switch. Adjusting the application volume is likely better handled by simply adjusting the overall audio volume on the phone, which can be done by using the buttons on the hardware device. However, I saw it as useful to include a way of quickly muting the audio in case the user is disrupted while using the prototype. When the user presses the button, the icon changes to add or remove the line through the icon, indicating that the state has changed.

2. Microphone on/off

The button for microphone control works much in the same way. It is a simple switch that enables the user to enable the audio input through the microphone in order to "let in" sounds from the surrounding environment, having it mix together with the application sounds. My idea was that this could be beneficial when using the app in a busy place, where it might be necessary to react to auditory cues. Furthermore, as mentioned in sec 4.4, the inclusion of environmental sound in the experience aims to strengthen immersion, and make the musical arrangement feel like a part of the environment, while also enabling users to experience elements of the audio augmented reality presented in the *Listening Mirrors* project [49]. When the user presses the button, the icon changes to add or remove the line through the icon, indicating that the state has changed.

3. Clear all nodes

In their augmented reality design guidelines, Google point to the need for providing users with a simple way of resetting the experience[52]. I therefore implemented a button which was intended to let users remove all their own nodes from the playspace.

[52] `https://designguidelines.withgoogle.com/ar-design/interaction/ux.html#ux-reset`

4. Connect with other user.

While the functionality had not yet implemented, I considered that the interface would likely require a way of lettings users connect to the shared user experience. I decided to model this after the Just a Line application, and added an icon in the lower left corner which would later be used to connect with another player. The similarity to Just a Line can be clearly seen by comparing the fourth icon in figure 4.15 with the icon in the lower left of the interface seen in figure 4.14.

5. App information button

While not essential to the functionality of the application, I decided to add a button that would display some general information about the application to the user. This was used mainly to disclose the use of ARCore in the application as required by the framework [53], and to give credit to the additional libraries used within the code.

Figure 4.15: The various Icons used in the user interface. The icons represent (1) audio on/off; (2) microphone on/off; (3) clear all own nodes from screen; (4) connect with another user; (5) display app info; (6) looping on/off;

4.5.7 Manipulating Nodes in the Playspace

Having established a fundamental user interface, I returned to designing the augmented reality user experience. An essential part of establishing a seamless AR user experience was making sure that the placement and manipulation of nodes can be performed with ease. The Sceneform library by default provides a controller interface with capabilities related to the selection and transforming objects within the AR space. I had made use of this in the prototype too, as this required little extra implementation, and supplied a intuitive and functional way of moving objects in 3D space. However, in addition to transforming and manipulating the content in AR space, I also required a way to remove individual nodes from the playspace. This was not provided by the framework, and I therefore began designing my own solution to this. While the possibilities were seemingly endless, I wanted a way of deleting nodes that felt somewhat organic in the context of augmented reality. In this I saw it as important that the user would not have to go "outside" of the AR experience, which would likely lead to a reduced or broken immersion. Therefore, using ex. buttons or other static UI-elements to select and delete nodes felt like a step in the wrong direction. While brainstorming for a solution, I considered how other applications approached the concept of deleting or removing content from the screen. My thoughts were immediately drawn to a specific mechanic in the Facebook Messenger application used to move or delete active conversations.

In the Messenger app, the current active chat conversation is displayed as a circular icon on the screen, referred to as a "chat head". This icon overlays any other active application on the screen, and is therefore often moved around to accommodate for other

[53] `https://developers.google.com/ar/distribute/privacy-requirements`

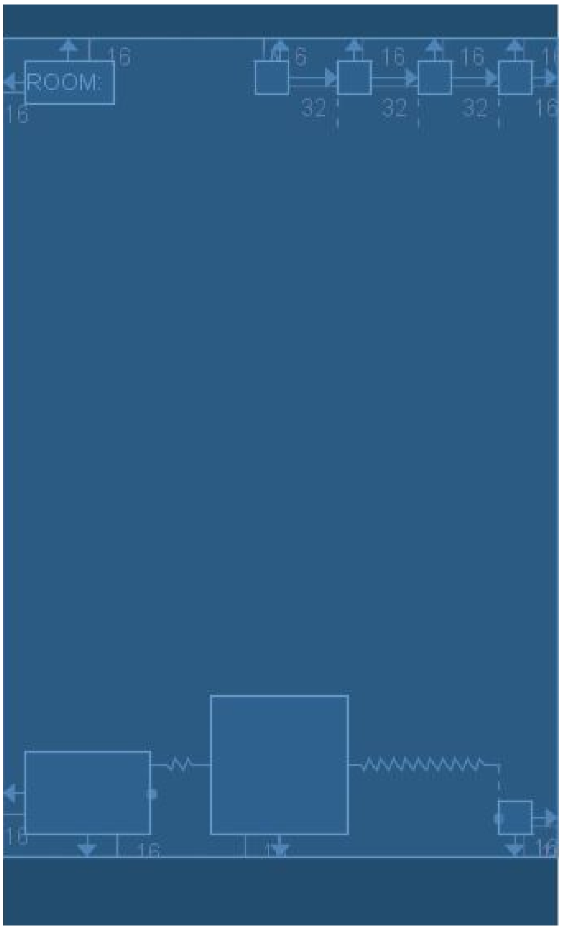


Figure 4.16: Main user interface as seen with blueprint view in Android Studio.

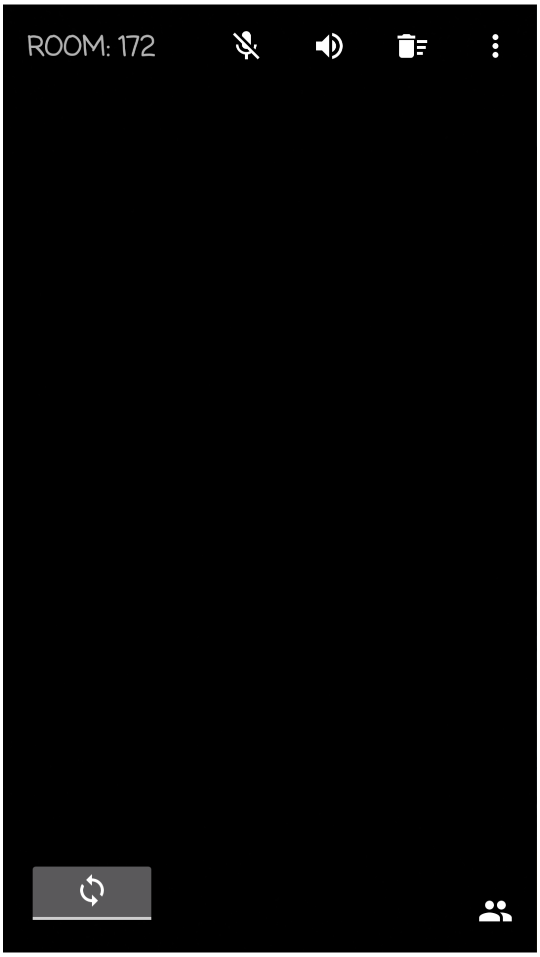


Figure 4.17: User interface as seen on a physical device.

content. To remove a chat head from the screen, users simply touch it and drag it towards the bottom of the screen. A circular overlay with an X will then appear, and by letting go of the icon when holding it within this circle the icon is removed from the screen. To fully illustrate this, I have provided screenshots showing a step-by-step example of this in figure 4.18.

Figure 4.18: Example of how a Facebook Messenger conversation overlay is removed from the screen on an Android device. The circular image represents the *chat head* icon which is removed from the screen by pulling it onto the circular region with a cross at the bottom of the screen.

Drawing on inspiration from this functionality, I had the idea of how a similar solution could be implemented within the prototype. In the prototype applications, users could now delete bouncing nodes from augmented space in a similar way by selecting them and pulling them down towards the lowermost part of the screen. In practice, whenever a users moving finger is detected on the touchscreen, the application will display a circular icon at the bottom of the screen, containing an icon of a garbage can. To delete a node from the playspace, users can simply touch a node and drag it onto this circular icon. The icon will then turn red, indicating that it has been activated. Releasing the node on this icon will then remove it from the local playspace. At the same time, the node will then be marked for deletion, and removed from the shared playspace. As the floating nodes dissipate naturally when their animation ends requiring no manual removal, I only implemented this functionality for bouncing nodes. The finished implemented mechanic can be seen in the three images in figure 4.19. During my initial testing of this new mechanic, I found this way of removing nodes to feel organic and natural, both since the movement was already familiar, and also as I did not need to break immersion from the playspace when performing the task.

Having now implemented a way of removing nodes, the prototype interaction was beginning to feel like a more coherent experience. However, despite having added a significant amount of functionality, I had not yet implemented support for binaural audio within the application. Currently, the audio was simply presented in ordinary stereo, not representing any of the spatial qualities. As this was critical to enable exploration of the second research question, I began considering how this implementation should be done.

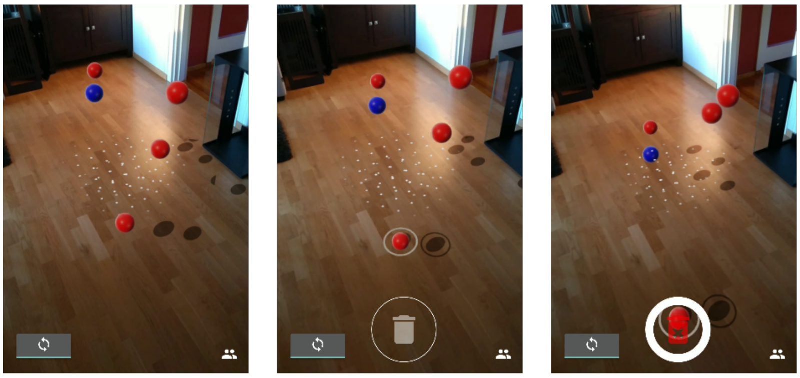


Figure 4.19: Removing a looping node from the shared playspace by touching it and dragging it onto the circular overlay at the bottom of the screen.

4.5.8 Implementing Binaural Audio Capabilities with Pure Data

Having implemented the basic functionality needed to create and interact with music within the playspace, I began to work towards realizing the binaural audio functionality. The inclusion of binaural audio was a deliberate design decision rooted in the theory discussed in section 2.3.3, which indicates that use of spatial audio can have a significant effect on both the level of immersion experienced by users within the augmented reality space, as well as improving the feeling of cooperation between them. I had become aware of this during the initial research before beginning the design, and it was finally time to actually implement this in practice. My initial idea was to use the binaural filtering to create the illusion that the sound created by each node was being emitted from its position in space. In other words, I wanted it to appear like sound was actually being produced by the nodes. To do this in practice, I decided to employ an external module for the Pure Data system and set about searching for alternatives.

I began my research by simply searching with Google, combining the keyword "Pure Data" with words like "binaural", "3D Audio" and "ambisonic". Through this process, I discovered the existence of several external modules for Pure Data with the necessary capabilities to produce binaural audio. Among these, I selected three different externals as candidates for use in the prototype, namely *HoaLibrary*[54], *earplug*[55] and *+binaural*[56]. This selection was based largely on discussions on the official Pure Data messageboard[57][58][59]. Each external was subsequently downloaded and evaluated for use within the prototype.

However, before an outside Pure Data external (in this case the binaural module) can be loaded from libpd within the host application, the pd-for-android library which contains libpd must be rebuilt from source to include the relevant external. Otherwise it will not be available for use in the embedded Pure Data environment, and the patch will simply attempt to load something that is not there. However, this process of adding

[54] `http://hoalibrary.mshparisnord.fr/en`

[55] `https://puredata.info/downloads/earplug`

[56] `http://www.soundhack.com/freeware/`

[57] `https://forum.pdpatchrepo.info/topic/7804/hoalibrary-for-pd-high-order-ambisonics-library/`

[58] `https://forum.pdpatchrepo.info/topic/9356/cw_binaural-external-crashing-pd-with-a-runtime-error/`

[59] `https://forum.pdpatchrepo.info/topic/2184/binaural-panning`

externals and rebuilding the library is possible for anyone to do without much trouble [60]. Anecdotal evidence shows that some externals are more difficult to incorporate than others[61], underlining the fact that this is a somewhat uncertain process depending on the external desired for use. This was the case for my project, and I experienced several complications with this leading to a lot of time spent debugging and configuring custom builds of the pd-for-android library with various binaural externals. This process of trial and error itself became an act of prototyping, and had to be done several times in the process of testing and evaluating externals for use.

Building pd-for-android with Externals

I initially decided to use the Pure Data external "earplug" due to its widespread availability on the internet, its ease of use, and its inclusion in the alternative Pure Data distribution pd-extended. This made it very easy to download and use the external in the desktop environment of Pd, and incorporating binaural effects within the prototype patch file became a relatively painless affair. In addition, the external provided control over both azimuth and elevation parameters, making it possible to position the sound both in a direction and at a certain height. While this initially seemed promising, when I attempted to include the external in a customized build of the libpd library I started encountering a lot of significant difficulties.

Firstly, when the earplug module was loaded within the Android application the patch suddenly began displaying extremely volatile behavior resulting in frequent and regular crashes. Furthermore, additional testing revealed earplug to be very resource demanding when used in the prototype, resulting in the outputting of glitched and noisy audio when manipulated. These failures can potentially be attributed to the external itself being quite old and outdated, having not been updated since 2009 [62]. This kind of behavior would be extremely disruptive in a future testing scenario, and I therefore decided abandon the earplug external, and explore other alternatives.

Moving on, I looked at the two other alternatives I had initially selected, namely *HOALibrary* and *+binaural*. A quick look at the former revealed it to be a very sophisticated library for binaural sound, with capabilities far outside what is required for this prototype. While the wide range of functionality could be beneficial in producing a convincing effect, it has a significant learning curve, and would likely require significant changes to the pure data patch to incorporate. However, the latter external appeared much simpler to use, and would require little modification to the existing Pd patch as it is functionally quite similar to earplug. In addition, anecdotal evidence pointed to it being less resource intensive[63], thereby potentially eliminating the previously described issues. I therefore decided to use the +binaural external in my patch, and by modifying the existing makefile I was quickly able to produce a new build of pd-for-android with the external implemented. As an example, the external and makefile used to create this build have been attached in appendix G. Compared to earplug, +binaural is simpler and lacks the possibility of manipulating the elevation of sound. However, as previously discussed

[60] `https://github.com/libpd/pd-for-android/wiki/Building-and-packaging-externals-for-Android`

[61] `https://github.com/libpd/pd-for-android/issues/67`

[62] `https://github.com/pd-l2ork/pd/tree/master/externals/earplug~`

[63] `https://forum.pdpatchrepo.info/topic/9356/cw_binaural-external-crashing-pd-with-a-runtime-error/3`

in section 2.3.3, humans are much better at locating sound according to azimuth than elevation or distance. I therefore considered this a minor sacrifice, and instead decided to move on to verifying whether or not it worked as intended within the prototype environment. Initial testing showed much more promising results this time around, and the external appeared to function well running as a part of the patch within the app.

Communicating Device Position to libpd

With the binaural external implemented in the libpd library, the next step was to incorporate binaural audio into the prototype experience itself. The binaural object responsible for filtering the sound requires the azimuth value as a floating point number in the range between -180 and 180. To further illustrate how it works, within this range an object directly in front of the device would have the value 0, while objects directly to the left and right would have respective values of -90 and 90 degrees. In the prototype code, I calculated the position of a given node relative to the device whenever the animation repeated, and converted this value to correspond with the azimuth range. Then, I passed the calculated azimuth value to the patch through libpd, just before sending the note. This enables the patch to quickly filter the sound according to the given value before receiving the note value, making it appear as if the sound is actually being emitted from the location of the node in augmented reality space.

The implementation has two significant limitations. The first (1) being that it does not simulate the difference in volume between left and right ear according to user orientation whenever a note is played. The velocity of each tone is based only on the calculated distance of the device itself, and is set after the azimuth value of a tone has been set. The second (2) limitation being that the binaural object does not support filtering according to elevation, and all nodes will therefore be perceived to have the same vertical position. It can therefore be argued that the prototype does not provide a fully spatial audio experience, and that it simply incorporates binaural sound with distance-based velocity. This was not considered a problem, and would still allow me to explore the second research question without any hindrance.

4.5.9 Achieving Shared Augmented Reality Using Cloud Anchors

Having implemented binaural audio capabilities in the prototype, there was only one missing piece needed to fully realize the design concept defined in section 4.1.2 - the possibility of cooperative interaction between participants in augmented reality space.

Creating a shared augmented reality space has traditionally been a complex process to achieve, often requiring bespoke software and hardware solutions, such as those used in projects like the *Studierstube* [80] and the *Augmented Groove* system [62]. However, in this project I was able to achieve the shared augmented reality space by making use of the Cloud Anchors API as discussed in 4.3.3. In this section I will describe how I have implemented this in the integration prototype, with a focus on maintaining consistency in the shared state across devices.

Before incorporating the Cloud Anchor functionality into the main prototype project file, I completed a brief training exercise [64] provided by Google to instruct developers in how to use cloud anchors. This lab exercise teaches how to host and resolve cloud anchors

[64] `https://codelabs.developers.google.com/codelabs/arcore-cloud-anchors/`

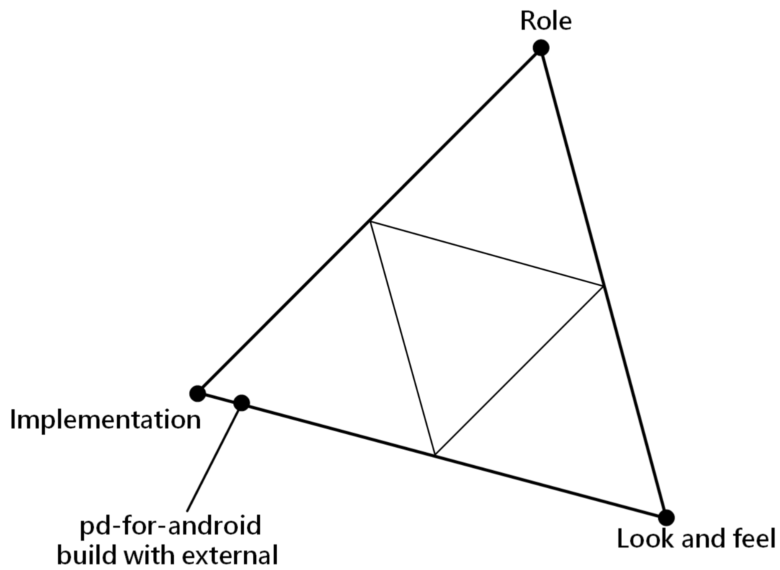


Figure 4.20: New pd-for-android build as seen through the Houde & Hill model

using the ARCore framework, and sheds light on some useful techniques for working with cloud anchors in practice. This was useful in providing a hands-on and guided technical instruction of how to use the API, and some of the concepts presented within this tutorial found their way into the prototype code. In addition to studying this tutorial, as a more advanced practical example I studied how shared augmented reality was achieved in the previously mentioned application Just a Line[65]. These two examples laid the foundation for my own implementation of the shared augmented reality functionality in the prototype.

Synchronizing Devices Understanding of a Room

Within the prototype application, a user initiates a new collaborative session by tapping the *New Room* button within the app. This initializes a new room in the database, and begins a process of creating and hosting a new cloud anchor though the Cloud Anchors API.

In practice, the application will attempt to silently create an invisible cloud anchor on the surface directly in front of the users device, without the user needing to provide any input. As cloud anchors largely function the same way as ordinary anchors in ARCore (described in sec. 4.3.3), this relies on the user having scanned and mapped the relevant surface, otherwise an error message is shown, and the user is prompted to retry. On the successful hosting of an anchor, the user is notified and presented with the newly generated *room code* to share with a co-participant. This room code is created by simply incrementing a value in the database each time a new room is created. On the successful creation of a new room, the cloud anchor id is then associated with the newly created room in the database. An example of the structure of these child nodes can be seen in figure 4.21. This generated room code can then be used by another participant to connect to the same room, fetch and resolve the cloud anchor on their device, and begin participating in

[65] `https://github.com/googlecreativelab/justaline-android`

the shared augmented reality playspace together with the other participant. In practice, each room is simply a new child node in the database, containing all the relevant data necessary to enable shared AR between two devices.

In order to resolve a cloud anchor and join an existing room, the device of the joining user needs to have a similar environmental understanding as the device that created the room anchor. This means that the joining user needs to have scanned and mapped the same plane as the host within the application, ideally from the same angle and position, so that both devices have identified the same clusters of feature points in the surrounding environment. When this has been achieved, the ARCore API will then be able to successfully resolve the associated cloud anchor, which then functions as the shared reference between devices.

The overall model and metaphor of the *room* in this context, as well as parts of the structure used in the database, were modeled and created based on the implementation seen in the application Just a Line[66]. Furthermore, elements of the implementation borrows from the aforementioned Cloud Anchors exercise[67] I completed before beginning development of this functionality, especially the code responsible for generating room codes.

Figure 4.21: An example of the structure of a room node within the database. This example shows the room with room code 404.

I decided to employ this process of invisible anchor placement to simplify the act of creating a new cloud anchor. Instead of users beginning the session by tapping to create an invisible anchor, everything happens seemingly by itself without the user knowing, which hopefully provides a better user experience. This means that users only tap to interact with the instrument interface, and not to perform other tasks such as anchor creation. Furthermore, as the cloud anchor itself only functions as a shared reference point for users, it does not need to be seen at any point during the interaction, and is therefore kept invisible throughout. However, during development i saw it as useful to sometimes be able to see the cloud anchor and its position in the room. I therefore included the possibility

[66] `https://git.io/fjnj5`
[67] `https://codelabs.developers.google.com/codelabs/arcore-cloud-anchors/#3`

to enable debug mode in the code. An example of a visible cloud anchor can be seen in fig. 4.5.9.

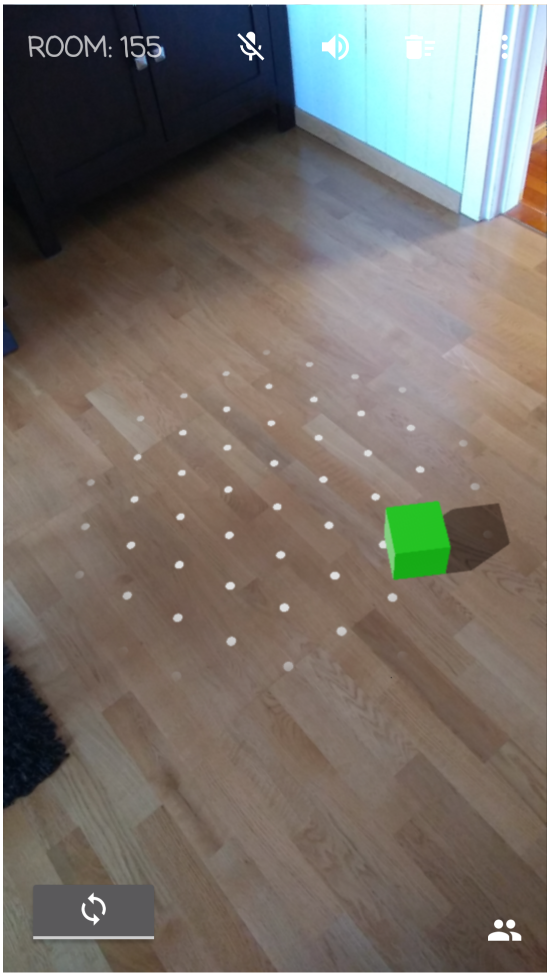


Figure 4.22: Screenshot of application running in debug mode. The green cube represents the cloud anchor, which is normally invisible.

Figure 4.23: Screenshot of prototype during a shared augmented reality session. The green cube being visible indicates that the prototype is in debug mode.

Maintaining a Shared State Across Devices

As both participants are interacting within the same playspace, it is critical to ensure consistency in the state of the room at all times. The room needs to react to the events of node *creation*, *manipulation* and *deletion*. When a node is created within the playspace on any given device, it is immediately uploaded to the cloud database and stored within the *nodes* tree in the database (as seen in figure 4.21). The uploaded data consists of all the information required to render an identical node with the same properties on the other participants device. However, the most important piece of information is perhaps the value which describes the position of the node, as it exists *relative to the cloud anchor*. These values can be seen in figure 4.21, attached to the "xPosition" and "zPosition" keys within each node. As a node is placed, its position in space relative to the cloud anchor is recorded and stored within the node object. Then, since both devices have the same cloud anchor as a shared reference point in the world, these coordinates enable any other device to render a node in the same place by using these coordinates. In short, we only really need one cloud anchor in order to enable the shared augmented reality, as all other entities we want to introduce can be placed in AR space relative to the position of the cloud anchor. However, in order to maintain the correct state across devices, the database needs to also react to the events of node movement and node deletion.

Whenever a node is moved by a user during a shared session, its values will be updated and changed, and the node will be given a new tone value, as well as a new position in space. Whenever such a change takes place within the playspace, the relevant node and its attached values are immediately updated within the database. The other device then notices that something has changed within the database, and downloads the new data to the device in order to reflect the change in state locally. This process repeats as users interact with and move the nodes around within the playspace. This way, the application maintains consistency with regards to node positioning across devices in a relatively simple way. However, to ensure full consistency the system also needs to react whenever a node is removed from the playspace.

When the user removes a node, either by clearing all nodes using the dedicated button, or by removing it individually, its data is removed from the active room and moved to a separate log-tree in the database. This way, the node will no longer be updated and rendered on any active devices, but its data will still be accessible for historical reference and post-deployment analysis. This is done using Cloud Functions for Firebase[68], a service within the Firebase system that allows for the execution of code on the backend in response to events happening in the Firebase database. In practice this is done by listening for changes within the database room. Each time the value of the field "isLooping" on a node (see figure 4.21) changes from true to false, meaning that a looping node has been deleted, a function that copies the affected node to the log database is executed, and the node is removed from the active room. This change is then registered by the client application, and the affected node is removed from the playspace on both devices. For the sake of transparency I have included the the JavaScript code written for the Cloud Functions service in appendix G.

After having implemented the shared augmented reality functionality, the prototype was functionally close to being ready for deployment in a field setting. The RtD model by Zimmerman & Forlizzi does not explicitly state when one should move to the evaluation phase, but asserts that this process can begin whenever the practitioner(s) have an artifact they like [87, p. 185]. At this stage, I considered it natural to assess both whether the prototype would be capable of exploring the research question, as well as whether it would be robust enough to handle deployment. This was done by performing a series of actions to finalize the prototype development.

4.5.10 Finalizing the Prototype

Finally, before I entered the evaluation and deployment phase, I performed some informal testing of the application. This was done by running through the general actions that would likely take place during a field deployment session, as well as by carrying out the pilot test described in section 3.3.8. My goal with this testing was to catch any breaking errors and bugs that might negatively impact or disrupt the deployment. In addition, I wanted

[68] Cloud Functions: https://firebase.google.com/docs/functions/

to verify that the prototype contained the sufficient functionality needed to explore the research questions. As pointed out in 4.1.1, the prototype is primarily designed to enable collection of data to explore the research question. It was not designed to be released to the public and is therefore not fully refined in all areas, but ultimately serves to enable the functionality essential to the prototype experience. Furthermore, the prototype has only been tested properly with the configuration mentioned in section 4.3.9. Therefore, I did not spend time adapting the application to run on different hardware configurations or alternate versions of the Android operating system. The initial testing of the prototype at this stage revealed the existence of a few bugs that were subsequently fixed before moving on to the deployment. In addition, after the deployment and evaluation phase had concluded, some additional work was performed on the project source code. This work consisted of cleanup, some refactoring, and the addition of some comments where deemed necessary. This work was performed only to increase readability, and to make the code more understandable, not to add additional functionality not present during evaluation.

Finally, as the prototype requires two or more people to deliver its intended functionality, I have created a video demonstration of the prototype in use to reduce the need for actual hands on testing during evaluation. The video has been uploaded to YouTube as a private unlisted video, and can be accessed through the link below. In addition, the video demonstration can also be found on the USB stick in appendix G.

`https://www.youtube.com/watch?v=M2rFNAagixo`

The video shows the application in use from the perspective of two users simultaneously, during two separate sessions. Tables 4.1 and 4.2 show timed lists of the actions performed in each of the sessions within the video.

00:12	User 1 and User 2 both open the application.
00:15	Both users scan the environment to get the required level of tracking.
00:20	User 2 creates a new room, and waits for the cloud anchor to finish hosting.
00:30	Room creation is finished. User 2 receives a room code, which is given to User 1.
00:36	User 2 inputs the room code and shortly after the application begins resolving the cloud anchor.
00:45	User 2 successfully resolved the anchor, and both participants begin playing.

Table 4.1: Timed list of actions performed by users in session 1 of the video demonstration.

03:04	Both users open the application.
03:07	Both users scan the environment to get the required level of tracking.
03:11	User 2 creates a new room, and waits for the cloud anchor to finish hosting.
03:24	Room creation is finished. User 2 receives a room code, which is given to user 1.
03:26	User 2 inputs the room code and shortly after the application begins resolving the cloud anchor.
03:32	User 2 successfully resolved the anchor, and both participants begin playing.

Table 4.2: Timed list of actions performed by users in session 2 of the video demonstration.

4.6 Summary

In this chapter I have described the whole prototyping process from beginning to end. The process began with an initial concept definition, wherein a broad design concept was established on the basis of my research questions as well as existing work in the field. Following this, I moved on to creating a low fidelity prototype using video as a medium to further explore the design concept. This helped reinforce my confidence in the concept, and enabled me to move onto the actual prototype implementation. Before I began development on the integration prototype, I went through a comprehensive process of selecting technologies for use in the prototype. After having established the technological foundation capable of delivering the functionality required to explore my research question, I began the iterative process of designing and developing the integration prototype. This process resulted in the creation of an application for the Android operating system, capable of collaborative creation of music within augmented reality space. The figure in fig 4.25 provides a simplified visualization of the final architecture and composition of this application. As encouraged by the methodological approach, all the work performed during each stage has been documented thoroughly along the way, in order to provide rich insight into every aspect of my process. Finally, the model in figure 4.24 shows the entire process seen through the prototyping model.

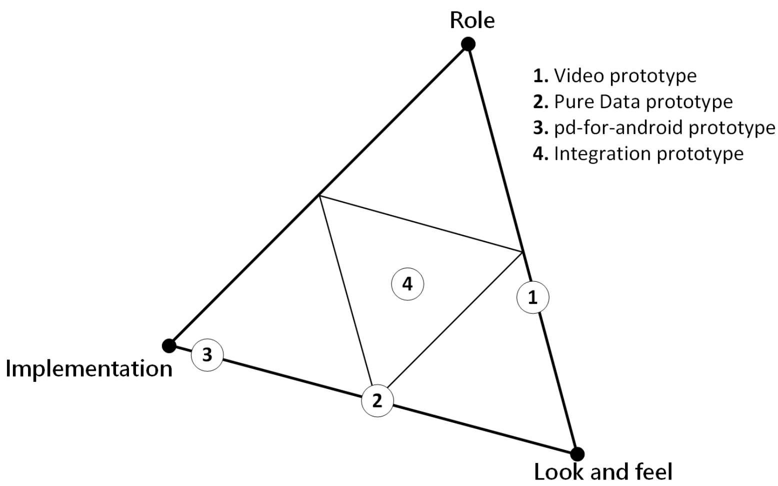


Figure 4.24: Entire design process as seen through the Houde & Hill model

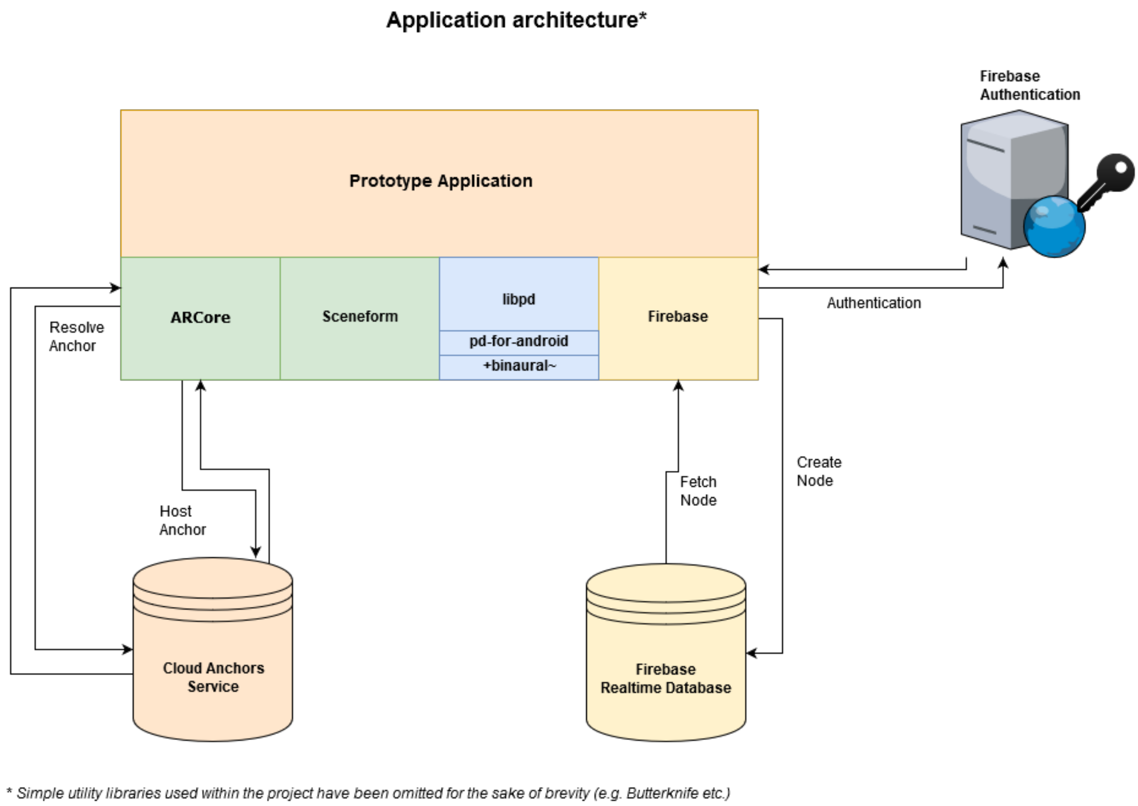


Figure 4.25: Diagram showing a simplified rendition of the prototype architecture.

Chapter 5

Results

> *"If something is boring after two minutes, try it for four. If still boring, then eight. Then sixteen. Then thirty-two. Eventually one discovers that it is not boring at all."*
>
> \- John Cage

This chapter presents the findings discovered as a result of the field deployment outlined in section 3.3. The overview in table 5.1 below shows which participants took part in which sessions, along with the age and gender of the participant. In order to protect the anonymity of participants, they have each been assigned a random name. This name has been created using an online name generator, and thus has absolutely no connection to the participants actual name [1]. However, by referring to participants with these generated names it is my intention to increase the readability and provide a more fluent and vivid account of both the observations and interviews.

Name	**Age**	**Gender**	**Session**	**Stated Musical Experience**
Elias	28	Male	1	Yes
Sverre	36	Male	1	Yes
Jonathan	32	Male	2	No
Nikolai	41	Male	2	Yes
Mikkel	35	Male	3	Yes
Edvard	34	Male	3	Yes

Table 5.1: Overview of the participants in the evaluation phase.

When speaking of participants throughout the chapter, I generally refer to them by name whenever I am speaking of one or two participants. When referring to three or more at the same time, I will generally employ a plural designation such as *the participants* to avoid confusion and fatigue in the reader from overuse of names.

[1] `https://www.fakenamegenerator.com/gen-male-no-no.php`

5.1 Open Coding

This section briefly describes how I performed open coding of the collected data. The coding process was performed according to the approach presented in section 3.4.2.

After having transcribed the interviews to text, I began the process of coding the material. This was done by going through each transcribed interview line by line, and assigning codes to individual paragraphs. As described in section 3.4.2, the concepts were generated through a combination of deductive and inductive coding. As an example, the concept **Spatial Sound** was developed inductively, while the concept of **Control** was developed deductively. Following the coding process, similar codes were sorted and put into categories according to the themes they explored. These categories were created both on the basis of concepts created during coding, as well as from the theory covered in chapter 2. To illustrate the form of the transcribed data, a fully anonymized sample of text from the transcription data can be found in appendix E. The sample is provided in both Norwegian and English, and is intended to show the form and presentation in the transcription. The below excerpt show an example of how the coding was performed in practice:

> **Elias:** "*I think we quickly felt [**Feeling**] that this was something percussive [**Percussive**].*"
> **Sverre:** "*Mm.*"
> **Elias:** "*It was something rhythmic [**Rhythm**] in nature.*"
> **Sverre:** "*Yes.*"

Originally, all interviews were conducted and transcribed in Norwegian. This was also the case for observation notes. Therefore, most of the codes developed in-vivo have been through a process of translation. This translation was performed continuously during the initial coding process, making it so that the entire process of coding was performed entirely in English despite the material being Norwegian.

5.2 Findings

In this section I present the findings established as a result of the analysis. The analysis led to the creation of three categories, namely *immersion*, *augmented cooperation* and *spatial instrumentality*. Each category represents its own set of concerns, and I present the related concepts throughout each of the following subsections.

5.2.1 Immersion

During the coding process, the category of *immersion* emerged from the collected data early on. This category was established on the basis of the concepts found in table 5.2, and grew out of material from both the interviews, as well as the observations.

Absorbed	Spatial Sound	Movement
Ludic Experience	Consonance	Colors

Table 5.2: Codes for the category of immersion.

In speaking of the prototype experience during the first interview, Elias reported that the sensation of interacting with the space while wearing headphones made him feel deeply immersed by the visual aspects, reportedly feeling as if the nodes were actually present in the room with him, bouncing up and down around him as he moved around. At a later point in the interview he returned to this idea, expressing "*...I had an experience of... disappearing into the screen. I was a little lost in the room then*". These immersive qualities were echoed strongly by Edvard during the third interview, who spoke of feeling as if "*...a new world opened up, behind the one we go around experiencing everyday"*. These experiences led to the creation of the concept *absorbed*.

> **Sverre:** "*It is a little mind-blowing that you can actually do this.*"

While discussing the immersive qualities of the prototype during the first interview, Sverre attributed much of his own feelings of absorption to the inclusion of binaural audio in the prototype interaction. Jonathan also highlighted this as an important quality, noting how this was one of the things he perceived as setting the prototype apart from other augmented reality applications. This led to the creation of the concept *spatial sound*. While the prototype provides a rather simple implementation of spatial audio (as described in section 4.5.8), the inclusion of it in the prototype was almost universally described as something positive by participants.

> **Sverre:** "*That was a really cool experience.*"
> **Sverre:** "*You would kinda be like - 'Where is that sound coming from?' - and then just go 'Oh, there it is. Alright!*"

Before beginning the evaluation process, I did not inform participants of the binaural audio properties. I was therefore interested in whether it actually provided the effect successfully - and to a large enough degree that they would notice this organically by themselves. The collected data indicates that it did, and during the field tests, participants would generally discover the binaural properties of the audio whenever they began moving around and exploring the room more actively. Despite the general consensus, during the first interview Elias reported not really noticing the binaural effect, which could indicate that the binaural external might have crashed at some point during the test. However, the implementation appeared to be otherwise successful in producing the intended effect, having been brought up and discussed as a positive feature by all other participants in the three interviews. Furthermore, in addition to seemingly providing an increased feeling of immersion, Jonathan recalled that spatial audio was a driving force in making him move around in the surrounding environment.

> **Jonathan:** "*...I understood that I could move around, but I generally stood still. And then when I noticed that movement affected the sound, It suddenly became more interesting again.*"

This sentiment was echoed by several other participants, with the binaural audio being a key motivator in encouraging exploration of the environment .As one can glean from the observation reports in appendix F, there was at times very little movement going on, and participants spent much of the time during each session standing still. This despite the mobility arguably being a key feature of the prototype form factor, allowing for the utilization of space to cooperate creatively. This led to the concept of *movement* being established. This concept is rooted in how the prototype design almost demands active participation from users to fully experience the interaction.

> **Edvard:** "*I liked the idea that you could construct a chord and then move back and forth, hearing the volume increase and decrease.*"

While the use of movement is a key element in the design, Sverre had a mixed experience of how it was utilized in the prototype, questioning whether it was really necessary to utilize a space the size of a room. Instead, he expressed interest in having the playspace established on a table.

> **Sverre:** "*...It was exciting to walk around in it. But - the actual musical experience would have been just as nice if.. Let's say you were to walk around. Then - the whole aspect of bouncing something (nodes) off of stuff, that could just as well have been done on a table too - that would have been a little interesting.*"

However, when discussing the use of space further later in the interview, Sverre and Elias expressed more interest in spatial possibilities of augmented reality, and professed an interest in separating the room into different sections where each section housed its own separate instrument or functionality. They shared a brief vision of something resembling an augmented reality rehearsal space, where they could house and play various instruments together.

> **Elias:** "*I think that would have been cool. Then you could turn around and have like, sections of the room where you would have different instruments?*"
> **Sverre:** "*Mm.*"
> **Elias:** "*Or something?*"
> **Sverre:** "*Yeah, 'cause part of the problem facing musicians is that they want... would like to be able to play different instruments.*"

This concept of making life easier for musicians was a returning point of discussion in the interview with Elias and Sverre. However, while approached from a slightly different and less technical perspective, this notion of separating the augmented reality space into various sections was also brought up by Jonathan and Nikolai during the second interview. In their discussion, they requested functionality making it possible to group and synchronize nodes by placing them in specific parts of the space, using parameters of the augmented reality space to more directly control the behavior of nodes. As they discussed this further, they began playing around with the idea of using your whole house as a shared musical space, and using augmented reality to create and present a different piece of music for each floor.

> **Jonathan:** "*But this would be - i don't know - a cool idea? You could make 'The music of my house' - and create a piece that people could walk around and experience in my house, that would be awesome.*"
> **Nikolai:** "*That would have been cool!*"
> **Jonathan:** "*Yeah! This (the prototype) is basically that - only in a slightly different form.*"

The use of a larger space was also a central topic of discussion during the third session, despite the significant change in environment when moving the interaction outdoors. A recurring theme during the interview with Mikkel and Edvard was the idea of interacting within a greater and more permanent shared augmented reality space, where you might discover peoples creative output as a part of your everyday life as you went about your day.

> **Mikkel:** "*It could be like... If you want to share it (the music), people would go there and be like - 'What do we have here? - and then go - 'Oh, look! Someone made this song here! - you know?.*"

The idea of creating something and hiding it in plain sight using augmented reality technology was a persistent theme in the third interview, perhaps as a result of it being set in an outdoor environment. These discussions led to the establishing of the concept of *ludic experience*. This concept was centered around the sense of playfulness that might arise in everyday situations, when users are free to experience and explore their surroundings through use of augmented reality.

> **Edvard:** "*It became - in a way - this little space that I had never seen before.*"

Edvard in particular seemed enthused by the idea of taking part in a larger AR space, basically describing what is often referred to as the AR Cloud[2]. He described how he envisioned people not only taking part in this prospected AR world, but how everyone could help construct it. As an example, he pointed to the video game Minecraft and how this enables people to cooperate on constructing a virtual world. He wanted to be able to do the same in the real world, using augmented reality technology as a vessel for enabling these experiences. This futuristic line of thinking was also present in the other interviews, and there was a general feeling of enthusiasm in discussing the possible future applications of augmented reality technology in general. In the interview with Jonathan and Nikolai, they were not afraid of dreaming out loud about a more advanced version of the current technological status quo.

> **Jonathan**: "*You just have to map out the whole room.*"
> **Nikolai**: "*Imagine this with particle effects!*"
> **Jonathan**: "*Mm!*"
> **Nikolai**: "*You'd be like 'Whoa!' - and then you would sit down and go like - 'OK, who put something in my drink?!'*"

However, despite the desire for more complex visuals and graphics, there was a general consensus across the board that the design of the visual elements in the current iteration of the prototype corresponded well with the auditory elements, leading to the development of the concept *consonance*. When asked about the combination of visuals and audio within the prototype, Jonathan stated that he "*...did not experience any dissonance whenever the node hit the floor, and the sound played*". This was also the case when discussing the visual representation of the arrangement through the looping structure. Inspired by the current looping structure, Mikkel brought up the idea of using different geometric shapes to represent different rhythms, to create an even stronger bond between the spatial and auditory dimensions within the app.

[2] A persistent and worldwide augmented reality world existing on top of the real physical world

> **Mikkel:** " *It would be nice if you could go even further with it. Like, maybe there could be other sounds? Maybe different shapes? Maybe you could integrate the rhythm into - like - if maybe there was a triangle that went like...*"
> [**- vocalizes a complex rhythm -**]
> **Mikkel:** "*So that you had different shapes producing different rhythms?*"

Similarly, while initially feeling no apparent dissonance between the visual and auditory experience, Jonathan also requested more advanced visuals, citing use of color, light and particle effects as desirable improvements. The concept of *colors* was subsequently established as a result of this. During the third interview, Edvard mentioned how he saw potential for the prototype to incorporate color to a larger degree in the music creation process. He mentioned how the prototype might make use of the camera to capture contextual information, to "*...perhaps create chords based on the colors around us*". His vision was similar to how contextual information is utilized by Ljungblad [53], only instead of creating lomographic images, he wanted to create music.

However, contrary to the wishes of Jonathan and Edvard, Sverre stated how the use of color in the application was generally less important to him, requesting instead that the balls had been given separate sounds to separate them in the playspace. In a similar vein, Mikkel and Edvard experienced confusion during their session due to how the application assigned the color red to its own nodes, and the color blue to the co-player.

> **Mikkel:** "*But, if we had each been assigned separate colors...*"
> **Edvard:** "*If he had been given red, and I had been given blue. Then it would have been easy. Then we could have looked at each others device and seen where we.. 'cause we ended up comparing, and then we saw that we both had been assigned the color red on our device.*"

As witnessed by the above discussions, the concept of *color* appeared to be one of the more divisive aspects within the category of immersion, especially in contrast to concepts such as *spatial audio* and *space*, where the participants expressed many of the same thoughts and attitudes throughout the interviews. However, it was apparent that the prototype had succeeded in providing a relatively high level of immersion in general, despite being presented in the mobile form factor.

5.2.2 Augmented Cooperation

The category of *augmented cooperation* arose quite fast during the categorization process as a result of discussions around collaboration and co-creation. This category describes themes relating to aspects of cooperation in augmented reality space. The individual concepts that emerged from this process can be seen in table 5.3 below.

Awareness	Privacy	Ambiguity
Context		

Table 5.3: Codes for the category of *augmented cooperation.*

The first concept arose out of participants expressing a need for mechanisms to support *awareness* in the augmented reality space. These comments generally emanated from situations where participants reported feeling overwhelmed or confused by what was going on in the playspace, seemingly as a result of not having sufficient awareness of what their co-participant was doing at a given moment.

> **Nikolai:** "*It was a bit like him throwing a snowball at me! Like - whoosh!*"
> **Jonathan:** "*Only it wasn't really my intention!*"
> **Nikolai:** "*No - it's not like..*"
> **Jonathan:** "*It wasn't like - 'OK, now I'm gonna place one right beside..'*
> **Nikolai:** "*No? I guess I wasn't really sure of that actually!*"

It is likely that this seemingly uncomfortable situation described by Nikolai could have been avoided if the prototype had been better at providing awareness information to users while they were immersed in the playspace. Furthermore, this lack of supporting awareness mechanisms also presented some challenges during the first testing session where Sverre reported feeling confused as to why Elias was appearing so inactive.

> **Sverre:** "*There were several occasions where I thought - 'Am I the only one playing?*"
> **Elias:** "*Yeah!*"
> **Sverre:** "*Because of stuff I didn't catch in the moment, like Elias standing over there..*"
> **Elias:**"*Mm.*"
> **Sverre:** "*And then I was doing something over here. And then I simply didn't register whatever he was doing over there.*"

As indicated in the quote above, despite having a clear view of the other participant through the augmented reality view the lack of awareness information gave Sverre a feeling of diminished cooperation within the playspace. Additionally, it is likely that these experiences were worsened by technical challenges facing the prototype, such as loss of tracking in the ARCore module, or network latency leading to participants nodes not being rendered immediately on both devices. In speaking of cooperative aspects, during the above discussion around awareness Sverre and Elias also requested a more structured and selective workflow in the application, with the possibility of previewing your created work before sharing it in the playspace. This was a common theme, and across all three sessions participants requested more *privacy* in the AR space, giving statements such as "*What might have been useful would be to be able to mute the other participant's nodes*" and "*Sometimes you just want a little space, simple as that*". The concept of privacy was established on the basis of statements such as these. In speaking of privacy in this context, the concept refers to a more general definition of privacy as spoken about by participants, not the design tension described by Gross [30]. In general the uncompromising and compulsory nature of cooperation provided by the prototype in its current state appeared to be seen as a challenge more than a convenience by participants.

> **Sverre:** "*I had the feeling of - and not because Elias did anything wrong - but I sometimes got the feeling that it would have been easier to just do it by myself.*"

On a related note, several participants reported feeling hampered when their co-participant exhibited a high level of activity in the playspace, not knowing when to step in and participate themselves. As put by Edvard during the interview, "*I suddenly found* *myself doing nothing, and simply paying attention to what was happening*". Several participants saw this feeling of *ambiguity* as connected to the application using the same sound for both participants, as this was reported as a common source of confusion during all three sessions. As a potential remedy to the issues related to awareness, Sverre proposed a more privacy-oriented model, with the inclusion of a mechanism to notify him whenever a node was being put into the playspace by another user, stating "*...that would really make things easier, because then you would know like - 'OK, something is coming now'*". Elias also requested more focus on supporting cooperation awareness in the prototype space, suggesting that the application simply notify you whenever the co-participant is active.

> **Elias:** "*I am immediately reminded of Messenger. This might be a little far out, but yeah - Messenger immediately comes to mind. Or Snapchat. It lets you know whenever anyone is writing you a message.*"

The discussions regarding a lack of awareness mechanisms emerged organically through the interviews, indicating the need to more strongly consider these aspects in the development of mobile shared augmented reality experiences. The inclusion of headphones in the prototype experience could also likely be seen as a contributing factor, as they undoubtedly had an effect on the overall awareness of each participant throughout the session.

During the first interview, Sverre and Elias also kept returning to the idea of extending the context of the prototype beyond the dimension of "same time - same space", in order to enable remote cooperation in augmented reality space. They essentially wanted the possibility to also use it in the context of "same time - different space", to allow for immersive and spontaneous musical interaction from their respective homes. Sverre remarked how it would likely be better to simply sit down on a computer to make music when in the same place together, instead of interacting through a tool such as the prototype. However, the prospect of having a shared immersive jam space in augmented reality seemed more attractive.

> **Sverre:** "*The dream is to be able to jam together whenever you want!*"
> **Elias:** "*Yeah*"
> **Sverre:** ""*I mean, how many times haven't we been just sitting at home bored like - 'Yeah let's just get a jam session going!*"

During the third interview, when reflecting on the cooperative aspects of the prototype Mikkel and Edvard were generally positive to the current context of the prototype, citing that "*...the fact that we were in the same space made it a little easier*". However, as discussed in section 5.2.1 they also envisioned extending the context of the cooperative work outside of the current state, dreaming of a collective augmented reality experience capable of supporting interactions in the "same place - different time" context.

Despite being successful in enabling simultaneous cooperative work in augmented reality, the prototype deployment revealed several potential challenges in developing cooperative experiences for mobile augmented reality.

5.2.3 Spatial Instrumentality

The level of control and customization over the musical aspects offered by the prototype was another returning point of discussion within the interviews, leading to the emergence of a category I dubbed *spatial instrumentality.* The concepts developed within this category can be seen in table 5.4.

Control	Looping	Expanded Sound Palette
Toy or Tool	Rhythm	Possibilities
Longevity		

Table 5.4: Codes for the category of spatial instrumentality.

The first concept of *control* was established due to participants repeatedly requesting more fine-grained control over the various aspects of the musical interface, regardless of their previous level of musical experience.

> **Elias:** "*I would really have liked access to more settings!*"
> **Sverre:** "*Yeah!*"
> **Elias:** "*Turning stuff on and off, tempo, BPM*[3]*. Give me all of these settings! Also, if I can't get BPM control, then give me a setting to control the animation time of each node. Give me a setting to define how high each node should bounce!*"

The first interview in particular was characterized by Elias and Sverre seeming almost frustrated at the limited control they had over the programmability of the prototype. The duration of the *looping* in particular was a common point of contention in the interviews, especially in interviews 1 and 3 where all participants had a musical background. Several participants reported feeling limited by the lack of control over the length and tempo of the loop, being reduced to work with the same rhythm every time. However, despite craving a more advanced looping mechanism, the fact that looping was included as a function received praise, with Sverre stating that "*If it (looping) had not been included, I don't think I would have enjoyed this very much*". Along with the loop duration, the lack of variety in sound was another common source of frustration during testing. Participants reported sometimes feeling confused by not knowing whether what they heard was a result of their actions, or those of their co-participant, and requested the addition of at least one more sound. This led to the establishing of the concept *expanded sound palette.*

> **Edvard**: "*Whenever he (Mikkel) is playing, it is difficult to play yourself and get anything out of it.*"
> **Mikkel:** "*Yeah - maybe if we had two different sounds? Or something?*"

During the interview, Sverre described feeling like they both had "*...two identical instruments*", and compared the experience to playing piano, citing "*When you have a piano, you don't really play together. You let the other person play first, and then you play afterwards*". This statement is also somewhat indicative of how Sverre perceived the prototype more as an instrument to be played, rather than musical interaction meant to be experienced. This difference was a recurring theme across all interviews, and led to the revealing of the concept *toy or tool* during the coding process. In general, participants appeared positive to the prototype as an interactive piece or an experience, but less so when speaking of it as an instrument or tool.

[3]Beats Per Minute, a common way of referring to the tempo of a song in electronic music

> **Elias**: "*But the question is - like - is it supposed to be more like a toy? Or should it be a real tool to create music?*
> **Sverre:** "*Yeah.*"

Sverre in particular appeared almost discouraged at times at how close it was to providing a comprehensive means of musical expression, but ended up falling short by the lack of advanced features. This was a central theme in the first interview, as both Sverre and Elias kept dreaming out loud about additional features such as tempo control and quantizing[4]. Another much requested feature was the possibility of saving the piece created through the prototype, either as a recording to listen to, or as an interactive piece to work more on at a later time. In the same vein, Edvard played with the idea of recording tidbits of the created music and importing it into an external piece of music software, in order to enable creation of larger and more comprehensive composition.

While the prototype is designed with a focus on harmony and melody, the concept of *rhythm* was identified after several participants requested additional functionality to support the creation of difficult and more advanced rhythmic sequences. Nikolai has a background as a drummer, and described during the interview that he was primarily driven by rhythm when interacting with music. He therefore felt the application was limited in how it only provided rhythm as simply a product of the looping mechanism, with little variation in what patterns could be created in the playspace. Similarly, Elias and Sverre underlined how their initial experience of the prototype was quickly shaped by what they perceived as qualities relating to rhythm.

> **Elias:** "*I think we quickly felt that this was something percussive.*"
> **Sverre:** "*Mm.*"
> **Elias:** "*It was something rhythmic in nature.*"
> **Sverre:** "*Yes.*"

In general, a significant amount of the themes discussed in the interviews were centered around the potential of extending the functionality further, leading to the creation of the concept *possibilities.* During the process of coding, I felt at times like the focus was mainly on what the participants considered to be flaws in the interaction. However, when I had become more familiar with the data, I also recognized how participants often spoke of the prototype with a feeling of promise and potential.

> **Sverre:** ""*But it seriously is a really cool, a really cool concept. I just want to say that, yeah... I'm going to be dreaming about this all night tonight."*
> **Elias:** "*Mm.*"
> **Sverre:** "*That I am sure of. Because there are so many possibilities there.*

The final concept of *longevity* was established as a result of discussions regarding how long it was viable to actually use the prototype. Observations during testing revealed that the participants engaged with it for a duration of between 30 and 40 minutes in general. However, this included the initial familiarization and learning stage, and would likely be shorter had the participants used it before. Therefore this does not necessarily reflect a general session duration.

[4]The process of correcting and locking musical notes to exact fractions of a beat.

> **Sverre:** "*So, if you had added one additional parameter to control, we would have kept going for another half hour!*"
> **Elias:** "*Yeah, but we would!*"
> **Sverre:** "*And then, the more parameters you enable us to control, the more half hours you get from us.*"

The above quote by Sverre and Elias nicely summarizes the overall attitude held by participants with regards to the aspects of control, and the experience of the prototype as a whole in its current form.

5.3 Summary

In this chapter I have presented the findings that were established through the open coding and subsequent analysis. These findings are rooted in both observational data recorded during testing, as well as the subsequent group interviews performed after each session. The findings were organized into the categories of *immersion*, *augmented cooperation* and *spatial instrumentality*. In general the findings point towards participants experiencing a significant level of immersion when engaging with the prototype interaction. Furthermore, the use of binaural sound was well received by most participants, and led to both a strengthened immersion and an increased motivation to move around in the environment. Furthermore, participants requested both access to more sounds, as well as more control over the various aspects of the experience to allow for an increased range of creative expression. The prototype also suffered from a lack of awareness support mechanisms, which led to a reduction in the coordination experienced by participants during the session. However, findings also point towards the prototype experience being a novel and engaging way of interacting with music mediated by augmented reality technology.

Chapter 6

Discussion

" The best musicians or sound-artists are people who never considered themselves to be artists or musicians."

- Richard D. James

In this chapter I discuss the project research questions in relation to the findings presented in chapter 5 and the related work presented in chapter 2. In addition, I will also review the methodological approach used throughout the project, and reflect on it in light of the theory on RtD presented in chapter 3. To reiterate, the primary research question for this project is as follows:

RQ 1 *How can mobile augmented reality technology be used to design cooperative musical interactions?*

In addition, the project has explored the following additional research questions:

RQ 2 *How does the inclusion of binaural audio affect the immersion of users in a cooperative mobile augmented reality space?*

RQ 3 *What awareness mechanisms are needed for ludic mobile real time cooperative interaction in augmented reality?*

In the following section, I begin by discussing the main research question. Moving on, I discuss the second research question, and provide some insight on how the use of spatial audio affects users immersion in augmented reality space, while also presenting some implications for design related to the research findings. This is followed by a discussion of the third research question, where I assess how cooperative augmented reality is affected by awareness mechanisms, seen in light of theory from the CSCW field. Following this, I evaluate my use of Research through Design as a methodology for this work, looking back at each phase of the project and reflecting on the work in the context of the five step RtD model presented in section 3.1.2. Finally, i provide a brief summary of the chapter, highlighting the main points of each subsection throughout.

6.1 Designing Cooperative Musical Interactions for Mobile Augmented Reality

While mobile augmented reality technology has been around for a long time now, it has yet to achieve widespread public adoption. Although many have been exposed to it through games and entertainment, we are still a ways off from the future envisioned in the science fiction works of yore [3]. However, as the new frontier of augmented reality is gradually explored, new experiences will unfold in virtual spaces. These will likely require new forms of interaction, which designers now have the possibility to influence and shape through experimentation and exploration.

Motivated by the project research questions, a high fidelity prototype in the form of a mobile application was designed and developed. The integration prototype, an application named *Petals*, was the outcome of a design process grounded in the Research through Design methodology, inspired by key exemplars of existing work in the field. These key exemplars were instrumental in guiding my design process, and were essential in the formation of my initial design concept. This design concept established in section 4.1.3, which has informed my creative approach throughout, was formulated as:

> *Exploring synchronized and ludic cooperative interaction with music in mobile augmented reality space.*

Looking back on the prototyping process described in chapter 4, I consider the above design concept to have been successful in keeping the creative process anchored in the fundamental aspects of my research questions, while at the same time providing freedom for creative thought and experimentation. Through the subsequent field deployment process, I was able to test and evaluate the prototype in three different environments, with six different people, which led to the collection of a rich and detailed data material providing valuable insight into how the prototype interaction is experienced by real users. The final integration prototype represents the conclusive embodiment of my vision for the design concept, and exists as an exemplar of how to enable cooperative musical interaction in a mobile AR space.

While the design concept is inspired by key pieces of work such as Bloom: Open Space [81], the Musical Moves project [12] and the work by Kiefer & Chevalier [49], the final prototype design emerged organically as a result of several key activities performed throughout the prototyping phase. As I moved through the various stages of design, I continuously evolved the design through each iteration, while also critiquing and evaluating each new resulting concept. However, as the prototype grew more sophisticated, I experienced it as somewhat difficult to decide when it would be ready for deployment. The RtD framework I used as my scaffolding simply professes that evaluation should happen "*When the team has an artifact they like*" [87, p. 185], which made it difficult to make a decision grounded in the methodology. Therefore, I chose instead to assess the state of the prototype against the research questions, asking whether it would be functionally capable of collecting the required data. After having implemented all the necessary functionality as described in chapter 4, and reaching a stage where I felt I had done the design concept justice, I went through a small phase of bug-fixing prior to putting on the finishing touches, before concluding the design phase and moving on to the field deployment.

6.1.1 Deploying an Immersive Interaction in the Field

Prior to the field deployment I had been somewhat uncertain of what results to expect from the participants. Having performed the pilot test as described in section 3.3.8, I was relatively confident that the prototype would function as planned without any major technical hiccups or breakdowns, and that I would henceforth be able to provide the intended experience to participants. Furthermore, rooted in my research questions I felt secure that employing qualitative methods through a field approach [50] was a suitable method of evaluation for this project. However, I did not know what to expect in terms of session duration, level of activity, or whether participants would actually appreciate interacting with the prototype in the way i had anticipated. As a starting point before the first deployment, I used the duration stated in Kiefer & Chevalier [49] as an estimate, and planned for the testing to last somewhere between 10 and 30 minutes, similar to what they had described [49, p 27]. It turned out that this estimation was not too far off the mark, with the first session having a total duration of 35 minutes, after which it organically dispersed. The next sessions both had comparable duration, lasting for 32 and 43 minutes respectively.

In discussing the prototype interaction with participants after each session, it seemed like the experience genuinely represented something fresh and of a different character, even to those with some experience using AR apps. Several participants reported feeling a significant level of immersion in the AR playspace during use, despite the interaction being presented in a mobile form factor with limited screen size and relatively simple graphical visuals. One participant reported feeling "*lost in the room*" while engaging with the prototype, indicating that he felt deeply absorbed in the experience. Another participant described feeling like he could suddenly see into "*... a new world*" through the augmented reality application, a feeling he linked to the experience of immersion in the space. During the observations, I saw moments where the participants appeared to be in a state of deep focus and immersion. Initially, these findings were somewhat surprising to me, as prior to the field deployment I had not expected that participants would attain this level of immersion through the experience. When seen through the Reality-Virtuality Continuum scale [57] displayed in section 2.1.2, the experience provided by the prototype can arguably be said to exist closer to the left side of the spectrum (e.g. closer to a real than a virtual environment), as the amount of visual augmentation taking place is somewhat limited. However, one might consider that this scale does not account for the influence of other possibly important factors present in the prototype interaction, such as the cooperative aspects, or the use of binaural audio along with over-the-ear headphones. These factors were likely significant in strengthening the feeling of immersion experienced by users. In speaking of augmented reality, it can often times be easy to ascribe the feelings of immersion as a result of the factors relating to the visual experience, such as graphical fidelity or spatial placement of virtual content. However, the findings developed from the analysis point towards audio being a powerful tool for designers who wish to encourage immersiveness in mobile interactions. In further support of this, the importance of auditory aspects when aiming to cultivate immersiveness can be seen in the findings reported by Kiefer & Chevalier. In their study, participants described experiencing feelings of deep immersion despite the complete absence of any visual augmentation at all [49, p. 27]. This immersion might be further strengthened by incorporating binaural or spatial audio as a part of the interaction. As they relate to the second research question, the effects of the binaural audio module employed in the prototype are discussed further in section 6.2.

While participants generally appeared to enjoy engaging with the immersive playspace, the interviews also revealed the existence of several weak points in the design, which were mainly related to users experiencing a lack of programmability and control. In describing their interaction with the prototype, participants reported experiencing feelings of excitement and fun, which were then subsequently followed by feelings of being creatively constrained in the interaction as a result of lacking control over the musical elements of the interaction. This was identified in the findings as a tension between the prototype being a *ludic experience*, or a musical instrument and *utilitarian tool*. To shed further light on this, I will assess these differences and discuss how they relate to the major design choices made during the prototyping process.

6.1.2 Petals: AR Experience or Musical Instrument?

My intention of creating an experience rather than a tool for musicians has been declared at several times throughout chapter 4. This choice was grounded in both the research question and the design concept, which are both focused on encouraging a form of ludic interaction with music, rather than mastery of an instrument. Furthermore, this view also reflects my main sources of inspiration [49] [81], who are by design both exemplars of *musical experiences* more than instruments. I also felt that this view was supported in part by the principles for designing computer music controllers [20], wherein one of the proposed principles allude at the possible pitfalls of creating a musical instrument simply for the sake of it, without a musical idea or piece to give the work direction and meaning. Therefore, the design phase drifted more towards creating something less *programmable*, and more *explorable*. Musical complexity was traded for ease of use and fluency in the interaction. While participants seemed to at first enjoy the experience of using the prototype, after becoming familiar with the basic mechanics they quickly seemed to want to go beyond the possibilities offered in the interaction. In one way, this could indicate that participants might have approached the prototype from a slightly different perspective than I had intended in my design. On the other hand, the nature of the interaction itself should signal the nature of the experience simply by design, so that users are put into the correct mindset when first starting the experience. To directly underline this tension I refer to the following quote:

> "*But the question is - like - is it supposed to be more like a **toy**? Or should it be a real **tool** to create music?*"

The above quote which I have recited from chapter 5 stems from the first interview, and summarizes much of the tension experienced by users as they explored the interaction. In a broader perspective, this is almost indicative of something like a dichotomy emerging between the playful and ludic aspects (toy) on one side, and the utilitarian (tool) on the other. Prior to the deployment, I had been wary of this difference myself, as repeatedly discussed throughout chapter 4. Oftentimes I had found it challenging to decide between increasing the amount of control available in order to increase the range of creative expression, or limiting it to simplify the experience and attempt to encourage the emergence of flow (as discussed in section 4.5.4). As mentioned, I repeatedly found myself drifting towards the latter, focusing on my overarching goal being to create an *experience* rather than a fully fledged *instrument*. This is reflected at key points in the development such as in section 4.5.4 where I performed a relatively simple mapping of parameters intended to reduce the learning curve and encourage simplicity, or in section 4.5.3 where I decided to implement a mechanic to mediate the selection of notes, to encourage synergy and cooperation during play. I saw these choices as essential in designing an experience that would feel rewarding for anyone, not only for musicians.

However, while these design choices received a mixed response, some aspects were met with more disapproval than others. One of the most criticized elements of the prototype was the limited control given over the looping mechanism.

While the inclusion of looping as a mechanic in general received praise, participants unanimously expressed feeling limited by the lack of control over the loop length. As discussed in section 4.5.2, this was something I myself had been wary of during implementation. At the time, I had decided against implementing control over the loop length, hoping that this constraint would be seen as creatively stimulating rather than limiting, by enabling improvisation within a set of boundaries. The limited loop length was also connected to the musical concept of ostinato, which as described in section 4.5.4 had been a source of inspiration during the design of this mechanism.

Another common source of frustration in the prototype was the lack of variety in the available sounds. As described in section 4.5.5, I had made the deliberate decision to assign the same sound to both participants within the playspace. This choice was made with the goal of encouraging the emergence of a collective and shared musical expression and blurring the lines of creative ownership in the playspace, where it would be less important *who* created a given node, and more important *what* it became part of. This was another element in designing towards an experience rather than an instrument. However, the results indicate that while participants at times enjoyed the feeling of contributing to a shared piece of music within the playspace, the lack of sonic variety was generally a source of confusion and frustration more than a benefit. Several participants mentioned how after having placed a node into the playspace, they were unsure of whether what they heard was a result of their own actions, or the activities of their co-participant. This confusion also relates to the level of awareness experienced within the interaction, which is discussed in detail in section 6.3. Nonetheless, the findings generally indicate that the interaction would have benefited from an extended vocabulary of sounds. Reflecting on this, one could maintain the idea of sonic collectivity somewhat by ex. giving both users access to multiple sounds, weakening the connection between a specific sound and a users identity, while also expanding the creative possibilities offered by the experience.

Despite the above points of contentions there were other aspects of the design that were more successful. One such aspect was the connection between the visual aspects and the sound. The findings indicate that participants did not feel any dissonance between the visual and auditory aspects of the prototype, indicating that I was generally successful in creating a coherent audiovisual experience as described in section 4.4. Furthermore, the inclusion of microphone input seemed to be a positive addition, as several of the participants experimented with vocalizing and singing while playing with the prototype. In general, the use of microphone also reflected the results reported by Kiefer & Chevalier, wherein the most performative participants were also those with seemingly less inhibition to experiment and play using their voice [49, p. 27].

6.1.3 Augmented Reality as a Space for Cooperative Musical Interaction

While the findings point towards the prototype having something of an identity crisis, there was also a distinct air of positivity to be traced in the participants throughout the interviews. While the application was at times unsuccessful in stimulating flow in the creative cooperation, the participants appeared enticed by the general concept of cooperative musical interaction in augmented reality, and appeared to see a great deal of potential in the medium as a space for creative expression. Several participants talked of ways they would like to extend the functionality, and how they imagined the interaction to work in a future iteration. One participant envisioned a world where this form of interaction was commonplace in our everyday lives, functioning as a way of creating and sharing content in a larger and more persistent augmented reality world. Another participant cited how the prototype allowed him to peer into a space behind reality, a sort of hidden world in the environment where magic could happen.

The deployment of the prototype has shed light on how we might interact with music in augmented reality, providing knowledge not only grounded in HCI, but also giving insight into how we might explore the challenges of sound and music computing as well.

The somewhat polarized reception as described by the findings could indicate that the project has developed a good design concept, which is somewhat hampered by its implementation. As indicated by the points made in section 6.1.2, the prototype interaction seems to exist in a space *between* being an experience, and being an instrument. It is too malleable to be considered an artistic piece, but at the same time too limited to be a satisfactory instrument, giving it an ambiguous identity. Evaluating this problem through the prototyping framework [36], the problem could be considered existing along the dimension of *role*. To shed further light on this, we might look back and evaluate the final prototyping model depicted in figure 4.24. This model shows role as the least explored dimension of all, with the dimensions of *look and feel* and *implementation* having received more attention throughout the prototyping, indicating that there has been a lack of focus given to the dimension of role. This asymmetry can likely be explained by having too strong a belief in the concept alone, and feeling like the sources of inspiration such as Bloom: Open Space [81] or the study by Kiefer & Chevalier [49] were proofs of how the broad conceptual ideas behind the prototype were sound. However, as the concept evolved throughout the design process, new features were implemented, which moved the concept away from the exemplars, increasing the need to re-evaluate its role. Many of the above mentioned issues could therefore likely have been mitigated by more actively assessing the dimension of role within the prototype, and concentrating more on what function each implemented change would serve in a larger perspective.

To reiterate and summarize, the prototype received a mixed reaction from participants during testing, mainly due to a lack of variety programmability. While it was successful in providing an enjoyable and ludic experience, the implementation is faced by several challenges that limit the overall potential for long-term engagement and revisitation. Similar applications might benefit from establishing themselves more clearly as either an *experience* or an *instrument*. In short, it appears that the design concept is good, with a somewhat flawed implementation.

> "*But it seriously is a really cool, a really cool concept. I just want to say that, yeah... I'm going to be dreaming about this all night tonight.*"

6.2 Supporting Immersion in Mobile AR Through Binaural Audio

As discussed in the previous section, the prototype appeared successful in evoking a strong sense of immersion during the field deployment. When speaking about the immersive qualities in the interviews, several participants reported that the use of binaural audio was a significant contributing factor. As described in section 4.5.8 the binaural audio elements were included with the intent of both strengthening immersion and improving the feeling of collaboration between participants in the AR space. In addition, the previously mentioned study performed by Kiefer & Chevalier shows how the use of audio alone can have a powerful effect on the immersiveness experienced by users, without the use of any visual elements [49]. While their system did not appear to employ binaural filtering, one of the patches used was stated to be "*...altering spatial qualities by presenting different band-pass filtered reverbs in each ear*" [49, p. 26], thereby still manipulating the soundscape and presenting the user with something spatial and dynamic, different than a uniform audio signal. The effect on immersiveness is as previously discussed supported by other studies [86] [75], underlining the potential of using audio as a tool to achieve increased immersion. However, prior to the field deployment I was not sure how these effects would manifest in my project, as it was rooted not only in augmented reality, but in sound and music computing.

> "*That was a really cool experience. You would kinda be like - 'Where is that sound coming from?' - and then just go 'Oh, there it is. Alright!*"

The above quote underlines the enthusiasm felt by participants as they explored the playspace. The findings indicate that the use of binaural audio both helped increase the perceived immersion, as well as stimulate a feeling of fun during use. I see the use of audio to strengthen immersion as a powerful tool, especially when working in the mobile phone form factor. However, implementing binaural or spatial audio comes with one significant drawback, which is that it requires users to wear headphones in order to produce the desired effect. Playing back binaurally filtered audio through device speakers does not produce the desired result, and simply results in a diminished experience. However, using headphones might not always be possible or desirable for users, and it is therefore important to - if possible - also provide an experience that is mindful of scenarios where audio is either muted or played through the device speakers. Furthermore, wearing headphones might also make users less aware of their surroundings, which could potentially introduce risks and negative consequences when using them in certain environments (ex. along a trafficked road or in crowded public spaces). It is important to be attentive to these issues, and also assist in keeping the user aware of their surroundings. Therefore, the context of use for the application should be assessed before deciding if binaural audio is suitable as a mechanic to support immersion in the experience.

6.2.1 Instigating Active Participation through Sound

While the main reasoning behind the inclusion of spatial audio was rooted in factors related to immersion, the findings indicate that the inclusion of spatial audio might also be an effective way of provoking users to move around and explore an augmented reality environment more actively. To illustrate this, I restate the following quote, also presented in chapter 5.2.1:

> "*...I understood that I could move around, but I generally stood still. And then when I noticed that movement affected the sound, It suddenly became more interesting again!*"

This effect was something I had not previously anticipated during the design process, and its effectiveness came as somewhat of a surprise. When envisioning the prototype in use, I had simply assumed that the participants would organically begin moving around and exploring the space on their own. However, during observation of each session I frequently noticed that users would regularly stand still when interacting with the playspace, despite the inherent spatial nature of the experience. This lack of movement could be rooted in the interface requiring a certain level of focus and precision during use, making it difficult to both move and operate the interface at the same time. In addition, the instrument interface was designed in such a way that users could access the entire range of available tones in the system without having to actually move around in the playspace, thereby not actually encouraging movement to play. It was therefore interesting to see how the use of binaural filtering counteracted this, and gave the participants an incentive to move when they became aware of the binaural properties. Furthermore, while both Google [28] and Apple [38] encourage use of audio in their respective augmented reality design guidelines, the format of the audio is not specified, and there is no mention of either binaural or spatial audio. The findings indicate that if designers aim to promote movement in their experiences, spatial audio might be a useful tool in achieving this.

However, while the binaural audio was seemingly effective in instigating active participation during the field deployment, it must also be considered that the motivation reported by users could stem simply from this instance-specific inclusion and implementation of binaural audio. In other words, it is difficult to assess whether the use of binaural audio would be as effective when generalized to other applications and contexts, or even in repeated uses of the prototype application. It is possible that the effect was tied to the novelty of discovering the binaural audio in the app, and that this effect would quickly wane over time. Nonetheless, this finding could be of significance when designing augmented reality experiences that require a certain level of active participation from users, to encourage movement and exploration to a larger degree. While the results indicate that the use of binaural audio was generally successful, leading to both a strengthened immersion as well as increased movement, the prototype had some limitations in how the binaural aspects were implemented. These are discussed in section 7.2.2.

6.3 Awareness Mechanisms for Ludic Cooperation in Augmented Reality

When seen through the CSCW-matrix presented in section 2.2.1, the experience provided by the prototype interaction resides within the context of *same time, same space* [42]. During use, participants engage in the interaction in a synchronous manner, and while they may move about in space and explore the environment, the cooperative aspects are intended to stimulate simultaneous creative expression within the same physical environment. Designing the experience to take place in this context was my intention from the start, and is rooted both in the main sources of inspiration [49] [81], as well as my own curiosity and drive to explore aspects of synchronous cooperative interaction in AR. In analyzing the findings from the field research, it became evident that while the prototype interaction was successful in enabling synchronous cooperative interaction, it was at times significantly affected by what participants perceived as inadequate support of awareness in the AR space. As discussed in section 6.1.2, several participants reported experiencing feelings of confusion or frustration due to lacking awareness information within the application, leading to sporadic breakdowns in cooperation, ultimately reducing the overall coordination experienced by participants. This lack of awareness manifested itself in several different ways, and was related to both visual and auditory factors. To illustrate how the lack of awareness mechanisms affected the participants, I will repeat the following quote from chapter 5:

"*There were several occasions where I thought - 'Am I the only one playing?'*"

Situations like this one occurred multiple times, generally as a result of what participants described as a lack of awareness information. As described in section 5.2.2, one participant even described feeling startled by the sudden appearance of a node next to his ear, feeling like he had just been "hit with a snowball". These issues can likely be seen as related to participants having a lack of information on each others activities, despite both of them being the same space and co-interacting. In addition, it is possible that this situation was exacerbated by the high level of immersion, and the use of headphones. When analyzing the material, I considered it crucial to assess this issue in detail, as a central concept in both the main research question as well as the prototype is to encourage harmonious cooperative interaction. I therefore see it as integral to discuss how we might better design for awareness in shared augmented reality spaces. To reflect on this, I would like to consider the findings in relation to the awareness related design tensions described in section 2.2.3, namely the concepts of *privacy* and *conventions*.

6.3.1 Awareness Design Tensions in Augmented Reality

In further probing how awareness mechanisms affect cooperative augmented reality, I will assess the findings in relation to the design tensions of *privacy* and *conventions* as discussed in section 2.2.3. In doing this I will attempt to identify any connections between augmented reality, and the remote cooperation scenarios generally described in these tensions [30, p.] While the nature of shared augmented reality within the prototype diverges somewhat from cooperation over a distance, it might still be useful to consider how these design tensions affect interaction between participants in the shared space. In approaching the design tension of conventions with regards to cooperative augmented reality, I see it as relevant to probe how the findings might shed light on what challenges are faced within this space. In deploying the prototype I was curious as to if any distinct conventions would emerge between the participants, and whether the participants would be able to make use of established patterns of behaviour related to musical cooperation when using the prototype, despite the novelty of the interaction. In citing Mark [55], Gross points to how conventions are built on the existence of common ground between people, which needs to be established before cooperation can take place [30, p. 458]. This common ground is generally shared between ex. people belonging to the same social group, or people who work together [30]. I therefore found this particularly interesting with regards to the first deployment session, wherein both participants had significant experience playing music together. In other words, as there existed a significant amount of already established common ground, I was curious how the cooperation would play out in augmented reality space. In general, the findings indicate that the interaction did not establish any identifiable conventions during the deployment, and despite a high level of existing common ground, participants found it somewhat difficult to predict and interpret each others behavior which subsequently led to a reduced coordination. During the first interview, Sverre was quoted as saying "*...I sometimes got the feeling that it would have been easier to just do it by myself.*", which could be indicative of how the lack of conventions made it difficult to coordinate cooperation, despite the participants being used to cooperating musically with each other. However, this must also be seen in light of this being a completely new form of interaction for the participants, and users might need more time to establish the most basic conventions of use before they can make use of existing ones.

In discussing the design tension of privacy, Gross points to how moving social interaction to remote channels introduces new challenges [30, p. 457]. The mechanics we are used to from face-to-face interactions might no longer apply, and we have to consider new ways of communicating this awareness information. While Gross focuses here on social interactions over distances, I was curious as to how the design tension of privacy would relate to the medium of cooperative augmented reality as well, particularly with regards to the prototype context as seen through the CSCW-matrix [42] depicted in section 2.2.1. Furthermore, gross points out that as the amount of information transmitted about one's actions increases, so does the potential for awareness among those receiving the information [30, p. 456]. At the same time, the increase in information transmitted also increases the potential for violation of privacy [30, p. 456]. Rooted in these ideas of privacy, I was interested in exploring how the participants experienced the privacy aspects of the prototype interaction.

With regards to the tension of privacy, the prototype interaction can be said to take an uncompromising approach, as everything users put into the cooperative space becomes part of the shared musical expression. This was a conscious design choice, and I saw the ludic nature of the experience as a motivator in shaping this aspect of the prototype. However, due to the large amount of information transmitted as part of the cooperative interaction, it is possible that users might still see this as breaching their privacy, despite the interaction being rooted in playfulness and fun. As one participant was quoted as saying, "*sometimes you just want a little space*". This quote in particular can be considered indicative of a need to keep some activities separated from the shared space, despite the playful nature of the work. In the same vein, one participant requesting the ability to mute others nodes, which can be interpreted as a request for more privacy in the shared space. However, while a few participants expressed some interest in more strongly supporting privacy mechanisms, they were generally positive to the cooperative nature of the experience, and did not express any feelings of breached privacy. This could likely be related both to external factors such as the nature of the deployment, as well as the playful design of the interaction. Furthermore, while the medium of augmented reality clearly communicates each participants' presence, the information of activities communicated through augmented reality space is sparse and only related to the ludic elements of the interaction, which could reduce the potential for users privacy being breached.

With regards to the awareness design tensions, the findings were generally inconclusive, and more research is needed to probe the space. However, framing our thinking in these already established conventions could be helpful in uncovering how we might better design for awareness in augmented reality. Viewed through the lens of theory, the challenges described in the quote presented in section 6.3 can be seen as indicating a lack of *cooperation awareness* in the experience. Looking back to the definition presented in section 2.2.1, cooperation awareness is characterized as "...users' mutual information on their activities — either as background information in a collaborative working environment, or as foreground information in a cooperative application" [30, p. 438]. In light of the challenges experienced during the deployment, I see it as useful to more closely assess how cooperation awareness, or lack thereof, affected the participants experiences.

6.3.2 Cooperation Awareness in the Musical Space

In the prototype, information on cooperation awareness was communicated largely through the presence of nodes in the playspace, as well as their color and sound. The presence of a node would indicate activity, while the color of a node revealed the source of this activity. In other words, when engaging in a shared session, users would see visual indicators of their co-participants' activities in the presence of blue nodes. Whenever a new blue node appeared alongside their own red nodes, it meant that their co-participant had been active in the shared space. Each node having its own color made it clear who had place what node, and it was then possible to ex. glean the activity level of another participant by counting the number of blue nodes present. However, while the nodes were given different colors according to which user created them, all nodes were assigned the same sound regardless of their origin. As discussed in section 6.1.2, this was identified as a central flaw in the design first with regards to programmability, but this likely also had a disruptive effect on the amount of cooperation awareness experienced by participants during the deployment. By also assigning different sounds to the nodes, one could register co-participants activity not only through visible cues, but through auditory signals. It is possible that by simply assigning each user a different sound, the level of experience cooperation awareness would also increase, as users would be able to fully separate their own activities from those of their co-participant.

However, the quote presented in section 6.3 also hints the need for additional - and more dedicated - awareness supporting mechanisms in the AR space, to indicate the activity of other users even when they are out of sight (or earshot). One potential solution to this could be to use a separate dedicated audio track within the application to communicate awareness information on users' activities in the background. One might envision something like a soft ambient background soundtrack increasing or decreasing in volume in accordance with the activity level of users in the playspace. This way, users would be able to hear that their co-participants were active and interacting with the space, despite being out of range from their actual output. In other words, with this mechanic implemented, a quiet space would mean an inactive space. This sort of mechanic could also be used in other less musical experiences, as the nature of the audio track could be varied according to the nature of experience provided. Communicating awareness information through sound like this is nothing new, as evidenced by the previously mentioned *ambientRoom* project by Ishii et al. where the volume and density of natural soundscapes are used to communicate awareness information [40].

Another possibility could be to employ a more direct visual indicator signaling whenever the other user was active. During the interview, one participant suggested implementing a mechanic similar to what is found in messaging applications such as Facebook Messenger and Snapchat, who both have a mechanic where the application lets you know whenever someone is currently typing a message to you. The prototype could implement a similar mechanism, which displayed a visual indicator or notification whenever a co-participant was displaying a consistent level of activity. One could also envision visual information being communicated more in the background, by perhaps slightly modifying the illumination of the AR space according to the level of activity displayed by another user, again building on concepts of ambient displays as seen in the work by Ishii et al. [40].

No matter which kind of mechanic is introduced, it is important to strike the correct balance in how much information is provided, and in what way. Here we might consider the design tension described by Gross [30, p. 453], which describes how awareness information can both aid in improving coordination and communication between users, but might also disrupt their activities if provided in abundance [30, p. 453]. While dedicated awareness mechanisms would likely reduce many of the above issues, in two of the interviews the participants also requested the addition of more privacy-oriented features, such as the ability to mute other participants contributions, or to preview your own work before selectively introducing them into the shared space. Inclusion of features such as these would give users a greater control and influence over the interaction, and would also likely reduce the scope of several of the issues described in this section. However, this would also move the interaction away from the initial concept and vision, significantly changing how the experience is used. The interaction is by design meant to provoke uncompromising cooperative interaction. A central part of the experience is having everything you do become part of a shared space. Introducing ex. privacy oriented mechanisms therefore represents a significant shift in perspective, and goes against what I have intended with my design.

However, it is evident from the findings that the experience would likely benefit from more strongly incorporating dedicated awareness mechanisms as part of the design. I had initially assumed that the nature of the experience (i.e. musical and creative improvisation) would require less focus on supporting cooperation awareness, as spontaneity and surprise were part of what made it interesting. Furthermore, as described in section 4.4 I had been somewhat aware of the potential for experiencing discomfort and disruption as a result of sudden sonic stimuli, which had helped shape the character of the sound during the design process. The findings however paint a different picture, and indicate a need for support cooperation awareness even in ludic experiences such as the one provided by the prototype.

6.3.3 The Context of Musical Collaboration

In discussing the experience during the interview, several participants expressed interest in expanding the prototype to include to other contexts beyond the currently used ”same time, same space”.

During the first interview, both participants expressed interest in engaging in synchronous *remote* cooperation within a shared augmented reality space. This was rooted in their perspective as musicians, and they played with the idea of using the AR space as a place to remotely engage in virtual jam sessions together. The idea was that by being able to instantly create a virtual jam space in AR, you would reduce the need for organizing and planning, as users could simply check in and out whenever they were free. While this too diverges significantly from the vision of the prototype, it shows how changing the context might enable other forms of interaction, perhaps moving the interaction closer to being a tool for musicians. On a slightly different note, during the third interview participants played around with the idea of extending the interaction to include the context of "different time, same space". This idea was more closely related to the current version of the prototype, and was centered around having the possibility to create compositions for others to find and enjoy in AR space. I see this concept as more in tune with the design concept, and the notion of creating a ludic and playful musical experience. Nonetheless, while participants played around with the idea of extending the interaction to other contexts, these feelings were presented more as "what if" scenarios, discussed more as visions of the future applications of AR, somewhat beyond the current scope of possibilities.

6.4 Research through Design as a Methodology for the Project Work.

In this section I discuss the project methodology, and assess both my use of methods as well as the potential for knowledge generation as a result of the methodology, as I described in section 3.5.2.

6.4.1 My Approach to Research through Design

The work performed for this thesis has been varied, challenging and at times chaotic, requiring me to simultaneously perform research, while developing prototypes and planning field deployments all at the same time. As I have described in chapter 3, I selected Research through Design as the methodology for this thesis as it allowed me to employ an explorative and concept driven approach, while also fitting well with the project research questions. In putting the methodology into practice, I have employed the five-step approach to RtD presented by Zimmerman & Forlizzi [87, p. 184] as my anchor, which has connected everything together by providing a structural foundation on which to construct the project. I also felt that this approach was beneficial in that it separated the work into clearly separate phases, which helped reassure me that I was performing the right activities at any given time. Furthermore, guided by the RtD framework proposed by Koskinen et al. [50], and inspired by the work of Ljungblad [53], I carried out a field deployment to test and evaluate the prototype artifact. Looking back at the process I consider the RtD approach to have been beneficial in several ways. Firstly, using RtD as my methodology provided me with a significant amount of flexibility and freedom in the design process, which I saw as playing a central part in allowing for new and creative approaches in the design process. This can be seen in ex. my choice to employ video prototyping as a way of evolving my idea and concept, as detailed in section 4.2. As shown in chapter 4, the various iterations of design evolved the design concept from a relatively simple idea to its manifestation in the final integration prototype, the application Petals. Secondly, the use of RtD as a methodology allowed me to direct a lot of focus towards the prototyping process itself, enabling me to realize the artifact I requite to enable exploration of the research questions in my envisioned way - through the construction and deployment of artifacts. This is perhaps best summed up by Carcani et al. who state that "*...by making things and placing them into the world, RtD can change the current state, creating new situations and new practices for anthropologist and researchers to investigate*" [12, p. 270]. While I enjoyed the level of freedom provided, it was also at times challenging to navigate within the methodology. Compared to many other methodologies, RtD is often regarded as less mature, even being described by some as still being in its formative stage [76]. As this was my first brush with RtD, this project required me to familiarize myself with a lot of literature, some of which gave slightly different approaches as to how it should be done. However, this problem is arguably not unique to RtD, and would likely have been a concern to some degree no matter which methodology I had chosen.

Despite meticulous and detailed planning, there were several tacit and silent processes taking place throughout this work, some which might not be immediately discernible in the end result, but which nonetheless had an important role in shaping the prototype. By having employed the RtD methodology, it is my hope that these processes might be illuminated, to produce explicit knowledge and unveil the work in its entirety.

6.4.2 Documenting Design Knowledge

As discussed in chapter 3, one of the most central concepts of RtD work is the act of documenting the design work performed to develop a given system or artifact. Within this project, I have intended to document and communicate this work mainly through the contents of chapter 4. My intention in writing the contents therein has been to provide detailed insight into each part of the design process, from the definition of the initial design concept, through the various iterations of prototyping, to the finalization of the prototype. This process has been verbosely described and annotated with relevant theory, technical information and my own thoughts and reflections along the way. In other words, embedded in the contents of chapter 4 is a detailed description of how I performed the design and development of the final prototype artifact - the application *Petals*. Documenting and communicating this process in a coherent and understandable way has been a difficult task, as the creative process is not always straightforward and obvious to outside observers. In writing this chapter, the paper on the *Musical Moves* system by Carcani et al. [12] has functioned as a touchstone with regards to the tone and style of writing, providing me with clear guidance on how to formulate and communicate my work.

In writing chapter 4, I have constantly strived to highlight the ideas I considered most essential within the project. I wanted to document the central parts of the design and development process in such a way that the focus was not only on technical details, aiming to also disclose the high-level concepts and ideas behind the implementation. The details of development kits and programming languages are constantly changing, but I see the bigger ideas and concepts as potentially having a significantly longer shelf life. As an example, the description of how I implemented the shared augmented reality functionality in section 4.5.9 should hopefully be somewhat applicable to other similar solutions, outside of the specific API(s) used in my implementation. However, I have also chosen to include some more technical information such as in section 4.5.8, where I provide insight into how one might create a custom build of pd-for-android to include an external. This was something I had to learn as I went along, and by including it in this chapter I have both provided insight into my process, while also disseminating this knowledge further.

In the end, it is up to the reader to consider whether I have been successful in producing and communicating coherent and valuable knowledge. However, it could be seen as contributing knowledge in presenting a descriptive and detailed narrative of how an AR based prototype might be designed and developed, from initial concept to final artifact. As a testimonial to the value of using existing work as a guide, I point to how I utilized the work by Ljungblad[53] as a scaffolding for my own RtD process, as described in section 3.1.2.

6.4.3 What Did the Prototypes Prototype?

The use of a dedicated framework [36] for prototyping was intended to clarify and communicate the purpose of the various individual prototypes developed throughout the project. By presenting and evaluating each minor prototype according to the dimensions within this model, I intended to explicate its purpose and place in the larger project context outside of the specific prototype instance. Furthermore, this model was intended to be of use not only to myself, but also to the reader of this work, as it helps examine my reason for developing the specific prototype. In hindsight I consider this to have been a useful tool, particularly when looking back on the development process as a whole. As discussed in section 6.1.1, the prototyping framework assisted in illustrating clearly the lack of attention given to the dimension of *role* in the final integration prototype.

6.4.4 Reflections on the Field Deployment

As my work adopted the field approach as defined by Koskinen et al. [50], the executive research performed was performed by way of field deployment. As previously mentioned, I chose this on the basis of my research questions, as I saw the field approach as most fitting to gather the data I needed to explore the problem space. Looking back at how my deployment was performed, it might seem at times less comparable to other typical field deployments such as the work of Ljungblad [53], where the artifact was provided to participants for them to explore and use freely over a certain time. My deployments were shorter, more focused, and the conditions were to a certain extent controlled by myself (e.g. devices used, time of day, physical location). However, I would argue that this was required in order to test this prototype, as organic deployment would be difficult due to several elements of its design. Firstly, the prototype was somewhat tailored for use on the deployment devices (Moto G5S Plus), as anecdotal testing on emulators or other devices showed limited compatibility. Secondly, the intended context of use required simultaneous cooperation between two participants, a scenario which would have to be arranged somewhat anyway. Lastly, while the prototype generally works as intended, errors are not uncommon, and my presence is required to a certain degree in case it breaks down during use.

In general, I consider the deployments to have been successful with only minor errors occurring, none of which appeared to have a significant effect on the general quality of the session.

6.4.5 Reflections on the Technology Choices for Prototype Development

This section provides some reflections on the various technology choices made for the prototype, as described in section 4.3.

Choice of 3D Rendering Engine

Looking back, selecting Sceneform as the rendering engine for the prototype brought with it some significant limitations that were not obvious before prototype development begun. While Sceneform performed well, and produced the necessary graphical fidelity required for the prototype, use of the more mature engine Unity[1] should have been considered more strongly as it would have enabled creation of more complex (and possibly more immersive) visuals. In the process of composing the technology stack I favored Sceneform due to its native compatibility with ARCore and Android, and its API being accessible in a familiar programming language (Java or Kotlin). The latter being important as to enable rapid and immediate prototyping, with the goal of maximizing productivity and reducing development time. However, with Unity being a very popular language there are plenty of good educational resources available (e.g. tutorials and literature), making it easy for a novice user to quickly produce results. Furthermore, during the project I discovered that Unity was used for the development of the application "Bloom: 10 Worlds"[2], which demonstrates its capabilities in delivering high quality audiovisual experiences.

[1] `https://unity3d.com/`

[2] `http://kitmonsters.com/blog/new-bloom-app-brian-eno-and-peter-chilvers`

Chapter 7

Conclusion and Future Work

> *"It would be egotistical of me to say my ideas were my own. I believe that there is a network of ideas, and the ideas come through me."*
>
> \- Robert Moog

7.1 Conclusion

In this thesis I have explored how we might design cooperative musical interactions for mobile augmented reality. Furthermore, I have examined how the use of binaural audio affects users experienced immersion in augmented reality. In addition, I have shed light on which awareness mechanisms might be needed for ludic interactions in a shared augmented reality space. Throughout the project, my work has been guided by the methodology Research through Design. My design and development process has been driven by its explorative and iterative approach, and its focus on constructing artifacts to probe the problem space. This resulted in the construction of a high-fidelity prototype in the form of the Android application *Petals*. This prototype was the result of an iterative design process guided by research and some key exemplars of related work, and represents current final manifestation of my design concept. As mobile augmented reality is emerging as an important computing platform for new experiences to take place, I was interested in exploring the ways in which it could mediate ludic and cooperative interactions with music between people. In order to shed light on this, I performed a field deployment of the prototype application, evaluating it with six different people in three different environments, leading to the collection of rich and detailed data in the form of field observations and interviews.

The findings suggest that the prototype was successful in providing a novel and immersive way to cooperatively interact with music and sound in augmented reality space. The prototype experience was effective in stimulating a sense of playfulness in the users, and enabled them to cooperatively take part in ludic interaction with music. Participants reported experiencing a significant level of immersion when interacting with the prototype, despite the limited screen size and the mobile form factor. This sensation was reported as being enhanced by the use of audio, and was connected to the use of binaural filtering. Furthermore, the findings revealed that use of binaural audio might be an effective way of provoking active participation from users when designing immersive augmented reality experiences. During the deployment, the spatial properties of the sound helped stimulate movement in the participants, and encouraged them to move about and explore the environment. However, the field deployments also revealed some considerable challenges in the implementation, mainly pertaining to lack of programmability and control, as well as a lack of support for awareness mechanisms. While the experience was generally perceived as enjoyable, participants repeatedly reported feeling creatively constrained by what they experienced as inadequate musical complexity within the prototype. This lack of programmability was identified as a central challenge in the implementation, and was brought up as the main drawback of the application. Furthermore, participants also experienced the prototype as lacking in mechanisms to more strongly support awareness in the augmented reality space. To probe for potential solutions to these challenges, the findings were assessed according to the design tensions of privacy and conventions. While this assessment was inconclusive, the awareness related challenges should be addressed in future work, to enable the creation of more seamless cooperative augmented reality experiences. To further support the immersiveness of the experience, It is suggested that this information is communicated in the background, by interweaving the awareness information into the augmented reality experience. Finally, in providing detailed insight into my design process as a part of this thesis, I aim to contribute knowledge not only as a result of the findings presented above, but also as a result of the documented work seen in the context of Research through Design. Through this I hope to inspire and motive others who might be interested in applying a similar methodological approach to their own work, while also contributing knowledge through the descriptive accounts of my design work.

In conclusion, the work performed for this thesis has shown how mobile augmented reality can be employed to enable cooperative interaction with music and sound. Furthermore, it has illustrated the effects of binaural audio on immersion in mobile augmented reality experiences, and has put forth some implications for designing towards immersion and active participation in mobile augmented reality. While some challenges were identified in the implementation, the findings pointed towards the design concept being generally well regarded in users, with challenges being connected mainly to details of the implementation.

7.2 Limitations

This section briefly presents and discusses some potential limitations of the study, as well as some technical limitations that might have affected the deployment of the prototype.

7.2.1 Limitations of the Study

The following aspects have been identified as possible limitations with regards to the study.

Number of Participants

The number of participants recruited for the field deployment could be considered a limitation. During the field deployment process, I recruited 6 people in total. The small number of participants could therefore constrain the generalizability of the findings.

Incentive for Participation

All participants who took part in the study were rewarded with a digital gift voucher for a cinema of their choosing, with a value of 150 NOK. Compensating participants for their attendance could potentially have impacted their attitude towards the study, and influenced them to react more positively than if not rewarded. This should be considered a possible limitation.

Existing Connections to Participants

Four of the six participants who took part in the field deployment were acquaintance of mine from before. This could potentially have affected their attitude towards the study, and should be regarded as a potential limitation.

7.2.2 Technological Limitations

In considering possible limitations regarding technological factors, the following aspects have been identified.

Loss of Tracking in ARCore

As previously mentioned in section 5.2.2, the challenges related to awareness might have been affected by the technical implementation of the ARCore platform, in which content rendered in AR space occasionally becomes deactivated and disappears if the device loses tracking of a surface. Similarly, users might therefore experience an imbalance in their awareness of the other participants activities if their devices have a significantly different understanding of the playspace. In other words, if they each have mapped different parts of the room, they might not see and hear each others work. However, while this might occur it does not negate the previously discussed awareness challenges, and this limitation should be seen as an additional design challenge rather than an explanation.

The Implementation of Binaural Audio

As described in section 4.5.8, the binaural external used in the prototype is limited technically in that it only positions audio according to azimuth, neglecting the dimension of elevation entirely. This means that the sound of a node is positioned in a given direction within the playspace, but remains at the same vertical position at all times. Prior to the deployment, I was curious as to how this would be perceived by the users during testing, and how the lack of elevation might affect immersiveness. However, during the interviews participants reported no dissonance related to the lack of change in sound elevation, and were instead commending the use of directional positioning. This can be explained by looking at existing work which shows that humans are much better at estimating the direction of a sound source compared to the elevation or distance [75]. It is therefore possible that users simply did not sense the lack of elevation in the sound, and were therefore unaware of the limitation. Furthermore, in using the prototype participants were likely interacting with multiple sound sources at once, which could have made it difficult to keep track of singular points of sound in space. It is therefore plausible that these two factors reduced the overall negative impact of not incorporating elevation into the experience.

Another limitation can be found in the implementation of binaural audio within the prototype code. While the user might continuously be moving around the playspace during use, the direction of sound was only changed whenever a node was triggered to play, and did not update until the next time it played. In other words, if the user drastically changed their position while a sound was playing, they would not perceive any changes to the position of the sound until the next time it played. This was done to limit the amount of processing required during use, as continuously updating the sound of every node in a playspace during use would increase the resource use significantly, potentially introducing disruptive latency and lag to the experience (similar to the issues described in section 4.5.8), ultimately reducing the immersiveness experienced by users during field deployment. As indicated by the findings presented in chapter 5, users did not appear to recognize this limitation during the deployment, reducing the overall negative impact on the experience. Finally, as noted in chapter 5 there was one participant who reported not experiencing the binaural audio at all. It is likely that this could have been the result of a potential software issue, such as the binaural audio external crashing or malfunctioning during use. It is unlikely that hardware was at fault, as the pair of headphones used by the participant (Sennheiser HD-25 II) are of a sufficient quality to produce the intended effect. Furthermore, these have also been checked and confirmed to be working at a later point, ruling out hardware malfunction. Despite several tries, I have not been successful in reproducing this error.

7.3 Future Work

This section describes some possible avenues of future work for experiences and interactions of a similar nature to the prototype application. These suggestions are rooted both in the research findings, as well as unimplemented ideas and concepts that emerged during the design process.

7.3.1 Cooperative Musical Interaction in Other Contexts

The prototype developed for this project was designed to operate only within a single context in the CSCW perspective, namely in the dimension of *same time, same space*. However, the findings indicate that users are interested in experiencing a similar interaction in other contexts. The experience could be extended to incorporating a remote collaboration aspect, by exploring the dimension of *same time, different space*. Furthermore, one might also want to explore the dimension of *different time, same space*, by allowing for the creation of persistent compositions that can be explored and revisited over time.

7.3.2 Probing the Use of Binaural Audio to Provoke Active Participation in Mobile Augmented Reality

The findings indicated that binaural audio was a significant contributor in creating and supporting immersion. Further research could explore the use of binaural audio in less musical and ludic context, to probe how binaural audio might be utilized outside of the prototype interaction. In addition, the results pointed towards binaural audio being an effective tool in provoking movement and active participation in participants. Further research could probe this problem space by designing and deploying a prototype more directly tailored towards exploring this phenomenon.

7.3.3 Cooperative Musical Interaction With an Extended Range of Sounds

As indicated by the project findings, participants requested an expanded range of sounds within the application. Further work could explore creating a similar interaction - or adapting the prototype experience - to include at minimum one additional sound. Extending the range of sounds could enable further inquiries into both the experience itself as a space for ludic cooperation, as well as its effects on aspects of awareness.

7.3.4 Repeating the Process Once More

A key element in the RtD model employed in the project is repeatedly returning to explore the same problem space [87]. This is directly reflected in step five in the model [87, p. 186], which is to repeat the five steps once more, incorporating everything that was learned through the first iteration. In the context of this project, this could be valuable to both assess the challenges discovered during the field deployment, and to further evolve the design concept into an even more completed form.

Bibliography

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

[2] Jeffrey Bardzell, Shaowen Bardzell, Peter Dalsgaard, Shad Gros, and Kim Halskov. Documenting the research through design process. In *DIS '16*, pages 96–107. Association for Computing Machinery, 2016.

[3] Lyman Frank Baum. *The Master Key.* Bowen-Merrill, 1901.

[4] Benjamin B Bederson. Audio augmented reality: a prototype automated tour guide. In *Conference companion on Human factors in computing systems*, pages 210–211. ACM, 1995.

[5] Nicola Bernardini and Giovanni De Poli. The sound and music computing field: present and future. *Journal of New Music Research*, 36(3):143–148, 2007.

[6] Mark Billinghurst, Hirokazu Kato, and Ivan Poupyrev. The magicbook - moving seamlessly between reality and virtuality. *IEEE Computer Graphics and applications*, 21(3):6–8, 2001.

[7] Ann Blandford. Semi-structured qualitative studies. *The Encyclopedia of Human-Computer Interaction, 2nd Ed.*, 2019. Retrieved at: `https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/semi-structured-qualitative-studies`.

[8] Joachim Bondo, David Barnard, Dan Burcaw, Tim Novikoff, Craig Kemper, Chris Parrish, Keith Peters, Jurgen Siebert, Eddie Wilson, Brian Wilson, et al. *iPhone user interface design projects.* Apress, 2010.

[9] Peter Brinkmann. *Making Musical Apps: Real-Time Audio Synthesis on Android and iOS.* O'Reilly Media, 1 edition, 2012.

[10] Wolfgang Broll, Eckhard Meier, and Thomas Schardt. The virtual round table - a collaborative augmented multi-user environment. In *Proceedings of the Third International Conference on Collaborative Virtual Environments*, CVE '00, pages 39–45, New York, NY, USA, 2000. ACM.

[11] Capgemini Research Institute. Augmented and virtual reality in operations: A guide for investment, 2018. Retrieved at: `https://www.capgemini.com/wp-content/uploads/2018/09/AR-VR-in-Operations1.pdf`.

[12] Klaudia Çarçani, Veronica Wachek Hansen, and Harald Maartmann-Moe. Exploring technology use in dance performances. In *International Conference on Human-Computer Interaction*, pages 268–280. Springer, 2018.

[13] A. Case and A. Day. *Designing with Sound: Principles and Patterns for Mixed Environments*. O'Reilly Media, Incorporated, 2018.

[14] Thomas Caudell and David Mizell. Augmented reality: An application of heads-up display technology to manual manufacturing processes. In *Proceedings of the Twenty-Fifth Hawaii International Conference on System Sciences*, volume 2, pages 659 – 669 vol.2, 02 1992.

[15] Kathy Charmaz. *Constructing grounded theory*. Sage, 2014.

[16] Corey I. Cheng and Gregory H. Wakefield. Introduction to head-related transfer functions (hrtfs): Representations of hrtfs in time, frequency, and space. *J. Audio Eng. Soc*, 49(4):231–249, 2001.

[17] B.W.C. Chung. *Multimedia Programming with Pure Data: A comprehensive guide for digital artists for creating rich interactive multimedia applications using Pure Data*. Community experience distilled. Packt Publishing, 2013.

[18] E.F. Churchill, D.N. Snowdon, and A.J. Munro. *Collaborative Virtual Environments: Digital Places and Spaces for Interaction*. Computer Supported Cooperative Work. Springer London, 2012.

[19] C. F. Codella, R. Jalili, L. Koved, and J. B. Lewis. A toolkit for developing multi-user, distributed virtual environments. In *Proceedings of IEEE Virtual Reality Annual International Symposium*, pages 401–407, Sept 1993.

[20] Perry Cook. Principles for designing computer music controllers. In *Proceedings of the 2001 conference on New interfaces for musical expression*, pages 1–4. National University of Singapore, 2001.

[21] Paul Dourish and Victoria Bellotti. Awareness and coordination in shared workspaces. In *Proceedings of the 1992 ACM conference on Computer-supported cooperative work*, pages 107–114. ACM, 1992.

[22] Clarence A. Ellis, Simon J. Gibbs, and Gail Rein. Groupware: Some issues and experiences. *Commun. ACM*, 34(1):39–58, January 1991.

[23] Daniel Fallman and Erik Stolterman. Establishing criteria of rigour and relevance in interaction design research. *Digital Creativity*, 21(4):265–272, 2010.

[24] U. Flick. *An Introduction to Qualitative Research*. SAGE Publications, 2009.

[25] Norwegian National Committees for Research Ethics. General guidelines for research ethics, 2014. Retrieved at: `https://www.etikkom.no/forskningsetiske-retningslinjer/Generelle-forskningsetiske-retningslinjer/general-guidelines-for-research-ethics/`.

[26] C. Frayling and Royal College of Art (Great Britain). *Research in Art and Design*. Royal College of Art research papers. Royal College of Art, 1993.

[27] William Gaver. What should we expect from research through design? In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 937–946. ACM, 2012.

[28] Google. Introduction to AR & ARCore - Augmented Reality Design Guidelines, 2019. Retrieved at: `https://designguidelines.withgoogle.com/ar-design/`.

[29] Mark Green and Dani Beaubien. Minimal reality toolkit version 1.2 programmer's manual. *Department of Computing Science, University of Alberta, Edmonton, Alberta*, 1992.

[30] Tom Gross. Supporting effortless coordination: 25 years of awareness research. *Computer Supported Cooperative Work (CSCW)*, 22(4-6):425–474, 2013.

[31] Jonathan Grudin. Computer-supported cooperative work: History and focus. *Computer*, 27(5):19–26, 1994.

[32] Maria Håkansson, Lalya Gaye, Sara Ljungblad, and Lars Erik Holmquist. More than meets the eye: an exploratory study of context photography. In *Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles*, pages 262–271. ACM, 2006.

[33] Anders Henrysson and Mark Ollila. Augmented reality on smartphones. In *Augmented Reality Toolkit Workshop, 2003. IEEE International*, pages 27–28. IEEE, 2003.

[34] Hal Hodson. Google's ingress game is a gold mine for augmented reality, 2012. Retrieved at: `https://www.sciencedirect.com/science/article/pii/S0262407912630589`.

[35] Kristina Höök, Peter Dalsgaard, Stuart Reeves, Jeffrey Bardzell, Jonas Löwgren, Erik Stolterman, and Yvonne Rogers. Knowledge production in interaction design. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '15, pages 2429–2432, New York, NY, USA, 2015. ACM.

[36] Stephanie Houde and Charles Hill. What do prototypes prototype? In *Handbook of Human-Computer Interaction (Second Edition)*, pages 367–381. Elsevier, 1997.

[37] IKEA. Ikea launches ikea place, a new app that allows people to virtually place furniture in their home. Retrieved at: `https://www.ikea.com/us/en/about_ikea/newsitem/091217_IKEA_Launches_IKEA_Place`.

[38] Apple Inc. Human interface guidelines: Augmented reality. Retrieved at: `https://developer.apple.com/design/human-interface-guidelines/ios/system-capabilities/augmented-reality/`.

[39] Apple Inc. Prototyping for AR, 2018. Retrieved at: `https://developer.apple.com/videos/play/wwdc2018/808/`.

[40] Hiroshi Ishii, Craig Wisneski, Scott Brave, Andrew Dahley, Matt Gorbet, Brygg Ullmer, and Paul Yarin. ambientroom: integrating ambient media with architectural space. In *CHI 98 conference summary on Human factors in computing systems*, pages 173–174. ACM, 1998.

[41] Bruce L. Jacob. Algorithmic composition as a model of creativity. *Organised Sound*, 1(3):157–165, 1996.

[42] Robert Johansen. *GroupWare: Computer Support for Business Teams.* The Free Press, New York, NY, USA, 1988.

[43] Peter Johnson-Lenz and Trudy Johnson-Lenz. Groupware: Coining and defining it. *SIGGROUP Bull.*, 19(2):34–, August 1998.

[44] Kent Jolly. Usage of pd in spore and darkspore. In *PureData Convention*, 2011.

[45] Brett Jones, Rajinder Sodhi, Michael Murdock, Ravish Mehra, Hrvoje Benko, Andrew Wilson, Eyal Ofek, Blair MacIntyre, Nikunj Raghuvanshi, and Lior Shapira. Roomalive: Magical experiences enabled by scalable, adaptive projector-camera units. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*, UIST '14, pages 637–644, New York, NY, USA, 2014. ACM.

[46] Alisa Bokulich (eds.) Juliet Floyd. *Philosophical Explorations of the Legacy of Alan Turing: Turing 100.* Boston Studies in the Philosophy and History of Science 324. Springer International Publishing, 1 edition, 2017.

[47] H. Kato and M. Billinghurst. Marker tracking and hmd calibration for a video-based augmented reality conferencing system. In *Proceedings 2nd IEEE and ACM International Workshop on Augmented Reality (IWAR'99)*, pages 85–94, Oct 1999.

[48] Hirokazu Kato and Mark Billinghurst. Marker tracking and hmd calibration for a video-based augmented reality conferencing system. In *Augmented Reality, 1999.(IWAR'99) Proceedings. 2nd IEEE and ACM International Workshop on*, pages 85–94. IEEE, 1999.

[49] Chris Kiefer and Cecile Chevalier. Towards new modes of collective musical expression through audio augmented reality. In Thomas Martin Luke Dahl, Douglas Bowman, editor, *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 25–28, Blacksburg, Virginia, USA, June 2018. Virginia Tech.

[50] Ilpo Koskinen, John Zimmerman, Thomas Binder, Johan Redstrom, and Stephan Wensveen. *Design research through practice: From the lab, field, and showroom.* Elsevier, 2011.

[51] S. Kvale, S. Brinkmann, T.M. Anderssen, and J. Rygge. *Det kvalitative forskningsintervju.* Gyldendal akademisk, 2015.

[52] Micheal Lanham. *Augmented Reality Game Development.* Packt Publishing Ltd, 2017.

[53] Sara Ljungblad. Designing for new photographic experiences: how the lomographic practice informed context photography. In *Proceedings of the 2007 conference on Designing pleasurable products and interfaces*, pages 357–374. ACM, 2007.

[54] Waves Audio ltd. Ambisonics explained: A guide for sound engineers. Retrieved at: `https://www.waves.com/ambisonics-explained-guide-for-sound-engineers`.

[55] Gloria Mark. Conventions and commitments in distributed cscw groups. *Computer Supported Cooperative Work (CSCW)*, 11(3-4):349–387, 2002.

[56] James McDermott, Toby Gifford, Anders Bouwer, and Mark Wagy. Should music interaction be easy? In *Music and human-computer interaction*, pages 29–47. Springer, 2013.

[57] Paul Milgram, Haruo Takemura, Akira Utsumi, and Fumio Kishino. Augmented reality: A class of displays on the reality-virtuality continuum. In *Telemanipulator and telepresence technologies*, volume 2351, pages 282–293. International Society for Optics and Photonics, 1995.

[58] Elizabeth D Mynatt, Maribeth Back, Roy Want, Michael Baer, and Jason B Ellis. Designing audio aura. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 566–573. ACM Press/Addison-Wesley Publishing Co., 1998.

[59] Elizabeth D Mynatt, Maribeth Back, Roy Want, and Ron Frederick. Audio aura: Light-weight audio augmented reality. In *Proceedings of the 10th annual ACM symposium on User interface software and technology*, pages 211–212. ACM, 1997.

[60] Kristine L. Nowak and Frank Biocca. The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, 12(5):481–494, 2003.

[61] Tomislav Pejsa, Julian Kantor, Hrvoje Benko, Eyal Ofek, and Andrew Wilson. Room2room: Enabling life-size telepresence in a projected augmented reality environment. In *Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing*, pages 1716–1725. ACM, 2016.

[62] Ivan Poupyrev, Rodney Berry, Mark Billinghurst, Hirokazu Kato, Keiko Nakao, Lewis Baldwin, Jun Kurumisawa, et al. Augmented reality interface for electronic music performance. In *Proceedings of HCI*, pages 805–808, 2001.

[63] Ivan Poupyrev, Michael J. Lyons, Sidney Fels, and Tina Blaine (Bean). New interfaces for musical expression. In *CHI '01 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '01, pages 491–492, New York, NY, USA, 2001. ACM.

[64] Miller Puckette et al. Pure data: another integrated computer music environment. *Proceedings of the second intercollege computer music concerts*, pages 37–41, 1996.

[65] Jiten Rama and Judith Bishop. A survey and comparison of cscw groupware applications. In *Proceedings of the 2006 annual research conference of the South African institute of computer scientists and information technologists on IT research in developing countries*, pages 198–205. South African Institute for Computer Scientists and Information Technologists, 2006.

[66] Jun Rekimoto and Katashi Nagao. The world through the computer: Computer augmented interaction with real world environments. In *Proceedings of the 8th Annual ACM Symposium on User Interface and Software Technology*, UIST '95, pages 29–36, New York, NY, USA, 1995. ACM.

[67] Interaction Design Foundation. Augmented reality – the past, the present and the future, Aug 2018. Retrieved at: `https://www.interaction-design.org/literature/article/augmented-reality-the-past-the-present-and-the-future`.

[68] Sound and Music Computing Network. Smc roadmap: Challenges and strategies. Retrieved at: `http://www.smcnetwork.org/roadmap#challenges`.

[69] Louis B Rosenberg. Virtual fixtures: Perceptual tools for telerobotic manipulation. In *Virtual Reality Annual International Symposium, 1993., 1993 IEEE*, pages 76–82. IEEE, 1993.

[70] Dariusz Rumiński. An experimental study of spatial sound usefulness in searching and navigating through ar environments. *Virtual Reality*, 19(3-4):223–233, 2015.

[71] Kjeld Schmidt. The problem with 'awareness': Introductory remarks on 'awareness in cscw'. *Comput. Supported Coop. Work*, 11:285–298, 09 2002.

[72] Kjeld Schmidt and Liam Bannon. Constructing cscw: The first quarter century. *Comput. Supported Coop. Work*, 22(4-6):345–372, August 2013.

[73] Nina Siegal. Brian eno wants to take you 'inside the music'. *The New York Times*, Mar 2018. Retrieved at: `https://www.nytimes.com/2018/03/19/arts/music/brian-eno-immersive-music-bloom.html`.

[74] Katie A. Siek, Gillian R. Hayes, Mark W. Newman, and John C. Tang. Field deployments: Knowing from using in context. In Judith S. Olson and Wendy A. Kellogg, editors, *Ways of Knowing in HCI*, pages 119–142. Springer New York, New York, NY, 2014.

[75] Jaka Sodnik, Saso Tomazic, Raphael Grasset, Andreas Duenser, and Mark Billinghurst. Spatial sound localization in an augmented reality environment. In *Proceedings of the 18th Australia conference on computer-human interaction: design: activities, artefacts and environments*, pages 111–118. ACM, 2006.

[76] Pieter Stappers and Elisa Giaccardi. Research through design. *The Interaction Design Foundation*, 2019. Retrieved at: `https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/research-through-design`.

[77] Pieter Jan Stappers, F Sleeswijk Visser, and AI Keller. The role of prototypes and frameworks for structuring explorations by research through design. *The Routledge Companion to Design Research*, pages 163–174, 2014.

[78] Dan Stowell and Alex McLean. Live music-making: A rich open task requires a rich open interface. In *Music and human-computer interaction*, pages 139–152. Springer, 2013.

[79] Ben Swift. Chasing a feeling: Experience in computer supported jamming. In *Music and human-computer interaction*, pages 85–99. Springer, 2013.

[80] Zsolt Szalavári, Dieter Schmalstieg, Anton Fuhrmann, and Michael Gervautz. Studierstube: An environment for collaboration in augmented reality. *Virtual Reality*, 3(1):37–48, 1998.

[81] Victoria Turk. Brian eno's ar experience lets you play synth in mid-air, Feb 2018. Retrieved at: `https://www.wired.co.uk/article/brian-eno-amsterdam-bloom-open-space-installation-reflection`.

[82] J.P. van Arnhem, C. Elliott, and M. Rose. *Augmented and Virtual Reality in Libraries*. LITA Guides. Rowman & Littlefield Publishers, 2018.

[83] Pierre Wellner. Interacting with paper on the digitaldesk. *Commun. ACM*, 36(7):87–96, July 1993.

[84] Lucy Yardley. Dilemmas in qualitative research. *Psychology & Health - PSYCHOL HEALTH*, 15:215–228, 03 2000.

[85] Noah Yoo. Fields wants to be the augmented reality app for experimental music fans and creators alike, May 2018. Retrieved at: `https://pitchfork.com/thepitch/fields-wants-to-be-the-augmented-reality-app-for-experimental-music-fans-and-creators-alike/`.

[86] Zhiying Zhou, Adrian David Cheok, Xubo Yang, and Yan Qiu. An experimental study on the role of 3d sound in augmented reality environment. *Interacting with Computers*, 16(6):1043–1068, 2004.

[87] John Zimmerman and Jodi Forlizzi. Research through design in HCI. In *Ways of Knowing in HCI*, pages 167–189. Springer, 2014.

Appendix A

NSD Application

The application for my project sent to Norwegian Centre for Research Data.

NSD NORSK SENTER FOR FORSKNINGSDATA

Meldeskjema 325229

Sist oppdatert

18.03.2019

Hvilke personopplysninger skal du behandle?

- Navn (også ved signatur/samtykke)
- Lydopptak av personer

Type opplysninger

Skal du behandle særlige kategorier personopplysninger eller personopplysninger om straffedommer eller lovovertredelser?

Nei

Prosjektinformasjon

Prosjekttittel

Collaborative Musical Expression in Augmented Reality

Prosjektbeskrivelse

Prosjektet utforsker hvordan mennesker skaper musikk, og samarbeider i en "augmented reality"-kontekst.
Foreløpige forskningsspørsmål som utforskes er

"How does the spatialization of music affect users creative expression in augmented reality?"
"How can mobile augmented reality mediate collaborative interactions between users?".

I prosjektet ønsker jeg å gjennomføre en kvalitativ studie ved hjelp av en prototype-applikasjon til Android telefoner som jeg har utviklet til dette prosjektet. Denne applikasjonen lar brukere improvisere og interagere med musikk i et delt virtuelt "augmented reality" rom. I studien ønsker jeg å gjennomføre testing av applikasjonen med 8 til 10 personer, satt i grupper på 2. Etterfulgt av testing skal det gjennomføres et semi-strukturert intervju, der temaer som "augmented reality", "samarbeid i virtuelle rom" og "musikalsk improvisasjon" diskuteres. For å få tilstrekkelig datamateriale ønsker jeg å ta lydopptak av intervjuene.

Fagfelt

Teknologi

Begrunn behovet for å behandle personopplysningene

Opplysningene som jeg samler inn gjennom gruppeintervjuene er kun relatert til temaer som utforskes i prosjektet. I tillegg til observasjon av testing ønsker jeg å gjennomføre parvise gruppeintervju. Disse ønsker jeg å gjøre lydopptak av, for å kunne bygge opp et rikt kvalitativt datamateriale som deretter kan fungere som grunnlag for analyse ved hjelp av Grounded Theory metoder.

Ekstern finansiering

Type prosjekt

Studentprosjekt, masterstudium

Kontaktinformasjon, student

Adrian Jensby Sandaker, adrian.j.sandaker@hiof.no, tlf: 95252896

Behandlingsansvar

Behandlingsansvarlig institusjon

Høgskolen i Østfold / Avdeling for informasjonsteknologi

Prosjektansvarlig (vitenskapelig ansatt/veileder eller stipendiat)

Susanne Koch Stigberg, susanne.k.stigberg@hiof.no, tlf: 69608344

Skal behandlingsansvaret deles med andre institusjoner (felles behandlingsansvarlige)?

Nei

Utvalg 1

Beskriv utvalget

Personer med interesse for mobilteknologi, musikk og kreativ uttrykkelse.

Rekruttering eller trekking av utvalget

Utvalget rekrutteres enten fra eget nettverk, eller fra utdanningsinstitusjonen gjennom plakater, direkte kontakt, informasjonsskriv eller innlegg i sosiale medier.

Alder

18 - 73

Inngår det voksne (18 år +) i utvalget som ikke kan samtykke selv?

Nei

Personopplysninger for utvalg 1

- Navn (også ved signatur/samtykke)
- Lydopptak av personer

Hvordan samler du inn data fra utvalg 1?

Gruppeintervju

Grunnlag for å behandle alminnelige kategorier av personopplysninger

Samtykke (art. 6 nr. 1 bokstav a)

Ikke-deltakende observasjon

Grunnlag for å behandle alminnelige kategorier av personopplysninger

Samtykke (art. 6 nr. 1 bokstav a)

Felteksperiment/feltintervensjon

Grunnlag for å behandle alminnelige kategorier av personopplysninger

Samtykke (art. 6 nr. 1 bokstav a)

Informasjon for utvalg 1

Informerer du utvalget om behandlingen av opplysningene?

Ja

Hvordan?

Skriftlig informasjon (papir eller elektronisk)

Tredjepersoner

Skal du behandle personopplysninger om tredjepersoner?

Nei

Dokumentasjon

Hvordan dokumenteres samtykkene?

- Manuelt (papir)

Hvordan kan samtykket trekkes tilbake?

Samtykket kan trekkes tilbake ved å kontakte meg, veilederen min eller personvernansvarlig ved Høgskolen i Østfold. Ved signering av samtykkeskjema gis deltakeren sin egen kopi som inneholder kontaktinformasjon til alle de overnevnte. Deltakere kan velge å ta kontakt via epost eller telefon.

Hvordan kan de registrerte få innsyn, rettet eller slettet opplysninger om seg selv?

De registrerte kan når som helst få innsyn i opplysningene ved å kontakte meg eller veilederen min. Ved signering av samtykkeskjema gis deltakeren sin egen kopi som inneholder relevant kontaktinformasjon.

Totalt antall registrerte i prosjektet

1-99

Tillatelser

Skal du innhente følgende godkjenninger eller tillatelser for prosjektet?

Behandling

Hvor behandles opplysningene?

- Ekstern tjeneste eller nettverk (databehandler)

Hvem behandler/har tilgang til opplysningene?

- Prosjektansvarlig
- Student (studentprosjekt)
- Databehandler

Hvilken databehandler har tilgang til opplysningene?

Nettskjema / Universitetet i Oslo.

Tilgjengeliggjøres opplysningene utenfor EU/EØS til en tredjestat eller internasjonal organisasjon?

Nei

Sikkerhet

Oppbevares personopplysningene atskilt fra øvrige data (kodenøkkel)?

Ja

Hvilke tekniske og fysiske tiltak sikrer personopplysningene?

- Opplysningen krypteres under lagring
- Adgangsbegrensning

Varighet

Prosjektperiode

18.04.2019 - 01.06.2019

Skal data med personopplysninger oppbevares utover prosjektperioden?

Nei, alle data slettes innen prosjektslutt

Vil de registrerte kunne identifiseres (direkte eller indirekte) i oppgave/avhandling/øvrige publikasjoner fra prosjektet?

Nei

Tilleggsopplysninger

Appendix B

NSD Assessment

Attached is the assessment made by the Norwegian Centre for Research Data granting me permission to perform the executive research.

NORSK SENTER FOR FORSKNINGSDATA

NSD sin vurdering

Prosjekttittel

Collaborative Musical Expression in Augmented Reality

Referansenummer

325229

Registrert

18.03.2019 av Adrian Jensby Sandaker - adrian.j.sandaker@hiof.no

Behandlingsansvarlig institusjon

Høgskolen i Østfold / Avdeling for informasjonsteknologi

Prosjektansvarlig (vitenskapelig ansatt/veileder eller stipendiat)

Susanne Koch Stigberg, susanne.k.stigberg@hiof.no, tlf: 69608344

Type prosjekt

Studentprosjekt, masterstudium

Kontaktinformasjon, student

Adrian Jensby Sandaker, adrian.j.sandaker@hiof.no, tlf: 95252896

Prosjektperiode

18.04.2019 - 01.06.2019

Status

11.04.2019 - Vurdert

Vurdering (1)

11.04.2019 - Vurdert

Det er vår vurdering at behandlingen av personopplysninger i prosjektet vil være i samsvar med personvernlovgivningen så fremt den gjennomføres i tråd med det som er dokumentert i meldeskjemaet med vedlegg den 11.04.2019, samt i meldingsdialogen mellom innmelder og NSD. Behandlingen kan starte.

MELD ENDRINGER
Dersom behandlingen av personopplysninger endrer seg, kan det være nødvendig å melde dette til NSD ved å oppdatere meldeskjemaet. På våre nettsider informerer vi om hvilke endringer som må meldes. Vent på svar før endringer gjennomføres.

TYPE OPPLYSNINGER OG VARIGHET
Prosjektet vil behandle alminnelige kategorier av personopplysninger frem til 01.06.2019.

LOVLIG GRUNNLAG
Prosjektet vil innhente samtykke fra de registrerte til behandlingen av personopplysninger. Vår vurdering er at prosjektet legger opp til et samtykke i samsvar med kravene i art. 4 og 7, ved at det er en frivillig, spesifikk, informert og utvetydig bekreftelse som kan dokumenteres, og som den registrerte kan trekke tilbake. Lovlig grunnlag for behandlingen vil dermed være den registrertes samtykke, jf. personvernforordningen art. 6 nr. 1 bokstav a.

PERSONVERNPRINSIPPER
NSD vurderer at den planlagte behandlingen av personopplysninger vil følge prinsippene i personvernforordningen om:

- lovlighet, rettferdighet og åpenhet (art. 5.1 a), ved at de registrerte får tilfredsstillende informasjon om og samtykker til behandlingen
- formålsbegrensning (art. 5.1 b), ved at personopplysninger samles inn for spesifikke, uttrykkelig angitte og berettigede formål, og ikke behandles til nye, uforenlige formål
- dataminimering (art. 5.1 c), ved at det kun behandles opplysninger som er adekvate, relevante og nødvendige for formålet med prosjektet
- lagringsbegrensning (art. 5.1 e), ved at personopplysningene ikke lagres lengre enn nødvendig for å oppfylle formålet

DE REGISTRERTES RETTIGHETER
Så lenge de registrerte kan identifiseres i datamaterialet vil de ha følgende rettigheter: åpenhet (art. 12), informasjon (art. 13), innsyn (art. 15), retting (art. 16), sletting (art. 17), begrensning (art. 18), underretning (art. 19), dataportabilitet (art. 20).

NSD vurderer at informasjonen om behandlingen som de registrerte vil motta oppfyller lovens krav til form og innhold, jf. art. 12.1 og art. 13.

Vi minner om at hvis en registrert tar kontakt om sine rettigheter, har behandlingsansvarlig institusjon plikt til å svare innen en måned.

FØLG DIN INSTITUSJONS RETNINGSLINJER
NSD legger til grunn at behandlingen oppfyller kravene i personvernforordningen om riktighet (art. 5.1 d), integritet og konfidensialitet (art. 5.1. f) og sikkerhet (art. 32).

Nettskjema er databehandler i prosjektet. NSD legger til grunn at behandlingen oppfyller kravene til bruk av databehandler, jf. art 28 og 29.

For å forsikre dere om at kravene oppfylles, må dere følge interne retningslinjer og/eller rådføre dere med behandlingsansvarlig institusjon.

OPPFØLGING AV PROSJEKTET
NSD vil følge opp ved planlagt avslutning for å avklare om behandlingen av personopplysningene er

avsluttet.

Lykke til med prosjektet!

Kontaktperson hos NSD: Elizabeth Blomstervik
Tlf. Personverntjenester: 55 58 21 17 (tast 1)

Appendix C

Consent Form

Attached is the consent form used to seek consent form participant before performing the field deployment. Both Norwegian and English consent forms are included.

Samtykkeskjema til Forskningsprosjekt

Adrian Jensby Sandaker
adrianjs@hiof.no

1 Introduksjon

Denne studien utføres som en del av min masteravhandling for graden Master i Anvendt Informatikk ved Høgskolen i Østfold. Prosjektet utforsker hvordan mobil augmented reality teknologi kan brukes til å muliggjøre kollaborativ skapning av-, og interaksjon med musikk. Videre ser studien også på hvordan mennesker samarbeider i augmented reality med et bredere perspektiv, samt hvordan teknologien støtter opp under mekanismer for awareness og tilstedeværelse i et virtuelt augmented reality rom. For å belyse dises spørsmålene ønsker jeg å utføre en kvalitativ studie ved hjelp av en applikasjon jeg har utviklet som en del av dette prosjektet. Jeg ønsker derfor å rekruttere deltakere som kunne være interessert i å delta i en slik studie. Det er ingen formelle krav til deltakelse utover at du er myndig, at du har kjennskap til moderne smarttelefoner, og at du er i stand til å bruke en smarttelefon med en viss grad av ferdighet. Videre så sees det som positivt hvis du har noen grad av interesse for temaer som teknologi, musikk eller kreativ uttrykkelse i generell forstand. Ettersom studien vil bli utført parvis er det en stor fordel at du er komfortabel med å samarbeide med en potensielt ukjent person i løpet av studiens varighet.

2 Hvem er ansvarlig for studien?

Studien er utformet og planlagt gjenommført av undertegnede. Jeg er en student ved Høgskolen i Østfold, på avdeling for Informatikk. Min veileder for dette arbeidet er Susanne K. Stigberg. Dette prosjektet utføres som en del av faget ITI40614 Master's Thesis, ledet av emneansvarlig Jan Høiberg, førstelektor ved høgskolen.

3 Hva slags informasjon blir samlet inn?

Om du velger å delta i studien vil du ta del i en praktisk testing av en prototype-applikasjon utviklet for mobiltelefoner, etterfulgt av et gruppeintervju sammen med meg og din meddeltaker. Den praktiske delen av studien involverer å benytte en augmented-reality applikasjon sammen med meddeltakeren i mellom 10 og 30 minutter. Den totale lengden av denne testingen avgjøres til dels av deres opplevelse, og hvor lenge dere velger å interagere med applikasjonen. Jeg vil være tilstede for å observere og ta notater under denne testingen. Du vil få anledning til å se over notatene mine før økten avsluttes om du skulle ønske dette. Direkte etter denne testingen vil det gjennomføres et gruppeintervju, der du og din meddeltaker vil bli spurt om ulike aspekter ved deres opplevelse av å bruke prototype applikasjonen. Dette intervjuet er estimert å vare mellom 30 og 60 minutter, og vil bli tatt opp ved hjelp av en diktafon. I løpet av dette intervjuet vil du bli spurt om dine meninger omkring temaer som augmented reality, samarbeid i virtuelle rom, kreativitet samt dine opplevelser med bruk av prototype applikasjonen.

4 Hvordan vil informasjonen om meg bli brukt?

Informasjonen som samles inn i løpet av studien vil kun bli brukt for formål relatert til denne masteroppgaven. Informasjonen vil bli samlet inn, behandlet, analysert og til slutt publisert som en del av min masteravhandling. Ingen personlig identifiserende informasjon skal bli inkludert i rapportens innhold, og alle forekomster av informasjon som kan brukes til å identifisere deg som person skal holdes konfidensielt. Dersom jeg skulle ønske å benytte direkte sitater fra intervjuet i rapportens tekst, vil du bli kontaktet for å gi samtykke til bruk.

5 Hvilke rettigheter har jeg som deltaker?

Så lenge du kan identifiseres i informasjonen som er samlet inn, har du full rett til å:

- Gjennomgå informasjon som er samlet inn om deg.
- Trekke tilbake informasjon som er samlet inn om deg.
- Korrigere informasjon som er samlet inn om deg.
- Be om at informasjonen som er samlet inn om deg fjernes og/eller slettes.
- Etterspørre en kopi av informasjon som er samlet inn om deg.
- Utstede en formell klage til Personvernombudet ved Høgskolen i Østfold eller Datatilsynet, angående behandling og bruk av informasjon samlet inn om deg.

Min rett til å bruke informasjon samlet inn om deg avhenger fullstendig av ditt samtykke om deltakelse som gitt ved signering av dette skjema. Du har når som helst rett til å trekke deg fra studien.

Hvis du ønsker mer informasjon, eller skulle ønske å benytte deg av dine rettigheter, ta kontakt med:

Susanne K. Stigberg	susanne.k.stigberg@hiof.no	+47 696 08 344
Personvernansvarlig v/ Høgskolen i Østfold	personvern@hiof.no	+47 696 08 009
Norsk Senter for Forskningsdata (NSD)	personverntjenester@nsd.no	+47 555 821 17

6 Samtykkeerklæring

Jeg erklærer herved at jeg har lest og forstått innholdet i dette dokumentet, og at jeg er blitt informert om hvordan informasjonen samlet inn om meg vil bli håndtert og behandlet. Jeg har blitt informert om mine rettigheter som deltaker, og har fått muligheten til å stille spørsmål før signering av dette dokument.

Jeg gir herved mitt samtykke til å:
☐ Delta i testing av en prototype applikasjon
☐ Bli observert under overnevnte testing
☐ Delta i et gruppeintervju etter testing av applikasjonen
☐ Få stemmen min tatt opp på diktafon under intervjuet

Jeg samtykker til at informasjon om meg blir lagret og behandlet frem til prosjektet er avsluttet, ca. 1. Juni 2019.

____________________ ____________________

Prosjektdeltaker Dato

____________________ ____________________

Prosjektansvarlig Student Dato

User Research Consent Form

Adrian Jensby Sandaker
adrianjs@hiof.no

1 Introduction

This study is performed as part of a thesis project for the degree of Master's of Applied Informatics at Østfold University College. The project explores how mobile augmented reality technology can be used to enable collaborative creation and interaction with music. Furthermore, the project examines how people cooperate within augmented reality in a broader perspective, and how the technology supports mechanisms of *awareness* and *presence* within an augmented reality space. In order to shed light on this, i wish to perform a qualitative study using an application I have developed as a part of the project. I am therefore interested in recruiting participants to take part in this study. There are no formal requirements for you as a participant, other than being 18 years old or more, being familiar with modern smartphones and having the ability to operate such technology with some level of proficiency. Additionally, it is seen as beneficial if you as a participant is interested in the topics of technology, music or creative expression in general. Additionally, as the study will be performed in pairs of two, you should be comfortable in collaborating with another potentially unfamiliar person for the total duration of the study.

2 Who is responsible for the study?

This study is designed and carried out by myself, Adrian Jensby Sandaker. I am a student at Østfold University College, at The Faculty of Computer Sciences. My thesis advisor is Assistant Professor Susanne K. Stigberg. The thesis work itself is part of the course ITI40614 Master's Thesis, led by Associate Professor Jan Høiberg.

3 What information will be collected?

Should you agree to participate, you will take part in a practical test of a prototype mobile application, followed by a group interview with myself and your co-participant. The practical test involves using an augmented reality application together with your co-participant for some period between 10 and 30 minutes. The total duration depends on your collective experience, and how long you both choose to engage with the application. I will be present to observe the session, and take notes whenever I find something interesting. You will be able to review these notes before the full session is concluded, should you wish to do so. Directly following this test session, a group interview will be conducted where you and your co-participant are asked about various aspects of your experience using the prototype application. This interview is estimated to take between 30 and 60 minutes, and will be recorded using a standalone voice recorder. During this interview you will be asked about your opinion on matters such as augmented reality, cooperation in virtual spaces, creativity, as well as your own experiences using the prototype.

4 How will my information be used?

The information collected in the session will be used only for the purposes of this study. Information will be collected, processed, analysed and published in a master thesis report. No personally identifiable pieces information shall be included in the contents of the report, and all occurrences of personally identifiable information shall be held confidential. Any direct quotes taken from the interview shall be cleared for use with the originator before being used in the report.

5 What are my rights as a participant?

As long as you can be identified within the collected material, you hold the right to:

- Review the personal information collected about you.
- Retract any information collected about you.
- Correct any information collected about you.
- Request the removal and deletion of information collected about you.
- Request a copy of any information collected about you.
- Issue a formal complaint to the Data Protection Officer (Personvernsansvarlig) at Østfold University College, or the Norwegian Data Protection Authority (Datatilsynet) regarding the handling and processing of your information.

My right to process your information is based solely on your willing participation and express consent given by signing this agreement. You hold the right to withdraw at any point during the process.

If you would like more information, or if you wish to make a claim, you can contact one of the following:

Susanne K. Stigberg	susanne.k.stigberg@hiof.no	+47 696 08 344
Data Protection Officer at Østfold University College	personvern@hiof.no	+47 696 08 009
Norwegian Centre for Research Data (NSD)	personverntjenester@nsd.no	+47 555 821 17

6 Declaration of Consent

I agree that I have read and understood the contents of this document, and that I am aware of how information provided will be handled and processed. I have been informed of my rights as a participant, and have been given the opportunity to inquire about any questions before signing. I consent to:

☐ Participating in a prototype testing session
☐ Being observed during the prototype testing session
☐ Participating in a group interview
☐ Having my voice recorded as part of the interview process

I consent to the collected information being stored and processed for the period until conclusion of the project, estimated to June 1st, 2019.

____________________ ____________________

Experiment Participant Date

____________________ ____________________

Experiment Administrator Date

Appendix D

Interview Guide

The interview guide used during the interviews.

Both Norwegian and English interview guides are included.

Intervjuguide

Adrian Jensby Sandaker
adrianjs@hiof.no

Denne intervjuguiden er laget for bruk etter en praktisk brukertest av prototype-applikasjonen. Hvert avsnitt begynner med en kort beskrivelse av hvilke tema spørsmålene har som formål belyse. Punkter uten nummerering representerer potensielle handlinger og ting å huske på, og de nummererte punktene representerer mulige spørsmål. Disse spørsmålene er nummerert kun med den hensikt å forenkle eventuelle referanser i ettertid av intervjuet, og rekkefølgen har ingenting å si i praksis. Jeg anser det som sannsynlig at flere spørsmål og temaer utover de som er foreslått vil dukke opp i løpet av intervjuet.

1 Introduksjon

- Ønsk deltakerne velkommen, og takk dem for deres deltakelse.
- Presenter målet med dette intervjuet, samt dets rolle i forskningsprosjektet.
- Informer om hvordan intervjuet vil bli gjennomført i praksis.
- Spør begge deltakere om de har eventuelle spørsmål før intervjuet starter.
- **Dobbeltsjekk at Diktafon-appen er slått på og aktiv**.

2 Generelt

Denne delen utforsker deltakernes generelle opplevelser og følelser omkring augmented reality, samt deres tanker omkring kreativitet og kreativ uttrykkelse.

- Fokuser på augmented reality på mobiltelefoner eller nettbrett.
- Forsøk å få innsyn i hvordan deltakerne ser på kreativitet og musikalske uttrykksformer generelt.

Spørsmål

1. Kan du beskrive dine følelser om augmented reality generelt?
2. Er du opptatt av å uttrykke deg kreativt?
3. Har du tidligere erfaring med bruk av augmented reality på mobiltelefoner?
4. Kan du beskrive din holdning overfor musikk i et generelt perspektiv?
5. Har du tidligere erfaring med å spille på instrumenter eller komponere musikk?
6. Hvordan vil du beskrive din holdning overfor samarbeid gjennom digitale plattformer?

3 Musikk og Augmented Reality

Dette avsnittet er fokusert på å utforske deltakernes opplevelse av musikalsk improvisasjon i augmented reality, og deres opplevelse med å bruke prototype-applikasjonen.

- **Vær obs på bruk av teknisk språk**, Forsøk å unngå avansert språk, og forklar underveis alle begreper som kan være fremmede eller vanskelig å forstå.
- Ikke vær redd for stillhet. La deltakernes følelser og synspunkter dukke opp på en naturlig måte.
- Betrakt de forslåtte spørsmålene som et utgangspunkt, og forsøk å stille gode oppfølgningsspørsmål. Ikke vær redd for å grave når du finner noe som virker interessant!

Spørsmål

1. Kan du beskrive med egne ord hvordan du opplevde det å bruke prototypen?
2. På Opplevde du at applikasjonen tillot deg å uttrykke deg kreativt i stor nok grad?
3. Hvordan opplevde du sammenhengen mellom visuelle og lydmessige komponenter?
4. Kan du fortelle litt om hvordan du opplevde selve lyden som produseres av nodene i applikasjonen?
5. Hvordan opplevde du de binaurale-komponentene i interaksjonen?
   Forklar litt om binaural lyd om nøvendig.
   Alternative formuleringer: 360°lyd, 3D lyd.
6. Hvordan opplevde du den loop-baserte arrangement-strutkuren?
7. Følte du deg drevet mest av *melodi* eller *rytme* under improvisasjonen?
8. Opplevde du at det på noe tidspunkt oppstod mønstre i plasseringen av noder?
9. Vil du si at du foretrekker å *lytte* eller å *skape*?

4 Roller og Samarbeid i Augmented Reality

Denne delen er fokusert på å utforske deltakernes opplevelse av samarbeid og kommunikasjon, samt deres opplevelse av egne og felles roller innen augmented reality situasjonen.

- **Vær obs på bruk av teknisk språk**, Forsøk å unngå sjargong, og forklar begreper som kan være fremmede eller vanskelig å forstå.
- Ikke vær redd for stillhet. La deltakernes følelser og synspunkter dukke opp på en naturlig måte.
- Betrakt de forslåtte spørsmålene som et utgangspunkt, og forsøk å stille gode oppfølgningsspørsmål. Ikke vær redd for å grave når du finner noe som virker interessant!
- Tilrettelegg for diskusjon mellom deltakerne også, ikke nødvendigvis med deg. Det kollektive perspektivet er av interesse!

Spørsmål:

1. Hvordan opplevde du kommunikasjonen med meddeltakeren under bruk av applikasjonen?
   (a) Hvordan forstod og tolket du denne kommunikasjonen?
   (b) **(Til den andre deltakeren)** Ble dette oppfattet slik du tenkte?
2. Brukte du på noe tidspunkt noder kun med intensjon om å kommunisere? Med andre ord, ikke med intensjon om å bidra musikalsk.
3. Hvordan opplevde du din egen rolle som deltaker i det improvisatoriske rom?
   (a) Hvordan påvirket denne rollen måten du interagerte med applikasjonen på?
   (b) Er dere begge enige i den opplevde rollefordelingen?
4. Følte du på noe tidspunkt i økten at du enten **ledet** eller **ble ledet** i interaksjonen?
   (a) Endret dette seg på noe tidspunkt?
5. Opplevde du at din opplevelse av aktiviteten endret seg på noe tidspunkt i løpet av økten?
6. Opplevde du at det i løpet av økten etablerte seg normer for oppførsel innad i interaksjonen?
7. På hvilken måte påvirket tekniske utfordringer samarbeidet i augmented reality rommet? **Eks. Nettverksforsinkelser.**
8. Opplevde du på noe tidspunkt i løpet av økten en følelse av *flow*?
   NB: Forklar *flow* om nødvendig!

5 Avslutning

Debrief deltakerne, og avslutt intervjuet.

- Takk begge deltakerne for at de deltok i intervjuet.
- Forklar hvordan lydopptakene vil bli lagret og benyttet i et forskningsperspektiv.
- Informer deltakerne om at de vil bli kontaktet dersom det skulle være aktuelt å benytte direktesitater fra intervjuet. **NB! Husk å notere kontaktinformasjon (tlf. eller epost adresse) før du avslutter intervjuet!**
- Spør om de har noen ytterligere spørsmål eller kommentarer.
- Avslutt.

Interview Guide

Adrian Jensby Sandaker
adrianjs@hiof.no

This document is intended to function as a dynamic interview guide after finishing a test session with the prototype application. Each section hereafter begins with a brief statement of my intent. Items marked with a *dot* represent suggested actions, while items marked with a *number* represent suggested questions. Questions have been numbered for reference during post-interview work, and this does not denote a specific order of asking. In addition to the questions prepared, it is likely that other questions and themes will emerge organically throughout the interview. This guide is mainly a tool for myself when conducting interviews, but functions also as tangible documentation of how the interview is to be conducted.

1 Introduction

- Greet both participants and thank them for participating.
- Briefly present the purpose of this interview and the associated research project.
- Inform them how the interview will be administered, i.e. how the conversation flow will be moderated.
- Ask whether participants have any questions before starting the interview.
- **Make sure that the recorder app is opened and recording**

2 General

Looking at participants' experiences and feelings towards augmented reality in general, as well as their feelings toward creativity and creative expression. Less focused on the prototype instance.

- Focus on mobile augmented reality being used on smartphones or tablets.
- Attempt to gain a perspective of how the participants view creative and musical expression in general.

Suggested Questions

1. ”How would you describe your attitude towards creative expression in general?”
2. “Do you have any prior experience using mobile augmented reality applications?”
3. “How do you feel towards augmented reality technology in general?”
4. ”Could you describe your attitude towards music in general?”
5. ”Do you have any previous experience playing or composing music?”

3 Music and Augmented Reality

This section is concentrated on exploring the participants' experience of improvising music in augmented reality, and their experience using the prototype application.

- **Be mindful of technical language use**. Try to avoid jargon, and explain terms that might seem foreign or difficult to comprehend.
- Don't be afraid of silence. Let the participants feelings and viewpoints emerge naturally.
- Consider the suggested questions a base, and supplement by asking follow up questions. Don't be afraid to probe when something catches your interest.

Suggested questions

1. "Could you describe in your own words how you experienced using the prototype?"
2. "How did you feel like the application allowed you to express yourself creatively?"
3. "How did you experience the connection between sonic and visual components?"
4. "Could you tell me how you felt about the sound produced by a node?"
5. "How did you experience the binaural qualities of the interaction?"
6. "How did you experience the looping arrangement structure when working?"
7. "Did you feel like you were driven more by *melody* or *rhythm* when improvising?"
8. "Did you feel any patterns of node placement emerge?"
9. "Would you say that you tend more towards *listening* or *creating* when using the application?"

4 Role and Collaboration in Augmented Reality

This section is focused on exploring the participants' experience of collaboration and communication within the augmented reality space, as well as how they experience roles.

- **Be mindful of technical language use**. Try to avoid jargon, and explain terms that might seem foreign or difficult to comprehend.
- Don't be afraid of silence. Let the participants feelings and viewpoints emerge naturally.
- Consider the suggested questions a base, and supplement by asking follow up questions. Don't be afraid to probe when something catches your interest.
- Facilitate for sufficient discussion between participants, not necessarily involving you. **Their collective experience of collaboration is what matters here**!

Suggested questions:

1. "How did you experience the communication with the other participant within the application?"
   (a) "How did you understand and interpret this communication?"
   (b) **(To other participant)** "Was this interpreted how you intended?"
2. "Did you at any point use nodes only to attempt direct communication? I.e. not with the intent of expressing music."
3. "How did you perceive your own role as a participant within the improvisational space?"
   (a) "How did this this role affect the way you interacted with the playspace?"
   (b) "Do you both agree on the separation of roles experienced?"
4. "Did you at any point during the session feel as if you were either **in charge** or **being guided** in the interaction?"
   (a) "Did this dynamic at any point change?"
5. "Did your experience of the activity change throughout the session?"
6. "Did any norms establish themselves in the playspace during use?"
7. "To what extent did technical challenges affect your cooperation in the augmented reality space?" **Ex: Network latency**
8. "Did you at any point during the session experience a feeling of *flow*?"
   Note: Explain *flow* if necessary!

5 Finishing up

Closing the interview, and debriefing the participants.

- Thank the participants for taking part in the interview.
- Explain how the recorded material will be stored and used in a research context.
- Inform them that I will be in contact should any direct quotes be extracted and used in the published material. **NB: Make sure to obtain a means of contact, such as a telephone number or email address before concluding the session.**
- Ask if they have any additional questions.
- Disband.

Appendix E

Transcription Example

The transcription excerpt has been included to show the tone and style of transcription used. As described in the report, the interviews were done in Norwegian and the translation was performed as part of the coding. Therefore, there exists no translated version of the interview transcripts.

Utdrag fra transkripsjon.

M: Meg.

I5: Informant 5.

I6: Informant 6.

M: «Jeg lurer litt på om dere – hva dere følte dere mest drevet av når dere brukte det? Om det var mer rytmen, eller om det var liksom melodi? I...»

I5: «Ble kanskje – ble kanskje rytmen i og med at den var litt sånn – den takten. Jeg prøvde litt melodi, men det var litt vanskelig å time da, kanskje.»

M: «Mm?»

I5: «Så det ble lettere å liksom lage en sånn – akkord aktig i en rytme – sånn arpeggio aktig. Ja.»

M: «Mm?»

I5: «Så det ble kanskje litt mer rytme – rytme og harmoni drevet da?»

M: «Ja, ja nemlig – så hvis du ser harmoni som litt separat fra melodi men..»

I6: «Ja.»

M: «...det er mer det å liksom få til en harmoni *i det*.»

I5: «Mm.»

M: «Nemlig.»

I5: «Mm.»

I6: «Ja altså jeg føler jo at det var rytme som drev det hele frem, også via det akkord – som han sa – så det med melodi kom litt i andre rekke for meg da.»

M: «Mm?»

I6: «Fordi det handlet om at da måtte du slå av loopen, og så måtte du jo treffe riktig da..!»

I6: *latter*

M: «Ja, ikke sant?»

I6: «Ja!»

Appendix F

Field Observation Notes

This section presents the observations done during each test session. These are included to provide a narrated and verbose portrayal of how the prototype was used during each test session. These observations are based solely on notes taken by myself during each session. The sessions are presented in chronological order.

Session 1 Observation

Elias is a personal friend of mine, and the testing was performed at his home. Sverre was recruited by Elias, and I had no acquaintance with him before the test session. Both participants in this session were trained and practicing musicians. Before beginning the testing, both participants were given basic instructions to familiarize them with the interface. This was aided by the instruction manual found in appendix x. The testing session had a total duration of 35 minutes.

Session start: 19:16

The session commenced seamlessly after each participant had received instructions on how to use the application. Both participants appeared immediately focused on the activity at hand, and there was no apparent communication taking place between them. I observed some expressions of enjoyment such as smiles and the odd chuckle from both participants. After a short while, it appeared as if the participants had begun collaborating on a looping arrangement. Following this was a period of little communication, and apparent focused use of the prototype. Around the **10 minute mark**, both participants began moving around the room. Up until this point, they had both held their initial position. After a brief period of movement, they stopped actively moving around. Both participants were now facing each other.

Around the **15 minute mark**, Elias began using his voice to sing along with the music generated by the prototype. This was followed by slight verbal communication between participants, centered mainly around planning the musical arrangement. Around the **20 minutes mark**, the prototype application crashed on the device used by Sverre. I decided to interrupt the session to restart the prototype application and re-connect the device. This was successful, and the crash resulted in no significant inconvenience. As the participants continued the session, I observed an increase in communication. Sverre began tapping his foot to the rhythm generated by the application, and there was some brief discussion regarding the rhythmic structure. Around the **25 minute mark**, both participants were increasingly using their voice to sing and vocalize over the generated music. There was again an increase in movement, and the participants appeared to engage more with the spatial qualities of the interaction. At the **30 minutes mark**, the participants began discussing what they experiences as a disruptive latency in the microphone input. This was followed by more vocalizing and singing from both participants. Slowly, the intensity of the activity appeared to gradually slow down, and the session dissipated naturally after **35 minutes**.

Session end: 19:51

Session 2 Observation

The session took place in a classroom at Østfold University College, and was performed by participants Jonathan and Nikolai. Jonathan is a friend of mine, and was recruited directly via instant messaging. Nikolai was recruited through Social Media call for participation, however I had made his acquaintance a few years earlier. Jonathan reported no significant musical experience, while Nikolai has prior experience as a practicing drummer. Both participants had significant software development experience, and Nikolai had a little experience working with AR technology. Before beginning the test session, both participants were given basic instructions to familiarize them with the interface similar to the approach taken in session 1. This was again aided by the instruction manual found in appendix x. The session had a total duration of 32 minutes.

Session start: 14:28

The session began after I had instructed the participants in how the application was used. Some difficulties were had getting connected, as the texture of the floor appeared less than ideal. After a short while (1-2 minutes), both users were able to connect to the same room. Both participants appeared interested in exploring the interaction, and immediately began engaging with the experience. I observed some verbal communication between the participants.

At the **5 minute mark**, the application crashed on the device used by Nikolai. I immediately engaged and reconnected the device, and no significant disruption took place. Following this, there was more verbal communication between both participants, mainly related to the prototype interaction.

Around the **15 minute mark**, both participants had significant difficulties in getting the augmented reality system to recognize the floor as a surface. This can likely be attributed to the uniform color and texture of the floor. To attempt mitigation of this and improve the tracking, I scattered the floor with whatever small trinkets I could find in my wallet and bag (USB-cables, loyalty cards, earbuds etc.). This was very successful, and significantly improved the tracking, which allowed the session to continue. The photograph in figure F.1 shows how the floor looked at this point.

At this point, both participants appeared to be exploring the possibilities offered by the prototype system. Jonathan appeared to be testing the limits of the system somewhat, by filling the space with nodes. The participants playfully engaged with the interaction, and appeared to be entertained by the interaction. Some movement around the room could be observed. There was some ongoing verbal communication from both participants, mainly related to each others activities in the augmented reality space. Around this time, Jonathan appeared to discover the binaural properties of the interaction. This led to a more active engagement with the spatial nature of the sound.

At the **20 minute mark**, Nikolai verbally expressed interest in directed collaboration, however Jonathan did not seem to reciprocate this. Around the same time, both participants noted a slight frustration related to the uniformity of the sound. This was followed by a instance of *battling* using the sound, where the participants playfully attempted to startle each other using sound. This was followed by a discussion between the participants on potential technical challenges of the system. Around the **30 minute mark** both participants seemed to slowly lose interest, and the session organically concluded after a few minutes. **Session end: 15:00**

Session 3

This session took place in the garden outside of the house where I currently live. Mikkel is my next door neighbour, and was recruited in person. Edvard is a good friend of Mikkel, and was recruited by Mikkel. I was unacquainted with Edvard before the test session. Both participants have significant musical experience.

Session start: 18:05

The session began immediately after participants had received instructions. At the start of the session there was little movement from both participants. They appeared immersed in the experience, and exhibited signs of enjoyment such as laughter and smiles. I observed little communication at this point. Edvard recognized that the tones were mapped along a scale, and asked me about that.

Around the **5 minute mark** there was still little movement. The participants appeared to expand the playspace, and began looking for each others nodes throughout the garden. Mikkel was quotes as saying "Did you place one all the way over there?". Mikkel discovered that walls could be tracked by the software as well. Edvard remained relatively still. There appeared to be some playful competition taking place within the playspace. Mikkel enabled the microphone, and encouraged Edvard to do the same. Mikkel inquired about changing the synth sound.

Around the **10 minute mark** I observed more tracking of walls. The movement increased somewhat, and both participants moved around actively. Edvard asked about the possibility of extending the length of the loop. The participants appeared to begin actively collaborating on a piece within the playspace. There was an increase in communication, and the participants exchanged tips about how to use the interface.

Around the **15 minute mark**, Edvard appeared to actively examine the interface for how the tones were mapped. The cooperation appeared to continue, and both participants seemed focused on creating a shared piece of music within the AR space. This moved over into what appeared to be a sort of competition within the AR space, and both participants appeared to partake in this with interest.

Around the **20 minute mark**, Mikkel mentioned he had created an arpeggio within the space. This was followed by some communication between participants, before they both moved into a more focused state. At this point, the application crashed on the device used by Edvard. I interrupted the session to quickly restore it to a working state. This was successful, and the session could be restored. Following this, Mikkel asked for a short break, and excused himself. During this time, I tended to his device to maintain tracking in AR, but I chose to not participate in the session and remained still. I did however capture the screenshot seen in figure F.2, showing nodes from Edvard within the space. Edvard continued to interact with the playspace, and appeared focused.

Around the **25 minute mark**, Mikkel returned and resumed his use of the prototype. Both participants appeared focused, and seemed to be in a state of focused work. This continued for a few minutes, until the participants began moving around and using the space a little more. This was followed by some light communication, with Edvard requesting that they collaborate more actively. This was followed by more exploration of the space, and they now expanded the playspace to include a signification portion of the surrounding area.

Around the **30 minute mark**, both participants appeared to be actively exploring the possibilities offered by the playspace, and engaged in what appeared to be a game of "hide-and-seek" using nodes. Both participants expressed amusement and laughter during this period. Mikkel was quoted as saying "There are balls (nodes) everywhere!". Edvard mentioned that he had constructed a chord within the space, and appeared to be playing alongside it. Both participants appeared entertained.

Around the **35 minute mark**, the participants returned to exploring the space. around them. Edvard noted some lagging in the system due to a large amount of nodes being present. Mikkel described the current playspace state as "chaotic", and suggested they clean up the space, and both participants appeared to start removing nodes. This was followed by a return to regular use, with what appeared to be directed and collaborative work taking place.

Around the **40 minute mark**, the activity level slowed down, and the participants appeared to be more focused on working separately again. The participants then gradually began to wind down, and the session slowly dissolved.

Session ended at 18:48.

Summary

All three sessions were carried out within the time frame of one week. To support the comparison of any findings across each session, I decided to not modify the prototype further between sessions despite receiving concrete feedback on possible improvements from participants. In addition, as described in section 3.3.2 all observations were carried out with a non-participatory approach. However, In each session i briefly engaged whenever an application crash occurred, in order to restore the prototype to a working state. Crashes were experienced in all three sessions. Despite some application crashes and slight technical difficulties, the testing sessions were carried out without any significant breakdowns.

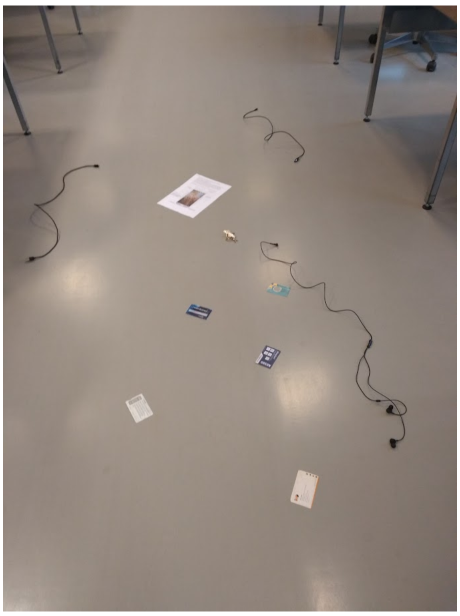


Figure F.1: Photograph from session 2 showing various trinkets scattered on the floor to improve tracking in the ARCore system.

Figure F.2: A screenshot of the playspace during session 3.

Appendix G

USB Memory Stick

The USB memory stick has been included with the report, and contains the contents listed below. To allow for easier access, the full contents of the memory stick have also been uploaded GitHub. Due to the sensitive nature of some of the attached data (API keys etc.), I have registered a dedicated account for this on GitHub, and created a private repository. The files can be accessed using the following account information:

> **Username:** `Username omitted before publishing`
> **Password:** `Password omitted before publishing`
> **Repo URL:** `Repo-URL omitted before publishing`

To download the files, you can run the following command using a terminal, and authenticate with the above account information when prompted:

git clone https://github.com/`repo-url-has-been-omitted`

1. **Project source code and APK file**

   The full source code for the application Petals. The project can be viewed as-is, but should ideally be opened in the IDE Android Studio[1] to ensure the best experience.

   In addition, I have included the APK-file used during the field deployment for easy installation and testing.

2. **Video demonstration**

   A comprehensive video demonstration of the application in use. This video has been included to reduce the need of actively testing the application in order to understand its functionality.

3. **Video Prototype**

[1] `https://developer.android.com/studio`

Two clips from the video prototyping performed in section 4.2. These clips represent the end result, after the video has been augmented with musical sound.

4. **Pure Data patch**

   The Pure Data patch used in the application. This patch is also embedded int he project source code, but has been included here for simplicity. In addition, I have included screenshots of the patch to remove the need to install the Pure Data software[2].

5. **Firebase Cloud Functions code**

   The code used to enable the back-end functionality of the Firebase Cloud Functions. The relevant code can be found in the file **index.js** inside the functions folder.

6. **Makefile for modified pd-for-android build**

   The makefile used to build the pd-for-android library with the +binaural external.

In addition, the memory stick and the GitHub repository both contain a copy of each of the appendixes from this report.

[2] `https://puredata.info/`

Appendix H

Petals Installation Guide

Petals v. 1.0 - Installation Guide and Instruction Manual

1 - Installing the application.

The application can be installed one of the following ways:

1. Build and run the project from Android Studio.
2. Install the release APK bundled with the project file.

The application requires that a recent version of ARCore is installed on the device.
If running on a physical device, ARCore can be downloaded and installed through the Google Play Store: https://play.google.com/store/apps/details?id=com.google.ar.core&hl=no

If running on an emulator, please download and install the emulator version from GitHub
https://github.com/google-ar/arcore-android-sdk/releases
Look for a filename like *ARCore_VERSION_x86_for_emulator.apk*

2 - Launching the application.

WARNING: The application has been known to display erratic behaviour on some phones leading to the output of warped and noisy audio instead of the intended sound. It is therefore advised that you disconnect the headphones and lower the device volume when first launching the application. Create a new room and place a few nodes to confirm/deny before connecting headphones.

The first time Petals is launched, it will make runtime requests to ask for four different permissions from the user.
All permissions must be granted for the application to function as intended.

The permissions are:

- External Storage Access
- Camera
- Location / GPS
- Audio Recording / Microphone

External Storage Access is required by the libpd library, and is used to extraction and loading of the zipped Pure Data patch which provides the sound for the application.

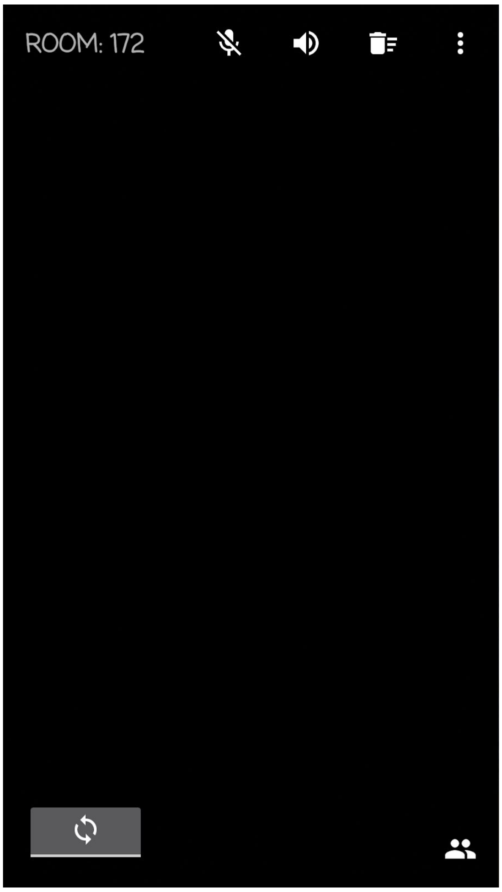
Camera is required to support the augmented reality functionality through ARCore and Sceneform.

Location / GPS is used to record user position at time of creation. *This data is currently not used for anything in the application, and should be considered as a leftover from the design phase.*

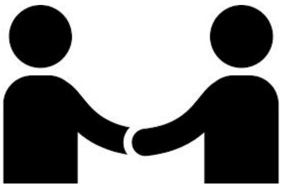
Audio Recording / Microphone permission is used to allow libpd to take control over the microphone, to provide the audio throughput functionality.

As long as these permissions are granted to the application, it should function as intended for any subsequent launches.

When successfully launched, you should see something like the interface displayed in the screenshot below - only with an actual camera feed instead of the black backgrond:



3 - Getting Started with Petals



Note: Due to the cooperative nature of the interaction, the prototype application is intended for use by two people simultaneously in the same environment. Using the application alone will enable the creation of music, but will not provide the full intended experience.

Note: A full demonstration of how to connect to a room and begin playing can be seen in the video demonstration available on the USB disk.

Step 1 - Get an understanding of the environment

If the application started successfully, you should see a moving icon like this one moving around on your screen:

This icon indicates that the Sceneform module is ready to begin mapping the environment. Move your device as indicated by the movement of the hand, and some dots should appear on the surface you scanned, indicating that it has been mapped by the ARCore library.

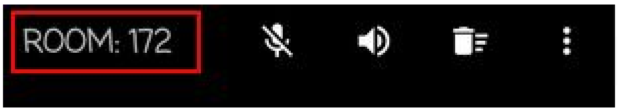
You are now ready to connect to a room!

Step 2 - Connect to a room

The prototype has been constructed in such a way that offline interaction is not possible. This means that you will have to connect to a room in order to begin playing. To connect to a room, locate and press the following icon in the user interface:

Look in the bottom right corner!

You will now be presented with a choice of either ***New room*** or ***Join Room***. If your partner has not yet established a session, choose new room. This will initiate a connection process in which a cloud anchor is placed, and a room is created. If successful, you will be presented with a ***Room code***. This room code enables your partner to connect to the same session, to enable cooperative interaction. Your room code will show up in the upper left corner, in the area indicated in the below image.



The value of code will likely be significantly larger.

Make sure your co participant has scanned the same area in the environment, and give him your code. He should press the icon in the bottom right corner and select ***Join Room***, where he will be presented with a dialog to enter the code. Upon successful connection, you will both be connected to the same room, and can now begin playing!

Step 3 - Begin playing!

When you and your partner have connected to the same room, you are now free to begin playing! By tapping on the scanned environment, you can place nodes into the shared playspace.

To swap between *floating nodes* and *bouncing nodes*, press the icon in the lower left corner.

🎵 Have fun and play nice! 🎧

4 - Bugs and Known Issues

The application is prototype software, and there are certain bugs that might show up during use. The below issues have been documented in testing the app. Additional and unknown bugs might occur.

No sound from nodes.

This bug occurs sometimes during the initial run, and is related to the libpd module. It is usually resolved by simply restarting the app.

Application refuses to start or crashes immediately.

This could be related to a missing our outdated version of ARCore. Try reinstalling ARCore from the PlayStore, or installing from https://github.com/google-ar/arcore-android-sdk/releases

No effect of binaural audio.

This is likely caused by the binaural external within the libpd patch crashing. Restart the application. If this does not resolve the issue, try deleting and reinstalling the application

Nodes appear in the wrong place.

This is likely caused by a misaligned anchor. Try reconnecting to the room. If this does not help, restart the application and create a new room.

Appendix I

Petals User Manual

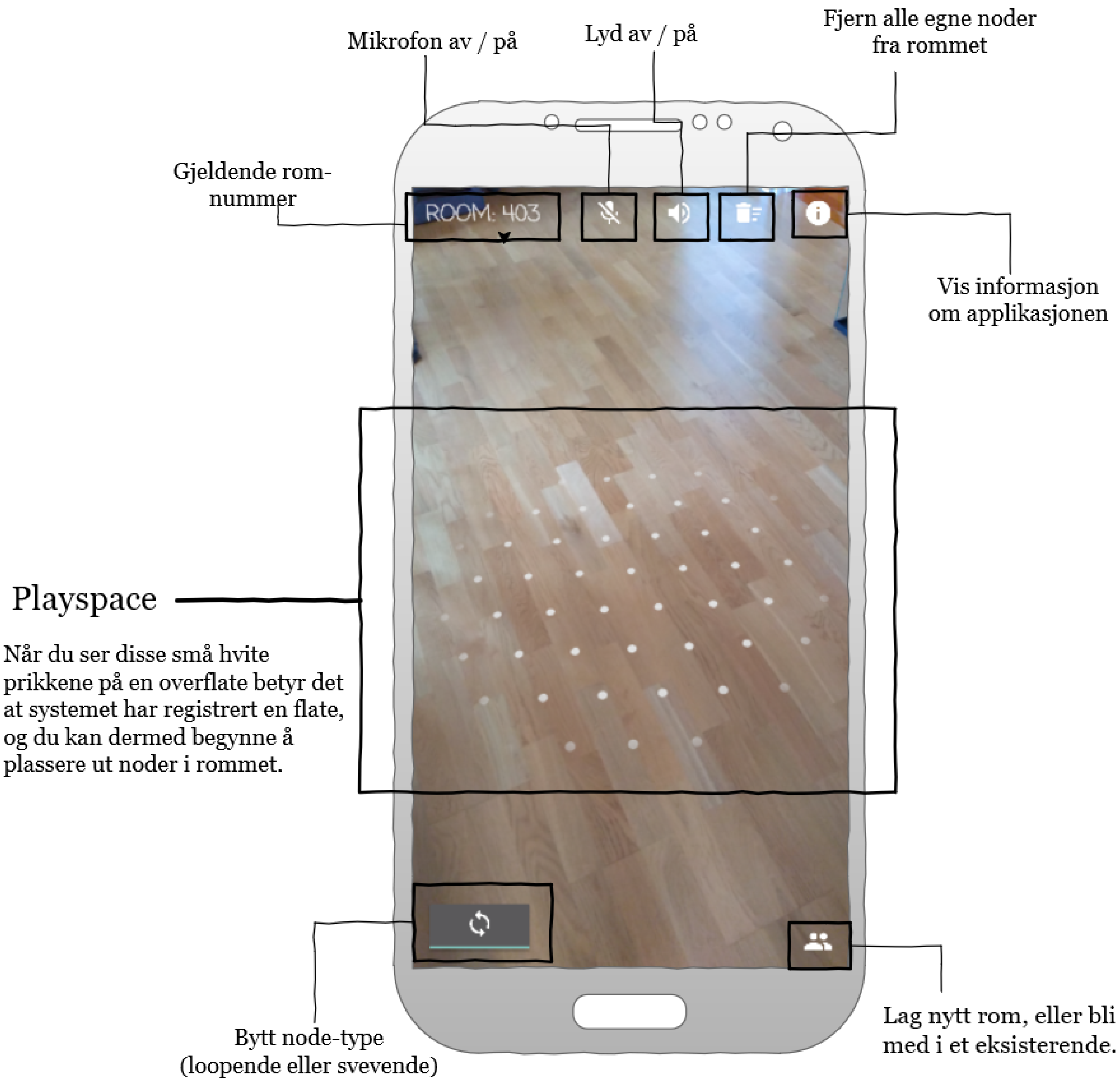
A simplified application user manual created to aid in explaining the application during the field deployment. As all sessions were carried out in Norwegian, I never created an English version.

(Norsk) Forenklet Brukermanual - Petals

Hvordan fungerer applikasjonen?

Det overordnede målet med applikasjonen er å skape musikk sammen med en annen bruker i et augmented reality miljø. Dette gjøres ved hjelp av virtuelle kuler som representerer musikalske toner i augmented reality rommet. Disse omtales som *noder*. Hver bruker kan fritt plassere ut noder i rommet så lenge applikasjonen har registrert en overflate. En hver node som plasseres på en flate vil umiddelbart bli synlig både for en selv og ens medspiller, og kan deretter flyttes og manipuleres av begge. Brukere kan i applikasjonen benytte to forskjellige noder - *repeterende noder* og *svevende noder*.

Hvilke funksjoner har de ulike knappene i brukergrensesnittet?



Hva er forskjellen på repeterende og svevende noder?

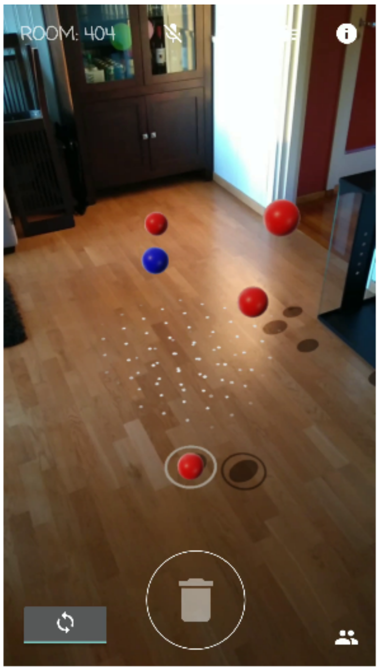
Repeterende noder vil bevege seg opp og ned på overflaten de blir plassert på, og vil avgi en tone hver gang de treffer overflaten. Svevende noder beveger seg opp og ut av rommet, og vil kun avgi en tone når de først plasseres.

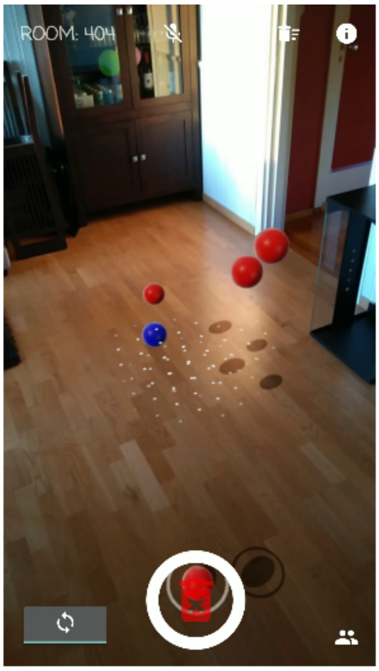
Hvordan kan jeg kontrollere hvilke toner som skal spilles?

Hvilken tone som tilknyttes en node avgjøres av hvor langt unna brukeren den plasseres. Toner som plasseres tett på din fysiske posisjon vil ha en lav tone, og toner som plasseres langt unna vil bli gitt en høy tone. Maksimal avstand er 5 meter.

Kan jeg slette enkelt-noder fra rommet?

Dersom du ønsker å slette en enkelt node kan det gjøres på følgende måte.





1

Finn en node i rommet som du ønsker å fjerne. Det er kun mulig å fjerne repeterende noder.

2

Ta tak i noden ved hjelp av touch-skjermen. Noden blir markert som *valgt*, og et ikon bli synlig i nederst på skjermen.

3

Dra den valgte noden mot ikonet , og hold den innenfor sirkelen. Ikonet vil endre farge til rødt, og ringen rundt vil uthevet.. For å slette noden, slipp den over ikonet, og den fjernes fra rommet.

Hvordan avsluttes en økt?

En økt kan avsluttes ved å koble enheten fra det delte rommet. Dette gjøres ved å trykke på ikonet nederst i høyre hjørne. Du vil dermed bli koblet fra rommet, og alle noder vil bli deaktivert.