VJ x VR

Exploring concert-like experiences through a collaborative approach to VJing inside of Virtual Reality.

Bachelor's Thesis

Theis T. B. Nielsen (thtn@itu.dk)

88.731 characters (in **49** pages)

Course code: BIBAPRO1PE

Supervisor: Stina Hasse Jørgensen (shaj@itu.dk)

Abstract

This thesis was written by Theis Theodor Bech Nielsen and handed in on April 17th 2021, for the Digital Design & Interactive Technologies bachelor's programme (BDDIT) offered at the IT-University of Copenhagen.

This project proposes an explorative design for creating concert-like experiences using a collaborative approach to VJing inside of Virtual Reality, dubbed "VJ x VR". The project explains contemporary practices of VJing, Virtual collaboration and Virtual Reality. The method of contextual inquiry and experience prototyping is used to capture tacit and volatile recollections from VR to inform the creation of a concert-like Unity prototype. A total of 6 VR participants tested this experience. Using thematic analysis in VR-situated evaluating interviews, seven themes spanning from relational to interactive and tangible concepts are constructed. These themes reveal a strong sense of relatedness in the ability to share visual effects but shows conflict as to what the focus of the experience is. These findings are then discussed in relation to existing concert-experiences, situating the visual dimension of Virtual Reality and VJing to the aural dimension of music and concerts. It is concluded that the project was successful in its creation of an interactive virtual and collaborative environment. The presence of music is however less present, and this should be investigated in future designs. This project is intended to contribute to the area of interaction aesthetics, to provide audio-visual designers with an approach that formulate collaborative and meaningful integration of contemporary immersive technology.

Table of contents

1. Preface and introduction

This report was written in April 2021 in a 20–week project at the IT University of Copenhagen under the supervision of Stina Hasse Jørgensen. I thank my brother Thomas T. B. Nielsen for extensive assistance in prototyping, as well as all the friends who helped test and evaluate the prototype.

During the project, the designer/researcher developed an executable Virtual Reality application with the aim of immersing users into a virtual concert-experience through collaborative audio-visual interaction together, drawing inspiration from the existing visual and digital artistry of VJing. This was created in response to the virtualisation of concert-experiences because of the Covid-19 lockdown. The program is developed in the Unity Game Engine and situates multiple VR users in a virtual scene in which their actions can reproduce visual stimuli. The program recognizes gestures and button-presses from multiple users simultaneously in a shared virtual scene. These interactions trigger a variety of visual output that is displayed in tandem with continuous music through scripts within the scene. Users can see each other and share the space of developing interesting visuals together.
The main research goal of this project is:

To explore concert-like experiences using a collaborative approach to VJing inside of Virtual Reality.

Secondly, this project aims to explore the affordances of Virtual Reality as a medium for experiencing interactive and collaborative audio-visual narratives.
(For a video showcase of the Unity prototype, see Appendix 1 in the attached appendix pdf)

1.1 Background and description of the problem

In 2020, the corona virus escalated into a pandemic, causing most countries to go into varying stages of lockdown. The virus is highly contagious and thus requires the practice of social distancing to avoid risk of infection (Sundhedsstyrelsen, 2021). Among many other sectors, this particularly resulted in most social events having to be postponed or cancelled entirely. Among these social events are concerts and festivals, traditionally gathering large crowds of

people. Without the possibility of physical concerts, initiatives have been taken to move concert-experiences online, into virtual performances.

This project culminates around the three fields: Virtual Reality, VJing and Virtual Collaboration. The following section will map each of these respective areas, as well as the description of existing works that also attempt to merge these disciplines.

2 State-of-the-art field

2.1 The Video Jockey

Video Jockeys are visual artists who use digital media to express themselves to an audience during a live audio-visual performance. The term VJ historically refers to Video Jockey, though VJ can be used as an umbrella-term for several various practices that utilise live visual media performances.

2.1.1 The history of VJing

The cultural history of VJing is comprehensive. While the technological prerequisites may differ wildly from contemporary practices, its predecessors have all aimed at combining storytelling, moving images, and different forms of technology. The emergence of cinema in the mid-1890s introduced cameras operated by hand cranks, which could dictate the speed of the playback motion, fast repetitive cuts, exaggerated motion, time based, and spatial montage. Rhythmic elements within the frame such as double exposures or overlays, slow motion, reversed motion, and stills were also used in conjunction with these techniques to elaborate montage processes.

It was the introduction of VHS tapes in the 1970's that introduced the practice of traditional VJing as we see it today. Unlike with cinema, the audience in dance clubs and musical performances no longer had their full attention focused on the screen, as they were dancing. This meant it was no longer necessary to have one focal point, allowing the VJ to further experiment with fast paced mixing and juxtaposing of various different types of content. (Lamb, 2010)

2.1.2 Contemporary VJing

Most people recognize the DJ as an artist who mixes music to enhance the audio. Similarly, the VJ is a visual artist who mixes different *visual media* to adapt visuals to music in a live

setting. The VJ uses images, composition, colour, and contrast to enhance the overall experience of the music. Adding video to music, it resembles the art of music videos where stories and moods from the music are being presented through visual perception, thus delivering empowered narrative and emotion to the experience. In a typical setting, the VJ is seated in the vicinity of the performance, and it is their job to control what is seen on monitors/projectors at the venue. Aside from the projector/screen, the most important tool of a VJ is the MIDI controller. The MIDI controller offers granular control with the use of analogue dials and potentiometers to blend media seamlessly. Chances are, when you have gone to a venue or concert, a VJ was present behind the scenes, pulling the visual strings. You might even see the DJ practicing VJing on a mixer alongside their audio-production. (Lamb, 2010)

2.1.3 Motivations for this project

This study is motivated by the researcher's volunteering participation as a VJ in a handful of local festivals throughout the past two years (see Picture 1). The initial nervousness of having to perform a personal aesthetic 'take' of the artist(s) music to a crowd, was quickly substituted by the adrenaline rush of timing, narrating, and experimenting with different visual impressions to tie into the music. The creative expressive capabilities of VJing makes it a personalized artform in which the motoric granularity of turning dials in tandem with the head-bobbing and loud, intense music provides instant gratification. This only gets amplified further when peeking through the blinds and seeing the happy faces in the crowd, with their eyes centred on a highly personalized creation, enjoying the concert and aiding the intended vision of the music artists. It is experiences like these which drive this study, and the reason why the VJing must continue to stay relevant.

Picture 1: "Blood Child's" performance during the 2020 "KU.BE YHC Vinter Festival" - Visuals were projected onto a canvas, with the researcher mixing video material live backstage.

The act of VJing is often considered a marginalised culture with creative practice to enhance the listening experience (Lamb, 2010), and should be included as concerts and festivals move online.

2.1.4 Existing virtual concerts

But why should the VJ be considered in contemporary concert-experiences? Aside from the satisfaction mentioned earlier, the practice of VJing has always been closely adherent to technological trends. VJs often find themselves evaluating technology primarily based on the expression through visual media. From this perspective, one can consider VJs to be pragmatic designers who appropriate technology. For this reason, the artistry should find a place within virtual performances. One might argue that VJing has already been virtually adapted in light of the contemporary works of entirely digitally constructed live concerts such as the 'Travis Scott x Fortnite' (Polygon, 2020) (see Figure 1) or 'Lil Nas x Roblox' (BBC, 2020) performances. Here we see 3D visuals, and a visual narrative that fits the mood and rhythm of the music. One could, however, argue that this practice differs from VJing, as while the visual material is video/computer generated imagery, the material is created *prior* to the experiences. VJ performances are inherently *live* performances of visual media (Malmström, 2012). Hook (2011) argues that the live aspect of a performance is a central part of expressive interaction among VJs, as it allows for improvisation, which in turn creates a unique experience to great motivation for the VJ and audience.

Figure 1: Screenshot from the Travis Scott x Fortnite "Astronomical" concert with over 27 million virtual attendees. (Polygon, 2020)

With the risk of being forgotten in these new virtual-concert practices, the practice of VJing must be adapted to stay relevant during the ongoing pandemic.

Additionally, Hook identifies key themes of expressive interaction for VJs, such as their motivation for performing and adherence to live improvised performances. Similarly, Lamb (2010) and Malmström (2012) isolate key factors that impact content creation, selection, and artistic concepts for live VJ mixing. These expressive interactions are important to notice as they might directly motivate people to be engaged while VJing. For use in this project, the researcher expects the expressive motivations and artistic concepts of VJing, as discussed by Hook, Lambo, and Malmström to also extend to novice or first-time VJ's. The reason for this interest in extending the VJ practices relates to existing collaborative initiative from other designers.

2.2 Virtual Collaboration

2.2.1 Outsourcing the VJ

The ambition of this project is not only to transpose traditional VJ practices to a different medium, but, more importantly, to alternate the role of the VJ, and thus the role of virtual concerts, into a more collaborative context.

Recent VJ movements have been set in place to 'outsource' the role of the VJ into the hands of the audience. An example of this participatory form is VJ Yourself (2016). This is a movement which has created multiple installations to make audiences the focal point. The audience is invited to participate in the creative VJ process with passive viewers becoming more active collaborators, whose actions and inspiration play a vital role in the real-time actualization of the performance (see Figure 2).

Figure 2: Screenshots from VJ Yourself's "Gamer Disco" installation. The audience is given game controllers to influence the visuals while music is playing.

Similarly, Freeman (2010) has created the "Layer Synthesis Device (LSD)" which substitutes the concept of VJ Yourself's game controller with the audience's smartphones for a MIDI-like level of control with sliders to collaborate visuals in which every member of the audience can now participate. These installations provide an interesting approach to an almost role reversal of the VJ, who now creates a sort of interaction framework for the audience to interact with, instead of the audience watching a single point of expressive visual production.

2.2.2 Collaborative Virtual Reality

This project will also explore the VR medium in a collaborative context (the VR medium will be elaborated on further in the 'Virtual Reality' section), also warranting an investigation into existing virtual collaborative practices. Most literature regarding collaboration in virtual surroundings seems to focus on one of two areas: Using online collaborative design to improve learning, or the field of enhancing cooperative gaming experiences. The study of shared, virtual experiences is thus seemingly undersaturated. One can transpose existing literature from cooperative gaming experiences as the terms 'game'/'experience' and 'cooperative'/'collaborative' can arguably coincide within the same field.

On the topic of Virtual reality, Liszio & Masuch (2017) challenges the collaborative capabilities of VR. They further that actions or events happening in the real world might destroy the illusion of immersion in VR space. This holds for any interaction with other social entities outside of VR which make the real world become salient again. Therefore VR is often perceived as being a solitary technology, which isolates a single individual in an artificial environment.

The situation of Virtual Reality in general should also be mentioned.

2.3 Virtual Reality

2.3.1 History of VR

"The term *virtual reality* was coined in 1987 by Jaron Lanier, whose research and engineering contributed a number of products to the nascent VR industry. Historically, the notion of Virtual Reality is however no recent invention. In 1838, the first pair of stereoscopic glasses were unveiled, using a pair of mirrors at an angle to the user's eyes, reflecting a picture. The purpose of the research was to demonstrate the human brain's ability to combine two photographs (one

eye viewing each) of the same object taken from different points to make the image appear to have a sense of depth and immersion (3-dimensional) (Bernard, 2019). This approach of eluding the human brain into constructing 3D imagery from stereoscopic imagery carries on into contemporary headsets.

2.3.2 Contemporary VR

In recent years, the interest and development of Virtual Reality (VR) has seen a resurgence. The affordances of VR are highly dependent on the display, tracking technology and controller input(s). Contemporary VR devices emerged with the introduction of the 'Oculus Rift' prototype in 2010 (Desai et al, 2014). From 2014 to 2017, the market expanded from PC-tethered headsets (e.g., the HTC Vive) to console-tethered headsets (e.g., Sony's PSVR). Untethered headsets (e.g., Oculus Go and HTC Vive Focus) arrived in 2018, making VR an independent platform. (GlobalData, 2020). In turn, companies like Google have discontinued their smartphone-based mobile VR headset in favour of this 'standalone' approach. (Koksal, 2020)

2.3.3 VR affordances

Malaika (2015) argues that there exists a congruent standardization for 1st Gen VR systems. It can be argued that this standardization is still present in the new wave of VR systems, as the controls-schemes have remained largely the same and new experiences are compatible with 1st gen headsets (see Figure 3). To situate VR affordances in an interactive environment, a high-level taxonomy of the technology can be used to explain its affordances and limitations.

This project will adhere to 'PC-tethered' headsets for their open-ended integration with 3rd party software to allow for prototyping. The term 'Head-Mounted Display' (HMD) refers to the headset itself: a screen which uses lenses to position the image naturally to the user's eyes. Contemporary HMDs include the HTC Vive, Valve Index, Windows Mixed Reality (WMR), and Oculus Quest. VR controllers come in pairs and provide 'hands' for the user inside of VR.

Figure 3: Different VR controller-interfaces: Trigger (red), analogue movement (blue) and grip (green) - interfaces offer similar functionality.

As mentioned, most contemporary headsets have standardized control schemes, but they differ in their tracking techniques used. The Vive and Index systems use the 'inside-out' technique to achieve six 'Degrees Of Freedom' (DOF). This means the user can freely move physically around their room with controllers and HMD also being tracked by tracking stations to achieve accurate tracking (See Figure 4). The Oculus Quest and WMR use an 'inside-out' system to achieve six DOF through tracking cameras inside the HMD itself. This can constrain movement from behind the cameras and can create blind angles', but overall works on the same premise as outside-in systems.

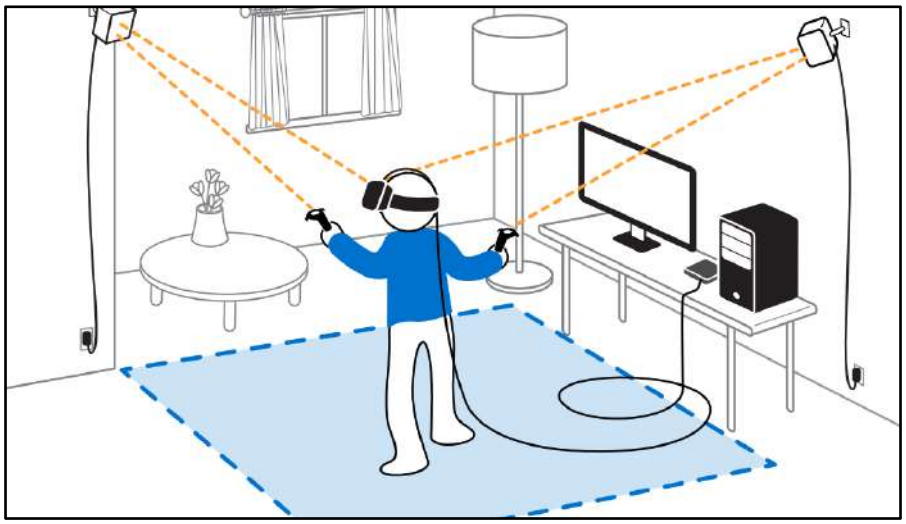


Figure 4: Example of an outside-in VR setup (HTC Vive). The blue square highlights the playspace in which the player can move around. The tracking stations translate the coordinates of controllers and HMD onto the PC.

The resulting impression of this technology offers a dichtonominal impression of *immersion* - defined as a technological quality of media engagement and *presence* - defined as the psychological experience of ‘‘being there’’ (Cummings & Bailenson, 2015). It is ultimately the goal of the VR to fully emerge the user into a virtual environment.

3. Theory

3.1 Aesthetics of Interaction

Aesthetics have historically been associated with the stimulus and beauty of visual works, meaning something aesthetically pleasing has often revolved around a level of appearance. Similarly, the VJ practice maintains a focus on visual aesthetics, while VR is concerned with interaction. To combine these fields, the study of ‘Aesthetics of Interaction’ aims at providing a sense of beauty, often associated with visuals, into the realm of interaction. The following texts aim at classifying the beauty and aesthetics in interactions. These concepts may help in making the statements in the participant evaluation and reasoning behind the concepts within this project more explicit.

Dalsgaard & Hansen (2008) argues that a central facet of aesthetics of interaction is rooted in how users experience themselves and the understanding and perception of their interaction. This is relevant as this project aims to incorporate a level of performance in the context of multiple people. Participating in an interactive experience also encompasses the experience of people around you, namely how you know others see and experience what you are experiencing. This relationship is discussed using a trichotomy of *system-user-spectators*, detailing how the system, audience, performer, and user are all part of a series of roles in interactive experiences (see Figure 5). The article also untangles the concept of *embodiment* as a foundation for context-aware computing.

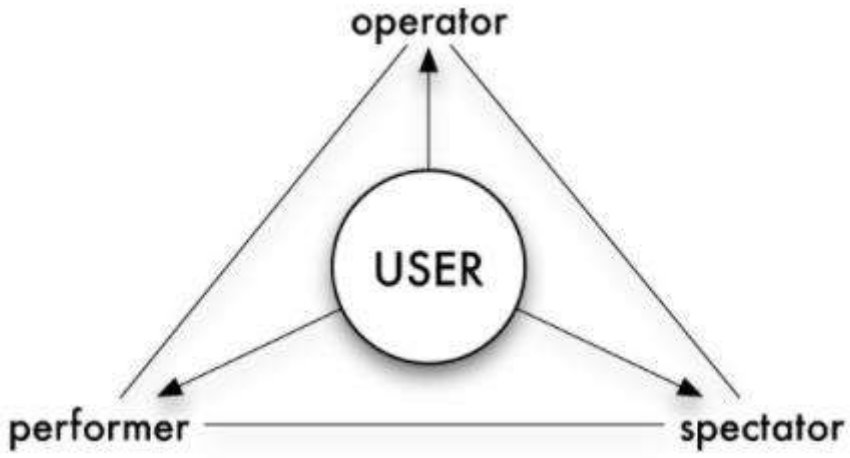


Figure 5: Dalsgaard & Hansen's "User roles".

Dalsgaard & Hansen also call for a certain terminology to be in place for describing interactive experience. The term "system" is used to denote the artefacts with which the user interacts. In this case, this refers to the virtual space in which the interactions are taking place. The term "spectator"/"observer" is used to denote a person somehow observing the interaction between user and system; this observation may be co-present or it may be mediated in various forms. In particular, the term "user" will be used for the individual testing an experience, and thus coincides with the term "operator" in this project. The "performer" is a role played by the user as the user is performing an act.

Lenz (2014) argues that with the rise of new technologies providing expanded opportunities for the possibilities of interaction, a new responsibility for designers is created. Besides the product's visual aesthetics, one must make choices about the aesthetics of interaction.

Based on a literature synthesis reviewing 19 approaches, Lenz provides an overview of today's state of the art, detailing two broad categories. One contains the detailed spatio-temporal attributes of interaction sequences, whilst the other group addresses the feelings and psychological attributes. These are categorized as needs adhering to one of three categories. The *descriptive/motor-level* (the How) focuses on a sensomotoric level of physical interaction to categorize a need for gratifyingly performing do-goals. Assigned motor-level attributes include the temporal and spatial feelings, action-reaction feedback, and presentation of information. The *descriptive/be-level* (the Why) focuses on emerging experiences and meaning on a psychological level. This level addresses psychological needs as well as emotions created and mediated through interaction. This category details the feelings of autonomy, competence and relatedness (and more) to describe the need for self-causation, capability, and intimate

contact with others. A group of attributes was not assigned clearly to one of the levels. For that reason, the same attribute could be assigned to different levels. We actually see the term *immersion* (described as an artifact's ability to "offer possibilities for quasi-physical immersion") through virtual reality technologies as a non-categorized attribute, as it addresses both be-level and prescriptive qualities, thus indicating a gap for further study into the interaction of VR-technologies.

The literature in this synthesis which has seen the most non-categorization relative to the concepts presented is the work of Löwgren (2009). This warrants investigation for the ambiguity of presented concepts. Löwgren argues for four aesthetic interaction qualities that seem to add some clarity to the task of unpacking the beauty of interaction. *Pliability* refers to the sense of malleability and tightly coupled interaction that makes the use of an interactive visualization captivating. *Rhythm* occurs when a product or service is experienced as distinctly rhythmical in the sense that it characterizes pacing of musical interaction and the flow of peripheral emotional communication. The *Dramaturgical structure* of a product should be acknowledged by designers to incorporate meaningful dramatic tensions in their designs through analysing the structure of a design. Lastly, Löwgren argues for the *Fluency* of an interaction to articulate the gracefulness with which we can cope with multiple demands for our attention and interaction in augmented spaces.

4. Method

4.1 Empirical Data-gathering

One of the initial steps for any design-process is to gather data. This section will discuss the methods used for inquiry, recruitment, and analysis.

4.1.1 Participant recruitment

For this VR-centred project, it is crucial to recruit participants who are capable of using and have access to VR technology. Whilst VR is expected to become increasingly popular amongst consumers as the technology progresses, consumer adoption of VR-headsets has not lived up to this prediction. (Larsen, 2020) Last year, only 22 percent of Danes aged 16-89 had tried smart home technologies, such as VR and smart-watches (Danmarks Statistik, 2020). The participant recruitment process would be tricky for an outside researcher to get involved with, whilst also accounting for foreseeable technical competence for setting up the intended prototype, due to this small percentage of consumer adoption. For this reason, the chosen

recruitment strategy was centred around the participation of insider acquaintances of the researcher, as it was known that at least a handful of these acquaintances owned VR-headsets with tracking capabilities. Through online messaging via media platforms such as Discord (Discord, 2021) and Messenger, (Facebook, 2021) the purpose of the project and formalities of their participation were handled. A total of six participants were recruited. The participants were all Danish males aged 18-22, residing in Region Hovedstaden. One of these participants was a relative of the researcher (more on this in the prototyping section). Five of the participants have interrelational ties, with the sixth member having no ties to the other participants. The participants have all owned headsets for a few years which should provide insights beyond the novelty of the technology. While working with friends obviously poses a different context for interviewing, it is important to not simply disregard this bias against the researcher as stated earlier by Lichtman. This subjectivity is an integral part of the interviews and must instead be considered an opportunity for more in-depth responses.

4.1.2 Semi-structured interviews

Semi-structured interviews were conducted prior to- and after testing the prototype.
A semi-structured design was used to enable a fluid method of data collection allowing for flexibility when responding to in-depth personal accounts, generating detailed and potentially rich qualitative data (Willig, 2013). Participants will be given names instead of labels to tell the stories more fluently from interview and analysis.

4.1.3 Contextual inquiry

Contextual inquiry is a part of the contextual design orientation from Holtzblatt & Beyer (2015), which reveals unconscious and tacit aspects of life. It guides designers in going out into the field and talking with people about their work and life while observing them. If designers watch people while they engage in their activities, then people do not have to articulate their practices. If they do blow-by-blow retrospective accounts of things that happened in the recent past, people can stick with the details of specific cases using artifacts and re-enactments to remind them of what happened. Contextual inquiry immerses designers in the user's whole life, including those aspects which the user does not know how to articulate.

4.1.4 Thematic analysis

Thematic analysis (TA) is a common and flexible strategy, developed as a way of identifying patterns across a data set, and interpreting these in a meaningful way. The strategy was first introduced by Braun & Clarke (2006). Braun & Clarke has since refined the TA practices to what they refer to as 'Reflexive Thematic Analysis'. The purpose of TA is to identify patterns of meaning across a dataset that provide an answer to the research question being addressed. Patterns are identified through a rigorous six-stage process of data familiarisation, data coding, and theme development and revision. Braun & Clarke are social constructivists and mention that relevant themes do not emerge from the dataset, as they are constructed by the researcher. The result of the analysis is a theme explaining people's experiences, perceptions, views, or representations of a given topic. For this reason, reflexive TA is a compatible method of analysis for evaluating an experience-driven VR prototype.

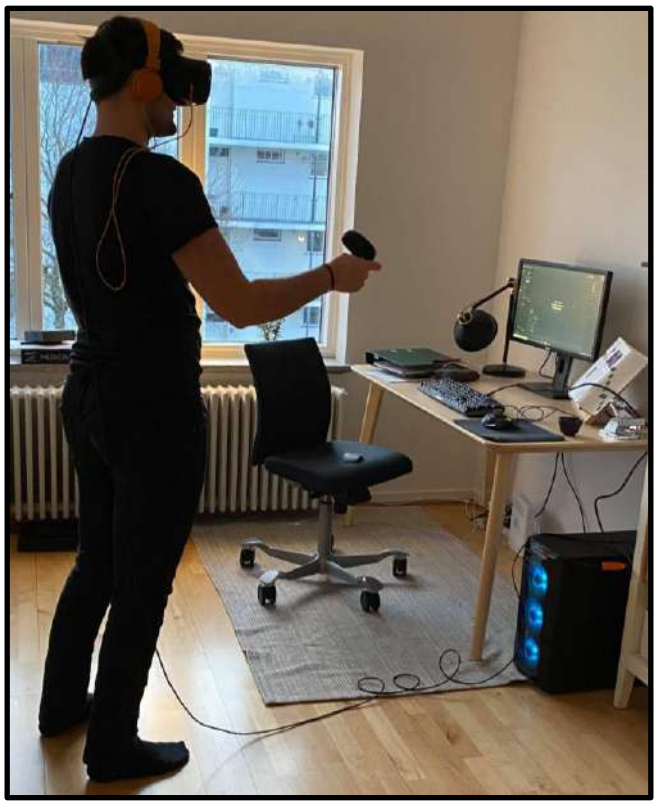
4.2 Ideation

With a scope of an audio-visual experience to include at least 6 participants, the next objective was to gather insight into the context of how participants utilize VR in practice to challenge the researcher's own perception of the medium to later inform the ideation and decision-making process.

4.2.1 Preliminary interviews

To help inform early design-decisions prior to shaping the prototype, a series of 5 one-on-one interviews has been conducted with the participants (see Appendix 4). Prior to these actual interviews, a test interview was conducted with a relative/participant who is also a VR user and will participate in the experience. Not only did this provide feedback for traditional semi-structured interview formalities, but it also revealed a key issue in gathering information: The relative/participant had not been inside of a VR headset in several months, and thus had difficulty expressing or remembering some of the experiences. The context of VR is inherently expressive and tacit, making it volatile to articulate. To combat this issue from happening in the remaining interviews, a series of cooperative gameplay-sessions were planned for the remaining participants to then be followed up by the actual interviews. This approach to interviewing adheres to the field of Contextual inquiry, where the researcher observes intended users in the context of which the digital artifact is present. In this case, the context is the VR experiences in which the participants are engaged (see Picture 2). The reason for not doing the

preliminary interviews within the context of the gameplay sessions is that it can introduce sensory overload to be questioned while attention is required elsewhere, along with the fatigueness that comes with physical exercise in these environments. (Malaika, 2015)



Picture 2: Participant inside of VR, wearing an Oculus Quest HMD tethered to desktop PC.

The gameplay-sessions were two-player experiences with participant and interviewer. The common denominator present within the games was a level of cooperative multiplayer integration to prime and inspire participants to think within the context of collaborative VR interaction. Aside from this requirement, the chosen games were improvised and thus varied from participants (such as 'Zero Caliber' (XREAL, 2021), 'Sam & Dan: Floaty Flatmates' (Pelican Party, 2021), etc). Within 4 out of 5 of these interviews the gameplay-sessions were implemented. This strategy seems to have been successful, as participants were now able to link overall experiences with what they had just experienced. A participant (Tage) explained the following:

> "For it [the interaction] to be able to play a sound would also be cool. As an example, when we played that game ['Sam & Dan'] earlier with the tablets or PDA's, when you pressed the dial button, it did nothing other than simply play a dial tune, but that was very satisfying to do".

This preference might have remained undiscussed if the participant did not have a recent memory of a similar experience, showcasing the quality of contextual inquiry. The following section provides a summary of remaining participant insights.

4.2.3 Insights from preliminary interviews

The preliminary interviews revealed insights into existing practices of VR interaction. While these interviews were not analysed, they provided inspiration and information for the following ideation.

Firstly, it was expressed that every participant goes through the 'SteamVR' platform to access content. Secondly, none of the participants had tried dedicated audio-visual experiences in VR before nor had they any experience with the practice of VJing, making it difficult to directly inquire about previous experiences in that regard. The participants had different motives for owning and using VR HMD's. Most prominent factors were aspirational towards the interaction capabilities of the technology, with participants appreciating the 'little details' that dictate the space in which a weapon is manipulated or how players can physically interact with each other. The virtual worlds themselves were also described as providing *immersion* and when discussing the social capabilities in VR, participants highlighted areas in which VR multiplayer experiences differ from the 'flat' media counterparts. The level of competitiveness was described to be lesser than usual, as the rules and objective of these games were shifted to socio-spatial gatherings in which participants were more concerned with experiencing the intricate details of the games together, knowing that their friends were situated in equally atypical and silly physical dimensions as themselves.

The participants also verbalized their expectations for the upcoming experience. A desire for virtual representations of concert experiences mimicking their real-world counterparts was mentioned among participants.

4.2.4 Sketching/idea generation

From the preliminary inquiries, several ideas and impressions popped up, detailing ways of creating an experience. The first idea was based around the idea of integrating Löwgren's *dramaturgical structure* of an interaction (see Figure 6, left). The dramaturgical structure should be acknowledged by designers to incorporate meaningful dramatic tensions in their

designs using varying levels of dramatic tension. Löwgren references the use of this within *parafunctionality* in critical design. Parafunctionality is design artifacts that clearly reveal their function – but on reflection you find that the function revealed is somewhat anomalous, paradoxical, inappropriate, or in some other way critical of our tacit assumptions or our genre expectations. Using anomalous interactions to provide visual effect, while also creating dramatic tension, seemed fascinating to incorporate into an audio-visual context. The VJ is often using the MIDI controller for their performances, and VR enables a more abstract range of interactions for visuals. Löwgren furthers that users experience parafunctional design by usually following a structure in three steps: 1) Simple recognition of the product and its intended function. 2) A period of frustration at the obvious inappropriateness of the intended function, and 3) A sudden insight (the »a-ha« moment) when you realize what the artist-designer wants to make you see.

The researcher will explore the possibilities of this in the prototyping section.

The second idea (see Figure 6, right) was mocked up to explore how existing VJ tools could be represented more realistically. The MIDI controller is again prominent, and the idea of a localized button/slider interface seemed interesting as a counterpart to the abstract nature and scale of the first idea, being more tangible overall. This idea was integrated as an 'Alien button station' as described in the 'Technical description' section.

Figure 6: Two idea-sketches detailing the overall layout of possible experiences - Room #1 (left) and Room #2 (right).

4.3 Prototyping

Prototyping is a key activity within the design of interactive systems. Prototypes are created to inform the design process and design decisions such as what the product should look-, behave-

and work- like, to explore and communicate propositions about the design, and its context prior to developing the final product.

To define exactly what this prototype should achieve, Houde & Hill (1997) argues for three different corners that can be saturated to various degrees when prototyping (see Figure 7). These corners consist of the 'look and feel', 'role', and 'implementation' of a prototype.

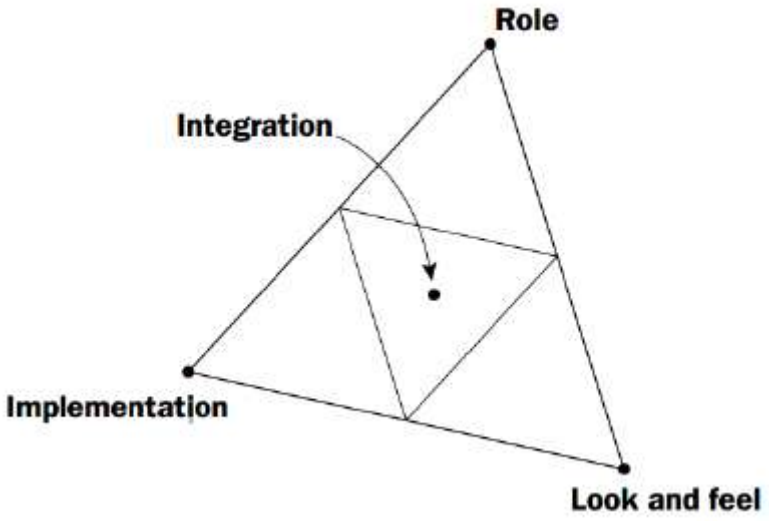


Figure 7: Houde & Hill's four principal categories of prototypes.

When all three corners are saturated, the prototype more closely resembles a final product. When this is the case, the prototype is referred to as an integration prototype.

4.3.1 Experience Prototyping

Suri (2000) argues for the "Experience Prototyping" strategy as a means of emphasizing the experiential aspect of representations which are needed to successfully (re)live or convey an experience with converging hardware and software, spaces, and services. An experience prototype is designed to understand, explore, or communicate what it might be like to engage with the product, space or system that is being designed. Suri relates this strategy to Houde & Hill's "look and feel" of a product, or system, yet Suri argues that experience goes beyond the "concrete sensory", and also asks questions about the 'role'. Suri also emphasizes the important influences of contextual factors, such as social circumstances, time pressures, environmental conditions, etc. Suri also argues that "A true Experience Prototype for users — providing a really relevant experience seems to require a level of resolution and functionality such that it can be 'let loose' into an everyday context and more fully integrated into people's lives.". This strategy is executed in three critical design activities: understanding existing experiences, exploring design ideas, and in communicating design concepts. The first of which was

addressed with the preliminary interviews, while the others will be at the centre for the prototyping-process.

4.3.2 Unity as a prototyping platform

Since this project aims to deliver an experience centred around a VR experience, one must find a medium to develop such experiences. The two leading free game development platforms, Unity (Unity Technologies, 2021) and Unreal Engine (Epic Games, 2021), have recently adopted and matured its support for VR game development. Unity was found to be the least complicated to get into due to previous experience from the researcher. The Unity platform is also well-established as an entry point for game development- and learning, as it features an inspector and 3D-viewport as a visual aid while scripting features (See Figure 8, right). It is possible to create an experience in Unity with little-to-no knowledge of scripting using the Unity Asset Store: an interface for purchasing or downloading free user-submitted assets to be used within a project (See Figure 8, left). This allows for a more rapid development process but can also backfire as some systems are not compatible with others, requiring more development time. A combined method of scripting and using premade assets were chosen to best accommodate this.

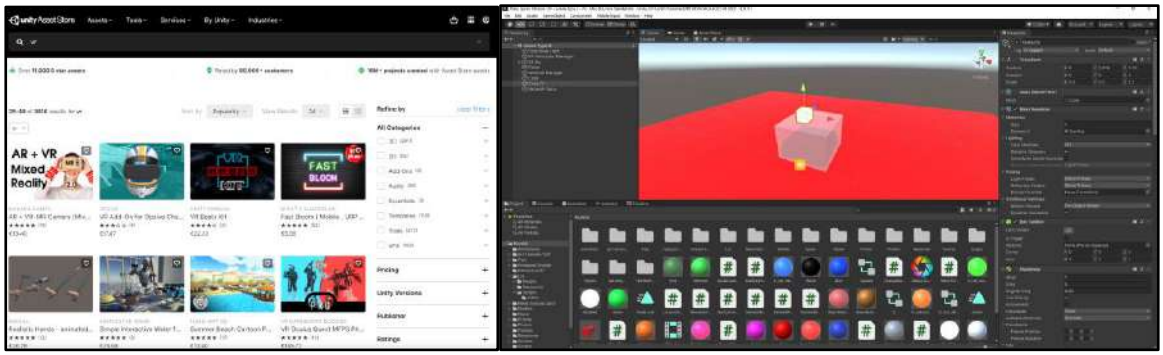


Figure 8: The Unity Asset Store (left) and a screenshot from the 3D viewport inside of Unity (right)

The researcher met all of the prerequisites for developing a VR prototype: owning a VR headset (the HTC Vive) and a capable windows-based computer. The following section will detail the development-process.

4.3.3 Experience prototyping together

In VR there is no substitute for observing and playtesting (Malaika, 2015). Ideally, it would be preferable to reach out and test new features with all participants. This would, however, put a great strain on the participants, as the need for testing arises often. It would also influence the

final experience to a point where it would be less meaningful. For example: Löwgren's 'dramaturgical structure' would suffer as participants would have already uncovered interactions, eliminating opportunity for an »a-ha« moment as it has might have already happened in testing. Fortunately, one of the participants was the researcher's brother, of which they share accommodation and an identical HTC Vive headset. This meant that the testing could happen with a higher frequency, but also initiate discussions between researcher and tester, which is what situates the prototype as experience prototyping. These discussions lead to exploring new design ideas while also communicating new concepts iteratively based on an experienced user.

A basic system was set up and tested prior to the multiplayer integration, to allow for full six DOF movements. The multiplayer process was guided by a 3-part YouTube tutorial (Valembois, 2021) that details introductory steps towards full multiplayer integration. A blue head and a pair of hands were set in place to visualize where the player HMD and controllers were situated respectively (see Figure 9). A multiplayer-management package was utilized to achieve this.

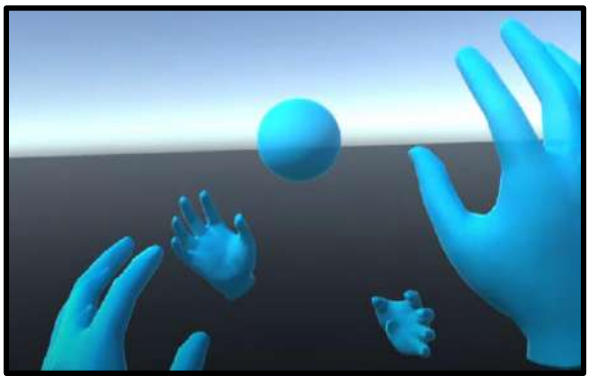


Figure 9: A screenshot of the first 'point of contact' between players inside of Unity. A blue head and floating pair of hands were chosen to represent player presence.

With the multiplayer-framework taken care of, the next step was to add interactable objects, so that the participants could throw objects and push buttons around the map. This is done in adherence to Lenz' described motor-level need for *Forces*, i.e., adjusting the effort required to interact with the system. The added technicalities of multiplayer-synchronization made it much more difficult than expected to implement new concepts, as there was often a mismatch of information which would/would not get sent from the client to the host. The third step was to add a movement system. A simple teleportation system was chosen to avoid motion-sickness often found with continuous movement, and to prevent participants from wandering out of bounds (Malaika, 2015). The fourth step was then to implement the abstract gesture system, as

depicted in the ideation phase. Models for the 'stars' were created using Blender, (The Blender Foundation, 2021) and an existing 2D gesture recognition script was edited to recognize drawings in 3D space to then enable spawning objects related to corresponding gestures, to fully utilize the affordances of the six DOF.

4.3.4 Technical description

The technical description of the prototype aims to describe the different available activities inside of the prototype.

The program is an executable file which can be run from any windows-based PC. Upon launching the executable, the participant is met with a lobby-screen to then 'log on' to the server and load in the main scene. As labelled in Figure 10, the participant loads in next to a crashed space-shuttle (1). A 'skybox' video-texture wraps around the scene which depicts the landscape of an alien planet (see Appendix 3). The participant is then free to explore the rest of the scene using the teleportation and interaction system. The intended route to take is: Equipping a hat from the 'hat' station (2) to identify participants from each other. Secondly, the participant would then pick up a 'powercell' from the adjacent station to collaboratively decide on which interfaces should be usable (3).

The participants would then have the option of exploring the following: (4) a button interface which transforms the properties of an alien figure (i.e., the green buttons change its material). Each of the six participants are required to place their respective power cells inside the interface to 'power' its functionality (cells can be placed inside the semi-transparent blue cylinders). The participant can also see a disproportionate 70's TV displaying a static image (5) and three smaller static TVs with an adjacent remote (6). Upon placing two power cells under the remote, participants can use the remote to change channels on each of the three TVs, cycling between looping clips of pop-cultural shows from the 80's and 90's. Lastly, the participants can look up to the sky and see stars indicating gestures which can be drawn to trigger visual effects ranging from small rotating props to scene-wide lasers and abstract shapes (7). After initial exploration the researcher would then turn on the song using a keyboard command. This would start the song 'Starman' by David Bowie inside each of the participant-headsets. Additionally, the largest TV would now display a live-performance recording of said song. After the song has played through, the experience is over.

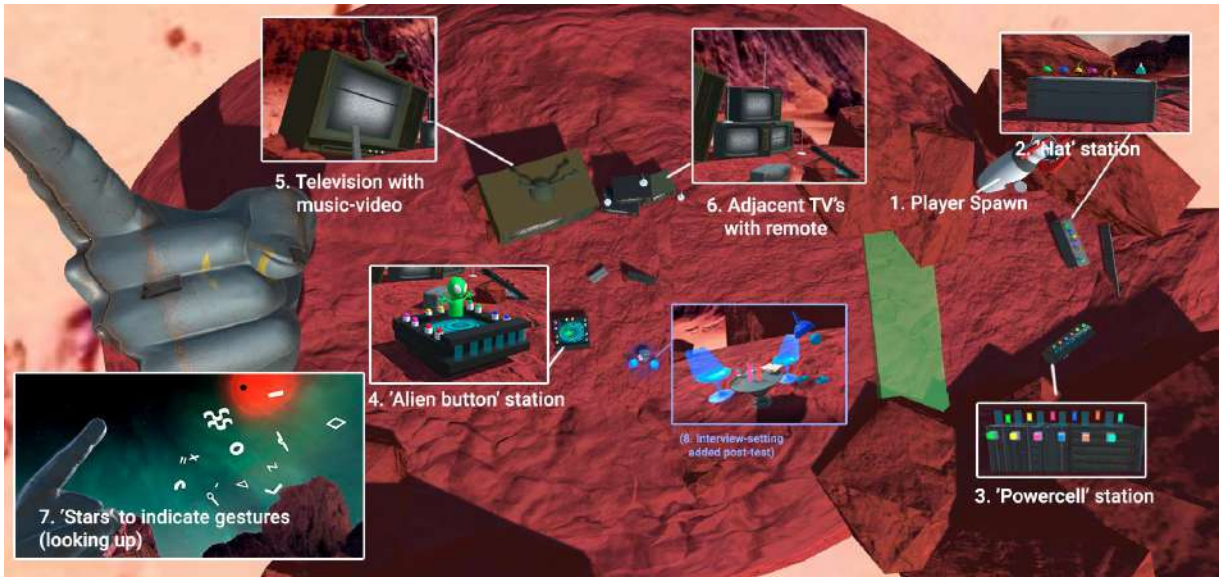


Figure 10: Top-down overview of the interaction possibilities of the prototype. The interview-setting (8) will be discussed in the 'Evaluation and Data Collection' setting.

5. Test

The final showcase on the prototype happened March 27th. A total of six participants were present for testing. Every participant could 'log onto' the application and explore the prototype as described in the technical description. There was a fair deal of technical abnormalities which had not shown up in prototyping. This will be covered in the analysis. The test lasted for approximately 45 minutes. The test was documented using screen-recording software as the researcher was also present in VR. The following sections will detail the process of evaluating and analysing the test. The tested prototype can be downloaded from the attached appendix (see Appendix 2)

6. Evaluation and Data Collection

Immediately after the showcase came to an end, the plan was to do a follow-up interview in Discord with some of the participants. The participants, however, wanted to discuss the experience in plenary inside of VR to express their immediate thoughts. This improvised interview was conducted inside of the researcher's 'SteamVR Home' hub-area in which players can connect in different settings represented by their avatar. Aside from a few remarks, this did not prove effective as the hub-area offered too many distractions for participants to direct their attention towards (see Figure 11). The verbal communication and presence afforded by the avatars, however, proved to be a very engaging approach to contextual inquiry as gestures and head movement aided communication. For this reason, the intended follow-up interviews were altered with this approach in mind, albeit with less distractions.

Figure 11: A display of the distracting interactive possibilities in the first interview using 'SteamVR Home'.

The final round of one-on-one interviews were conducted post-test. As opposed to the preliminary interviews which were conducted *after* the gameplay-sessions, this evaluating round was conducted entirely *within* the VR prototype, with the participant and researcher both wearing their respective HMD's. The purpose of doing this is primarily to allow for the participant to remember and express insights that may have otherwise been lost, by pointing to- or re-enacting details in the VR scene. This is also in line with what the contextual inquiry data-gathering technique aims to reveal: unconscious and tacit aspects of life of which the participants cannot describe, like inner motivations such as the need to express a particular identity. This is especially prevalent given the timing of these interviews, as they were conducted within three weeks of testing the prototype.

The VR prototype has also been altered to better suit an interview-setting. While existing systems stayed in place, a new set of props were added to the scene. Two chairs and a table were placed in the middle of the scene, accompanied by more novel items such as bottles and cups, to mimic the spatial dimensions of a regular one-on-one interview (See Figure 12, left). The participants were instructed to sit in the virtual chair and align it to the position of a real chair so that participant and researcher were now sitting across from each other while also sitting down 'for real' to further cement the setting. Additionally, two clipboards were modelled and inserted which contained the two halves of the interview-guide, with the text only visible to the player holding it (See Figure 12, right).

Unlike the preliminary interview-guide which was mainly concerned with participant practices and experiences within VR, this interview was constructed to acquire information about the prototype and to evaluate the effect that the experience had on the participants. The interview-guide was constructed using a mixed approach of observations from the test and existing theoretical frameworks regarding aesthetics of interaction. The framework aids the researcher

into asking more in-depth questions of tacit experiences to avoid and elaborate on more superficial answers. Additionally, the observations help pinpoint the exact events that require questioning or explanation.

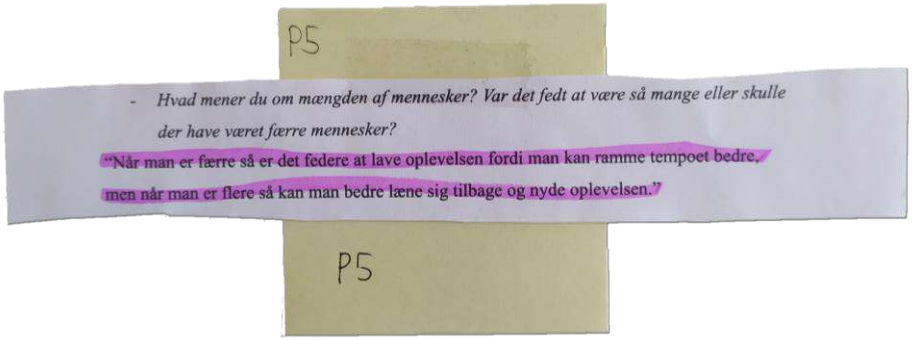
Figure 12: The interview-setup (L) and the interview in progress with another participant (R)

The interviews were all recorded using a screen-recording tool capturing the VR headset footage from the researcher's perspective mirrored to a screen. The interviews were then transcribed using the 'intelligent verbatim transcription' (Bailey, 2008). This transcription method omits filler-words, pauses and grammatical structure, whilst still retaining the content of the interviews. This has the added benefit of making an already substantial amount of data more compact and easier to manage, but also requires the researcher to be aware of the context when analysing, as it is less obvious what is verbally expressed. A total of 5 interviews were conducted post-test and this produced approx. 38 pages of transcribed interview-data (see Appendix 5). This data will go on to be analysed in the following section.

7. Analysis

As mentioned in the methods section, reflexive thematic analysis was chosen as the method for evaluating the final round of interviews. This analysis adheres to the six-stage process for doing TA as suggested by Braun & Clarke (2006). The interviews were all transcribed in succession in a timeframe leading right up to the analysis, to maintain this familiarity with the dataset.
While this thematic analysis is informed by an interaction-framework, it is a mixed approach that does also consider the observations and preliminary interviews. The subsequent stage to familiarization with the data is to identify important features of the data that might be relevant to answering the research question. These features are to be used for generating succinct labels called codes. Interesting quotes were highlighted with a purple marker to then be cut out and collated in the next stage as codes. The quotes were cut out in larger chunks, to retain the ability

to read them separately in the correct context. In addition to cutting out the codes, they were also taped to post-its featuring their respective participant numbers (see Picture 3).



Picture 3: The appearance of one 'code' - participant numbers were substituted with fictive names

The selection process was based on the familiarity of the data and the importance relative to the overarching research question. As an example: the quote "VR is still a niche thing so I do not know how many will be able to participate if it becomes very demanding for a PC to handle" (John) was sorted out at this stage, as it did not serve a purpose in relation to the research question. Quotes such as "It was great knowing that people see what I am doing, and I can see what people are doing" (Oscar) revealed an insight into the user-relationship within the experience and was thus interesting enough to continue with. Braun & Clarke also notes that through theoretically informed case studies, it makes more sense to emphasize relevant themes than to provide a rich description of the dataset in its entirety. A total of 135 codes were generated from this process (See Picture 4)

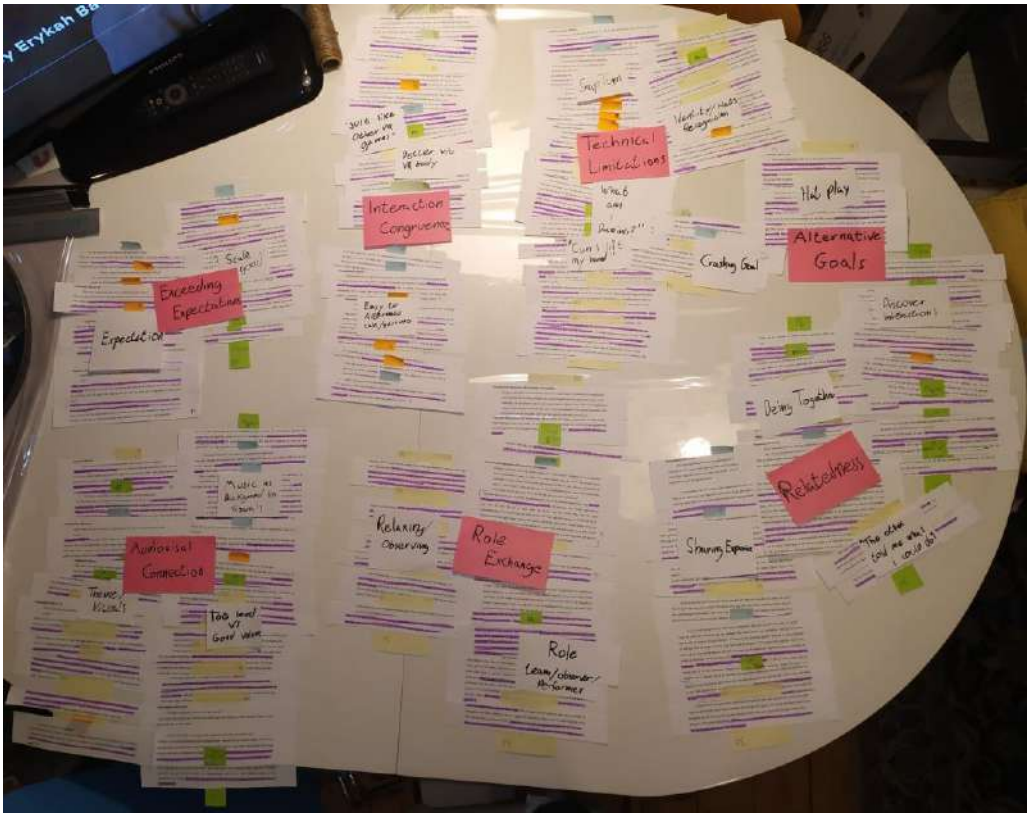
Picture 4: Second stage. highlighting codes, cutting them out and labelling the participant number.

The third stage was then to generate initial themes from categorizing these codes into broader patterns of meaning to become potential candidates for themes. As there were a lot of codes on the table (literally) the assigned initial themes sought to condense some of codes into matching statements, layering, and grouping said statements into clusters. Example: The "Relaxing/Observing" cluster includes statements from Tage and Dennis that summarize a desire to relax for different reasons (such as the number of people, intensity of the music etc.). A total of 21 initial themes were constructed in this stage (See Picture 5). While there are defined stages to this process, the process is inherently non-linear and iterative, and sometimes codes swapped places to reveal new initial themes, in later stages of the analysis.

Picture 5: Second round. grouping the codes by what they reveal, forming descriptive labels.

The fourth stage aims to review these 21 candidate themes to further determine if a convincing story can be told through discarding, refining, and merging the different candidates. More in-depth analysis revealed many points in which the candidates were related to or overlapped with each other to tell a coherent story. Example: the "Relaxing/Observing" candidate theme from earlier shares a similar story to what is being mentioned in the "Role - team/observer/performer" candidate theme. Dennis and Tage's story of relaxing in the company of others ties in perfectly with the experience of Arne: transposing the experience to New Year's Eve, describing how it would be tiring if he had to do all the fireworks [visuals] himself. As the theme candidates both discuss some sort of relational shift in the number of people performing and observing, a fitting theme to encompass this would then also be the

'Role Exchange". This approach was true for the other categories for the fourth and fifth stage of reviewing and defining themes (see Picture 6)



Picture 6: Third and last round. Grouping the descriptive labels into themes.

None of the 21 candidate themes were discarded, they were instead merged into 7 final themes. The 7 final themes were: *'Relatedness', 'Alternative goals', 'Audio-visual connection', 'Exceeding expectations', 'Role exchange', 'Technical limitations'*, and *'Interaction congruence'*. The stories from these themes will be explained in the next section.

8. Results

This following section detail the stories and insights that follow each of the constructed themes, relating back to the focus of this project in the context of virtual/physical concert-going and aesthetics of interaction:

8.1 Theme: Relatedness

The 'Relatedness' theme is one the most saturated themes. This story is central for this project as it details the relational factors between participants as they interact in the prototype and share the experience.

Relatedness is a reference to the 'descriptive/be-level' term in Lenz' synthesis, set in place to explain a psychological need for the "feeling of intimate contact with people who care about you." He furthers that, because of greater focus on the "product" in an interaction, "relatedness" from a product-centred view is often misunderstood as referring to the relation one has with the product itself rather than referring to the experiences of relatedness it may create or mediate. With this focus in mind, initial inquiry revealed that all the participants believed that their experience had been more of a shared experience than a solitary experience with their friends present. Despite having no relational ties to the other participants, Dennis explains:
"It was cool because we were all experiencing it at the same time. They, of course, knew each other, but it actually did not really matter because we were together, experiencing this". He compares the experience to that of a concert, where even though he was not talking much with the others, they still shared the same experience. The participants explained that the 'fun' and the entertainment-value they achieved from the prototype would have been much lesser if they had experienced it by themselves, even to the point of being 'boring' or them feeling 'lonely'. Participants also explained that 'sharing the experience' was a major factor in experiencing it together. John exaggerates that it was simply the thought of the others being able to see the effect he drew which created the community behind the interaction. Arne says that if he was not able to share the experience, it would be like doing the experience by himself, which would negatively influence the experience as noted before. Sharing of contemporary physical concert-experiences are arguably different in this regard, as concertgoers are motivated by a sense of self-realization, such as a status symbol when proclaiming statements such as "I was at Woodstock" (Charron, 2017). Dennis goes into great depth explaining his reasons for sharing VR experiences:

> "It is one of those feelings which are unattainable on 'flat media'. It is like you are there. [...] This happens because the experience was authentic for me in a way. This experience reached that level and that is something special." (Dennis)

For him, it is the VR technology itself which can make experiences authentic, thus increasing the incentive for sharing it. Oscar explains a phenomenon in the preliminary interviews, attributing the likeliness or sharing the experience to presence of his friends as the event unfolds:

> "It is like going for a bicycle ride where you experience a sunset or something. If there had been something unique about this bicycle ride, then it would always be something to talk with your friends about if you had experienced it together. It is almost impossible to explain it to others with all these impressions. It is easier to digest with others." (Oscar)

This need for sharing experiences can be attributed to Lenz' psychological or *be-level* need for meaning: "Feeling that you are developing your best potentials and making life meaningful". The act of sharing a memorable experience seems to affect how meaningful an experience is.
Dalsgaard & Hansen brings up Goffman's (1966) classification of relations in situations as being either focused or unfocused. Focused interaction is "the kind of interaction that occurs when persons gather close together and openly cooperate to sustain a single focus of attention, typically by taking turns at talking". Unfocused interaction is when people are in the same situation but without interacting, even though they are still somewhat aware of each other's presence: "In this realm of unfocused interaction, no one participant can be officially "given the floor"; there is no official centre of attention."
Goffman's account for focused and unfocused interaction proves to be important when understanding how interactions take shape from the social situatedness. The notion of 'sharing the experience' initially exhibits strong modus for unfocused interaction. As John mentions, the thought of people seeing his visuals was enough to keep his idea of a community thriving. Dennis's metaphor of a concert-like connectedness similarly implies that it is the implicit and tacit acknowledgement of being present together that shape the social sense of sharing the experience, seemingly transferring it to something underlying and subconscious.
Dalsgaard & Hansen argue that Goffman's concepts contribute to an understanding of interactive systems as something that is *always* about focused interaction. One is always supposed to engage with the system, making it the official centre of attention (see Figure 13). For example, Tage mentions that he was made aware of the gesture-system by some of the others mentioning it.

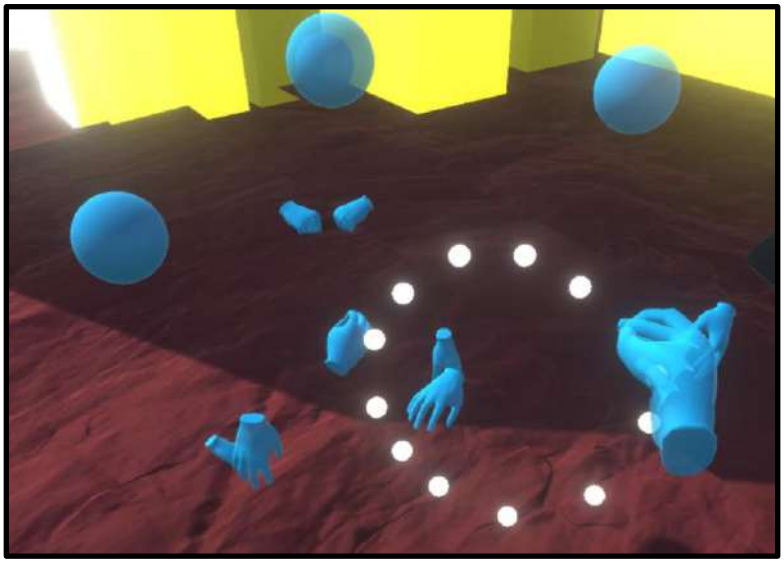


Figure 13: Screenshot from the experience showing three participants in dialogue with the researcher

Dalsgaard & Hansen further that some systems can point out how it may be true that you are in a focused relation to the system, but this focused interaction is part of a focused interaction with the surroundings. This transforms the user experiences into a double focused interaction. The user is focused on both the system use and their relation to other people present, meaning interacting with a system is no different than interacting in other cultural settings. Oscar furthers that even though he felt as if the others had no influence, they still influenced a way of feeling of doing something together: anticipating something to happen. As Oscar states, this feeling of being together has likely also accumulated outside of the prototype, anticipating the experience, chatting with the other participants leading up to the 'big reveal' of the researcher's prototype which had been discussed for multiple weeks prior. This temporality lends itself to the anticipation-factor of concert-like experiences, which can be considered a positive reinforcement in this situation, as anticipation can harbour excitement.

In summary, this theme reports a strong indication that the psychological need for relatedness was saturated. VR affords authentic experiences comparable to that of a concert. This, combined with a focus on creating visual effects individually, incentivises sharing the experience, thus promoting this feeling of relatedness.

8.2 Theme: Role exchange

As briefly mentioned earlier, the role exchange tells an interesting story of how the roles fulfilled by the other participants determined the role of the individual.

This theme is based on inquiry about Dalgaard & Hansen, particularly their trichotomy of system-user-spectators. They conceptualize this trichotomy holistically, implying that the *embodiment* term maintains a constant reciprocal relationship with the context (with meanings ascribed to both users, system, and physical surroundings). In the interviews, this trichotomy was discussed when participants were asked to situate their *role* in the experience, referring back to the user roles discussed in the literature.

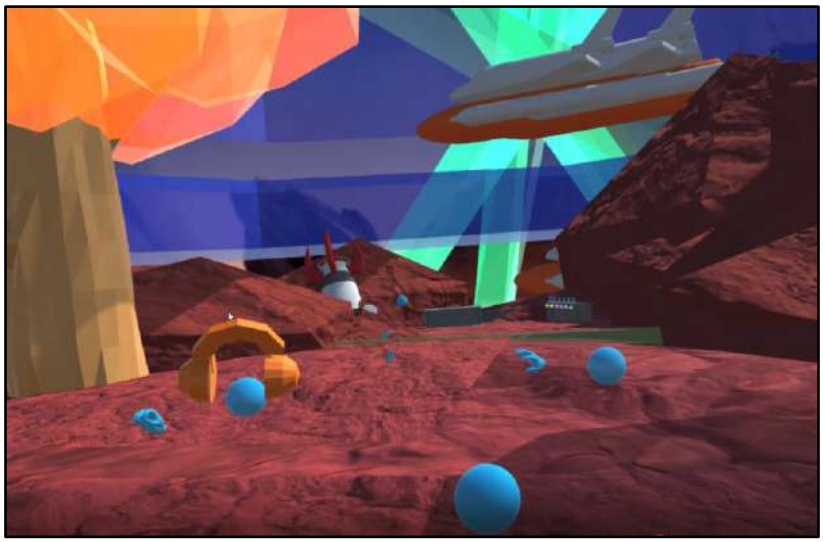


Figure 14: Participants change between viewing and creating visuals.

None of the participants felt like they were exclusively adhering to one of these roles (spectator, observer, operator). The term 'operator' simply refers to the users interacting within this interactive system and thus remains a stagnant point of discussion throughout, as every participant is inevitably operating within the experience. Participants all expressed some level of ambiguity and alternation in their respective roles as either performer or observer. Dennis explains that he was not just a listener, but that he played a part in the experience as well (see Figure 14). In some cases, he would perform for the others and verbally direct their attention towards his creations. He would, however, also utilize the tracking capabilities of VR to physically lay down on the floor, while observing the effects created by other participants, relaxing and refraining from using gestures. Tage had a similar experience to that of Dennis:

> "It was great to not having to worry about drawing because there were so many people. In this way, I can just admire [the effects] and instead listen to good music. In the beginning i felt like it was nice to draw and create effects, but there was a time where i felt like it was nice that i did not need to do anything" (Tage)

Tage started the experience as a performer. He then exchanges his role to an observer, based on the engagement from the other participants, filling his role as a performer.
Arne described his performer/observer relation using a metaphor:

> "It was a bit like New Year's Eve, to see the show in front of you. This makes it so that you decide yourself if you want to go down and fire off rockets. I did this as well and made lots of visuals that also received attention. I liked the idea of people seeing what I'm doing and vice versa. The attention went both ways, which definitely enhanced the experience." (Arne)

This tells a similar story of alternating roles and provides an interesting perception of what the participant considered his own role to be. It is implied that the attention given/paid positively reinforces his motivation. To gain an insight into the collective role of the participants, Oscar explains:

> "I feel like it was almost like sitting backstage at an ongoing concert, where you sit and control the light effects. The actions you were taking by yourself were something you presented to other sound/light-artists. The short answer is it was like being a crew" (Oscar)

Dennis shares this view of the experience, exclaiming that the relation was like a team building a setting for an experience. Compared to the traditional concert audience, the experience differs, as sitting backstage with a crew obviously testaments the shared nature of the experience. It also implies a lesser focus on perceiving the musical experience.
The design of the spectator experience is necessarily a question of not just how spectators perceive the performance, but also of how users simultaneously perform their own perception and perceive their own performance (in relation to the other performances). In this case, the user is both a performative spectator and a spectating performer.

8.3 Theme: Technical limitations

Technical limitations and errors played a varied part in how relations and interactions were shifted, but also how they contributed to reveal underlying insights. This theme is quite comprehensive, but this is mostly due to a consistent repetition of technical limitations from

every participant. The candidate themes within this theme help sum up some of the key insecurities/confusion that the participants felt throughout the experience because of technical limitations with the Unity prototype.

8.3.1 Autonomy and insecurity

Every participant commented on their frequent inability to distinguish what exactly they were drawing using the gesture system. Oscar says: "You knew there was a gesture, but you were not completely certain if it was yourself that had made it, because 6 other people were creating similar gestures".

This lack of identity can be attributed to two technicalities, the first being the generic implementation of visual effects. To elaborate, each of the participants inherit the same interaction possibilities as they load into the prototype. It is insufficient in the way visual effects are portrayed identically independent of the participant activating them, as the effects all spawn from the same coordinates and in the same colour. This results in a 'layering' of effects making individual effects difficult to identify. This would not have been a severe problem; in fact it was intended to promote collaborative patterns using this layering method. This identity problem was made worse by another technical difficulty, which was that the visual effects would not always disappear after their designated intervals of 1-20 seconds, some would even stay indefinitely. This made the problem of identifying visual effects a far greater obstacle, which only worsened as the experience progressed. The origin of this error was not discernible after the fact but it was deemed likely due to a network instability. Using Lenz, this had a noticeable effect on the be-level need for *autonomy* of the experience. The autonomy refers to the feeling of being in control of one's own action rather than feeling pressured by external forces. These errors directly inhibit the fulfilment of this need as John describes:

> "In terms of the audio-visual, it all started out great with the gestures and all. It then turned into visual noise when we realized that things were not despawning very fast. I feel like I had a greater impact in the start than towards the end [of the experience]" (John)

8.3.2 Gesture annoyances

Ultimately, the sky was flooded with visual effects, which was not the intended result. This was not seen as an immediate problem to every participant though: Tage took a liking to the 'trippy' aesthetic which had emerged which also relieved him of having to perform himself (as mentioned earlier), while Arne and Oscar created an alternative goal for themselves to spawn as many visuals as possible (more on this in the 'Alternative Goals' section)

The participants suggested the implementation of a motor-level action-reaction feedback-loop, using sound/visual cues to let them know if they performed the gesture correctly. This approach was argued to introduce more autonomy to the interaction.

Figure 15: Screenshot from the test showcasing the 'trippy' visuals

8.4 Theme: Exceeding expectations

This theme will discuss the effect of juxtaposing participant expectation for the experience.

The participants all had some sort of expectations as to what this experience entailed prior to experiencing it themselves. This theme can actually be traced back to the preliminary interviews in which John and Oscar expected a realistic concert-like experience. John expressed several desires that mimic the concrete physicality of concerts: "I would like to see glow sticks in VR, or being able to press a button and then fireworks are displayed".

Similarly, Oscar had also set up expectations for a representational concert experience:

> "I am thinking that it [the audio-visual experience] would be some form of concert-experience, it is what makes the most sense to make a projection of. [...] The visual is

> often part of defining the genre, so if you were listening to EDM, there are most likely some laser beams present" (Oscar)

Oscar was surprised over the scale and 'focus elements' that he had not expected. He mentions that he now sees that VR has the capability of creating a much more impressive visual experience, but that the sensory experience of physical concerts (such as sweating, smelling alcohol, freezing, and touching) remains unmatched by VR.

Participants were most positively surprised by the amount of interactions present, along with the scale of the experience. The scale mostly refers to the spatial surroundings, such as the massive finger, and oversized Television, but also in a more epic sense. Dennis attributed some of this feeling of scale to the volume of the music, matching that of a loud concert overwhelming the senses. This reveals a comparable trait of traditional and virtual concerts in terms of this notion of scale.

Löwgren's notion of dramaturgical structure (discussed as part of the 4.2 ideation phase) was also used to increase dramatical tension of the interactions, set in place to surprise participants when they figure out the intention behind the anomalous interaction of drawing visuals to fill the sky. Participants seemed to express a dramaturgical course of action which replicates the intended design. Dennis even mentions the >>a-ha<< moment associated with parafunctional design:

> "I ended up understanding some of them [the gestures]. Some of them I did not end up understanding, like the popsicle because I thought it was a tree. It was exactly a type of a-ha moment when you then figured it out" (Dennis)

In this case, the dramatic tension may have enhanced the act of discovering gestures by making the participant experience this sudden insight that leads to a meaningful interaction.

In summary: Creating abstract worlds and abnormal points of interactions may have a positive effect on participants, as it exceeds, or at least juxtaposes expectation, allowing for a more meaningful experience than what was expected.

8.5 Theme: Alternative goals

'Alternative goals' tells a story of the unforeseen and improvised interactions that aspired from the participants' own initiatives. Throughout the prototype, the researcher noticed strange

abnormalities. Participants were losing their designated hat-identifiers, and the sky was immediately flooded with an overwhelming number of visual effects. While the researcher suspected this to be attributed to technical errors, it is revealed that there is participant causation for at least some of these abnormalities.
John explains that:

> "[A participant] chose to drop his blue hat and then I went to pick it up. This then turned into a game of tag with the hat. He ran off and I teleported after him, picked the hat up, and tried my best to get a hold of it. This happened while the others were drawing. We were playing a game throughout [the experience]." (John)

In Löwgren's terms, one may attest this to the *pliability* of the prototype as the interaction feels coupled and responsive enough to invite the participants into further experimentation, creating another play-space inside of the prototype. VR hand tracking in general suits the quasi-physicality used to great pliability effect as it introduces new dimensions of malleable objects, such as the hats that can now be thrown, hidden, and traded among participants.
Oscar also explains that teasing other participants and "stealing their hats and throwing them out of bounds of the map" (see Figure 16) is something you "Have to do. It is a part of it." implying that it is not exclusive to this experience, but a reoccurring initiative to interact with other participants by teasing or annoying them as a sort of goal to be fulfilled in tandem with the intended objectives.

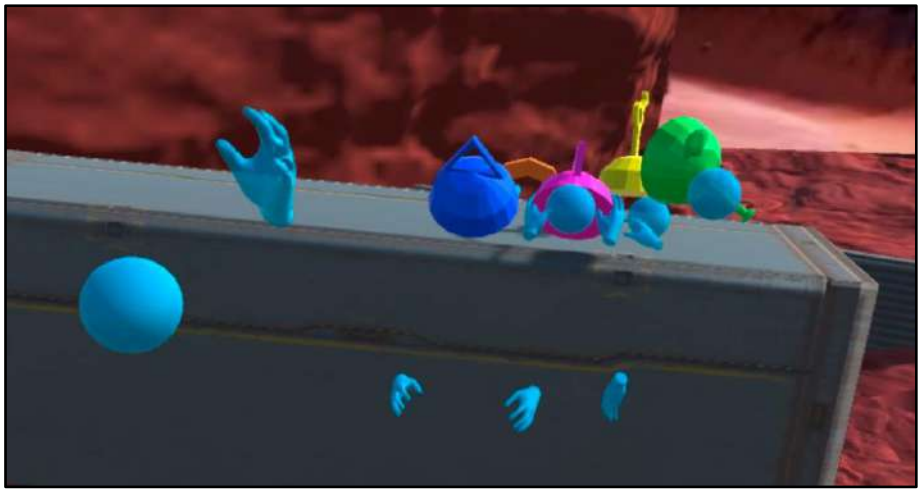


Figure 16: Screenshot from the test showing participants stealing and exchanging each other's identifiers (hats).

The second abnormality, seeing the sky being filled with objects, may be somewhat attributed to the limits of the server, but it also contributed to a second alternative goal which materialized from Arne and Oscar's collective efforts. Arne explains:

> "We set a goal for ourselves to see how much we could overload the server. I felt like we had millions of those arches going. [...] We were really not able to do it [crash] but it was fun to set a goal for yourself and have fun with these things". (Arne)

Using Lenz, this also reveals how participant *autonomy* affects participant motivation. The level of freedom and feeling that the participants were manipulating the system, and not pressured by external forces, led to a collective pursuit of this underlying goal. On the flip-side, John found this 'spamming' of visuals to affect their own autonomy which, in turn, made them feel less in control of their actions.

John underlines that this way of testing the system is not exclusive to this experience, nor is it exclusive to experiences with others. "It is something I normally do with every game, trying to break it and test it to its limits. [...] I see it as a standard practice. I am thinking 'How many arches/spaceships can I get in here before it [the system] crashes."

With regards to this User <=> System relation, Dalsgaard & Hansen argue that, when discussing the embodiment that takes place in an experience, meaning is not a constant, it arises from user interaction with the system which implies that the designer cannot control what the system means to the user, but can only influence the construction of meaning furthering that what a user means by engaging in some action may have little to do with what the designer imagined.

8.6 Theme: Interaction congruence

This theme tells a story of how the presented interactions were perceived by the participants in relation to other VR experiences. The participants were questioned in regard to how they grasped, moved, and manipulated the virtual world around them. This theme lends itself to the "descriptor/motor-level" attributes described by Lenz, i.e., the sensory and tangible needs of the VR prototype. Hansen and Dalsgaard's trichotomy would contribute this discussion to that of the User <=> System relation, Dalgaard & Hansen argues that one can deconstruct the systemic concept of embodiment to gain an understanding of some of the tensions the User <=> System relation as a relationship characterized by the user's exploration of the meaning of the system. The researcher questioned participants about their perceived interaction-aesthetic, starting with the implementation of the VR "rig" which replicated these bodily functions. Within seconds of seeing the participants entering the 'scene', they started

teleporting around and grabbing objects without the researcher having issued a single instruction. Participants found the virtual representation of their bodies to work 'flawlessly' and 'intuitively' as they explained that the interaction was almost identical to other first person VR titles (see Figure 17), indicating a high level of congruence and standardization of existing practices, the researcher had, however, overlooked a detail which should have allowed incremental snap-turning to avoid pivoting the head when rotating. Despite its relative niche, every participant noticed this missing feature, further detailing a somewhat existing standardization set in place for VR movement. (Malaika, 2015)

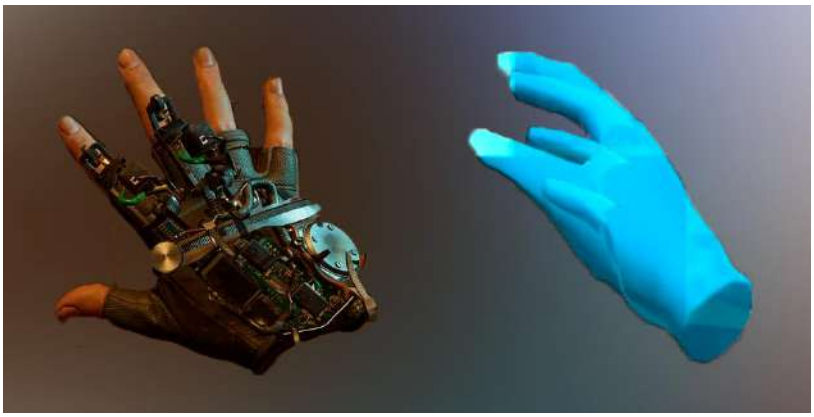


Figure 17: The hands represented in VR bestseller Half Life: Alyx present similar interaction possibilities to the hands in the prototype.

Hansen and Dalsgaard further that when tangible computing systems appear "natural" or "intuitive" to users, it is often because they offer users ways to "uncover, explore and develop the meaning of the use of the technology as it is incorporated into practice."

In summary: The chosen taxonomy for creating the interaction-system is congruent with existing games, and the controls should adhere to a standard for instant recognition among users.

8.7 Theme: Audio-visual connection

The theme of 'Audio-visual connection' conveys statements regarding the overall fit of the music in relation to the visuals. Additionally, conflicting perceptions of what the focus of the experience is discussed.

Participants generally liked the choice of aesthetic for the scene, as well as its relation to the choice of music. John explains:

> "Considering David Bowie's music, I felt like it was a very trippy musical period with Ziggy Stardust etc. So it matched the visual effects and location. It was like it had been

taken out of 'Life on Mars'. The theme was consistent throughout. If one had chosen Whitney Houston instead, it would not fit as well." (John)

While none of the participants felt a direct narrative in the experience, they all believed that the song, 'Starman' played into the setting and thus formed an audio-visual theme. John called it a "grand coherent experience which lends itself to a time period in accordance to the visuals" (see Figure 18)

Figure 18: Screenshot from the test where the live performance of David Bowie is shown on a large TV.

Participants had alternating statements regarding if the music or the visuals were the focus of the experience. Arne and John experienced that the music was a passive factor with the primary experience being the gestures in the sky. They were performing visuals with the music as a background-element. Oscar had a contrasting view, seeing the gestures as an extra addition with the music being the main attraction.

The participants were also questioned about their experience of rhythm in the experience. Löwgren describes rhythm as the temporal manipulation of beats and sequences. Tage explains that while the music is not synchronized to the visuals, it did not matter as much because the chosen song was not rhythmical but categorized as more of a trippy experience. He furthers that the other participants made the experience more trippy. As they layered the visual effects, he expected to get sick, which was instead made into a feeling of having a trip. John also described the rhythmical experience as 'asynchronous' but not inherently negative. It thus seemingly did not fulfil this rhythmical pleasure, but also may seemingly not have created a need for it.

In summary, the rhythmic setting seemed like a non-issue for the music and participants were more conflicted as to what the focus of the experience was.

This conflicted focus provides an interesting angle to discuss. The contradictory musical concert-experience and more visual elements of the VR/VJ experience might also shed light on the situation of the prototypes as a concert-like experience.

9. Discussion

This discussion will be framed around the problem description surrounding concert experiences: How does this VJ-centred collaborative experience compare to the existing traditional and virtually adopted concert experiences? *Can* this project substitute concert-experiences as they existed before the Covid-19 lockdown? Or is it better situated within an entirely different categorization?

9.1 Terminology of concerts

The term 'virtual concerts' must be properly addressed, as it can refer to a wide number of practices. While not an exhaustive definition, one can argue for spherical/360-degree video capture of traditional concerts is one practice, capturing live footage and watching it in VR in a panorama-like setting (Hosseini & Swaminathan, 2016). Then there is the act of performing concerts inside of 'metaverses' established in virtual hubs and video-games, exemplified by the 'Travis Scott x Fortnite' appearance. Lastly, live-streamed traditional live concerts are a practice of virtual concert-going, using services such as Twitch.tv (Twitch, 2021) or YouTube Live (Google, 2021).

9.2 Authenticity in virtual environments

The theme of 'Audio-visual connection' revealed a conflicted perception of the focus of the experience, as either visually centred or based around the musical performance.

To discuss and compare the physical concert to the VJ-experience, a common ground on which to compare the two must also be found. The researcher suggests that comparing authenticities might reveal further.

Gilbert (2016) argues that the perceived realism of virtual environments (as found in VR) depends on the notion of authenticity to be saturated. If exposed to a highly immersive environment, Gilbert exclaims that something is missing from the existing dichotomy of immersion (objective, system-focused) and presence (subjective, user-focused), namely the notion of *authenticity*. Authenticity relates to the extent to which the virtual environment reflects the expectations of the world that it is attempting to represent: the VJ-experience must

thus draw on two streams of thought to identify its authenticity: *expectations* and *motivations*. Gilbert positions expectations comparative to the theme 'Exceeding expectations', explaining that designers who understand observers' assumptions for the real world could emphasize only features that leverage those expectations and leave out irrelevant features. The VJ-experience exceeded expectations by shifting reality to a more abstract experience, thus saturating the concept of expectation. As furthered by Gilbert, Lovata (2007) argues that specifying origin and legitimacy simply state a nominal authenticity, and that a richer, more complex sense of authenticity is context dependent, and depends on the motivations of the observer. These motivations are designed with the intention of engaging an audience. Motivations are thus fulfilled when the purpose of a design is suitable for the audience. The audience was quite clearly defined in this project, and the motivations for participation were thus equally saturated. The authentic proposition of the virtual environment within the experience is only half the story. One must also consider the authenticities of concert-experiences to fully situate the VJ-experience in this context.

9.3 Aural and Visual Authenticity

To discuss their aural and visual differences, the researcher finds it imperative to explore the authenticity of VR and concerts as different authenticities. Jones (1993) argues that this exact discourse of authenticity should be investigated, particularly *whose* authenticity that will be attained or challenged by VR. In this investigation, the author situates both the aural and visual within a 'sense of space'. Both can create space, yet he furthers that VR technologies tend to privilege the visually stimulating, here being a *visual reality*. It is exactly this quality of VR that situates it well within a VJ context, and why the interactivity lends itself well to offering a sort of VR/visual authenticity.

But it also reveals a level of incompatibility with the sense of space between the two (see Figure 19).

Figure 19: At points, TV showing the live performance was almost swallowed by visual effects.

Studies situate the authenticity of musical experiences within a nominal context, as previously mentioned. For a musical performance to be authentic, it is argued to be legitimized through historical development (Tetzlaff, 1994), indigenous origin (Bigenho, 2002) etc. For the authenticity of a concert, we are more concerned with the concert performance and less with the musical attributes. Jones adheres to temporal immediacy or the live situation as distinctive to the authentic concert-experience. He argues that when incorporating recording, even the most perfect reproduction of a work of art is lacking in one element: its unique existence at the place where it happens to be. He furthers that, in virtual audio, a social concert space may be perceptible but not physically present. The transference of social (physical) to a more private (VR) concert space thus separates the performance situation. The researcher argues that it is this transference that makes VR adaptation of concerts less authentic with regards to the concert performance. Existing virtual concerts have seemingly found ways of coping with this. The metaverse in which the 'Travis Scott x Fortnite' experience took place situated the artist as a 3D avatar whose lips mimic the actual artist, thus spatially situating it as an authentic almost pseudo-performance. Jones further describes the displacement of sense of space as when someone listens to a recording, they do not claim to 'be there', and yet recording production is invariably concerned with recording a 'there' at which the consumer *could* 'be' as a kind of aural realization. The created visual space in VR is immediate and present while the aural presence is shifted in its realization of the 'be' not 'being there', but rather a suggestion of being there. As a result, the aural impression is less present than the visual, which provides an explanation as to why the VJ-centred experience was perceived as less of a concert-like experience, with the music being reduced to a background element, while also being very

prominent in its visual affordances. For the experience to be more concert-like, one should thus adhere more closely to the live situatedness of traditional concerts or adapt the concert performance into the visual focus.

9.4 User-roles in traditional and virtual experiences

As this experience was also aimed at exploring collaborative qualities, the experience of being together must also be discussed.

Dalgaard & Hansen's trichotomy has been used to describe the user-roles of system-user-spectators as they appear in the VJ-experience. The act of performing visual effects was seen as central in this regard, due to participants' ability to share their effects and observe the effects created by others. This contributed to an exchange of the role of performer and observer, which related participants to each other's actions, which ultimately reinforced the sense of 'doing together'. This provides an interesting reference for comparing the user-roles of traditional concerts. As most concerts are not interactive per say for concertgoers, the notion of the 'performer' is placed upon the performing artist who also shares the role operator with the adjacent stage artists, leaving only the role of observer to be delegated to the audience.

9.5 The sonic agency of listening

Yet somehow, the space of concerts is still perceived as highly sociable. One could speculate how relations play into this social space?

Labelle (2018) explores a similar question regarding the existence of modalities of a sounded subjectivity that can support new formations of coming together and in support of emancipatory practices. Labelle argues that from an auditory position of listening, comes nurtured modes of engaged attention, for listening is often relating us to the depths of others, extending across bodies and things, persons, and places: "sound is a medium enabling animate contact that, oscillates and vibrates over and through all types of bodies and things. From such conditions, assemblages and conversations may be fostered based on caring and empathizing." (Labelle, 2018). Labelle seems to contribute empathy and caring to the act of listening, thus situating it in a similar type of relatedness as the VJ-experience. He furthers that the networked space listens in tandem with *Overhearing*. It is pursued as a logic of network culture, in which attention and rootedness are always susceptible to interruption, capture, to being unhomed. Virtual experiences such as the VJ-experience is thus less focused and difficult for sound to

travel through. This further debate the aural relational compatibility of virtual concerts, while implying a visual sense of relatedness that is comparable to the aural.

10. Conclusion

10.1 The future of virtual concert-going

Looking forward, one may ask: where is the future of concerts heading, and what practices can be considered from this project?

State-of-the-art virtual concerts are seemingly getting closer to achieving authenticity in both experiences which are both sociable and concert-like. An emerging example of virtual concert-going 'The Weeknd's - Blinding Lights, The Tik Tok Experience' aims at doing just this. In accordance with previous descriptors, this virtual concert most closely resembles the live-streamed concert. The performance was entirely virtual, but still inherently live, as the streaming-capabilities allow for viewers to write messages to appear inside of a render of the artist, thus providing an authentic experience in both visual and aural aspects (see Figure 20). This conscious mix of virtual practices combines the authenticity of live performance (further emphasized in using a true-to-life avatar) with convincing visual effects. With a hopeful vision into the future of contemporary immersive technologies, these sorts of experiences may see even higher levels of presence, immersion, and authenticity in the context of such technologies.

Figure 20: The Weeknd's - Blinding Lights (The Tik Tok Experience) shows a live-streamed virtual concert performance in which the audience can write comments to appear on billboards inside of the virtual space.

10.2 Concluding this project

This project was set out with the goal of creating a concert-like experience to explore collaborative engagement in the context of a VJing. From addressing the state-of-the-art, participants were recruited and interviewed in the spirit of contextual inquiry to centre the project in actual experiences. From these experiences, a VR prototype was constructed in Unity with a similar focus of user experience through experience prototyping. Using concepts from the field of aesthetics of interactions to explain tacit interaction needs, thematic analysis revealed that participants were fulfilled in their intent of using the application as a collaborative VR environment. Participants felt the notion of sharing the experience of creating visuals to an imperative factor in the situation of relatedness. The concert-like authenticity seemingly requires more attention to create an authentic concert experience which has equal aural factor as that of the dominant interactive and visual VJ-centre. This project thus reveals further opportunity for shifting the practice of VJing into a more collaborative musical context using contemporary immersive technology such as Virtual Reality.

11. Literature

Primary literature:

Dalsgaard, P. & Hansen, L. (2008) Performing Perception—Staging Aesthetics of Interaction, University of Aarhus

Lenz, E. (2014) Aesthetics of Interaction - A Literature Synthesis

Houde & Hill (1997) What do Prototypes Prototype? Apple Computer Inc.

Löwgren, J. (2009) Towards an articulation of interaction aesthetics, to appear in: The New Review of Hypermedia and Multimedia.

Lichtman, M. (2014) *Chapter 10 & 11*. Qualitative Research from Social Sciences.

Suri, J. (2000) Experience Prototyping, ResearchGate.

Braun, V. & Clarke, V (2006) Using thematic analysis in psychology, University of Auckland & University of the West of England

Holtzblatt. K. & Beyer, H. (2015) *Contextual Design*: Chapter 3 Field Research: Data Collection and Interpretation,

Goffman, E. (1966). Behavior in Public Places; Notes on the Social Organization of Gatherings. Free Press of Glencoe, New York.

Labelle, B. (2018) Sonic Agency: Sound and Emergent Forms of Resistance. London: Goldsmiths Press

Gilbert, S. (2016) Perceived Realism of Virtual Environments Depends on Authenticity, MIT Press Journals

Lovata, T. (2007). Inauthentic archaeologies: Public uses and abuses of the past. Walnut Creek, CA: Left Coast Press

Jones, S. (1993) A Sense of Space: Virtual Reality Authenticity and the Aural, Critical Studies in Mass Communication Vol 10

Bailey, J. (2008) First steps in qualitative data analysis: transcribing, *Family Practice, Volume 25, Issue 2*

Secondary literature:

Lamb, N. (2010) VJING: A NEW ZEALAND CASE STUDY, Victoria University of Wellington

Malmström. M. (2012) Development of Artistic Concepts for VJing, Faculty of Engineering and Sustainable Development

VJ Yourself (2016) *VJ Video projection & Virtual Reality customized for events*, vjyourself.com, retrieved 12-02-2021 from:
http://vjyourself.com/events

Hook, J et al. (2011) A VJ centred exploration of expressive interaction, Proceedings of the International Conference on Human Factors in Computing Systems

Freeman, T (2013) Outsourcing The VJ: Collaborative Visuals Using The Audience's Smartphones, Leonardo Electronic Almanac VOL 19 NO 3

Goslin, A (April 2020) *Fortnite's Travis Scott event drew over 27 million players*, Polygon.com, retrieved 06-05-2021 from:
https://www.polygon.com/fortnite/2020/4/24/21235017/fortnite-travis-scott-event-concert-astronomical-12-3-million-concurrent-players-record

BBC (2020) Lil Nas X performs his first virtual concert on Roblox, bbc.co.uk, retrieved 16-05-2021 from:
https://www.bbc.co.uk/newsround/54951670

Koksal, I (2020) *Google Discontinues the Daydream VR Headset*, Forbes.com, retrieved 14-03-2021 from:
https://www.forbes.com/sites/ilkerkoksal/2020/10/03/google-discontinues-the-daydream-vr-headset/?sh=1c9afb833af3

GlobalData Technologies (2020) *History of virtual reality: timeline*, Verdict.co.uk, retrieved 14-03-2021 from:
https://www.verdict.co.uk/history-virtual-reality-timeline/

Malaika, Y (2015) Interaction Design in VR: Valve's Lessons, GDC

Desai, R, Desai, N., Ajmera1, K. (2014) A Review Paper on Oculus Rift - A Virtual Reality Headset, IJETT

Bernard, D. (2019) *History of VR - Timeline of Events and Tech Development*, Virtualspeech.com, retrieved 03-04-2021 from:
https://virtualspeech.com/blog/history-of-vr

Larsen, T. (2020, August 5) *VR og AR: En ny virkelighed vokser frem*. Prosa.dk. retrieved 13-04-2021 from:
https://www.prosa.dk/artikel/vr-og-ar-en-ny-virkelighed-vokser-frem/

Danmarks Statistik (2020) *It-anvendelse i befolkningen*. DST.dk. retrieved 13-04-2021 from:
https://www.dst.dk/Site/Dst/Udgivelser/GetPubFile.aspx?id=29450&sid=itbef2020

Liszio, S. & Masuch, M. (2017) Designing Shared Virtual Reality Gaming Experiences in Local Multi-platform Games, 15th International Conference on Entertainment Computing

XREAL (2021) *Zero Caliber VR*, xrealgames.com, retrieved 16-05-2021 from: https://www.xrealgames.com/zerocalibervr/

Pelican Party (2021) *Sam & Dan: Floaty Flatmates*, floatyflatmates.com, retrieved 16-05 from: https://floatyflatmates.com/

Cummings, J. & Bailenson, J. (2015) How Immersive Is Enough? A Meta-Analysis of the Effect of Immersive Technology on User Presence , Taylor & Francis Online

Hosseini, M,. & Swaminathan, V. (2016) Adaptive 360 VR Video Streaming: Divide and Conquer!, IEEE

Tetzlaff, D. (1994) Music for Meaning: Reading the Discourse of Authenticity in Rock, Journal of communication inquiry 18-1, SAGE Journals

Bigenho, M. (2002) Authenticity in Bolivian Music Performance: Chapters 1 & 4, Palgrave

Valembois, Q. (2021) *ValemVR*, YouTube.com, retrieved 16-05-2021 from: https://www.youtube.com/c/ValemVR/about

Sundhedsstyrelsen (2021) *COVID-19*, sst.dk, retrieved 16-05-2021 from: https://www.sst.dk/da/corona

Google (2021) *About Youtube*, YouTube.com, retrieved 16-05-2021 from: https://www.youtube.com/about/

Twitch (2021) *About | twitch*, Twitch.tv, retrieved 16-05-2021 from: https://www.twitch.tv/p/da-dk/about/

Discord (2021) *About Discord | Our Mission and Values*, Discord.com, retrieved 16-05-2021 from: https://www.twitch.tv/p/da-dk/about/

Facebook (2021) *Messenger Features*, Messenger.com, retrieved 16-05-2021 from: https://www.messenger.com/features

The Blender Foundation (2021) *About*, Blender.org, retrieved 16-05-2021 from: https://www.blender.org/about/

Charron, J. (2017) Music Audiences 3.0: Concert-Goers' Psychological Motivations at the Dawn of Virtual Reality, Autonomous University of Madrid

Unity Technologies (2021) *Unity Real-Time Development Platform*, unity.com, retrieved 16-05-2021 https://unity.com/

Epic Games (2021) *Making something Unreal*, unrealengine.com, retrieved 16-05-2021 from: https://www.unrealengine.com/

12. Appendix

Appendix attached separately - (see "Appendix_thtn.pdf")

Contents of appendix:

Appendix 1: "Video showcase of Unity Prototype Test"

Appendix 2: "The Unity Prototype Executable + instructions"

Appendix 3: **"**360-degree video-render texture taken from a science documentary to be used as a 'skybox' in the Unity prototype"

Appendix 4: "Excerpts from preliminary interviews" (approx. 3½ pages)

Appendix 5: "Interviews for evaluation" (approx. 38 pages)