UNIVERSITY OF CALIFORNIA SAN DIEGO

“Redefining the Virtual Reality Music Video: Designing Toward Music-Centered, Narrative-Based Virtual Reality Experiences”

A Thesis submitted in partial satisfaction of the requirements
for the degree Master of Arts

in

Music

by

Timothy Gmeiner

Committee in charge:

Professor Shahrokh Yadegari, Chair
Professor Mehmet Akten
Professor Tamara Smyth

2024

The Thesis of Timothy Gmeiner is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2024

DEDICATION

Dedicated to every child who dares to follow their curiosities and pursue their imaginations. Dedicated to everyone who recognizes this as a basic right for all of humanity. When that right is achieved for all, what a beautiful world we might all be rewarded with.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to thank my collaborator Eito Murakami for his creativity, patience, and insight through the creation of “Pigments of Imagination”. I would like to thank my committee chair Shahrokh Yadegari for his guidance and reassurance throughout my creation and research process. I would like to thank my other committee members, Tamara Smyth and Mehmet Akten for their insight and thoughtful questions that helped to better guide and organize this thesis. I would like to thank Gabriel Zalles-Ballivian and Shahrokh Yadegari’s Sonic Arts R&D at UC-San Diego for providing insight, and resources toward this project’s development. I would like to thank UC-San Diego Professors King Britt and Steph Richards as well as Zachary Konick and Nick Tolford for their musical contributions to this piece. I would also like to thank Keenan Parry and Aamir Asadi for their visual contributions. I would like to thank my wife, son, family, and friends for their patience and motivation to complete this project and thesis (especially my wife). Lastly, I would like to thank my mom for our family trip to the planetarium five years ago. It would be our last adventure together. The look of joy you had on your face staring at the virtual moon and stars above was the emotional kernel that propelled this project to completion.

Chapter 3, in portion, is a reprint of the material as it appears in *In Proceedings of the 29th ACM Symposium on Virtual Reality Software and Technology (VRST '23)*. Association for Computing Machinery, New York, NY, USA, Article 71, 1–2. The thesis author was a co-investigator and co-author alongside Eito Murakami.

Chapter 7, in portion, is a reprint of the material as it appears in *In Proceedings of the 29th ACM Symposium on Virtual Reality Software and Technology (VRST '23)*. Association for Computing Machinery, New York, NY, USA, Article 71, 1–2. The thesis author was a co-investigator and co-author alongside Eito Murakami.

VITA

2024 Owner, operator, lead designer, SoundRzn Design, San Diego California

2024 Lab Technician, San Diego City College

2024 Associate of Science, Digital Music Technology, San Diego City College

2024 Bachelor of Arts in Interdisciplinary Computing and the Arts – Music, University of California San Diego

2024 Teaching Assistant, University of California San Diego

2024 Director Assistant, Sonic Arts @ Qualcomm Institute, University of California, San Diego

2024 Master of Arts in Computer Music, University of California San Diego

PUBLICATIONS

Timothy Gmeiner & Eito Murakami. 2023. Pigments of Imagination: An Interactive Virtual Reality Composition. *In Proceedings of the 29th ACM Symposium on Virtual Reality Software and Technology (VRST '23)*. Association for Computing Machinery, New York, NY, USA, Article 71, 1–2. https://doi.org/10.1145/3611659.3617201

FIELD OF STUDY

Major Field: Computer Music
Studies in Audiovisual and spatial relationships in music-centered virtual reality works
Professor Shahrokh Yadegari

ABSTRACT OF THE THESIS

"Redefining the Virtual Reality Music Video: Designing Toward Music-Centered, Narrative-Based Virtual Reality Experiences"

by

Timothy Gmeiner

Master of Arts in Music

University of California San Diego, 2024

Professor Shahrokh Yadegari, Chair

Virtual reality as a platform provides untapped potential for the popular music video through its affordance of multisensorial convergence, which aligns with and greatly expands upon the evolution of music as an audiovisual medium. Current popular models for music video representation in virtual spaces utilize many of the unique features afforded by virtual reality including 360 video, audiovisual reactivity, spatialization and interactivity, however, most still act as repeatable experiences tethered to design conventions for two-dimensional digital spaces and ideals related to the archetypal performer-audience dynamic.

This thesis explores the potential for music artists to create toward virtual reality by examining the intersection between narrative-based music production and existing music-centered virtual reality experiences, proposing a framework in which emotion, narrative and both mediums iteratively inform each other to create a set of altered design principles for meaningful creation toward this intersectional space.

A series of music-centered virtual reality case studies will be presented to support this argument. Exemplified will be the interactive virtual reality composition “Pigments of Imagination”, a musical piece created by the author with virtual reality as a destination, and thus presented as a narrative journey weaved by and dependent on the technical relationship between compositional, visual, and interactive elements in a virtual space.

Chapter 1 : Introduction

There are few things more satisfying than when the emotional intent of an artist's work resonates with an audience in a meaningful way. The art, its story, and the format it's delivered through all play a role in the conditions of its receipt. As a music artist, I have released albums on both digital and physical platforms (CD and vinyl) for over 20 years, experiencing first-hand the fluctuating relationship between art and the format it's released through. I helped launch and co-operate a vinyl record shop for six of those years, discussing with patrons their divergent relationships between a digital stream and the physical nature of vinyl. Though many would subjectively assert that music 'sounds better' on a physical platform, I was more intrigued by their declarations that the ritualized and multimodal nature of tactile gatefold jackets and sprawling album art gave the music a richer, more engaging, and focused experience. It was this 'ritual' relationship between platform and art that interested me, something that I felt was lacking in current digital streaming and video formats. This feeling was reinforced by research that spotlights streaming platforms like Spotify, which focus on function-based playlists that detach music from its original context (the album) and relegate it as something to be heard in the background [12]. In the era of the algorithm-as-curator, journalist Paula Pelly describes "a new type of music listener, one who thinks less about the artists or album they are seeking out, and instead connects with emotions, moods and activities, where they just pick a playlist and let it roll" [12][33].

In the audiovisual space, a similar displacement of artist intent occurs in the decontextualizing of music across YouTube videos and TikTok clips, where songs are distilled down to the virality of their most lucrative moments. There is plenty to love about how fans

creatively repurpose these songs, however I've always felt a bit of sadness when the level of care artists put toward expressing underlying emotions and narratives through their music is not being received in the manner they intended. Even the music video (a logical evolution of the album art), once propelled forward by a host of innovative artists (Michael Jackson, Prince, Aphex Twin, etc.) and directors (Spike Jonze, Hype Williams, Michel Gondry, Chris Cunningham, Chris Milk) looking to extend musical ideas in visual spaces through monocultural mediums (MTV, BET, VH1) has struggled to remain as relevant in a stratified online ecosystem. What is worth noting however, is that both album art and the music video (mediums designed to enhance a music-centered experience) intersected with the emergent technology of the time, so in this lineage of popular music-centered multimodalism, what might logically follow?

1.1 Virtual Reality

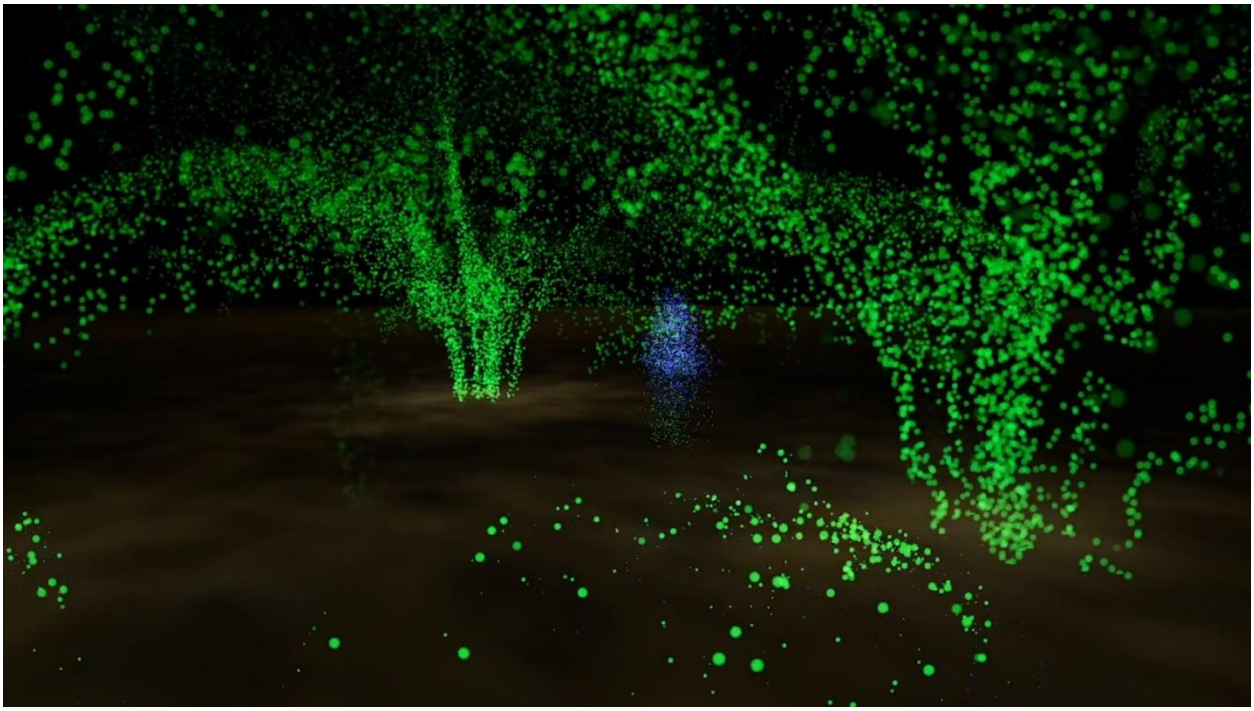


Figure 1-1: Screenshot from "Becoming: An Interactive Journey in VR" (Yadegari, et al)

Three years ago, I was invited by friend and collaborator Eito Murakami to visit the Audio Spatialization Lab at UC-San Diego's Qualcomm Institute to experience the prototype of an audiovisual VR composition he was contributing to, titled "Becoming: An Interactive Journey in VR". This was my first true experience in virtual reality and, much like those magical youthful memories of record shop patrons immersed in their favorite vinyl album, this experience altered me. It was as if I could reach out and touch the music; as if I was inside of it. As an artist with synesthesia (a sensory overlap condition in which I see colors in sound) [51], I saw my own synesthetic experiences brought to life and represented through audiovisual mappings, spatialization techniques and haptic interaction. Consequently, the subtleties and nuances of the music and visual narrative had room to flourish in ways I'd never digested through prior two-dimensional music-centered mediums. When I took the VR headset off, I was spellbound; I'd experienced music that took hold of every fiber of my body. Even more incredible, this was someone's imagination at work, and I was positioned directly inside of it.

In this experience, I saw the potential implications for music composition, production, narration, engineering, interactivity, and audiovisual relationships; how all these elements might re-engage with each other given the freedom of a virtual space. I saw potential for how artists might pull stories and meanings from their music that would otherwise not exist, therefore creating new ways for experiencers to engage with it. Unlike film or theater in which it might take scores of teams and money to achieve a vision, I saw an accessible medium for small independent collectives (or single artists) to create monumental work. On a more existential level, I saw the opportunity for these artists to represent their imagination in a three-dimensional space, allowing for exploration of its meaning in an artistic and potentially quantifiable way, reminding us that we are all small containers of uncountable infinities.

Poetics aside, only in the medium of virtual reality does the platform and hardware itself, particularly in its direct proximity to the eyes and ears, offer this level of immersion by demanding

of the experiencer full sensory attention (and even the potential for direct engagement and interactivity), at the express benefit of the experiencer no less. There is no phone notification, browser tab or distraction in the periphery of sight or sound to shake, even for a moment, the experiencer's presence within the desired artistic experience. For the modern music artist, who by now has spent the better part of 20 years shoehorning their music into platforms unconcerned with the emotional, narrative or technical proficiencies of its architecture, virtual reality offers an opportunity for a fantastic emergent medium to once again wrap around these musical principles and not only prioritize them to the experiencer, but to build upon them in imaginative new ways. This opens new doors, not just in how one consumes the experience, but in how one might create the experience as well.

1.2 Audience

Consequently, the primary audience for this thesis is the multidisciplinary music artist who creates narrative-based fixed-media recorded music and feels confined by existing mediums, both in creation and presentation, of which to evolve the narrative and emotional center of their work. By first examining some fundamental parameters of popular music composing, producing, and mixing methods, we can then illustrate how each stage of this creative process might be reimagined when integrated with or designed toward a virtual reality experience. This ideally allows the artist to utilize virtual reality more confidently, both as a supplementary creative tool, and as an alternative medium to existing popular two-dimensional platforms in which those core musical, narrative and emotional elements might feel diminished in the eyes of their creator. Thus, this thesis aims to create guiding philosophies and altered design principles targeted toward the interdisciplinary music artist exploring virtual reality as a co-medium to both influence and further advance the emotional resonance and narrative components of their compositions.

In exploring the nuances of this multimodal relationship, we conversely gain the immense benefit of reimagining existing fundamental design principles relegated solely to music-centered VR experiences and narratives: not just in how they might be altered when positioned against existing principles followed by a multidisciplinary artist in an adjacent field, but in how they might reimagine standard principles from the field they're now intersecting with. For example, there are guiding principles toward how we think about sound spatialization in a virtual space versus how we think about the stereo-field mix of a song designed for headphones or a speaker set-up in a physical space when attempting to convey a song's emotion and/or narrative. How might these sometimes-conflicting philosophies inform and alter each other when creating a music-centered virtual reality experience that relies on principles from both mediums to articulate its purpose? We will explore this case when discussing spatialization philosophies and techniques in virtual reality. But first, it is helpful to A) understand the space within the field of virtual reality in which the artist's creations might be organized into, and B) present a proposed artistic process in how those projects might come to form.

Therefore, in this thesis I will examine the intersection between narrative-based music production and existing music-centered virtual reality experiences (VR music videos and surrounding VR fields) by proposing a framework in which emotion, narrative and both mediums iteratively inform each other to create a set of altered design principles for meaningful creation toward this intersectional space. I will first offer a guiding philosophy of artistic practice centered around a work's *emotional resonance* of which all design principles will be evaluated. I will then explore the relevance of this emotional center toward both *music narrative* and *virtual reality (textual & visual) narrative,* defining the uniqueness of each along the way. From this foundational understanding of narratives and their role in delivering an emotional truth to the experiencer, I will explore existing *design principles* in both fields and offer examples of how they

might be reconsidered for this intersectional space currently termed the “virtual reality music video”.

Chapter 2 : Defining the VR Music Video

Figure 2-1: VR Music Videos: Björk's “Vulnicura” and Taryn Southern's “Life Support”

First, we must present exactly what this intersectional space is currently defined as versus what I propose it could evolve to. In the field of virtual reality, a linear fixed-media (deterministic rather than generative) narrative-based music-centered experience is currently defined by the term "Virtual Reality Music Video". My collaborator and I have stated in our prior research that “current popular models for music video representation in virtual spaces utilize many of the unique features afforded by virtual reality including 360 video (filmed with a 360 camera, photorealistic but non-interactive), audiovisual reactivity, audio spatialization and interactivity. Rare but notable progressive pieces such as Björk's Vulnicura [67] VR album or Taryn Southern's “Life Support” [28] envelop a range of these features to form a narrative beyond the *performer/audience dynamic* [39], but still ultimately act as repeatable *passive-observer* or *active-observer* experiences” [43]. Our aim in dissecting the design principles of these existing works is to “push the relationship between narrative, emotion, audiovisual connectivity and sound spatialization into a more singular, multimodal experience for the [experiencer] and present a narrative journey weaved by and dependent on the technical relationship between musical, graphic and interactive elements” that “may act as a collaborative model for future developers in the field of VR music videos” [62]. Multimodalism, for the purpose of this paper, is defined as the use of “multiple sources of sensory information derived from several different modalities including vision, touch and audition” to experience an artist’s work [37].

Essentially, I argue that the “VR music video” title disregards the potential for popular music of any genre to be foundationally reimagined inside the enormous affordances of virtual reality and instead repurposes two-dimensional ideas into three-dimensional spaces. Thus, the current conventions of VR music videos do not take full advantage of the capabilities this emergent medium offers. A reimagination of fundamental principles, both in the music and VR space, must take place to allow a more meaningful exploration of its potential.

Chapter 3 : Redefining the VR Music Video

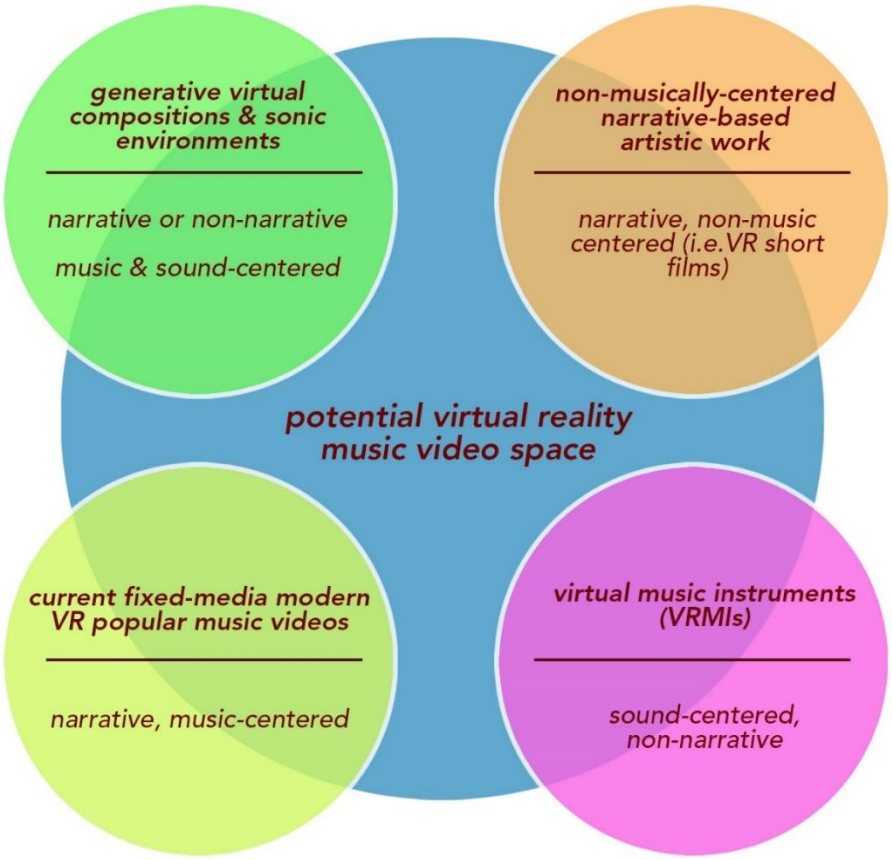


Figure 3-1: Diagram of a reimagined "Virtual Reality Music Video" field informed by four surrounding VR music-incorporated VR fields.

To better redefine the potential of virtual reality, I will first offer an overarching framework to the artist devised to help prioritize their artistic process in this new space, defining key virtual reality terms and concepts along the way. Throughout this framework, I will present a series of case-studies from existing works in virtual reality that co-exist with the VR music video in the greater music-dependent virtual reality field and can therefore offer tools and guidance to the music artist. Here, I aim to extract the relevant information from each case study that might inform an altered set of principles specified for these music-centered narrative-based fixed-media works. Thus, in addition to the field of *fixed-media modern VR popular music videos*, we will

highlight attributes of *virtual music instruments*, *generative virtual compositions & sonic environments*, and existing *non-musically-centered narrative-based artistic work* (i.e. short VR films). The idea here is to take note of elements from these existing but not often interacting fields and reimagine a tangible set of design principles toward VR music videos that push this emergent field closer to what some define as 'true virtual reality' [24], while maintaining key principles from the physical world of recorded sound. Preliminary descriptions of each category are as followed:

3.1 Virtual Reality Music Instruments

Figure 3-2: "Coretet: Virtual Reality Musical Instruments for the 21st Century" by Rob Hamilton [52]

In *virtual music instruments (VRMIs)* we will look at novel approaches to sensory mapping, inspired by the condition chromesthesia (sight/sound synesthesia, an overlap of senses). The idea here is to explore practical and imaginative approaches toward mapping often abstracted and subjective perceptual experiences into a more objective non-narrative/conceptual virtual space. Ideally these examples, such as the "Synesthsizer" and "Synesthesia Laboratory", can act as a deeper inspiration toward representing musical elements of a composition in a visual way across a narrative timeline.

3.2 Generative Virtual Compositions & Sonic Environments

In *generative virtual compositions & sonic environments*, we will look at how user-influence of music/sound generation offers a more immersive environment. In this study we can

compare existing principles toward the state of the VR music video, analyzing what is gained and lost in applying these principles to this medium. Works examined include “Becoming: An Interactive Journey in VR”, “Crystal Vibe” and “12 Sentiments for VR” among others.

3.3 Non-Musically-Centered Narrative-Based Artistic Work

In existing non-musically-centered narrative-based artistic work (i.e. short VR-films), we can explore how fixed media (or mixed/fixed media) pieces take advantage of the affordances of VR - how they intentionally develop toward a three-dimensional space, specifically in visual and sound (haptics and interaction as well). The idea here is to show that the liminal space between ‘true virtual reality’ and ‘VR music video’ is already being explored through other fields, and how we might use some of the principles from these more successful works. Works exemplified include “Notes on Blindness” [61] and “The Key” [50].

3.4 Virtual Reality Popular Music Video

In deterministic fixed-media modern VR popular music videos, we explore the existing spectrum of works and determine if there are any commonalities that might turn into a design principle. We run these works against the above studies to see where prospective improvements might be made. Cases include works by Björk, Run The Jewels and Travis Scott.

3.5 “Pigments of Imagination”

Figure 3-3: "Pigments of Imagination: An Interactive Virtual Reality Composition" - by Timothy Gmeiner & Eito Murakami

Included in, and central to these case studies, will be my contribution to this redefined space, "Pigments of Imagination: An Interactive Virtual Reality Composition". “Pigments of Imagination” is a music-centered virtual reality experience that presents the creative process as a narrative relating the inner-workings of one's unique imagination to the universe around us as told through the story of a small child's journey to the moon [62]. In this journey, the user ascends from a small room and treks a path toward the moon, painting the elements of a world they will descend through before returning to their room. In this thesis I will examine the reasoning of all relevant components in this piece as informed by the aforementioned design principles, inspired by an intersection of these two mediums layered around narrative and an emotional core.[1] Also highlighted from the author’s work will be speculative approaches to the upcoming music-centered VR work “Do You Smile In Your Dreams?”

[1] “Pigments of Imagination” was presented (and won an award) at the *ACM Symposium on Virtual Reality Software and Technology (VRST) 2023* and San Diego, California’s *Without Walls (WOW) Festival* in 2024. Though not formal studies, reactions from VR researchers (VRST) and patrons (WOW Festival) helped to inform this thesis’s analysis of “Pigments”.

Chapter 3, in portion, is a reprint of the material as it appears in *In Proceedings of the 29th ACM Symposium on Virtual Reality Software and Technology (VRST '23)*. Association for Computing Machinery, New York, NY, USA, Article 71, 1–2. The thesis author was a co-investigator and co-author alongside Eito Murakami.

Chapter 4 : Communication Diagram: Building Effective Communication Between Emotion, Narrative and Medium(s)

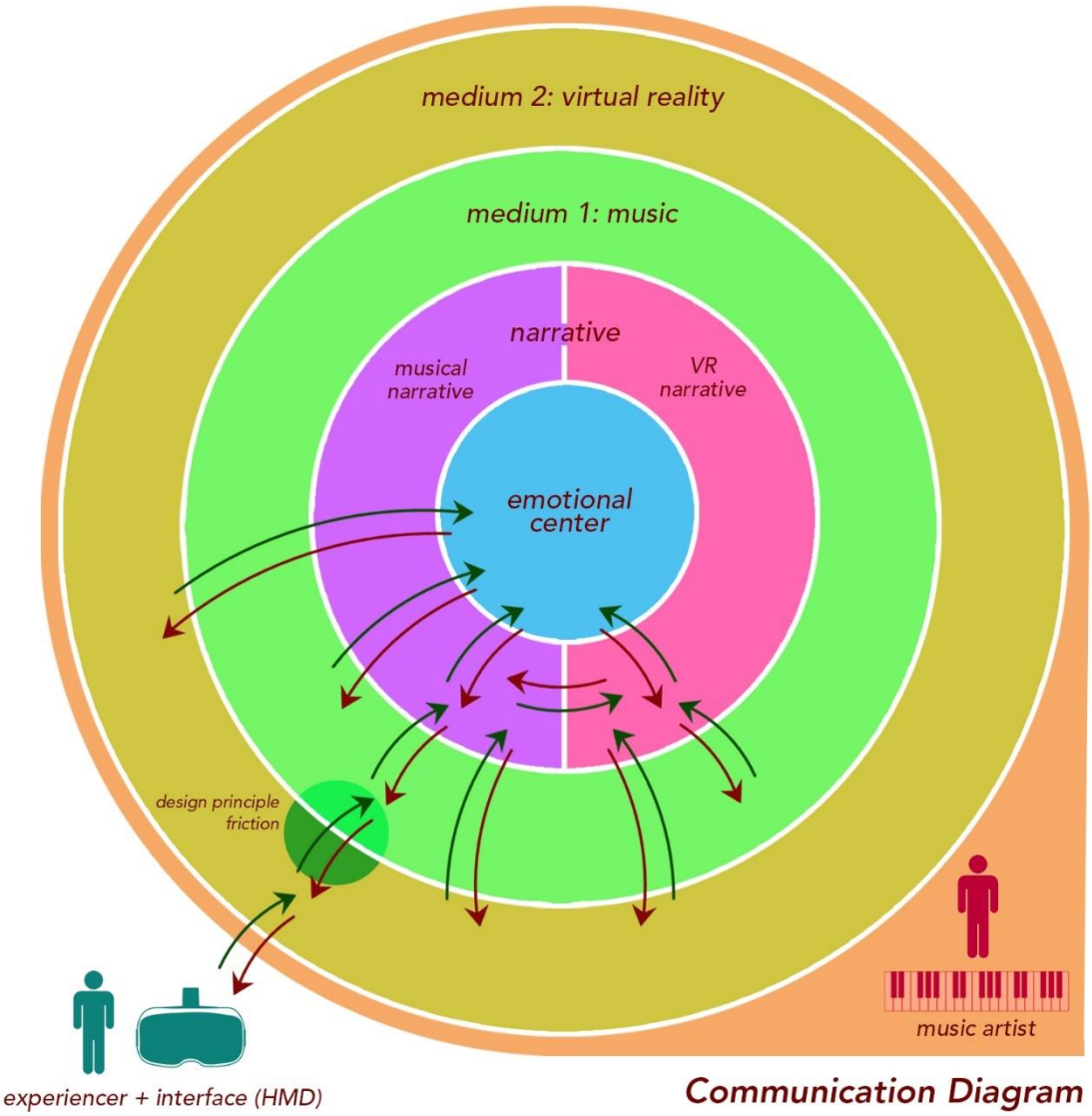


Figure 4-1: A proposed Communication Diagram that centers the artist's emotional intent through musical and VR narrative towards the mediums of music and virtual reality, and ultimately interfacing with the experiencer.

To inform the artist's decisions on how adjacent or conflicting cross-disciplinary design principles might intersect, I present a guiding philosophy that views the creation and presentation

of a music artist's work through the lens of emotion, narrative, and mediums. This **Communication Diagram** illustrates a recommended artistic practice in which the emotional center of the work defines the principles of every component that wraps around it.

Think of this diagram as a nesting doll expanding outward. At the center you have an *emotional core* that the entire work is built off. Wrapped around the emotional core is the piece's *narrative*: both *musical narrative* and *VR (textual and visual) narrative*. Principles from each of these narratives engage with and inform each other, but foundational to their interaction remains the emotional center.

The collective narrative (guided by the altered principles of either *musical*, *VR* or both narratives in collaboration) interfaces with the first medium, *music*. Think of this as the medium in which the artist transfers the theoretical dialog between narrative and emotion into physical output. In this medium certain music and sound design principles exist that would apply if music were the primary medium in which the listener would receive the artist's work.

However, in this diagram, the experiencer receives and interprets the intended emotion wrapped in narrative and transmitted through the first medium by way of a second medium, *virtual reality*. Once the experiencer puts on the VR headset, they are no longer listening to a song playing, they are receiving a virtual experience in which that song plays a primary role. In this way, music as Medium #1 is interfacing with virtual reality as Medium #2. And just as with music, there are certain design principles that play toward the affordances of virtual reality. Sometimes the design principles between both mediums align and sometimes there is friction. At these points of friction, the mediums must dialog with each other, informed by the guiding philosophy of the musical work's emotional center, to create a set of altered principles.

However, the mediums are not the only aspects of this diagram that communicate. Every component is in some way constantly interfacing with the others: decisions made in how the mediums of music and virtual reality interact might reframe an element of the narrative or distill the emotional core into something more concrete. This is expressed via the various inputs/outputs that exist between components (represented above by green and red arrows), closing with a single input/output exchange between virtual reality as Medium #2 and Experiencer. It is important to note that an affordance almost exclusively found in virtual reality is the ability for the experiencer to interact with and therefore influence the virtual reality medium and through this engagement, potentially influence musical, narrative and emotional components. It is important to note, however, that the emotional center always has the last word in these dialogs, including how the Experiencer's interactions affect each component of the work. It is in this constant communication, all taking place within the decisions of the music artist, that the form of the piece truly takes shape.

In the following sections, I will examine each component of the above diagram, the necessity of how each wrap around the next and then, under the guiding philosophy of maintaining an emotional truth.

Chapter 5 : Emotion

The *Communication Diagram* (see diagram 4.1 above) centers around emotion because at a visceral level, we all understand emotion. If not always intellectually or logically, we all feel emotions and at a core humanistic level acknowledge a shared spectrum of internal emotional responses. Before diving into particular principles of narrative structure, composition and virtual reality world building, my hope is to center us at a starting point in which the benchmark for comprehension is to simply acknowledge that I, the author, am capable of feeling emotions, just as you, the reader, are capable of the same. From here we can begin to explore emotion in context to our proposed mediums and how it directly communicates with narrative.

5.1 Emotional Center

"Emotion" is one of those words that is difficult to define without using the actual word itself. Oxford Languages defines emotion as "a natural instinctive state of mind deriving from one's circumstances, mood, or relationships with others" [55]. This thesis is not interested in investigating the affective science of emotions. This definition is offered simply to verify their tangible nature, such that they can inspire a physical output that can then be received and interpreted.

When we discuss a project built around an emotional center, we are referring to the artist's emotional truth, meaning the artist's awareness of their own emotions and how they might attempt to convey that outwardly. The artist's exact awareness of these emotions is not expected to be fully formed or articulated; this is what the process of creation is for; to sculpt something resonant out of the initial vagueness. However, an awareness of emotions is understood enough

by the artist such that they can make informed decisions in how these emotions engage with narrative and medium.

5.2 Emotion Related to Narrative

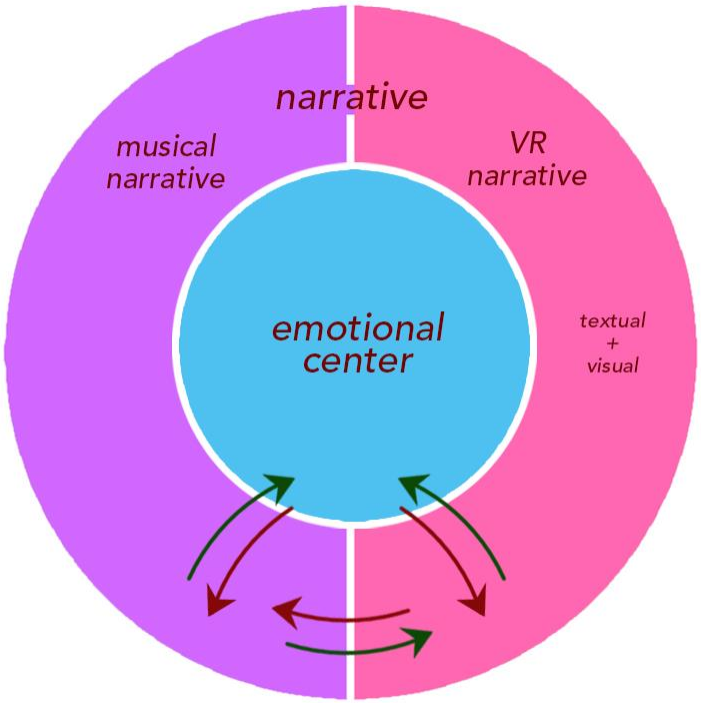


Figure 5-1: A closer look at the Communication Diagram's iterative relationship between emotion, musical and VR narrative.

Emotions are defined above as being derived from one's circumstances; occurrences that are experienced in a certain sequential fashion. It is in this way that we can think about the relationship between emotion and narrative. Tilmann Habermas argues in his book *"Emotion and Narrative"* that emotions are a form of communication, both to others and ourselves: "Because emotions react to and evaluate events", he states, "to understand them we need to understand the sequence of events and their implications" [60]. A narrative format thus offers us a way to make sense of emotions and how they are being communicated. "The emotion process has a narrative

quality. And narrating the emotional experiences has the power to transform them and the emotions they engender."

When composing a musical piece from an emotional space, the music artist is preparing to communicate and thus is already intrinsically thinking in narrative. Additionally, music itself, regardless of genre, is a sequence of events, especially in a digital space. When Habermas states that "emotions react to and evaluate" a sequence of events, we can logically bridge emotion to both musical and VR narrative and understand how each component (emotion, musical narrative textual and visual narrative) is iteratively informing the others. Since a narrative format offers us a way to make sense of emotions and how they are being communicated, it makes sense that a collaboration of musical and VR narrative 'wrap' around a core emotional center.

In John Bucher's *"Storytelling for Virtual Reality"*, Rob Bredow, Chief Technology Officer at LucasFilms, pinpoints the importance of emotional resonance within a textual narrative. "Resonance", he states, "suggests that two similar things share qualities. Emotional resonance would then speak to two individuals sharing qualities or emotions" [29]. Through the lens of Habermas, we might say these two individuals are journeying across a single narrative arc founded on a shared emotion. In our above schema (figure 1.1), these two individuals are the artist and the experiencer. The artist's intent is to share their emotional truth via narrative through a host of mediums outputted to the experiencer in the hopes that it resonates; that the experiencer shares a similar emotion (or emotion with resonant traits) and can thus engage in the artist's narrative in a meaningful way.

5.3 Emotion Related to Mediums

Just as with narrative, emotion acts as a guiding philosophy toward the mediums used to convey it. Virtual reality, by technological design and in how a participant interfaces with it, is "the ultimate emotion mediator…an emotion machine", offers VR pioneer Larry Rosenthal. It's "a

venue to feel, empathize, and process experiences, both their own and others [35][45])'". This concept is reinforced further in *"Storytelling for Virtual Reality"*, where Bucher interviews a host of established writers, directors and artists producing immersive narrative-based content for virtual reality. From these interviews, he highlights the affordances of the virtual reality medium toward emotionally guided narratives and then develops a set of methods and principles for achieving these immersive storytelling ideals.

From an interview with VR filmmaker Jessica Brillhart, Bucher distills the philosophy that "*energy*, the emotional journey of the audience through the experience, and *perception*, how the viewer is experiencing the world, should drive the technical decision making in VR storytelling." [29]. In this quote we see the foundational connection between an emotionally guided narrative and one of the core principles of virtual reality, perception. We then see how the behavioral and aesthetic principles of that virtual world are guided by the relationship between these two ideas. Bucher continues, "While technology provides the immersion for the viewer in VR, one of the chief concerns of and challenges to the technology is not breaking the immersion that it is providing…the emotional experience of the audience must remain at the pinnacle of our goals as creators." We will discuss the relationships between immersion and narrative further in this thesis, however what remains important to understand here is the necessity of maintaining an emotional journey and the support immersion plays in this maintenance.

Whereas Bucher highlights the relationships between immersion and narrative in virtual reality, authors Kristofer Blom and Steffi Beckhaus explore another emotionally centered correlation. In *"Emotional Storytelling"*, Blom & Beckhaus discusses the interplay between interactivity, the experiencer's interaction with the virtual world (a function of virtual reality that will later be discussed in-depth), and narrative, particularly in how to best determine the desired emotional outcome of the experiencer. Whether in creating a single or forked narrative (i.e.

multiple outcomes or routes to a single outcome), interactive storytelling "should explicitly parameterize the emotion of the user and use this as a guiding feature for on-line construction of the story" [31]. The relationships between both musical and VR narrative principles and virtual reality design principles will be discussed in greater detail as well, however prioritizing and parameterizing the emotion of the experiencer, and therefore emotional resonance of the artist's creation, is necessary in how narrative and design principles will reconcile each other.

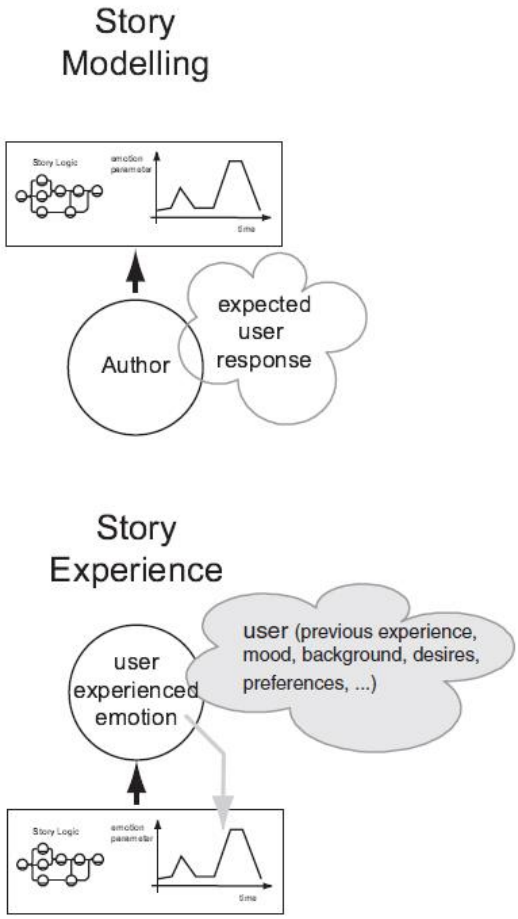


Figure 5-2: Blom & Beckhaus's Diagram of Story Modelling and Story Experience as relating to emotional resonance between creator and experiencer [31]

Blom & Beckhaus also reinforce what music artists may already have a keen awareness of: emotional reaction varies "not only with the person and culture, but also with the emotional state

of the user at this time" [31]. When we discuss an emotional center, we are referring to what the artist hopes to emotionally achieve with their work. Any interpretation of that emotion is beyond the artist's scope. The victory for the artist then, is not always in the experiencer feeling the precisely intended emotion, but that they feel anything at all that might hold some resonant trait to the artist's initial intent.

> Storytelling and its various derivatives have an implicit goal of eliciting emotion in the audience. To our knowledge, all interactive storytelling systems and theories follow this convention of implicit focus on eliciting emotions. Since the end cause of storytelling is to elicit emotion in the user, Interactive Storytelling systems should take the emotion of the user explicitly into consideration…An emotional parameter can provide the system with a structure for programmatically maintaining a better story arc and thereby enabling it to create a better user experience. - Blom, Beckhaus "Emotional Storytelling" [31]

The public description of "Pigments of Imagination" includes the statement "[Pigments] is an observation on the beauty, fear, adventure, sadness and loneliness of self-discovery and the artist's ultimate recognition of process as goal, as told through the story of a small child's journey to the moon." There is no guarantee that the experiencer will garner this precise emotional response, but in maintaining an intention as the artist that these emotions are fueling narrative, compositional and VR world-building aesthetics, the chances of at least some form of emotional response heightens. A similar suggestive emotional resonance occurs in "12 Sentiments for VR", a multi-act music-centered experience in which one act, titled 'melancholy, weak' uses VR principles like gaze, physics and time manipulation alongside generative compositional techniques to elicit the intended melancholic emotion [24]. This work will be explored more when we discuss VR design principles.

Blom' and Beckhaus's work has gone as far as to create *Emotional Path Graphs*, a way to plot out a forked narrative based on various prospective emotional responses by the experiencer. Though their work moves beyond the scope of this thesis, it's important to highlight, as this presents a new way for the music artist to rethink compositional strategies by way of virtual reality

design principles; less as a two-dimensional linear piece and more as a balanced binary tree in which the composition might potentially unfold into different directions and resolve in a singular space. These approaches will be examined in greater length when we discuss the relationships between music and virtual reality design principles, however the inclusion of this idea now reiterates an emotional guidance and opens the artist toward creating for a spectrum of emotional midpoints or outcomes not just within virtual reality but within their composition as well.

Chapter 6 : Narrative

Narrative, for the purpose of this thesis, is defined in two ways: *musical narrative* and *VR narrative*. Musical narrative accounts for the linear way we experience a fixed (non-generative) piece of music from a start point to a predetermined endpoint. VR narrative builds atop the musical narrative either through lyrics included within the composition (textual narrative component) and/or via the prospective world created in virtual reality (visual narrative component). In our discussions on the dialog between emotion and narrative, we see an iterative negotiation between emotion, musical, textual, and visual narrative forming itself to output toward a physical medium. In the music artist's case, this first physical medium in dialogue with narrative is music through its subcomponents of *composition, production,* and *engineering*.

6.1 Musical Narrative

> In the art of music, as in life, motion toward encounters obstacles, reverses, disappointments, involves great distances, detours, expansions, interpolations, and, in short, retardations of all kinds. Therein lies the source of all artistic delaying, from which the creative mind can derive content that is ever new. Thus, we hear in the middle ground and foreground an almost dramatic course of events.
> - Heinrich Schenker [5]

The idea of narrative in music has been chronicled since Russian Formalism and Structuralist studies applied the dramatism and plot devices examined in literary theory toward the similarly "identifiable [elements] arranged in partially predictable pattern" found in Haydn or Mozart. Rondos, for example, include variations on repeating refrains creating sets of patterns, and sonata forms include introductions, sequential groups, transitions and closings [18]. And though the commonality of both literature and music existing as a sequence of events may not alone be a strong connective tissue, when the listener begins to understand these musical events as actions or gestures, the relationship between narrative and music grows closer.

This connection is not just relegated to Russian literature and composed western classical music; it ventures into the improvisational. It is also not solely the listener who distinguishes a narrative connection to the music, it is the artist attempting to figure out the story within it. During a studio rehearsal of John Coltrane's seminal work "Giant Steps", Coltrane can be heard telling his collaborators "I don't think I'm gonna improve this…I ain't gonna be sayin' nothin' [I'm just gonna be makin' changes]…I ain't gonna be tellin' no *story*" [65]. The changes that Coltrane is referring to are harmonic and rhythmic in nature, and in this banter, he illustrates that the idea of a driving narrative force must catalyze those changes.

Even through the recording and production of modern music, from electronic to rock to hip-hop, the music itself is created in a DAW (Digital Audio Workstation), computer software that "conceptualizes sound as recorded material that's temporally mapped onto crucial points of what's essentially a linear narrative structure" [47]. Just as the sonatas mentioned above are composed on sheet music and contain introductions, transitions, patterns and closings, such is the layout of a DAW that views the linearity of music on a two-dimensional plane reading left to right, presents a suggested start and endpoint of a composition, and invites the creator to build or 'copy/paste' repeating patterns or loops. The narrative functions within literary and musical theory transcend genre, era, region, and creative approach which only strengthens the argument for an overarching connectivity of music and narrative. However, music does not exist within a void, it exists to be experienced, and in our case, experienced through the channels of an adjoining medium. Thus, it cannot be the sonic elements alone that determine musical narrative. To support this idea, we can think of music in context to the world it exists in.

6.1.1 Contextualizing Musical Narrative

When exploring musical narrative and narrative functions, one of the first things we might establish is the idea of music and context. In establishing this idea, we first make the decision to

not view music as an "autonomous object, expressing nothing but itself [15], but as an "ongoing process…in constant interplay with the surrounding world" (Wingstedt). In Johnny Wingstedt's *"Narrative Music: Towards and Understanding of Musical Narrative Functions in Multimedia",* Wingstedt offers the idea that music does not exist in a vacuum, there is always context to be derived from it which allows it to hold meaning greater than its existence. That said, Wingstedt adds the key caveat that though musical narrative functionality is dependent on context, this does not mean that music's sole purpose is to be a supportive accompaniment of some 'pre-existing' narrative. Providing context is crucial to understanding musical narrative functions that not only serve to accompany existing narratives through adjacent media, but that also guide these collaborative media. Since we are discussing virtual reality experiences in which the fixed-musical composition's narrative acts as the primary container of emotional resonance, this is an important concept to consider. Using Wingstedt's above framework, we can look at the key contextual dimensions that allow us to examine musical narrative and then highlight from a set of classes and categories within these functionalities, the groupings that reinforce the recursive conversations between emotion, narrative and mediums.

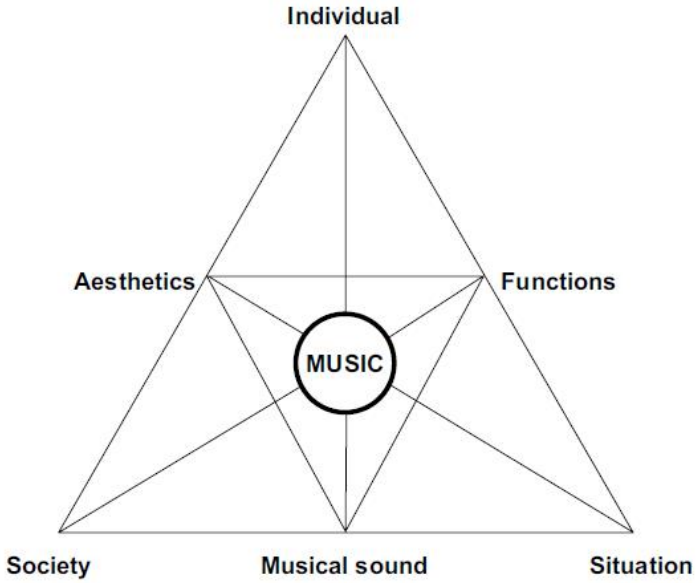


Figure 6-1: Wingstedt's six dimensions of musical narrative context.

Wingstedt splits musical context into six dimensions (we will focus on five), however it is in how these dimensions interact with each other that create a compelling representation of the musical context necessary to derive narrative functionality. This is why it is necessary to view music as a process rather than an object, as the exchange of these contextual dimensions is not a single discrete event. Thus, "the functional and aesthetical dimensions can be said to represent the essence of the musical experience."

The first dimension of musical context is the *individual*. This individual might be the creator, the performer, the listener, or possibly all three at once. It is in the individual's history, gender, culture, identity, education, experience, and mood in which their 'taste' or expectations of a piece of music develop. When we look at our original *Communication Diagram*, the individual is at times the creator, experiencing and performing in their own work, representing the full scope of the individual dimension in a musical context. When the music artist has shared

this experience with someone else, that person is in a sense the listener/experiencer, performer within the VR experience, and prospective co-creator depending on levels of interactivity.

The second dimension is *situation* which dictates 'the space and time where the musical sound is experienced'. It should be noted that beyond just the acoustic environment, the situational dimension also includes musical conventions, such as the expectations of a particular genre. This will become relevant as we later explore relationships between musical narrative and conventional principles of music composition/production found in particular genres. The situational dimension is where musical narrative functions are most closely related because this dimension includes the ways we use and understand music through collective experience.

The third dimension, *society*, provides a broader view of the societal, historical, and cultural conditions that create the act of music occurring at that moment. In our case, a technological condition might also be included, particularly in how society currently understands and engages with virtual reality as an expression of a musical work. This is less a dimension contextualizing the process of music occurring in the moment, and more so the overarching perspectives on these musical sensibilities within society.

The fourth dimension is *aesthetic* dimension, which "comprises concepts that are often thought of as difficult or abstract (and not seldom controversial), such as beauty, art, creativity or interpretation". Though abstract in nature, when factored into a constant dialog with dimensions like individual, societal and situational, we can see how this slightly more subjective dimension still factors toward a musical work's narrative functions.

The final dimension is *musical sound,* the elements that comprise the form, movement, and sonic elements of a musical work. Because of its objective existence in physical or digital space

through definable qualities of amplitude, timbre, rhythm, and harmonic sensibilities, we might view this as a foundational dimension of which all others are contextualized against.

This is the dimension in which we can place the above examples of compositional and improvisational narrative functionality, whether in the “formal musical patterns” of Mozart, the improvisational explorations of John Coltrane, or the suggested linearity of modern DAW-influenced music production. As the Russian Structuralist literary theorists stated above, literary, and musical narrative functions (across many genres of music) revolve around ideas “arranged in a partially predictable pattern”. The recognition of these patterns by the listener depends on a constant fluctuation between redundancy and entropy (Wingstedt). “Structuring a message according to known [musical] codes or conventions will lessen the entropy and increase the redundancy” of a musical work. Redundancy in this case may include repeating melodic or rhythmic patterns, symmetrical metric periods or repeating movements within the overall form. In more popular forms of music, it may present itself through the conveyance of repeating melodic hooks, lyrics, or riffs. Because redundancy in over-use might flatten the musical narrative (i.e. boring, predictable, etc.), it can be countered with entropy, which might be viewed as a form of randomness, whether compositionally or performatively algorithmic or artistically improvisational in nature. An iterative dialogue between these two ideas are what allow composers of any genre and music producers to create a “narrative stability and continuity for the audience” (Wingstedt). It is in this dynamic that we can then begin to examine how the interplay between redundancy and entropy communicates with the other contextual dimensions of musical narrative and then ultimately reconciles with a virtual reality-based visual narrative that might have the same dependencies.

6.1.2 Multimodality

Wingstedt speculates on these interdisciplinary interactions by examining how narrative meaning is not derived solely from these contextual dimensions, but in how this musical narrative works multimodally in correlation with other media. Through his lens, multimodality is a principle of musical composition, whether through combining instruments to create varying timbral effects or the building of complex polyrhythmic percussive layers, so to extend this idea out to other forms of media is quite natural. “Each of these elements communicates on its own ‘channel’ – or in its own mode. But all these modes also interact – mutually affecting each other – so that the meaning appears because of this interplay rather than as a result (or even the sum) of the individual narrative modes.” Viewing these other channels as extensions of existing music multimodal practice is the preferred way to think about how musical and VR narrative functions engage with each other, and how those engagements not only form the greater ‘meaning’ (what we refer to as the artist’s emotional truth), but will then inform and engage with the design principles we’ll soon explore both in the mediums of music and virtual reality.

6.1.3 Classes & Categories

But how might music go about expressing this ‘meaning’ or ‘emotional truth’? Wingstedt offers a series of classes and categories that allow us to better understand how the above dimensions of context are interacting with each other to more explicitly convey this narrative function and therefore a work’s meaning. With regards to the conversation between medium, emotion, musical and VR narrative, we are interested in three particular classes: the *emotive class*, the *guiding class* and the *temporal class*. Here we will highlight…

The most logical class related to this thesis’s focus is the *emotive class*. As Wingstedt shares, this is a foundational principle of musical narrative structure, and is reinforced by music psychology that shows “specific structural musical factors – such as articulation, melodic interval,

tempo, loudness, modality, register or rhythmic factors – correspond to certain human emotions" (Gabrielsson and Lindström, 2001).

Following the emotive class is the *guiding class*, which "includes musical functions that…[direct] the eye, thought and mind". This class is of interest because it contains the *indicative category*, which states that "synchronizing musical events to specific features or actions on screen" can lend a descriptive quality aligned with the aesthetic dimension, such as beauty or melancholy, to the musical narrative. In our case the 'screen' is a fully immersive spherical world that surrounds the experiencer, and the guiding class adds necessary description through synchronization that reinforces emotional intent.

"Pigments of Imagination" takes advantage of this concept across both musical and VR narrative functionality. For example, the use of object-based sounds mapped to specific virtual objects and instrument visualizations and then placed alongside a multichannel soundbed, are utilized to direct the experiencer's gaze to particular narrative elements. Concepts of 'gaze' within virtual reality as well as approaches to narrative-influenced sound-design will be discussed further when we examine binaural sound-spatialization in virtual reality versus stereo-field mixing techniques for music released on current conventional music platforms (i.e. streaming, vinyl). Similarly, the use of lyrics that reference visual narrative guides within the virtual space are in essence a dialogue between musical and textual narrative, but solely within the music medium. The textual element of lyric then extends in a multimodal way toward the 'channel' or medium of virtual reality.

Lastly, we integrate the *temporal class*, which objectively views music in the time domain. "It is difficult to imagine music that doesn't somehow represent or organize time"(Wingstedt), and since we're focused on fixed-media compositions as foundation to a particular type of virtual

reality experience, this is a relevant class toward viewing how the temporal sequence derived from music's form can "affect the perception of time" and create or subvert the experiencer's expectations. These features, Wingstedt states, have "strong emotional and informational connotations, but also [affect] the narrative structure" of the full experience. As we begin to examine the points of friction between virtual reality and music design principles, we will see how iterative congruence between the emotive and temporal classes establish principles not possible solely within virtual reality's lean toward generative music, specifically in which the experiencer dictates its compositional evolution. We will see the need to compromise the fixed-media nature of conventional music and the generative nature of virtual reality music such that musical narrative functions remain a guiding force.

6.2 VR Narrative

In our above examination of musical narrative functionality, we discussed multimodality as a principle of music composition. For instance, in thinking about how the music artist might expand the narrative of their piece, they might write for certain instruments or sample additional source material that adds the necessary timbres or crescendos to achieve these goals. Thus, in an abstract sense, there lies a multimodality within music before ever collaborating with other media. It is Wingstedt's argument that any collaboration with other media, then, is simply an extension of this already occurring multimodal behavior. This is a fantastic way to think about how, in the liminal space currently defined as "virtual reality music video", musical narrative can guide virtual narrative. It should be noted that musical narrative is ripe for alteration by VR narrative as well, however, since we have established that for the purpose of the music artist, emotional truth is born within the musical narrative, its ideals remain the primary narrative guide.

One reason this music-centered communication of narratives can occur is because virtual reality as an emergent medium for storytelling has yet to establish concrete narrative functions in the way that music has.

> Janet Murray, author of 'Hamlet on the Holodeck' (2017), believes that all the shortcomings experienced while inventing new storytelling forms and trying to evolve new genre schemas were necessary for fostering further exploration and refinement, fueling ideation, and the subsequent creation of new content that allows storytellers to continually (re)invent the narrative technology. 'Expanding human expressivity into new formats and genres is culturally valuable but difficult work…To begin to do this, the immersive actuality storytellers must continue the task of inventing and exploring immersive and interactive story-forms to meet the long-term communicative needs of the networked society while also embarking on the long, patient, slow work of building institutional infrastructures, developing audiences and making a culture.' [25]

Virtual Reality is still an emergent medium, especially in context to music and in comparison to existing adjacent mediums like film or video games. Since our target audience is the multidisciplinary music artist, it's worth exploring some core definitions, histories, and concepts of virtual reality before we expand into its speculative narrative functions and design principles.

6.2.1 What is Virtual Reality?

We have been discussing virtual reality as a medium in which the music artist might engage to both reimagine their creative process and ultimately present their work to an experiencer. But we've yet to really define what virtual reality is, both as a concept and a prospective collaborative medium. Merriam-Webster defines virtual reality as "an artificial environment which is experienced through sensory stimuli (as sights and sounds) provided by a computer and in which one's actions partially determine what happens in the environment" [66]. That works well enough, however it's in his 2017 book "Dawn of the New Everything" where author, musician and VR developer Jaron Lanier (the man who coined the term "virtual reality") provides *52 definitions* of what virtual reality is and can be. It is in his more imaginative,

philosophical, and existential definitions where we might better articulate the medium in which a music artist is looking to extend the imagination and emotion of their musical works. A few definitions of note:

> [Virtual Reality is] "A mirror image of a person's sensory and motor organs, or if you like, an inversion of a person."
>
> "Hope for a medium that could convey dreaming."
>
> "Imagine the universe with a person-shaped cavity excised from it. What can we say about the inward-facing surface that surrounds the cavity? [Virtual Reality is] the substitution of the interface between a person and the physical environment with an interface to a simulated environment."
>
> "The Technology of noticing experience itself."

And my favorite, "A twenty-first-century art form that will weave together the three great twentieth-century arts: cinema, jazz, and programming". It is in this quote particularly that we can see the physical and speculative interlocking relationships between music and virtual reality have been a priority within VR development longer than most music artists might imagine, and in quite an adventurous way at that.

6.2.2 Relevant Moments of VR History

It is helpful to contextualize these relationships between music and modern virtual reality by revisiting a bit of virtual reality's origins. In "The VR Book: Human-Centered Design for Virtual Reality", author Jason Jerald offers a thorough account of the history of VR. The purpose here is not to recount the entire history, but to highlight relevant moments that will lead the music artist to understand concepts particular to VR that will be necessary when negotiating narrative and design principle affordances across mediums.

> "When anything new comes along, everyone, like a child discovering the world, thinks that they have invented it, but you scratch a little and you find a caveman scratching on a wall is creating virtual reality in a sense. What is new here is that more sophisticated instruments give you the power to do it more easily" - Morton Heilig [16]

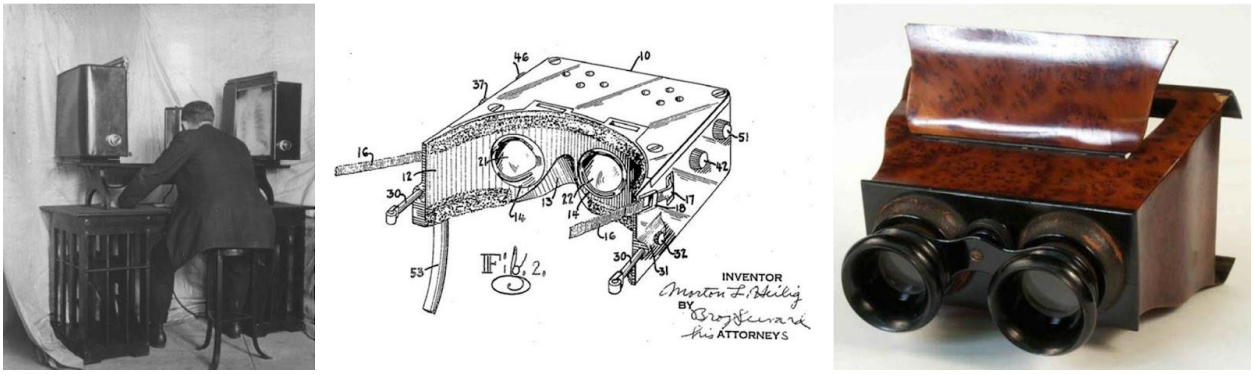


Figure 6-2: VR Headset early stages: Charles Wheatstone Stereoscope, Morton Heilig HMD and David Brewster Stereoscope (Jerald)

The prehistoric version of today's modern mechanisms for virtual reality began in 1832 with Sir Charles Wheatstone's static (as opposed to today's dynamic) stereoscope, which used mirrors angled at 45 degrees to reflect images into the eye from both sides. This was built upon by David Brewster in 1851, who made the technology hand-held and more consumer friendly. The success of his design inspired the first public "3D craze" and in fact is the technological model for modern, low-cost forms of virtual reality like Google Cardboard.

Notable to this thesis's interest in the affordances of virtual reality is the development of the VR headset or head-mounted display (HMD) by Morton Heilig in the 1950s, which not only widened the horizontal and vertical field of vision but introduced earphones and other cross-sensory systems. By the early 1960s, engineers were developing rudimentary head-tracking technology. Most notably, in 1965, Ivan Sutherland, et al. developed the Sword of Damocles, the first HMD that introduced computer-generated imagery alongside head-tracking (Jerald).

In the early 1980s, Atari Research explored and developed human-computer-interaction (HCI) concepts that would become foundational to modern virtual reality technologies through

its application across various industries. The surge in technological evolution during this time led to a VR bubble in the 1990s, the first of many in which predictions were that virtual reality would shift the global paradigm, only to fall massively short of expectations. As a result, public-facing research and development on virtual reality quieted through the first decade of the 2000s, however its use in academia, government, medical and military continued to strengthen (Jerald).

Figure 6-3: The Oculus Rift Head Mounted Display (Oculus)

A series of low-cost developmental innovations that widened the user's field of view was crucial to giving user's both affordability and the sense of 'presence' that was missing from prior technologies. This led to the 2012 Kickstarter-funding of the Oculus Rift. The wild success of this fundraiser caught the attention of media and corporate investors, reaching critical mass in 2014 with a widely publicized $2 billion purchase by Meta (formerly Facebook) [8]. Thus, resources began to pour back into virtual reality, leading into our current bumpy road of a virtual renaissance (Jerald).

6.2.3 What is Virtual Reality pt. II?

This brief historical account brings us back to our first question, *"What is Virtual Reality?"* This time, however, we will examine the question through more practical and cybernetic lenses that transition us from the above standard and existential definitions of 'virtual reality' into more technical terms and explanations that will help to demonstrate more explicitly to the music artist

how to think about collaborating music and virtual reality, both through narrative function and as mediums, toward a more immersive and resonant work.

6.2.4 Extended Reality Spectrum: AR/VR/MR/XR

Real - "the real world that we live in. Although creating real-world experiences is not always the goal of VR, it is still important to understand the real world and how we perceive and interact with it in order to replicate relevant functionality into VR experiences" - Jason Jerald [27]

Virtual - "virtual environments are artificially created without capturing any content from the real world. The goal of virtual environments is to completely engage a user in an experience so that they feel as if they are present in another world such that the real world is temporarily forgotten, while minimizing any adverse effects." - Jason Jerald [27]

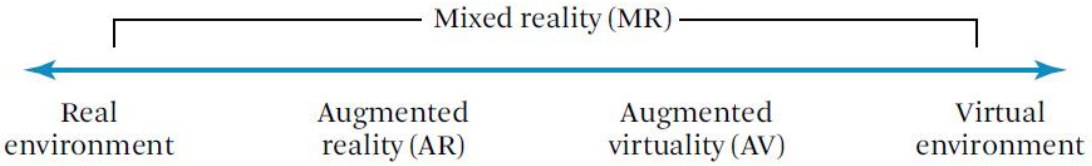


Figure 6-4: Virtuality Continuum [27][49]

In 1994, a group of researchers presented the idea that the interactions between physical reality and virtual reality exist on a spectrum of *extended realities*. This spectrum ranges from a real environment (as 'real' is defined above in figure 6-4) to mixed and augmented environments to a fully virtual environment (as 'virtual' is defined above in figure 6-4). Below is a quick definition of each [40]:

- Virtual Reality (VR): *A virtual world completely replaces all elements of a real environment.*
- Augmented Reality (AR): *Virtual elements overlay elements of the real world; however these elements do not interact.*
- Mixed Reality (MR): *Virtual elements are mixed in with the real world and do interact.*

- Extended Reality (XR): *A broad term designed to define this spectrum, however, is now grouped into the all-inclusive acronym-alphabet-soup AR/VR/MR/XR*

For the purposes of this thesis, our focus will remain on fully virtual environments (VR), however it is worth knowing the spectrum VR sits in to understand the prospective dimensions of collaboration between music and extended realities that are not covered here but are ripe for exploration.

6.2.5 Reality Systems

Just as there exists a spectrum of extended realities, there exists a spectrum of platforms, defined here as *reality systems*, of which to present these extended realities to the experiencer. According to author Jason Jerald, a reality system is "the hardware and operating system that full sensory experiences are built upon. The reality system's job is to effectively communicate the application content to and from the user in an intuitive way as if the user is interacting with the real world." When we speak of virtual reality as a medium for which narrative, emotion and music interact, the reality system is the physical structure and structural design philosophy in which the virtual reality medium operates from. It is helpful to understand the breadth of these reality system classifications A) to offer insight toward a range of prospective platforms that the music artist might be inspired to engage with, and B) to provide reasoning behind why we will focus on one platform.

6.2.6 Reality System Classifications

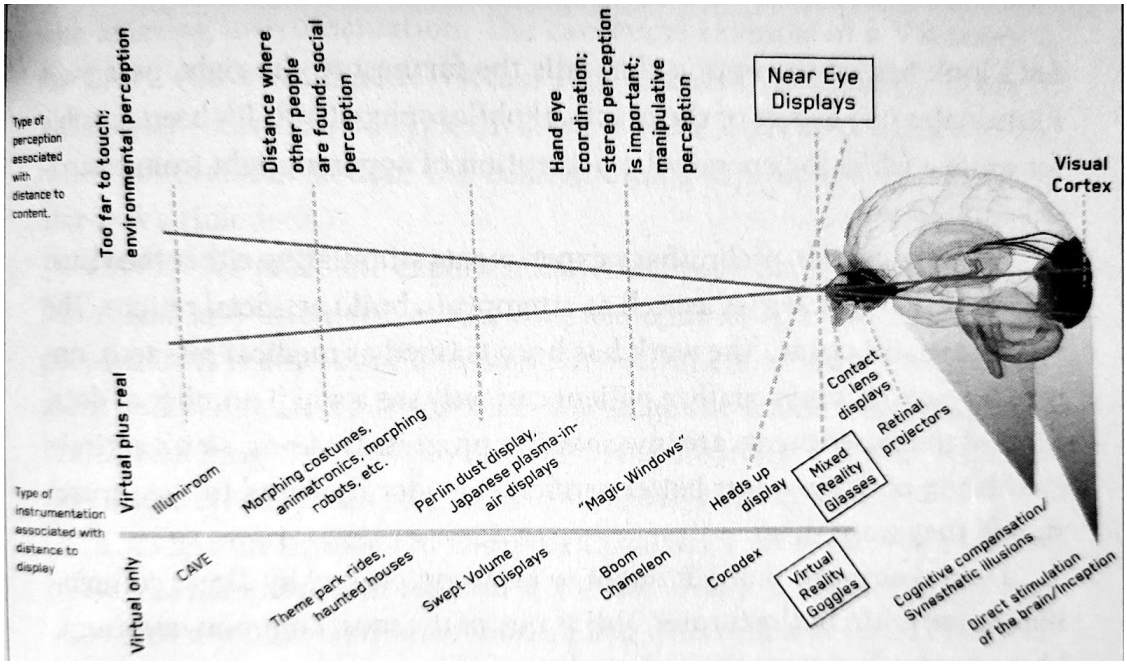
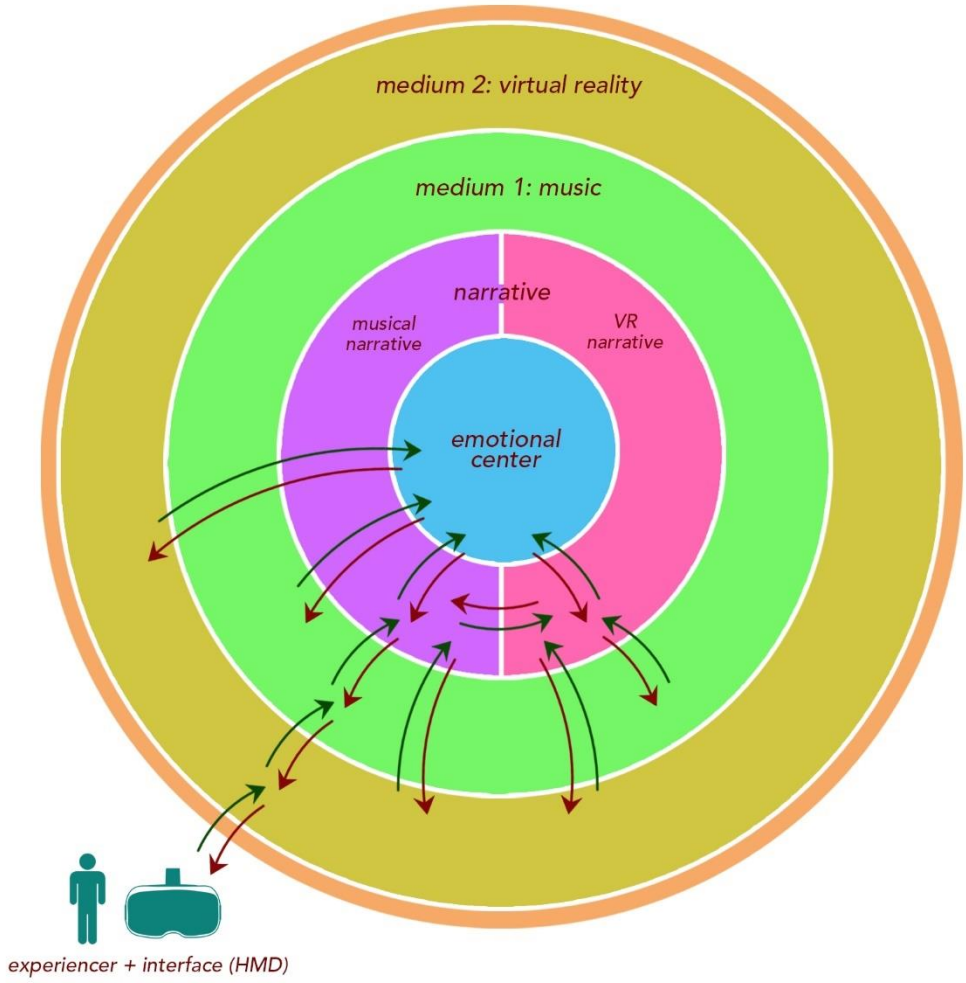


Figure 6-5: Reality system classifications based on distance from the user's eyes [26]

The chart above, created by VR developer Jaron Lanier, offers a set of reality systems classified by their distance from the experiencer's eyes and mind. These systems range from the outermost layer of virtual reality CAVEs (immersive environments, perhaps a set of screens that surround the experiencer in a physical space) to the innermost layer, such as contact lens displays, or cognitive synesthetic illusions. The most recognizable system in the above series is the VR Headset, or Head-Mounted Display (HMD). The HMD is a visual display that attaches to the experiencer's head and, depending on the user's pose, maps and renders in real-time the appropriate virtual coordinates for visual and sound object locations and behaviors toward that user's eyes and ears. HMDs, such as the Meta Quest series or Apple Vision Pro, tend to be the most popular form of reality system because they offer the greatest amount of immersion for the user [27].

Thus, HMDs are the most widely available extended reality system to develop toward and will most likely continue to remain the most affordable platform for budget-conscious artists for the near future. Even more important, current research and development funding is mostly pointed toward technologically progressing the VR headset. In discussing the progress on this research and development, Lanier highlights an ongoing issue that we as music artists may potentially assist in rectifying. Lanier states that VR headset development progress prioritizes vision over all other senses, and deems this to be “wrongheaded, since the other sensory modalities are no less important”. It is necessary in the development of virtual reality systems, specifically the HMD, for developers to focus on other senses beyond visual, and the more interest that music artists show in creating toward the medium of virtual reality, the greater the chance that VR developers prioritize the technology toward acoustics, spatialization, multimodal music creation, and so forth. This not only assists the music artist, but this also assists virtual reality system development. These platforms are still emergent and extremely young in the scheme of artistic mediums, and the music artist can sculpt the priorities and success of the system by how they interact and engage with it.

6.2.7 Reality System Design



In our initial *Communication Diagram* (figure 4-1 reshown above), we refer to the communication between emotion, narratives, mediums and experiencer. This constant dialog between components operates on a series of *inputs* and *outputs* (represented above by the green and red arrows). The idea in using these expressions is that there are perpetual lines of feedback occurring at any given time between these core components, both inputting and outputting to various other components. The music artist might think of this as a form of signal flow with an ongoing occurrence of feedback loops. Though the idea of emotion and narrative communicating

through a series of inputs and outputs can on surface sound a bit existential, the reality system model in which these communication methods are derived is anything but.

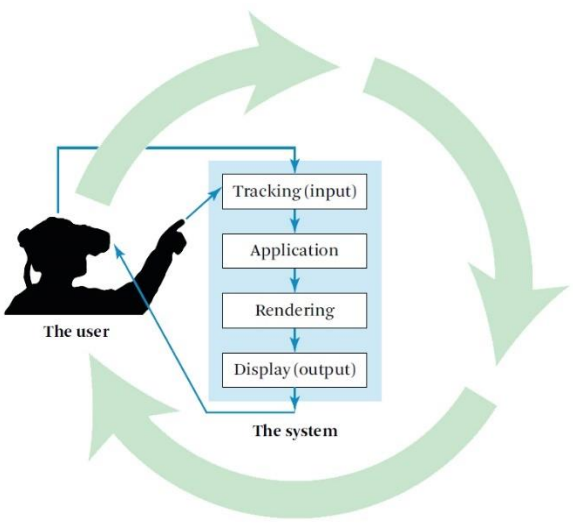


Figure 6-6: Jerald's Reality System feedback diagram

Above we define a reality system as the physical structure and structural design philosophy that the medium of virtual reality operates on. Like our *Communication Diagram*, this structural design philosophy necessitates inputs and outputs, however in this system, with something more specific in mind. Hardware devices (in our case the HMD or VR headset) offer the transfer of input/output data between the physical and digital space, converting human output to digital input or digital output to human input. We might think of this as a consistent and concurrent system of perceptual and sensorial mappings [27]. Input is treated as information coming from the user into the system, whereas output is information returning to the user from the system. The figure above shows an example of how this feedback loop occurs. Below is an explanation of these components:

Input - collects sensory and location data from the user

Application - the unseen behaviors, dynamic shapes, and physics of the virtual world

Rendering - The presentation of the application as converted to a user-ready representation, meaning something that more readily re-creates a perceptual world for the user's various senses (specifically sight, sound, and touch/haptics)

Output - converts back from the digital to physical space via perception of pixels, soundwaves or haptic vibrations seen, heard, and felt by the user.

6.2.8 Fundamental Concepts of Virtual Reality

With a greater understanding of physical reality systems, specifically the ongoing feedback exchange of inputs and outputs, we can now explore a few fundamental and interlocking concepts unique to virtual reality that are worth discussing because of how they affect ideas around VR narrative and catalyze the very principles of the virtual reality medium we will examine shortly. These concepts are *immersion, presence, embodiment, interactivity,* and *perception*, and they can be thought of as a sequential series of events initiated by the developer of the virtual experience that then occur within the virtual reality system, the experiencer, and the constant informational feedback loop between each.

6.2.8.1 Immersion

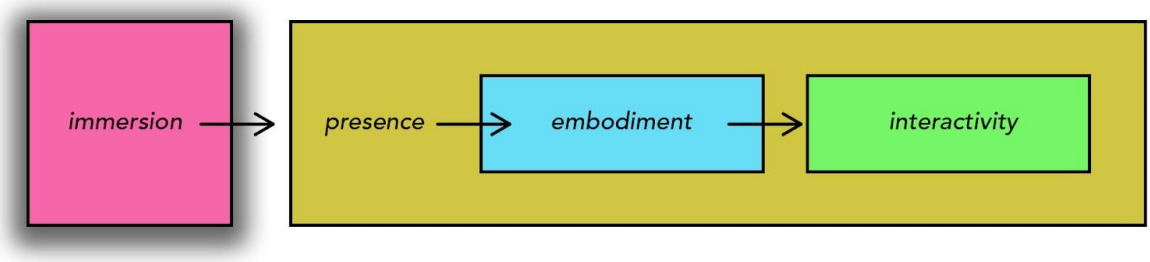


Figure 6-7: Immersion as related to presence, embodiment, and interactivity.

Immersion is objectively stated as the degree to which a virtual reality system "projects stimuli onto the sensory receptors of a user" [41] with no additional input received from external environmental sources [11]. Immersion is a technological approach using existing reality systems,

for instance the HMD, to create a psychological sense of presence within the user. Key ideas that strengthen immersion include:

Extensiveness - the range of senses a VR experience is being mapped toward (i.e. visuals, audio, haptic)

Matching - "congruence between what the senses perceive (i.e. shifts in sound and visual coordinates as the user's head turns)

Surroundedness - the length to which visual and audio cues are spatialized around the user

Intractability - the degree to which a user can affect elements of a virtual world, and how that world reacts as a result

Plot - the story, or the "portrayal of a message or experience, unfolding sequence of events, behavior of the world and its entities" [27]

6.2.8.2 Presence

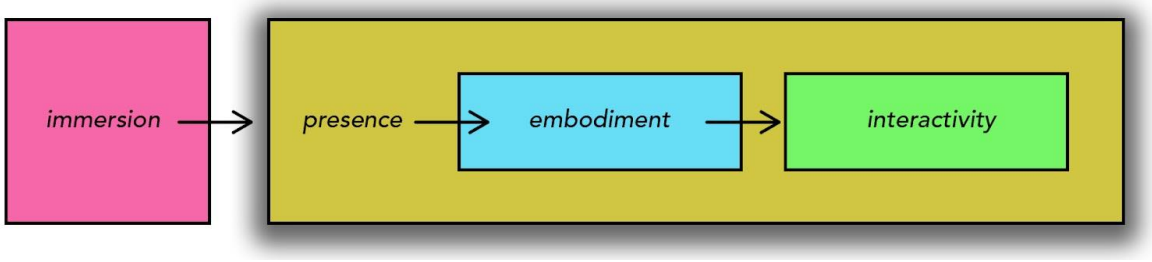


Figure 6-8: Presence as related to immersion, embodiment and interactivity.

Presence is the intended outcome of immersion; the experiencer's state of being, *feeling* as if they are physically inside of a virtual world. As defined by scholars at a 'presence conference' in 2000: "Presence is a psychological state or subjective perception in which even though part or all of an individual's current experience is generated by and/or filtered through human/made

technology, part or all of the individual's perception fails to accurately acknowledge the role of the technology in the experience." [23]

In the world of sound studies, we might think of immersion to presence as acoustics are to psychoacoustics: a tangible set of physical properties designed to first be received by our senses and then interpreted psychologically into a perceived experience.

6.2.8.3 Embodiment

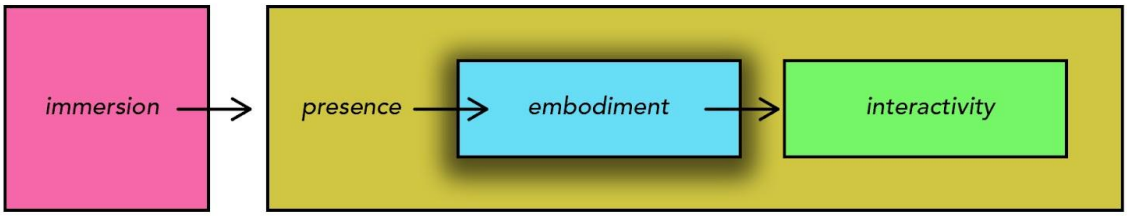


Figure 6-9: Embodiment as related to immersion, presence, and interactivity.

In "Philosophies of VR Design: Doing vs Being" Atherton discusses the relationship between presence and embodiment, particularly that for a user to be present in a virtual world, they must feel and behave as if an accurate representation of their physical body exists within it. "[Embodiment] is a mediation of the body, a way to experience our own bodies through the lens of a medium. Embodiment is a potent effect that has lasting impacts on users' psyche" [54].

6.2.8.4 Interactivity

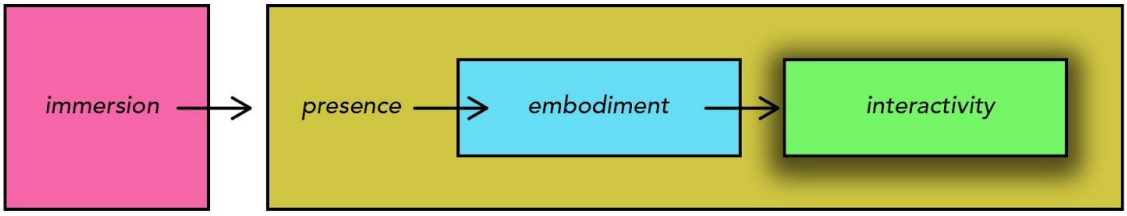


Figure 6-10: Interactivity as related to immersion, presence and embodiment

In comparison to aforementioned music platforms (vinyl, CDs, streaming and music videos), Interactivity is a unique affordance of virtual reality as a medium for music artists. As discussed in our *Communication Diagram*, not only do mediums, narrative and emotion iteratively inform each other to present a final experience to the user, the user themselves (once a sense of presence and embodiment is established), can influence the experience through their engagement with elements of the virtual world.

VR Interactivity can be thought of as an extension of Human-Computer-Interaction (HCI). In our day to day lives, we interact with computers all the time, whether our thumb swipes across our mobile phone screen or we click on an internet browser tab using a desktop mouse. In these cases, the computer is interfacing with us through the mobile phone screen, the mouse and the monitor. In virtual reality, particularly the scope of VR that we're covering here, the physical interface is the HMD and accompanying controllers, and the virtual interface is the immersive world presented to the user. Thus, the design principles behind HCI as translated to VR interaction are even more crucial to understand, as are some distinct differences exclusive to the VR medium. Notable HCI researcher Donald Norman offered six principles toward building interfaces designed for user interaction [56] and though not written solely for VR, this thesis aims to adapt these design principles toward VR interaction:

Visibility - the user should intuitively know how to navigate the virtual world they exist in using the interfaces presented.

Feedback - the user should receive feedback based on every action performed. Even something as simple as moving their head around should accurately portray a shift in sensory experience (i.e. sound and visual coordinate shifts), reinforcing a sense of presence and embodiment.

Mapping - the controls of the interface will map to the desired effect. For example, as we have integrated into "Pigments of Imagination", using the 'grip' button of a VR controller might simulate the holding of an item, in our case a paintbrush or trumpet.

Constraints - restrictions in interactions, implemented so as to not overwhelm a user, and therefore prioritize for them the most important interactive qualities of the interface. This is especially salient in a virtual world. As we will see, narrative and interactivity are in constant trade-offs with each other in VR. Understanding these principles behind general HCI will add nuance to discussions on the narrative-interactivity relationship.

Consistency - creating patterns of interactivity within the interface reinforces engagement behaviors and therefore user intuition. This idea can have the same desired effect in a virtual world.

6.2.8.4 Perception

In our discussion on music narrative, we proposed that music does not exist in a vacuum, it is context, particularly the context of the individual, which gives a composition its narrative weight. This applies in virtual reality as well. Just as an objective reality is perceived through one's subjectivity, such is an objective virtual reality. Just as a developer may create an objective approach toward immersion, the experience of presence or embodiment will always be subjective to the user. This individual subjectivity is defined here as perception; the way we "perceive our external environment within our mind" [27]. Specifically for the purposes of this thesis, we're interested in how an individual perceives sight and sound (especially sound) in a virtual space. This is important because when we dive into design principles of VR that act as guidelines when analyzing existing artistic works in the VR music video medium, we need a foundational understanding of how the objective nature of a virtual system is designed toward the subjective experience of a unique user.

As we begin to understand how perceptual processes work, we can create more objective virtual realities that articulate our subjective perceptions and imaginations. These perceptual understandings can then be used in conjunction with the physics, math, and algorithms used to create the underlying systems, behaviors, and environments of a virtual space.

6.2.9 VR Narrative and Immersion

In gaining an understanding of these fundamental concepts (immersion, interactivity, presence, embodiment, and perception), we can now begin to explore how they help determine the parameters of virtual reality narrative and design. From this understanding, we can more precisely identify the points of narrative and design congruence and conflict between VR and music and iteratively negotiate accordingly.

As we discussed above, within the relationship of music and virtual reality mediums, one reason this music-centered communication of narratives can occur is because virtual reality has yet to establish concrete narrative functions in the way that music has. In fact, in his paper *"The Next Innovation in Immersive [Actuality] Media Isn't Technology—It's Storytelling"*, author Michael Ogden makes the case that narrative development should be prioritized over technological development in the VR field. One linchpin question evoked is crucial toward reimagining the potential for music-centered narratives, or VR music videos. Specifically, Ogden questions whether the role dynamic between the storyteller and user might shift, particularly with the above ideas of presence and interactivity introduced into the fold [43]. In the VR music video field, we might restate this as the *performer/audience dynamic*, which reinforces the idea that a virtual space, no matter how artistically abstracted, must adhere by the same tendencies of the known physical space, in which we (the audience) virtually watch and listen to the performer, much as we would at a physical concert.

Figure 6-11: A still-shot of Björk via 360 video illustrating the audience / performer dynamic [67]

For instance, In Zack Bresler and Stan Hawkins "A Swarm of Sound: Audiovisual Immersion in Björk's VR Video 'Family'", the authors analyze work from the iconic artist Björk's VR album "Vulnicura", and in doing so define the purpose of a VR music video as "which is to conceptualize experiences of audiovisual immersion in music." By extending ideas on 'virtual acoustic space'" [63], they state, "this model emphasizes the relationship of visual imagery to sound and how this enhances agency for the *performer* and *viewer*." They continue "part of what constitutes aesthetic space [in VR] lies in the perceived proxemic relationship between *performer* and *viewer*, since the viewer is in constant interpretive negotiation between their own subjectivity and that of the artist" [67]. This is illustrated by the Björk video itself which hybridizes visual-only 360-camera interactivity with a more interactive 3D-world development, however central to it all is Bjork, the performer, and us, the audience, leaving little room for overarching narrative. Numerous other VR music videos, particularly those created for more well-known artists, tend to fall into this trope. This *performer/audience dynamic*, however, does not always lend itself to the

goals of the artist's emotional and narrative intent within a work. Moreover, this dynamic is not written in stone. It was derived from European observational cinema [53] and though it lends itself well to objective storytelling, neither of these attributes could be seen as necessary when analyzing the VR music video space through this thesis's lens and proposed *Communication Diagram*, and therefore this dynamic is open for alteration.

One reason this dynamic may be worth exploring comes from the challenges in balancing immersion and narrative. Immersion not in support of a prioritized narrative, Ogden argues, lends itself to the 'view' at the expense of the 'voice'. The initial spectacle of the immersive VR experience offers the user a sudden sensory and visceral emotional reaction, but the voice of the artist looking to connect their narrative, and therefore longer-lasting emotional response, to the experiencer suffers as a result. Immersion and narrative are not in conflict with each other; however, they must be in a uniform dialog with each other to work toward a cohesive whole. Ogden proposes a framework for collaborating "narrative technology" and "design technology", what he defines as "an attempt at merging the aesthetics of storytelling with the technological forces of immersive "storyworld" building by balancing the dilemma of user/participant agency with narrative authority (or structure) by fostering "narrative transportation" [36] such that the user is immersed and engaged with the virtual world (and it's interactive elements) while remaining "in the flow" of the narrative [43].

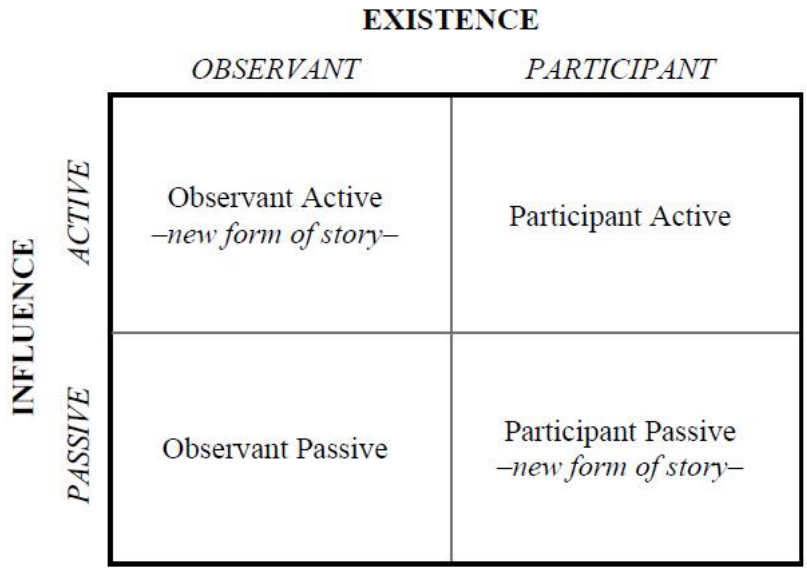


Figure 6-12: Approaches to VR Narrative through the lens of the experiencer categorized via the Passive and Active Observant and Participant chart by Dolan and Perets [13][43]

Dolan and Perets visually articulate this idea by diversifying the performer/audience dynamic into four quadrants [13] based on the user's existence and influence in the virtual world: *Observant-Active, Observant-Passive, Participant-Active and Participant-Passive*. *Observant-Passive* and *Participant-Active* exist on the extremes: from passive observation of a story to complete narrative control by the participant. *Observant-Active* and *Participant-Passive,* however, offer new forms of story. In *Observant-Active* forms, the user has no embodiment in the world, but is given the control to determine its narrative outcome. Conversely, a *Participant-Passive* user exists in the virtual world but is not asked of to dictate narrative decisions. Thus, each of these quadrants allow for various levels of user engagement or passivity. This is valuable because it re-contextualizes the performer/audience dynamic less as a binary relationship and more as a multidimensional spectrum of potential engagement between the two parties. If we revisit the *Communication Diagram*, the output generated from the *experiencer* toward *mediums 1 & 2 (VR and music), narrative (textual, visual and musical)* and *emotion* could be decided based on which quadrant the artist's proposed experience dictates works best to maximize emotional

resonance. This means the artist, guided by their work's emotional center within the musical space, can make intentional compositional and virtual-world-development choices from the outset as to which quadrant(s) the experiencer should exist in and negotiate the relationship between immersivity and narrative accordingly.

For instance, "Pigments of Imagination" exists partially in the quadrants of *Participant-Active* and *Participant-Passive*, as dictated by what I as the artist determined to be the emotional center of the work. Specifically, the experiencer, in taking on the role of a small child who ascends from their lonely room through the sky to the moon above, is invited to explore the room and its objects during the first act of the song. However, as guided by an intentionally arranged musical narrative, the experiencer shifts from an active to passive participant as the virtual world gently 'ascends' them out of the room toward the moon, akin to a slow rollercoaster ride to and from a peak point. The choice to alternate between active and passive participation is guided by a dialog between immersion and narrative in which each enhances the other. The allowance of the experiencer to determine and explore their presence and embodiment (as dictated by immersion) within the virtual world through a deliberate generation of their surrounding space hopes to avoid Ogden's pitfalls of early sensory spectacle, instead leaning on the musical elements to inspire a curiosity toward the overall experiential arc. The transition to passive participation ideally allows the experiencer to conclude their perception of presence and embodiment as an accepted 'truth' of the experience and situate toward the musical components that help reinforce narrative and emotional resonance. As the experiencer concludes their journey and returns to the room, they transition back to a state of active participation, this time with the experience and emotional impact of their passive track. The hope is that their sense of presence is not only a sensorial response, but a feeling further weighted by the emotional impact of their virtual experience.

This form of participation and observation is inspired by *non-musically-centered narrative-based artistic works* like "Notes on Blindness" and Celine Tricart's "The Key", VR short films that rely primarily on passive observation and participation but allow for active participation from the user at vital moments. The narrative for both works is ultimately deterministic (as is the accompanying score), however it's in these pieces' slight pivots to active user participation that allows immersion to reinforce narrative.

6.2.10 VR Narrative, Immersion, and Interactivity

Interaction, when used effectively, can enhance immersivity. It is important to consider then, as the creator, all the ways in which you can develop an interactive world that does not directly alter the architecture of the narrative. In our above exploration of emotion's relation to narrative, we examined *"Emotional Storytelling"*, and foreshadowed authors Kristofer Blom and Steffi Beckhaus's views on the interplay between emotion, narrative and interactivity. Interactivity, he states, "complicates the development of stories not only from a philosophical level, but also from an architectural and structural level." Blom and Beckhaus's research shows that "interaction should not be severely limited in scope by the system" as this can significantly "diminish the user's level of agency, thereby reducing their immersion in the story's world." Conversely, allowing the user a fully autonomous range of control might return that sense of agency, but it "complicates the generation of a meaningful story" [31]. It is for this reason that Blom & Beckhaus reinforce Ogden's sentiment that research in virtual environments is shifting from the technological to narrative and experiential-based structural innovations, and the reconciliation of interactivity and immersion is a key component in developing these new narrative planes [43].

In her 1999 piece *'Immersion vs. Interactivity: Virtual Reality and Literary Theory'* [38], Marie-Laure Ryan compares the medium of virtual reality with literature and notes that this trade-off

between immersion and interactivity toward a cohesive narrative is not a problem exclusive to VR. In the literary field, a greater allowance of interaction (for instance, autonomy of the reader in a choose-your-own-adventure book) lends more attention to the medium itself (the book) than the story. Ryan offers that to successfully maintain a balance of immersion and interaction; the experiencer must be able to "reasonably anticipate what will happen as a result of their actions in the reality." Thus, an ongoing problem is "how to create a coherent overarching narrative wherein the experiencer can make significant decisions that affect the world at any time (or even only at very specific times)" [24].

For this quandary we might stretch a bit further back into history to the Bardic traditions of storytelling. "In the bardic tradition, the audience exerted some level of influence on the Bard's story as (they) told it. However, the bard still retained control over the storyline, modifying only local pieces of the story without impacting the overall storyline" [31]. This is a good historical reference point for how music artists might think of structuring songs for limited interactivity with the experiencer - keeping the general architecture of the work within the music artist's control while allowing fringe alterations from the experiencer (the Bardic 'local pieces') that give them a certain sense of control. The music artist might reference similarities in a live improvisational jazz performance, in which the musicians agree to musically 'meet up' at certain points in the performance and then explore each musician's local boundaries of the song structure on their own before returning for another agreed upon 'meet-up'. This gives the overall performance a sense of narrative while still allowing the excitement of live exploration to coexist. Referencing the similarities between the Bardic model and the improvisational jazz model toward the narrative capabilities of the VR music video, the music artist can begin to think about composing in ways to which they maintain control over the structural integrity of the composition (and therefore narrative and emotional center) while still allowing the experiencer to contribute musically in such a way that they do not alter the intent of the narrative.

These approaches to interactive and immersive narrative-driven composition are especially intriguing when contextualized against other mediums of music release. Though we are comparing music interactivity within the field of VR music videos against other music-centered virtual reality experiences to properly define and explore the field's capabilities, we are also comparing them against the expectations of the experiencer compared to their other methods of receiving recorded music (i.e. streaming, social media, vinyl) which, accept for physical mediums, exist two-dimensionally and non-interactively. How might this look for a fixed-media recorded work? Perhaps there are gestures the experiencer makes within the musical world that alter the 'mix' of the song. Perhaps the creator leaves moments of openness in the original composition and guides the experiencer to create something in that local space, just as the Bards did.

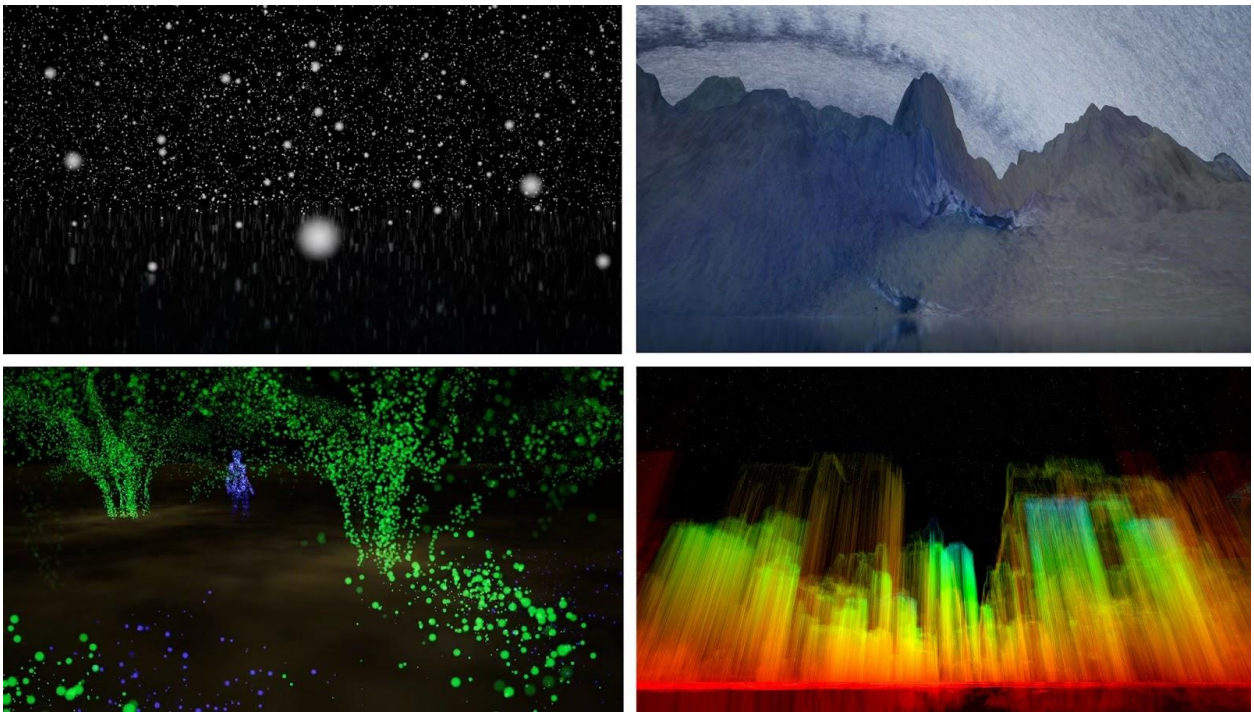


Figure 6-13: Examples of different environments from "Becoming: An Interactive Musical Journey In VR" that mark compositional transitions [57].

The music-centered virtual reality experience "Becoming: An Interactive Musical Journey In VR", an operatic narrative piece based on a Persian poem by Mowlana Rumi, illustrates this

idea in how it approaches evolving multimodal environments that actively engage with both the composition and user. “Becoming” collaborates physical recordings of vocals, strings, and percussion with synthetic sounds that at times support the “acoustic malarial” and at others the overarching narrative [57]. Compositional transitions between each environment respond to visual cues and act as pillars toward the work’s storyline. However the relationship between delayed lines of repeating percussive, musical and synthetic elements allow the participants to “localize” the story by changing or adding to the mix of the music [57]. Some fixed musical tracks are even attached to the user’s controllers, allowing moments of *participant-passivity* and *participant-activity* that still allow for a general storyline to guide the experience.

In “Pigments of Imagination", a limited spatial and temporal interactivity exists on the fringes of the narrative and is used to highlight the emotional center of the piece. Therefore, the opportunity for the experiencer to contribute their ‘local pieces’ is miniscule by comparison to more interactive music-centered VR experiences like “Becoming” but is still worth noting. Referring to the Experiencer's *active-participant* role in the child's room, nearly every item could be held with haptic response or tossed around the room with a logical physics-based response (i.e. velocity, gravity). Many items in the virtual room are instruments which the user can pick up, but not play. However, an ability to play these instruments, perhaps launching predetermined sounds or tones in key with the primary musical piece, might further immerse the experiencer while still maintaining a core narrative structure. However, one such object, a paintbrush, is highlighted through visual cues and onboarding instructions as a tool for the experiencer to pick up and paint with. The paintbrush reacts to the experiencer’s grip and painting movements both visually via an emergent interactive and colorful particle system, and sonically via synth notes mapped to the user’s movements that launch at varying lengths and pitches in key with the composition.

6.3 Negotiating Musical and VR Narrative

"If you look at Virtual Reality and you think that it will evolve the same way that other mediums have in the past, then you're thinking about it the wrong way…The language of cinema evolved out of a format that was birthed from a technology that was invented. The format and the technology were birthed at the same time, and the language grows out of that after many decades. In Virtual Reality, you have a technology, you have a format, and then you have a language. The language is unknown, but the format is also unknown." - Chris Milk, Music Video & VR Director [29]

In learning about the affordances of virtual reality, we see the challenges faced and opportunities that arise with a medium whose narrative architecture is still ripe for innovation. The challenge remains in the delicate balance and prospective trade-off between immersion and interaction toward a cohesive resonant narrative. The opportunity is in the lack of structure and therefore how that structure can therein be guided by an adjacent medium's existing narrative architecture, in our case music. Thus, we can revisit Wingstedt's classes, categories and concepts of musical narrative functionality and multimodality as a series of guiding principles that can uniquely navigate the challenges of immersion and interactivity to better inform virtual reality narrative development. Particularly, we can leverage the affordances of music as the narrative center, relying on centuries of accepted narrative traits that we as listeners subconsciously recognize (exchange of redundancy and entropy, musical context, multimodality with multi-instrument compositionality) to fill in the gray-areas of virtual reality narrative once the negotiation between immersion and interactivity are introduced. But how is this done?

By using music as a foundational medium in which the experiencer is already trained to seek narrative functionality, the artist gives themselves leniency within virtual reality, where the expectations of narrative functionality are far less apparent to the experiencer. This allows the artist to better fine-tune their approaches to the reconciliation between immersion and interactivity while still maintaining a defined narrative. As Wingstedt said, music does not exist in a vacuum, rather it is an "ongoing process…in constant interplay with the surrounding world" [30]. In our case, that world is a virtual world. Therefore, we can think about processing an

existing musical work through the lens of Wingstedt's *emotional, guiding*, and *temporal* classes to A) get a sense of the music's emotional center and B) determine how it might guide visual accompaniment across the time domain. If expressed successfully, the virtual world that will surround this music, an extension of Wingstedt's theory of musical multimodality, will begin to form emotionally and psychologically within the listener before they ever put a VR headset on. With a composition's narrative functionality as centerpiece, the artist can then better decide how they view the experiencer's engagement within the narrative: *passive-observant, active-observant, passive-participant or active-participant*.

Conversely, the abstract nature of music allows for interpretation of musical narrative functions, which means they can be stretched and reimagined, more than other narrative-based artistic mediums (i.e. film) that virtual reality narrative wraps around. As filmmakers face challenges mapping the narrative rules of a two-dimensional third-person medium to a three-dimensional first-person medium, there's little room for the existing affordances of virtual reality to influence these narrative structures, creating notable constriction. Music, however, already exists in the first-person three-dimensional space and can interplay with narrative accordingly, allowing key components of VR narrative to in turn inform musical narrative.

For instance, how does the interplay between redundancy and entropy within the musical narrative align with a VR narrative that might have the same dependencies? How does this correlation inform the redundancy/entropy dynamic? In "Pigments of Imagination", we turn the conventional structure of a (redundant) second verse (vocalized lyrics, a standard in hip-hop to be discussed later) into a dance between synths and flugelhorn that both thematically and directly (through audiovisual reactive mapping) interact with the textual narrative. We then use the vocalized lyrics as suggestive exposition to foreshadow visual cues within the VR world and in turn use visual cues to foreshadow lyrics. The idea is that the experiencer, coming to terms with their presence and level of interactivity in the virtual world, may not digest lyrics on the first

experience, thus these lyrics operate on a more subconscious suggestive level while the musical tones, textures, spatial decisions and shifts in movement interface more directly with the user's perceptual and emotional cues. This is in opposition to stereotypical musical narrative, particularly in popular music genres, where the lyrics are designed to be the top priority for listener focus. However, as the experiencer repeatedly revisits the virtual experience, their expectancy of presence and interactivity are set and compositional components, such as lyrics, can move to the foreground and ideally provide an increasingly enriching experience (i.e. replay value).

Thus, the role of lyrics in this experience oscillates between *diegetic* and *nondiegetic*, shifting with each recurrent listen. In this sense, the term 'diegetic' is key to unfolding whether narrative guideposts in the VR experience are received from 'within' the virtual world as beyond it. Anıl Camci's "Exploring the Effects of Diegetic and Non-diegetic Audiovisual Cues on Decision-making in Virtual Reality" offers the case study of a virtual room in which the user is given cues on how to navigate to the proper exit door. *Diegetic* clues include elements within the virtual world: a poster on the wall or note on the table with instructions. *Nondiegetic* cues might include audio or written instructions presented before the VR game starts, or overlaid on top of the virtual world for the user to read or listen [3]. To reinforce how VR-narrative cues can engage with music-narrative strategies in unique ways, "Pigments" plays with these diegetic cues in the examples above: the suggestiveness of the lyrics exist within the world but also narrate the world. Though more nuanced than explicit audio instructions, through metaphor and imagery, the cues are presented to the experiencer.

In the *non-musically-centered narrative-based artistic work* "Notes on Blindness", the narrative revolves around the life-long recordings of writer John M. Hull, whose life of deteriorating vision eventually led to blindness (Robey). The recordings of Hull's voice are accompanied by spatialized field recordings of various environments that act as the centerpiece

of this VR work, in which Hull recounts his perception of these environments using the acoustic properties of sound to build the shape of sight (people, animals, wind, trees, cars, etc). His notes, spanning decades, function as the emotional center. The user's subtle interaction with the accompanying visuals and field sounds are minimal; designed to engage the viewer toward the affordances of virtual reality through participation, but not so much that the importance of Hull's omnipresent recordings is ignored. Thus, the poetry, melancholy and joy expressed in the words are only brought to life; reinforced by a three-dimensional world in which visuals are abstracted to particle systems as a simulation of sight deterioration. Music's role in this piece is minimal: 1) a brief, lightly spatial generative synth pad transitions two scenes that are dependent on user interaction (the generative music allows user activity to span different times with no determined end), 2) static fixed scores for the credits and a scene transition, and 3) a diegetic and beautifully spatialized pianist and choir singing in a concert hall. The key takeaway here is to see how non-music-centered works are still effectively using sound and music both within the VR world and in narration of that world, all while illustrating stories in which the VR medium itself can expressly amplify the foundational emotional underpinnings.

Ultimately, the experiencer will build their own interpretation of narrative, regardless of how the artist attempts to frame it across these mediums. "VR creators will never be able to take the subjectivity out of a story experience. Users will always distort any stimuli presented to them based on their filters of the values, beliefs, memories, etc." [27]. The hope for the artist is that by modeling their emotional and narrative objectives through the lenses provided, they can more accurately connect the emotional intent of their work to the experiencer through the mediums, and within the mediums, a practical set of unified design principles to help facilitate these goals.

Guiding Philosophy pt. 3: Extending the Narrative Through Mediums

Chapter 7 : Mediums

So far, we have created a theoretical framework for the multidisciplinary music artist to explore the virtual reality space as a medium to convey their musical work. We have presented a *Communication Diagram* designed to strengthen artistic practice by defining an emotional center and conveying its resonance through the intertwining of musical and VR narrative. We have also discussed the way these components (artist, emotion, narrative, mediums and experiencer) iteratively interface with each other throughout the creative process to ultimately sculpt a final project. In our exploration of mediums, specifically the key design principles of both music and virtual reality, we can begin to more practically apply tangible results based on this diagram.

In this section, we will offer a quick synopsis of each medium, set expectations of what we are attempting to achieve within each, and then explore a selection of accepted design philosophies and practices based on those expectations. In exploring these practices, we will bridge discussed narrative frameworks into design philosophy and practice. We will also continue to examine projects from our case study Venn-diagram shown at the beginning of this thesis to provide the artist with a sense of how others are navigating music-related works in adjacent VR fields. In doing so, we will also look at specific components and approaches in “Pigments of Imagination” through the lens of existing and proposed altered design principles. The hope in doing this is to offer artists the suggested tools and artistic approaches to flourish in a meaningful way as music artists in the VR space, thereby allowing them to redefine the parameters and expectations of what is currently termed the “VR music video”.

7.1 Music Medium

Now that we have a greater understanding of the VR space music artists are looking to create toward, an exploration of each medium's design principles and case study analyses can follow. Wrapped around and in dialogue with the narrative is the medium of which the art will form, perhaps as theater, film, painting, recorded album, or virtual reality experience. If chosen faithfully to the emotional core, this medium (or combination of mediums) will inform the narrative, add depth, nuance and a revealing quality to the work's core truth/emotional center. For our purposes, music is the assumed primary vehicle to convey the contours of narrative and emotional seed.

In Johnny Wingstedt's *"Narrative Music: Towards and Understanding of Musical Narrative Functions in Multimedia"*, the author argued that multimodality existed within musical composition which made extending musical functions toward visual mediums a natural progression. This assumes an interdisciplinary seed within the composer to ultimately either create or collaborate further within the visual space to achieve the vision of their musical work. In the modern world of music creation, this extension of Wingstedt's modalities finds room to grow within the music field itself. In the last 20 years, DAWs (Digital Audio Workstations) and home studios have allowed the 'bedroom producer' to blur the lines between composition, music production and audio engineering, writing, recording, producing and mixing their own works, often before uploading to a streaming service for the world to hear.

Each of these subsets of music creation (composing, producing, and engineering) and release exist in the physical space, and just as virtual reality offers design principles toward building virtual worlds, modern popular music creation has developed principles as well. For the scope of this thesis, I will focus on principles that might engage directly with and possibly transform because of their relationship to adjacent principles of virtual reality design. Similarly,

I will focus on particular areas of virtual reality design principles that might be reconsidered when engaged with adjacent principles of an alternate medium. Whether the audience of music artists choose to use these principles, tools and examples to explore the potentials of their own compositions in a virtual space, or to more confidently and adventurously seek collaborations with existing VR-developers, the hope is that this thesis presents a new lens toward creating at the intersection of these mediums.

7.1.1 Design Principles of Music

This paper groups the essential design principles of music into four categories: *composition, production, engineering*, and *reception,* and revolves explicitly around application techniques. The author chose these categories to cover the sequential process of a music artist 1) composing work, 2) producing the musical elements of that composition, 3) refining the work and 4) releasing the work publicly.

7.1.1.1 Composition

Compositional design principles are unique to genre, culture, and intent; what Wingstedt calls the *societal dimension* of context when examining musical narrative functionality. Thus, they are subjective and constantly evolving. Rather than make a futile attempt to cover them all, this thesis will approach composition from two perspectives:

1) *Genre Study:* Single out a popular music genre for its existing conventions and provide case studies of existing works in the music-centered VR space that reinforce or deconstruct those conventions. This examination can then be extrapolated to other popular music genres that rely on similar compositional and narrative structures and may find their own unique design conflicts with a virtual reality medium.

2) *Composing Toward Visual Mediums:* Review conventions and relational history of sound/music composition and image. In offering examples of music-image relationships outside of VR, this section hopes to provide context to how these mediums engage, and these conventions may shift when composing toward VR.

7.1.1.1.1 Genre Study

"Pigments of Imagination" is "guided by a spatialized fixed media composition rooted in hip-hop but spanning genres of ambient-electronic, soul and blues as well. It was composed and produced with narrative and visual aesthetic in mind through a blend of field-recording, synthetic sounds, sound-design, vocal and acoustic recordings" by the author and a host of collaborators [62]. Part of what makes "Pigments" a fascinating case study is its foundation in hip-hop, a popular music genre whose structural conventions do not easily align with the affordances of VR (though there are some beautiful overlapping concepts of non-linearity that unfortunately exceed the scope of this thesis [64]. This occurs for a few reasons: 1) With a heavy emphasis on lyrics, hip-hop is a fundamentally narrative-driven genre 2) Beyond a few exceptions, which will be reviewed, hip-hop has not heavily embraced virtual reality as a medium either for sole visual accompaniment or deeper musical and narrative exploration, 3) Like many popular genres, modern hip-hop music is often composed, recorded and produced in DAWs (Digital Audio Workstations) that view the linearity of music on a two-dimensional plane reading left to right (more on this below), and 4) the prioritized loop-based, rhythmic and lyric-driven structural elements of hip-hop in particular [64] can conflict with basic procedural music-centered virtual reality design principles [24]. These reasons make hip-hop an especially rich genre to examine, as we can see more explicitly how virtual reality might deconstruct and rebuild an original musical work in radical ways. Similarly, we can see how a genre who's core elements appear incongruent to the strengths of virtual reality might influence its design principles as well.

The first convention "Pigments" defied to adapt to VR was to redefine the rhythmic component. "Pigments" stripped away the syncopated elements (i.e. drums) and instead relied on the subtle rhythm of ambient keys to both ground vocal performances and define the space of a virtual world. The second convention broken was in creating a full lyric-driven narrative (a song defined to contain lyrics through most of the work). Though an introductory verse and hook exists, the space for a second verse is replaced by solo synth and flugelhorn, both of which are thematically *depended upon* to push VR narrative forward and *dependent on* the 3D visual accompaniment to better represent their intent within the narrative. The third convention defied is in arrangement. Conventional hip-hop records offer 4/4 time-signature looped sample-based music production with minor variation (a strength, not a shortcoming). On top of this production are a set of 2-3 verses at 16 bars in length with an 8-bar hook between. To better engage with the VR narrative, "Pigments" departs from the convention of redundancy for distinct entropic movements of differing time-signatures, tempos and instruments that push immersion and narrative forward. Essentially, framing this as a narrative-based VR piece early on motivated and redirected some of these key musical elements and arrangements, allowing "Pigments" to maintain some founding hip-hop conventions while organically fusing with other genres (adding a wider cultural enrichment to the work) that naturally align with VR affordances.

Other hip-hop-oriented VR music videos have yet to explore altering these conventions. In fact, many from this genre have yet to fully adopt standard VR affordances that could exist without altering musical conventions. For instance, the critically acclaimed and forward-thinking hip-hop duo Run The Jewels released "The Crown" in 2016, a VR music video shot on 360 camera [68]. The song maintains the same compositional structure as the audio-only album version, meaning there is no tangible user interactivity and narrative is mostly reliant on the lyric-driven song itself. Potential immersive components such as audiovisual reactivity and sound spatialization are not utilized either. However, for what this VR video lacks in interactivity and

immersion, it makes up for in accessibility, as it was created for the New York Times VR mobile app, suggesting it's meant to experience on lower-cost HMDs like Google Cardboard. It should also be noted that this app no longer exists, nor do the additional VR videos promised to be launched by the group, suggesting either artist/fan disinterest or a niche market that did not warrant the artist and company pursuing the medium further.

On the more inventive side was hip-hop artist Travis Scott's 2020 collaboration with popular online gaming platform Fortnite for the "Astronomical" live VR Concert. Song "performances" were single audio files of the original recordings, which means there was not much room to deconstruct compositions and genre conventions, or to even offer spatialization. Additionally, the concert format ensured a standard *passive-observant* audience. However, creative audiovisual mapping, haptic feedback and beautifully illustrated surreal environments help reinforce this work's multimodality, therefore driving the user to the emotional resonance within the music and placing this popular VR experience at the top of what might inspire future generations of artists to explore this intersection. Also, the live-performance aspect of this experience "gives access to cues such as the artists' facial expressions, posture, and movement" [19][44][59]. Though live performance is purposefully disregarded for the scope of this thesis, studies of these multimodal features are worth exploring and potentially integrating into narrative-based music-centered VR work. Given the success of this collaboration, it is not difficult to imagine how using the convention-defying framework of "Pigments of Imagination", Scott and his producers could create music toward VR, building a deeper immersive and possibly narrative-driven experience.

7.1.1.1.2 Composing Toward Visual Mediums

"The happy marriage between image and music is a fascinating example of when the whole becomes something much bigger than the sum of the parts." - Johnny Wingstedt [30]

The origin of the English word 'music' is the Greek *mousiké*. In ancient Greece, this term defined not just the configuration of sound in harmonic and rhythmic structures, but 'the arts of the Muses'. "The use of music (in the modern sense) in ancient Greece was not that far removed from the [multimodal] 'media music' that we have today" [30]. It is clear then that the union of sound to a visual medium has existed well before the emergence of virtual reality, and it's worth exploring a brief history of this relationship to gain a sense of how the interaction between VR and music both fit within this history and uniquely expand upon it. It's especially important to observe landmark collaborations in which music composition (defined in the modern sense) is the guiding inspiration or narrative container in which visuals react.

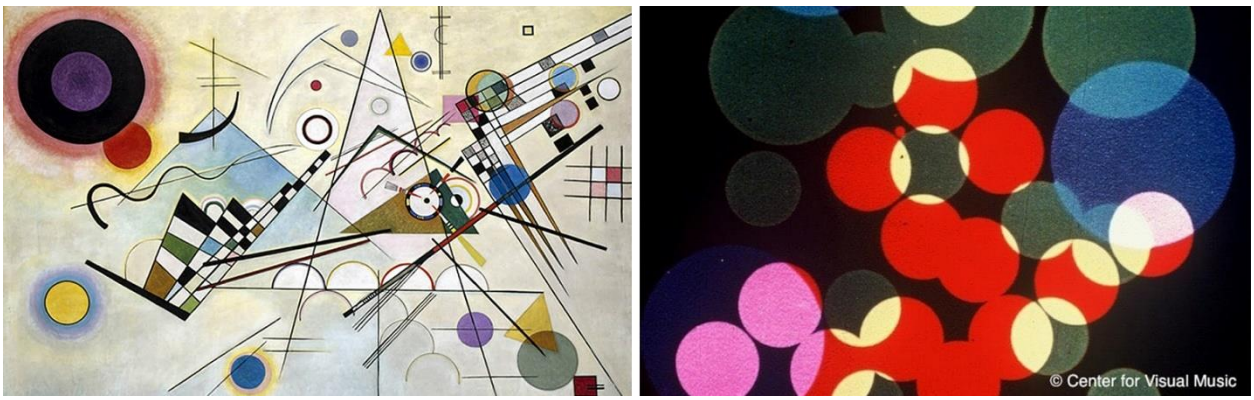


Figure 7-1: Russian painter Wassily Kandinsky's oil on canvas work Composition 8 (Reuters) and Still from Fischinger's Kreise (Circles), 1934. Courtesy of the Centre for Visual Music

For instance, Steven Holtzman's "Digital Mantras: The Languages of Abstract and Virtual Worlds" examines the careers of Arnold Schoenberg, a classical composer, and Wassily Kandinsky, an abstract painter inspired by Schoenberg's democratization of the chromatic scale. Schoenberg's distillation of western classical music's approach to intervallic musical structures redefined tonal relationships between notes. Seeing the relationship between sound and art, Kandinsky was relieving his painting of tonality as well [58]. In his letter to Schoenberg, he wrote of a common need to boil art down to pure emotional expression and recognized that force in each

of their works. Kandinsky dissolved the ideas of realist painting by championing 'color over form' and letting go of 'subjects' within a piece. Kandinsky was more intent on capturing the feeling of creating a piece rather than a particular subject for painting. Color to Kandisnky was what tone was to Schoenberg and both were concerned with centering their work around an emotional center, regardless of medium.

Graduating to moving pictures, filmmaker Oskar Fischinger was interested in the idea of 'abstractions-in-motion', specifically within the relationship of sound and visual [42]. He spent the 1920s creating various cinematic studies of abstracted visual effect creation, many of which might be seen as the precursors to how we understand visual effects via film, television, and digital media today. These early works led to another series of studies in the 1930s that focused on the syncing of visuals to preexisting music and sound sources. In these works, the abstracted visuals followed the rhythmic, melodic and timbral shape of the orchestral soundtrack they were designed around. As a way of drawing more attention and coherence to the studies, Fischinger tightly synced these visuals to more commercialized music. This gave the audience a launching point to better understand the nature of his visual work. Just as with Schoenberg and Kandinsky above, Fischinger was focused on a departure from the accepted reality of his medium and sought to create in a visual language guided by musical structures. Though he did not contribute to the final product, a great commercial example of this philosophy is the Disney film "Fantasia", a music-centered narrative work filled with surreal animation often emulating the orchestral movements and timbral attributes of the original score.

In the contemporary landscape of TV and film, the roles of sound and visuals have been mostly inverted; prioritization of a linear visual narrative and accompanying design principles supported by sound-design and music score to achieve a more immersive experience. Thus, the standard convention here worth deconstructing has less to do with a particular approach to

composing against visuals, and more with the overarching prioritization of the linear visual medium itself, in which nearly every compositional decision is made in reaction to its visual counterpart. Virtual reality, in its first-person construct, dissolves ideas of linearity most attributed to these other digital visual mediums (i.e. TV and film) and therefore invites a recalibration of prioritization between mediums. Scoring and sound-designing from "beginning to end" are no longer inferred, and in that epiphany, all conventions of composing to a linear narrative are up for reevaluation. Though the music artist, specifically within the scope of this thesis's focus on fixed-media recordings, might still plan to compose toward VR with a certain sense of linearity, a spectrum of approaches exists to stretch that definition across non-linear terrains, therefore opening the doors to new compositional techniques.

Thankfully, evolutions of approaches to scoring and sound-design offer worthwhile principles that *can* be carried into virtual reality. In Andre Knight-Hill and Emma Margetson's "Art of Sound: Creativity in Film Sound and Electroacoustic Music", a group of emergent sound-artists and composers are asked to discuss their creative processes. From these interviews, a broad set of principles manifest, including procedural techniques toward sound discoveries, curiosity toward new forms of sound expression and eliciting emotional response, among others. This set of principles works because it isolates the sonic components from the visual medium and can therefore be applied to mediums that do not live by the same rules as film, such as the music-centered virtual space. As discussed above, this works to significant effect in the VR short film "Notes on Blindness" in which the emotional crux lies less in visual cues and more in the narrator's story and the surrounding field recordings designed to enhance his narration [61]. The visual aesthetics of the piece then act as the final layer of enhancement, the most direct component to interface with the user, but one that acts more as a path toward the true emotional center found in sound.

I propose that modern music artists first explore these principles in depth, then revisit the initial works of Kandinsky and Fischinger as catalysts toward the reversion of music-centered audiovisual artistic work, specifically in the non-linear VR space, and finally approach ideas of composition accordingly. This means prioritizing the music's emotional core and narrative as discussed, letting the visuals (and other sensorially prominent components afforded by VR) react and then allowing both to negotiate toward a uniform solution.

7.1.2 Production

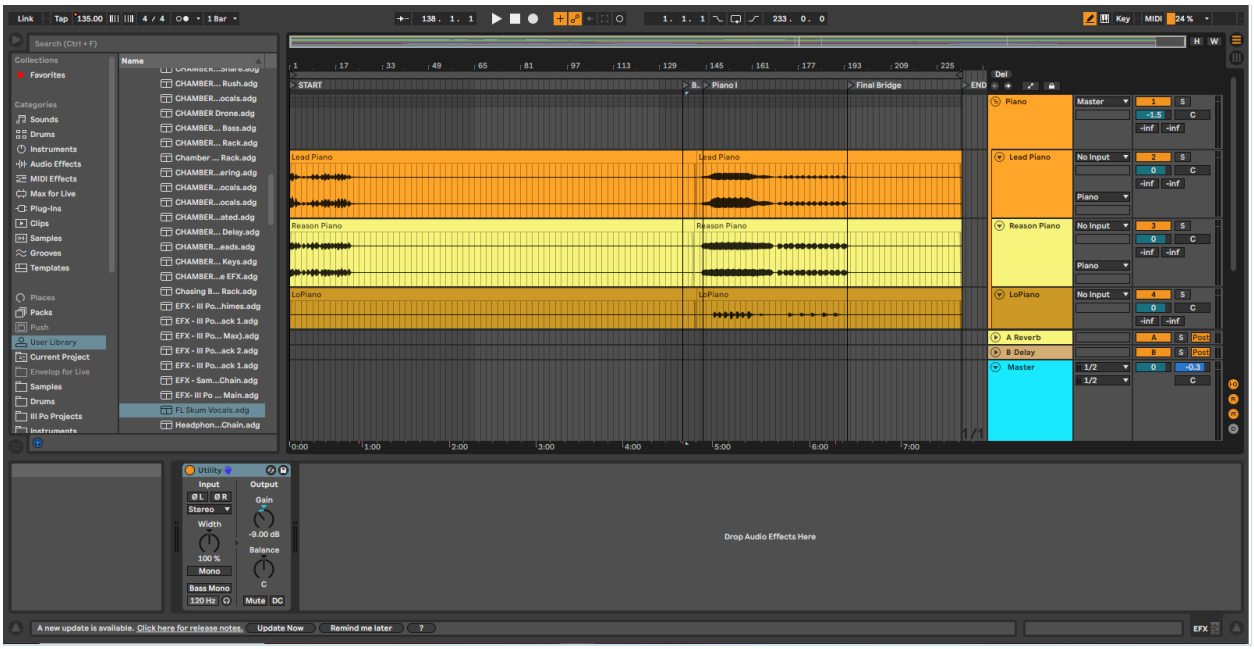


Figure 7-2: Example of a Digital Audio Workstation (DAW) featuring examples of recorded audio read left-to-right across a linear time domain.

For the modern 'bedroom producer', regardless of genre and especially when composing alongside visuals, the blurred line approach of composing, recording and mixing as one continuous process is often captured and manipulated on computers via a digital audio workstation (DAW). "DAWS are software packages designed to recreate and replace many of the technologies found in a traditional recording studio" [4], which, given their comparatively affordable nature, is what makes them so ideal for the contemporary music artist. Centering

modern music production (production being the full tangible construction of a musical work) around the DAW is an important concept to highlight because the DAW contributes vastly to how we perceive linearity in music. In a DAW, audio and MIDI tracks are either recorded, imported or directly composed and read across the time domain in "a horizontal timeline recording layout that emulates the linear temporality afforded by previous media forms" [47]. As we have discovered in examinations of music and VR narrative, VR opens the doors to non-linear approaches to composing and production.

Concepts of nonlinearity are by no means new to music. Composer John Cage was famous for his use of indeterminacy in music, leaving openings for chance and randomization in a piece (often represented through abstracted graphic scores versus left-to-right standard western notation). Similarly, Cage's one-time-student Brian Eno, a pioneering electronic musician, innovated the idea of non-deterministic electronic 'ambient' music, following similar principles through a sound-synthesis medium. These approaches have long existed in the physical music-making space, well before the suggestive linearity of DAWs housed modern music creations. However, these nonlinear approaches have also existed in the discretized digital space, and it's important to understand the co-existence of these two worlds when thinking about reconciling the pros and cons of linear production toward VR. In "Push: Software Design and the Culture Politics of Music Production", author Mike D'Errico explores these process-oriented approaches to music production: "While many commercial DAWs present the user with a linear, timeline-based interface best suited for music recording, [visual programming environments like Max/MSP and Pure Data] specialize in process-oriented [creation], presenting the user with a blank screen" on which the user can patch objects to build procedural sound and image-based media. These two innovative programs, developed by Miller Puckette, were designed to "liberate musicians from…creatively limiting features of commercial DAWs [that] conceptualize sound as recorded material…temporally mapped onto crucial points of…a linear narrative structure" [47].

We will soon explore fully procedural approaches to music and sound-design in virtual spaces when reviewing VR design principles, but the key takeaway here is that programs exist to completely depart from the potentially confining linear structure of DAWs. However, as we have examined, complete non-linearity, especially when reliant on interactivity or reactivity, can potentially dissolve narrative intent. Our goal is to carve a compromise between linearity and nonlinearity. I argue that it is in virtual reality where these abstract concepts can be represented in a three-dimensional visual space and therefore better negotiated.

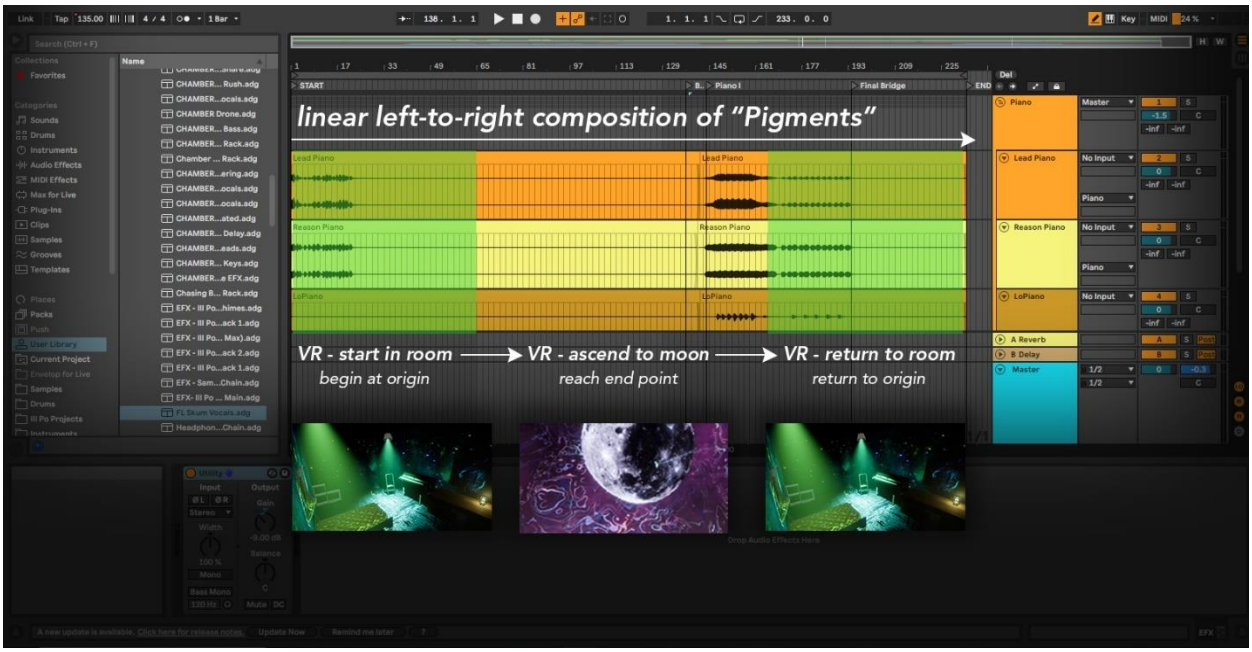


Figure 7-3: A looped narrative example in "Pigments of Imagination" in which, though the composition is a deterministic left-to-right linear composition, the VR narrative counters by reaching its destination at the midpoint of the composition and returning back to

For instance, "Pigments of Imagination" was created in Ableton in a linear, deterministic way through a collage of analog recording and MIDI-based production techniques. Once this production is introduced to the experiencer in VR, they can place themselves "inside" of the composition, shifting their positionality from *left-to-right* linearity (as composers and producers are trained to think in) to *in-front-of/behind* or *past/future* linearity. Though the arrangement of

the music is still produced in a linear way, the journey of the experiencer to and from a destination (the moon) aims to shift the standard 'beginning and ending' of a linear work into past (what exists behind the experiencer) and future (what the experiencer can anticipate in front of them). Therefore, the journey of the experiencer to and from a single location in their room, as experienced from a first-person perspective, gives the narrative a simultaneous conclusion *and* looped arc, as if the original production in Ableton was produced from left-to-right until a midpoint and then produced right-to-left. This might sound abstract to DAW-centric producers, but it's not much different than how first-person video games approach the topic. This *past/future* idea might be most recognizable in the popular VR game "Beat Sabre" in which the user rhythmically slices virtual blocks mapped to the BPM and musical elements of a fixed composition. Though marketed as a game, it achieves a similar structural departure of two-dimensional left-to-right linearity. Examining projects like this and "Pigments" aim to show how expanding into virtual reality as a supplementary medium can truly expand not just techniques on music production, but on foundational concepts of linear arrangements.

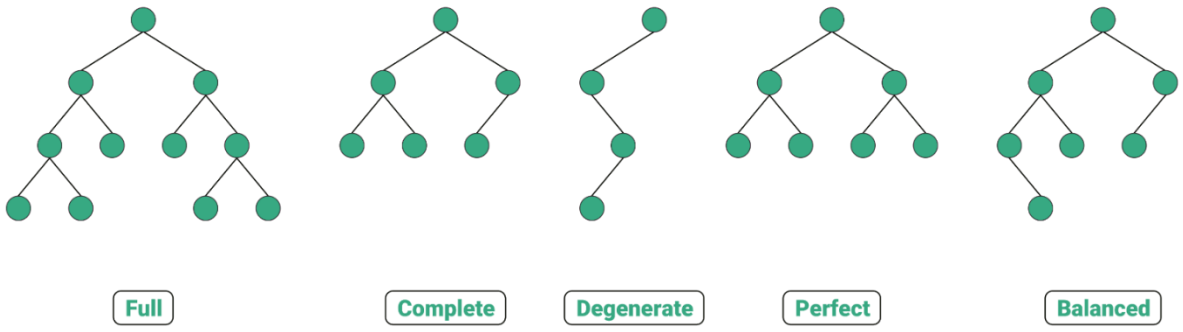


Figure 7-4: Examples of binary tree formations, of which the balanced binary stree structure presents a potential forked-path narrative and compositional approach that allows for a single beginning and ending.

As suggested earlier, the author's speculative in-progress music-centered VR work "Do You Smile In Your Dreams?" aims to further explore ideas of composing and producing with these affordances in mind, offering forked-path narratives in which numerous miniature linear paths exist in a larger, nonlinear balanced binary tree structure (exemplified in the figure above). In this way, the composition and production of the musical piece, which guides all other attributes, is recorded as linear deterministic fixed-media that can then be re-purposed alongside generative musical elements and split into a multiverse of sub-narratives that reach a singular conclusion.

For the music producer or recording artist tied to the use of DAWs (and their implication of left-to-right linearity), extending production techniques to virtual reality allows the artist to A) visualize the potential fragmentation of assumed compositional linearity in the fixed-media space, and B) collaborate fixed-media with generative (dynamic) audio to expand on nonlinear ideas in a practical and meaningful way. As seen above with examples of Mas/MSP, Pure Data, John Cage and Brian Eno, these approaches are not new, but in VR the music artist has the opportunity to better visualize them in a more tangible three-dimensional and interactive way.

7.1.3 Engineering

Whereas the properties of virtual reality can help crystallize theoretical nonlinear music production techniques, it can in a very practical way reimagine standard audio engineering techniques, particularly in its ability to accurately model the acoustical properties of physical spaces. Though recording audio in a virtual space would be fantastic for VRMI (Virtual Music Instrument) studies, it extends beyond the scope of our intention as the music in our studies has already been recorded. Therefore, this thesis aims to explore fundamental concepts and design principles of post-production mixing of said audio. Below are three key features of mixing that relate directly to spatial audio approaches in virtual reality and can therefore be altered.

7.1.3.1 Principles of Loudspeaker Stereo

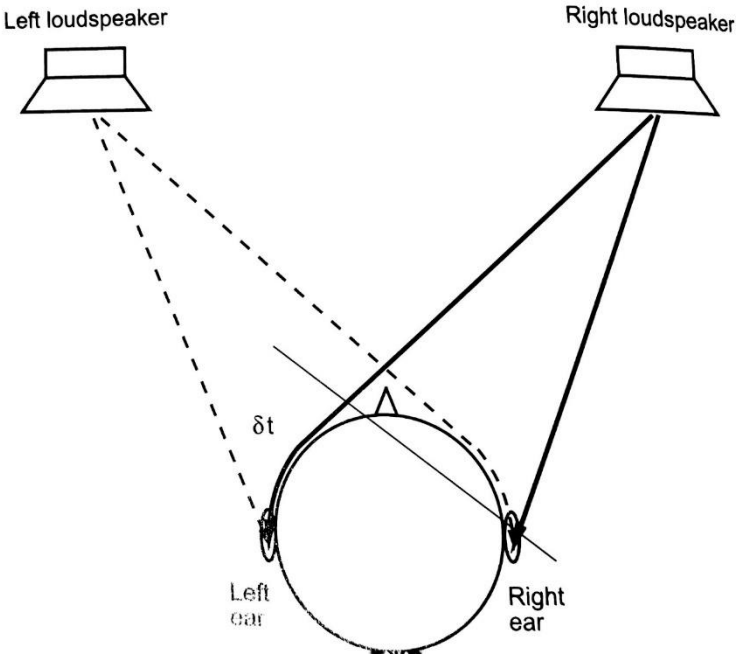


Figure 7-5: Two-channel stereo reproduction as received through a listener's left and right ears [17]

In the physical world, two-channel stereophonic reproduction (two front channels and no surrounding channels), or "stereo", is "the most common way of conveying some spatial content in sound recording and reproduction" [17]. In an optimal listening set-up, an equilateral triangle is formed between the listener, left and right loudspeaker such that an equal amplitude of signal from each speaker reaches both ears with equal delay approximation. This means the signal from the left speaker reaches the right ear at a slight delay conversely equivalent to the other ear and loudspeaker. The result offers the illusion of a three-dimensionalized sound, though this is a modest illusion since sound is only being reproduced in the front quadrants [17]. In fact, one could visually analogize stereo as a two-dimensional movie screen and multichannel audio spatialization as a full three-dimensional experience (akin to virtual reality).

As it relates to mixing in stereo, it's of typical practice to place the most prominent elements of a song directly in the center of the stereo-field. One key reason for this is prioritizing their survival when being played in a mono system (laptop or phone speakers, or a shared earbud for instance). Therefore, "kicks, snares, bass and lead vocals will normally inhabit the center" [48]. Beyond experimental outliers, this has become common practice across nearly all genres of music involving any of these elements and it will be highlighted as a design principle worth negotiating with when we encounter approaches to spatial audio in virtual reality. Another point worth thinking about is the use of panning automation, the subtle movement of musical elements across the stereo field over a fixed time. In a stereo-mix designed for speakers or headphones, this mixing choice can add movement to the mix and enhance the mix's stereo-width. However, when porting to a binauralized VR setting, this automated panning can become counterintuitive to the natural panning that occurs based on the experiencer's continual head rotations.

7.1.3.2 Reverberation

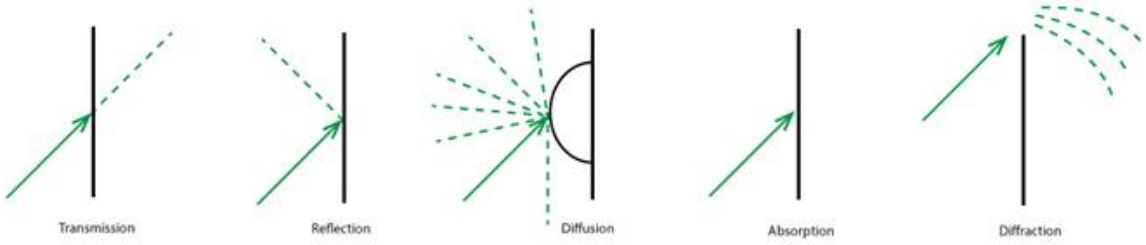


Figure 7-6: Basic examples of how sound reflections respond to surrounding surfaces [9]

Free fields are "acoustical areas in which there are no reflections" [17]. In this field, which rarely exists in real life, sound only radiates from its source, and none is reflected back. In reality, sound waves, as containers for potential and kinetic energy, are constantly reflecting and absorbing across surfaces. Depending on the surface type, a certain percentage of sound may be absorbed, while the remaining sound waves repeatedly reflect, possibly in multiple directions,

termed diffusion. Depending on the size of the soundwave and the size of the object containing surfaces, the waves may also diffract, or bend, before eventually tailing off at varying frequencies.

In the world of audio engineering, reverb can be represented either in the recording process or the mixing process. When capturing reverb in the recording process, the engineer creates the preferred acoustic conditions within the studio and records accordingly what might be considered a natural reverb. In the mixing process, artificial reverb is applied to the recording, typically using analog equipment or virtual plugins that offer a series of parameters to control the conditions discussed above: absorption, room size, delay time, diffusion, and more. The usual goal for the engineer is to create a shared sonic space for the recorded, sampled, or electronic musical elements of a song to 'live' inside and therefore glue together. Subtle uses of reverb can enhance or abstract the illusion of three-dimensional space in a stereo setup.

Challenges may arise if these musical elements are destined to soundtrack a virtual space, especially one in which acoustic conditions of the world have been applied. The music artist in this case runs the risk of applying artificial reverb set for an imaginary stereo-field room to a single audio/music stem and then placing that stem in a virtual space where natural reflections will occur. This buildup of reflections may cause unwanted 'smearing', both of the sharpness of the dry signal and its spatial coordinates. This doesn't mean artificial (or naturally recorded) reverb cannot be ported into a virtual environment, however the music artist/engineer should be aware of this potential conflict and work with nuance. It should be noted that these same principles apply to delay or echo, which are often administered artificially in post-production. In fact, reverb is just a form of delay with a heightened number of reflections at shorter time intervals.

7.1.3.3 Filters

Equalization (EQ) allows the mixing engineer to subtract or boost frequency ranges across the frequency spectrum. Within the field of EQ exists the high-pass and low-pass filter which provide “fixed attenuation slopes at various frequencies” [17]. These steep slopes allow a broad section of either high-end or low-end frequency ranges to be cut while maintaining the integrity of the mid-range. When working on a fixed-media composition (i.e. a song or film score), EQ and filtering are commonly used to help sculpt each element’s position in the overall song mix as well as cut any problematic or unwanted sound artifacts (for instance, a high-pass filter may be applied to a track with interfering low rumbling). Automation, in which the parameters of equalization and filtering adjust throughout the song, may also be useful in a fixed-media piece. For instance, when scoring for film that must prioritize dialogue and sound-design, an engineer may automate a low-pass filter or apply a frequency dip in music to carve space for voices. The issue in this arises when transitioning to a medium such as VR that must react to user engagement.

In a scene from “Becoming”, the experiencer is immersed waist-deep in water. If the user bends down low enough to submerge their head in the water, a low pass filter sweeps over the music and sound-design elements to simulate the real-life loss of high-frequency hearing in the same circumstance. Attempting to apply this filter preemptively would not work, as it depends on user interaction; the coordinates of the experiencer’s head must reach below that of the water’s surface, and only then can the low-pass filter apply [57].

A similar filter-based concept occurs in acoustic environments containing physical objects, occlusion. As discussed above, reverberation of a sound source is a series of rapid surface reflections. Some of these reflections diffract off smaller surfaces in the space, bending around the object. Other reflections transmit through the object. A large enough object may occlude, or eliminate, the direct sound signal, while diffractions may still reflect and reach the listener.

Typically, “the high-frequency sound mostly travels in a small region around the direct path, and thus is blocked more fully by a small amount of occlusion, whereas the low-frequency sound occupies a larger volume around the direct path and therefore requires more occlusion to be affected” [34]. A mixing engineer may simulate this effect by automating a low-pass filter, but again this would be difficult in a virtual space where the experiencer’s location and engagement guides diffraction and occlusion. For example, in Sonic Arts’ acoustical simulation system Space3D, the VR user experiences “high quality spatialization results over speaker arrays or headphones [that] simulate high-order reflections, diffraction and transmission” among other properties [34]. In this experience, the user listens to a piano played in various rooms, halls, and cathedrals. Each environment offers objects and surfaces for reflections to diffract or transmit. If the user stands behind a pillar or picks up a large item, the direct sound is noticeably and accurately occluded from the listener’s ears in real-time. This level of real-time spatial filtering, achieved through advanced ray-tracing techniques, would simply not be possibly with a standard EQ or filter device.

7.1.3.4 Negotiating Mixing Principles

In “Pigments of Imagination", all musical components were first mixed in the DAW Ableton Live. They were then bounced as individual stems for spatialization and visual mapping in Unreal Engine 5 (VR development engine). Post-processing effects such as compression and reverb were used sparingly. Reverb parameters specifically were applied to each stem in nuanced amounts depending on that stem’s role as either a static stereo-field or binauralized spatial musical element (key concepts of binaural audio will be discussed shortly). Vocal delays that would exist in a standard stereo mix of the song (perhaps for vinyl or digital streaming) were cut and reapplied in Unreal Engine where they could more accurately react in real-time to the listener’s head rotation. Though “Pigments” was not so dependent on real-time virtual acoustics,

future work by the author looks to build on these affordances. This future work aims to deepen the negotiation between post-processed stereo-field versus spatialized reverb and delay.

7.1.4 Reception

The *Communication Diagram* is a feedback loop, meaning the affordances of the relationship between music and VR mediums must be explored both for the artist and the experiencer. So far, we've discussed these relationships for the artist, but it's also necessary to understand these relationships to the experiencer as well.

7.1.4.1 Interactive Listening

In his 1997 article "Digital Revolution is a Revolution of Random Access", artist and author Grahame Weinbren explores film's transformation of platform from celluloid in theater to VHS and early digital [21]. In the emergence of these platforms, he interprets our ability to skip through films as we'd like as our ability to manipulate the order and thereby intentions of the original artwork, termed 'random access'. This turns the intended work into an interactive piece in which the filmmaker becomes a 'designer of patterns' for the viewer to move through, allowing for their own interpretations of the work's meaning to possibly supersede the filmmaker's.

Weinbren discusses his own interactive cinema installations, designed with intention to explore this nexus point between artist and experiencer/explorer with the goal of asking and answering questions that arrive from this experiment. Specifically, he touches on random access's ability to break down film structure into nonlinear dream-states where all aspects of the space (film or dream) exist at once and it's up to the experiencer (a 'burden' as he terms) to piece the meaning and structure together. However, he states, there are ways as a filmmaker to allow this interactivity while still providing definitive closure no matter the path the experiencer chooses,

citing the idea that regardless of nonlinearity within random access, there is a time-based linearity we as viewers (and humans) will always live by.

The concept of random access interactivity might be analogous to how the listener experiences music in the modern streaming era. Whereas streaming is akin to digital media and vinyl similar to cinema in its physicality, streaming lacks vinyl's 1960's-era rituality of opening a gatefold, doing a drug, dropping a record needle, laying on the couch eyes closed and playing "Dark Side of the Moon" with the intent to experience that album in a sequential order. In this way, the streaming era of music soundtracks the listener's experience rather than *acting as* their experience; songs can be accessed and playlisted to wrap around the listener's experience regardless of the artist's intention. Neither approach is better or worse, they are just two diverse ways to experience a piece of art: a sacrifice in artist-curated experience in exchange for listener freedom of choice. This is like the dialog of immersion and interactivity discussed when tackling narrative in virtual reality. Thus, with the VR headset, here viewed as a music experiencing platform alongside vinyl, CD, mp3 and streaming service, there may exist a space for the artist's intended sequence of musical events (emotion-by-way-of-narrative) to supersede the modern era of user-curated random access. And perhaps, whether with music or cinema and regardless of era, the progression of technology can alter one's relationship to a prior artform (i.e. celluloid versus digital, VR headset versus streaming versus vinyl), but it almost always leaves one the space as an experiencer in how they choose to consume that art. In that regard, their choice to interpassively experience a musical work through VR, vinyl, or streaming on a smartphone is its own form of interactivity. This thesis is not interested in convincing the reader that VR music videos are the new sole platform for most of the recorded music engagement, rather another platform that offers a different set of conditions, and therefore results, than existing ones.

7.2 Virtual Reality Medium

The components of music creation we have covered contain fundamental conventions that can challenge or be challenged by certain design principles of virtual reality. So far, we've seen how these conventions can be redefined when extended to the virtual reality medium. Now, we will see how existing VR design conventions can be reimagined in this collaboration. Virtual reality is the final medium the music artist extends their emotional center and narrative to, however it is the first medium that interfaces with the experiencer. It is also the medium that allows the experiencer to dialogue with every component of the *Communication Diagram*: emotion, musical & VR narrative, and the medium of music. The resultant discussions on VR design principles ideally bring this diagram to life.

7.2.1 Design Principles of Music-Centered Virtual Reality

The purpose of researching design principles is to better understand how the artist can achieve implementing the emotion and narrative of their musical work within a virtual reality experience. There are numerous books and papers on general virtual reality design principles, and though we've included some above to provide context, we will primarily focus on the design principle structure of Jack Atherton and Ge Wang's 2018 *"Doing vs. Being: A philosophy of design for artful VR"* because A) it's guiding philosophy (Doing vs Being, explained shortly) aligns with the needs of the music artist within the scope of this thesis, and B) its principles are specifically aimed toward music-centered VR projects and can thus be explored and modified as it relates to our specific goal of extending and intertwining a narrative project from music medium into the virtual space. These are the principles we will aim to negotiate with those discussed above of music composition, production and engineering designed for physical space. It is also here where we will explore a wider array of case studies to A) reinforce our design principle alterations, B) explore what specifically we can utilize from these case study approaches and attributes, and

C) aggregate this all into a defined artistic practice and successful reimagination of the VR music video field.

Atherton and Wang's paper explores the concept of designing 'artfully' for virtual reality, specifically focusing on music-centered virtual reality experiences. Their goal is to offer lenses through which to craft "expressive musical experiences in VR" and a set of design principles framed by each lens. However, before we explore these lenses and principles, it is worth examining the methods and underlying philosophies that helped form them. We do this not just to clarify the authors' reasoning, but to highlight the philosophical differences that arise when examining virtual-reality-driven music-centered works (as the authors do) vs. music-driven works expanded into the virtual reality medium (as this thesis does). This subtle difference, in which the potential for emotion and narrative are born either from virtual reality or from music, is foundational to altering these proposed principles for the music artist.

To contextualize their primary 'artful' design approach, Atherton and Wang explore a diverse array of perspectives on the medium of VR and use them to introduce this new set of lenses for designers, specifically the balance between being and doing (immersion and interactivity). They then introduce the role of Virtual Music Instruments (VRMIs) as well as existing music-centered VR projects as case studies examined through these lenses. Key case studies and VRMIs will be explored in this thesis for context and comparison alongside current case studies, particularly "Pigments of Imagination". Similarly, this document will only deconstruct pertinent lenses, sharing the principle itself, defining what it means, and then explaining A) why as written it relevant for the music artist or B) why this principle could be challenged and altered through the *Communication Diagram* framework. Before diving into the thesis's lenses, it's worth remembering a few concepts within virtual reality: *immersion, presence, embodiment and*

interactivity. It might also be good to add Atherton and Wang's definition of virtual reality to our existing list:

virtual – to be a virtual reality, the reality must be simulated (e.g. computer-generated)

This is a relatively straight forward definition, but when compared against some of the more existential definitions we explored (i.e. Jaron Lanier), I argue the potential of what constitutes a 'true music-centered virtual reality experience' can grow, and therefore so can approaches to design.

7.2.1.1 Key Perspectives

Atherton & Wang structure their ideas of artful design around a diverse set of perspectives spanning musicology, philosophy, communication, art and design-theory. A few important highlights below:

Philosophy

Wang and Atherton first reference a piece by David Chalmers who offers a founding philosophy that *virtual realism* is in a sense real, particularly in perception: that the user might feel very real emotions and visceral reactions based on the virtual world [10]. As related to the music within the virtual world, the idea is that we can then translate this into what feels real about music creation, specifically, as related to my interest: emotional connection, expression, storytelling, and reflections of our humanity.

Literature

As we discussed in VR narrative functionality, Atherton & Wang cite Marie-Laure Ryan's 1999 piece *'Immersion vs. Interactivity: Virtual Reality and Literary Theory'* [38], where she

compares the medium of virtual reality with literature and discusses the trade-off between immersion and interactivity toward a cohesive narrative. The important takeaway here is that VR is still an emergent medium, specifically in how it deals with narrative. The introduction of fully immersive three-dimensional worlds and interactivity in which the user can affect the narrative structure of the world is still relatively new, such that there are no structures in place the way there might be in film or book. When developing a music-centered VR experience, in which the music holds some level of narrative functionality (i.e. informing the narrative of VR), these affordances (immersion and interactivity) are necessary to consider.

Design

In exploring perspectives on design, Atherton & Wang references VRMIs (Virtual Reality Music Instruments) which, through development, must consider concepts like haptic feedback, audiovisual representation and user gestures. Some developers, such as Rob Hamilton, even make the case that the virtual world itself, with virtual acoustic modeling in place, can be thought of as immersive virtual instruments [24].

Well-designed VRMIs show an example of immersivity and interactivity working in congruence. Though no overarching narrative is really at play in these virtual instruments and virtual worlds, we can think about this as a philosophy of design that can be engaged with and altered in a narrative music-centered virtual world. As Atherton and Wang say:

> "Clearly, the field of VRMIs has much to offer in terms of how to design well for multimodal VR with a focus on audio: some aspects of creating artful virtual realities that are focused on sound and music can benefit from the lens of instruments. Still, our notions of artful VR can and should expand beyond this lens." [24]

Artful Design

Artful design unites the above philosophies and invites developers to not just design toward specific user needs at a surface level, but from an underlying set of human values. In doing

so, artful design merges practical design needs with emotional, social and moral aesthetics. My focus is rooted in artistic practice and hinges on the way the artist's intended emotional truth communicates with narrative and medium (in our case music and virtual reality). For this reason, artful design as a blend of designing toward practical needs and expectations as well as emotional aesthetics resonates with me.

7.2.1.2 Case Studies

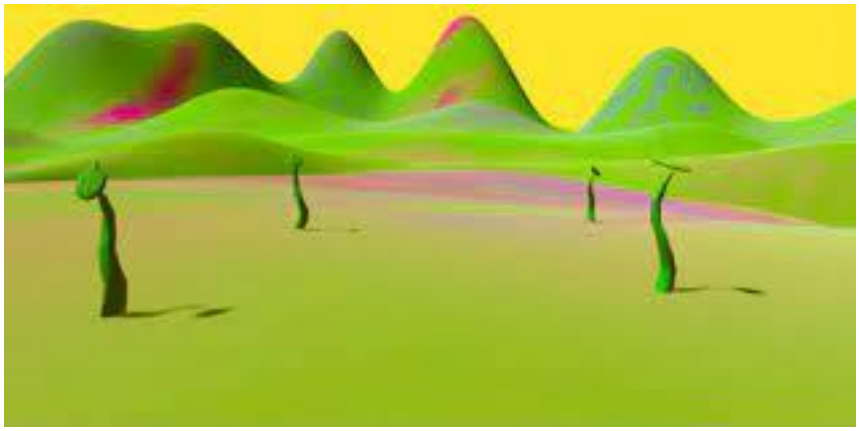


Figure 7-7: "12 Sentiments for VR", a music-centered VR piece that balances narrative, immersion and interactivity [24]

Atherton & Wang present a series of case studies that cover a range of music-centered virtual reality projects, from VR music videos to visual music programming worlds and VRMIs. Of these, we will reference some but examine one with depth, *"12 Sentiments for VR"*, which A) is the author's work and therefore explored at a deeper level and B) conceptually and practically balances immersion, interactivity and narrative functionality.

> 12 Sentiments for VR is an extended exploration of the emotional life cycle of a plant. Each of its 12 movements uses musical interactions to explore a different phase of a plant's life through a unique aesthetic (outlined in Figure 2). This scaffold gave us the opportunity to explore our open questions on how to design artfully for VR through 12 different, yet interconnected, experiences that are created with aesthetic-driven design. - Jack Atherton [24]

This work is successful because it is a *generative virtual composition/sonic environment* that allows the user to embody an immersive world and directly affect the musical attributes within it, for instance chord progressions, BPM, arpeggiations and timbral qualities, while maintaining an overarching narrative and emotional resonance.

7.2.1.3 Lenses and Design Choices

> The design principles presented below are…contextual ways of thinking about design choices. They are design critiques and are critical reflections of the designs as they stand. That is to say, design critiques are not a way to explore 'what was the designer thinking when they built this?', but instead 'how is the design experienced, and what values does it reflect. [24]

Atherton and Wang present these lenses and principles as 'contextual ways to think about design choices' guided by the way the design is experienced. This flexibility strengthens their overall framework by allowing VR works existing under different conditions (such as projects we have discussed throughout the thesis) to maneuver and alter accordingly.

Lens 1: Audio First

Within the scope of Atherton and Wang's paper, and the definition of 'virtual reality' by which they subscribe, this first lens of centering audio is founded on three principles:

- *audio should be dynamically generated.*
- *audio should be immersive.*
- *audio should be interactive.*

Principle 1: Audio Should Be Dynamically Generated.

Generative audio can be defined as a self-creating and reactive sonic or musical work. This form of music is typically born from a series of parameters, or conditions, set by the artist that

allows the music or sonic work to dynamically evolve. One condition that might be especially pertinent to virtual reality is the allowance of user engagement and control, such that the compositional and/or sonic properties of the piece react to the user's gestures, or the unfolding of the virtual world as dictated by the user's presence within it.

Specific to this principle, Wang and Atherton highlight two key nuances:

> **Alignment to medium.** "Dynamically generated audio can benefit from some of the unique affordances of the medium. In VR, audio can be synthesized or generated in real time. For example, a VR experience using physics simulations to move musical objects could use the physics data to inform the synthesis of audio for those objects and avoid the distracting repetition of playing back static recorded audio files." [24]
>
> **Responsiveness.** "Dynamically generated audio responds to the unfolding of events in the world. Under this notion, a virtual reality would not have a single audio file that plays back linearly and deterministically. Immersive animations that run alongside the playback of a static audio file are instead more like music videos or demoscenes than virtual realities." [24]

It is here where we can highlight and debate key statements as points of divergence: A) what constitutes a 'true virtual reality' experience, B) what constitutes a 'single audio file' (fixed media) and C) how we might explore this oversight as a liminal space to sculpt a VR medium between what they define as 'music videos' and 'true virtual reality experiences'.

This design principle argues that audio in virtual reality should be dynamically generative; that worlds scored by deterministic fixed media do not constitute a 'true virtual reality' experience [24]. Fixed-media music-centered experiences are instead defined as "music videos", implying a medium in which a two-dimensional design (the standard popular music video) is shoehorned into a three-dimensional space with little reverence for the unique affordances of that space. However, as we've seen, fundamental definitions of virtual reality, both practical and existential, do not specify the determinacy of an experience's musical structure.

More so, I argue that the field of 'virtual reality popular music videos' should be reconsidered as an emergent virtual reality medium containing fertile creative grounds for musicians interested in developing their work from two-dimensional to three-dimensional spaces. *Specifically, I argue that there exists between the space of 'true virtual reality' in which all sound and interactivity is dynamic and affected by the user, and the virtual reality popular music video, in which a determinate fixed-composition is processed into a three-dimensional space in after-thought of the song's creation. I argue there is a liminal space in which the potentials of creating toward virtual reality allows the emotional and structural integrity of the original fixed-composition to maintain in context to the user's interaction with it, while simultaneously reiterating itself in imaginative ways both compositionally and spatially, but at the creator's (not user's) discretion.*

I argue that a dynamically generated sound-environment, created solely through DSP sound-synthesis, can actually be counterintuitive to extending the human experience into virtual space simply in that it eliminates the presence of humanity in the creation of the music, an emotion immersion that's hard to quantify but easy to acknowledge. Though generative music composed through user interactivity in virtual reality does innovate the relationship between the music and the listener, it comes at a price: to be generated in real-time within a virtual space, the music is comprised of discrete digital samples that ultimately synthesize to tones, chords, patterns, and sequences in a solely electronic medium. Meaning at heart, only electronic music can exist in a 'true virtual reality experience'. We are missing out entirely on how other modes of human music creations across genres and generations might exist and be altered through user interaction. And the only way these modes can be represented is through that audio file. According to Atherton and Wang, the moment a deterministic audio file is introduced, the experience is no longer 'true virtual reality'.

The *Communication Diagram* was created to refute this idea. Through our scope, artful design is an extension of artistic practice. In aiming at an audience of multidisciplinary music artists, we've presented a framework in which the artist's emotional center is wrapped in narrative(s), extended first to the music medium and then virtual reality. In our model, music holds the core emotional weight and narrative. A fully dynamic generative musical work altered by the user has the grave potential of destroying the original emotional intent and narrative as dictated by the artist. We've discussed and will continue to discuss numerous dynamic and non-linear approaches to music-centered VR world development, however the stipulation that all music, at risk of losing emotional intent, narrative and human connection, must be dynamically (and therefore synthetically) created seems counterintuitive to flexible artistic practice, specifically in a medium whose narrative architectures are still so ripe for exploration and development.

The virtual design principles discussed in Atherton and Wang's paper are foundational to music-centered VR experiences thus far. The VR music video is different because it is reliant on music narrative, composition and production transferring from a living, physical world to virtual space. In the VR music video, we see the potential for an instant inclusivity of the full spectrum of music, not just synthesized electronic works, but genres rooted in live musicians, vocalists, samples, field recordings, and more. How might these genres fundamentally shift by composing toward a three-dimensional space? How might this three-dimensional space shift in return? How might emotional and humanistic immersion, musical narrative functionality and interactivity enrich when the potential of a deterministic audio file (the only current container of the entirety of live musicianship) is seen as a collaborative asset to virtual reality? I argue that the VR music video is the perfect space to explore these questions.

> **Principle 1.2: Audio should be immersive.** Audio in VR has a unique opportunity to create a rich tapestry that convinces the user they are in another reality. Designers can make many aspects of the reality sonic or even musical, surrounding the user with a landscape of many sounds. It can be fruitful to think beyond sound objects and consider fashioning sonic environments, as suggested by Hamilton (2009) [24]

This principle reinforces that music as a priority in a virtual space can be immersive and one key way to express that immersivity is via sound spatialization:

> **Spatialization.** Sound spatialization is an essential and surprisingly oft-overlooked affordance of VR. Surrounding the user with an environment full of objects that make localized sounds will help give the sense of being in a living world. [24]

One of the greatest affordances of virtual reality is audio spatialization. Spatial audio applications exist across various real-world industries, however virtual reality is especially well-suited to both take advantage of spatial audio's best features and highlight those features to the user. There are various approaches to sound-spatialization: ambisonics, vector-based amplitude panning, wavefield synthesis and more. Each method approaches spatialization differently and most are fantastic for spatialized experiences in physical settings. The most common form of spatial audio in VR is termed binaural and it's worth discussing some key related concepts and terms, both for general understanding and to better compare and contrast against stereo-field mixing practices in physical spaces.

Binaural Spatial Audio Overview

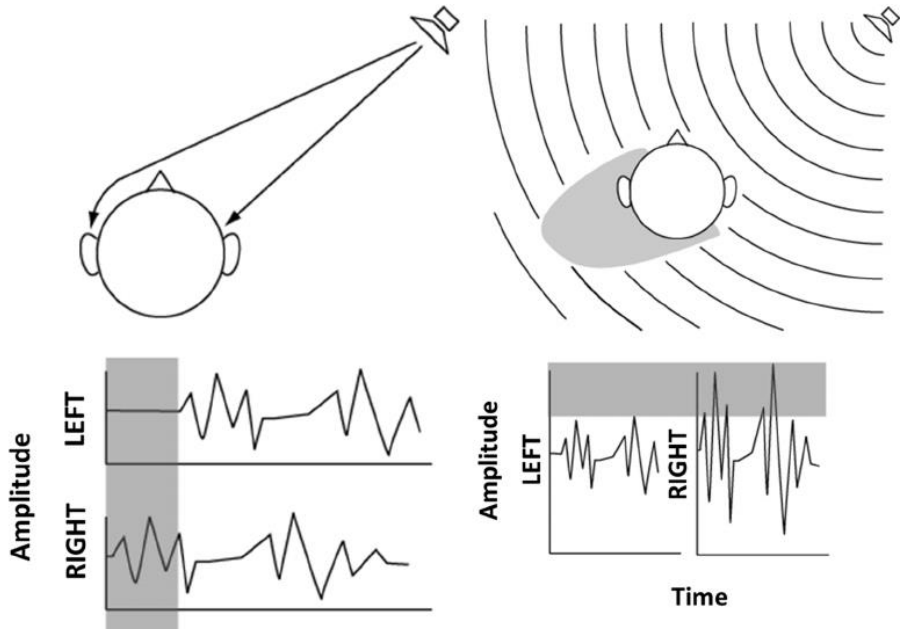


Figure 7-8: Visual representation of Interaural Time Difference and Interaural Level Difference, in which a listener determines the directionality and distance of a sound source [32]

In addition to loudness, pitch and timbre, humans can perceptually process a sound's spatial attributes, particularly the sound's direction and distance [7]. To determine these attributes, a person relies on two factors: the interaural time difference (ITD) and interaural level difference (ILD). A music artist/engineer may think of these in the mixing world as delay (ITD) and amplitude (ILD). ITD refers to the time difference between when a waveform reaches the left and right ears accordingly while ILD refers to the amplitude of the sound source across the frequency spectrum as received by both the left and right ears. Both help to localize the direction of the sound source by accounting not just for the source, but the body itself [7]. For instance, head rotation may result in a difference of time between when the sound reaches each ear (ITD). Similarly, head rotation may affect the amplitude of a sound source specifically in the higher frequency range. However, each person's body is unique, and a sound source does not just travel directly into our ears, it diffracts and reflects off the torso, head and pinnae (outer ear). These act

as cue points on the soundwave's journey to our inner ear that inform a more accurate localization of the sound source (for instance the ITD and ILD). From this knowledge, a virtual model of sound spatialization can occur using head-related transfer functions (HRTFs) to help determine the filtering effects of a listener's unique body-type (torso, head, pinnae) [7]. Binaural audio then uses these HRTFs models to re-create a more accurate spatial experience in virtual reality.

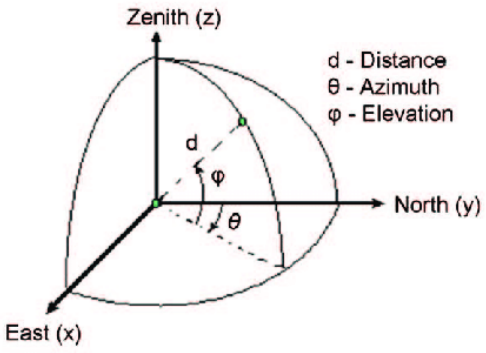


Figure 7-9: A view of how distance, azimuth and elevation angles exist across the X, Y and Z axes [2]

For a music artist accustomed to mixing in and toward a stereo-field space, it is worth reviewing a few key spatial audio terms to help the artist more intentionally think about how to compose and mix toward a spatial virtual world. Terms and concepts expanding into reverberation are not covered within the scope of this thesis but should be noted as well.

- *Azimuth angle:* the position of a sound-source across a horizontal plane, helping to determine if a sound is located in to the left or right, ahead or behind the listener.
- *Elevation angle:* the position of a sound-source across a vertical plane, helping to determine if a sound is coming from above or below.
- *Distance*: the measured distance between the sound-source and listener.
- *Spread:* a blurring of sound locality, as related to distance

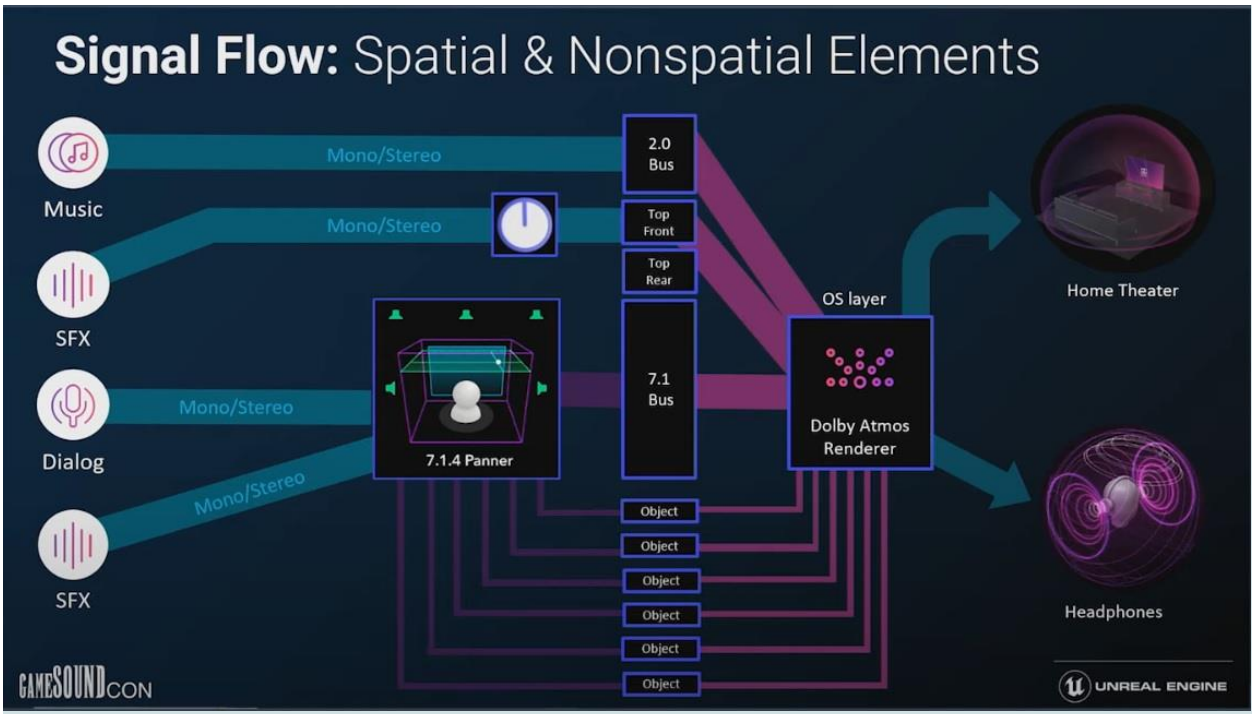


Figure 7-10: Dolby three-tier system of sound mixing. Numerous plugins and approaches were trialed, but the Dolby's harsh responses to shifting azimuth and elevation angles were softened when integrated with static stereo-field elements. [14]

In "Pigments of Imagination", spatial audio is delivered by headphones using the Dolby Atmos plugin for Unreal Engine 5.2 [14]. This plugin is unique in that it allows each element of a song to either act as an independent object-based sound source (sounds mapped to coordinates through meta-data within the virtual world), part of a 7.1 multichannel soundbed (sounds mixed more traditionally through audio channels and spatialized accordingly) or pass through and remain a static stereo-field sound. Object-based sounds and sounds within the 7.1 multichannel soundbed are effectively binauralized as individual HRTFs are applied to each sound-source. This makes the plugin extremely versatile and for the purposes of the music artist looking to integrate the physical and virtual mixing space, an excellent tool to do so.

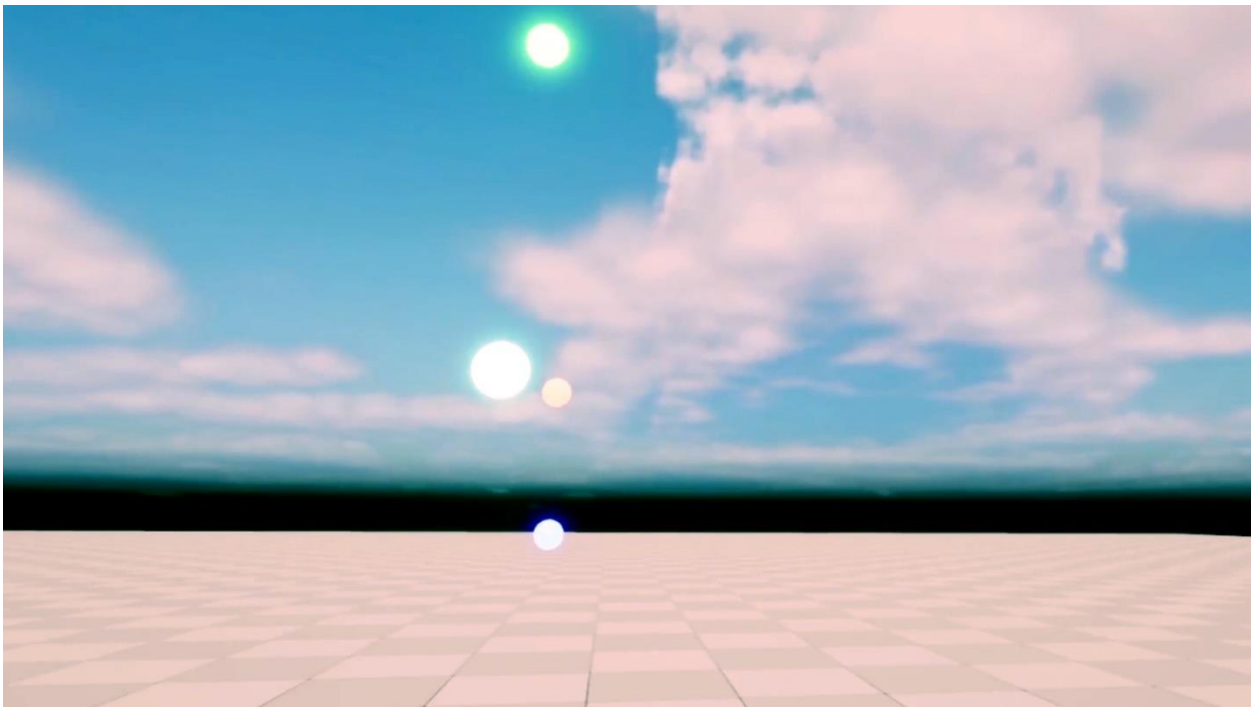


Figure 7-11: A study of sound elements mapped to emissive spheres in a 7.1 soundbed around the user's head at various distances, azimuth and elevation angles.

Key musical elements of "Pigments" relative to narrative and form, such as lead voice and bass, remain static to the listener regardless of the direction their head turns. This means these musical elements maintain a priority in the mix, which is conceptually in line with audio mixing principles in the physical stereo-field space. Other musical and sound-design elements, such as keyboards, synths, strings and backing vocals, are spatialized into the 7.1 soundbed through virtual speakers that maintain fixed distance, azimuth and elevation coordinates relative to listening location (see figure above). This allows the experiencer, no matter their global position, to move their head to listen to the composition from different angles and gain a unique listening experience each time. Additionally, there are dynamic sound sources that coexist with pre-composed music, such as birds and user-generated musical notes; these sounds are attached to their corresponding visual components and follow them around the environment, regardless of listener location or orientation. These sound-sources are not relative to listening location and therefore give the entire sonic world a sense of shape and depth. HRTFs are applied both to the

soundbed and dynamic object-based sound sources to ensure the listener's head rotations and movements produce an accurate spatial response. The key takeaway in integrating physical-world stereo-field mixing and binauralized multichannel/object-based spatial audio is that the musical composition is open to find new points of mix convergence: new ways for the music artist to think about 'gluing' the elements of their musical work to find a balance of static and dynamic that provides the experiencer immersion and pushes the narrative forward. For "Pigments of Imagination", the low frequencies of the bassline and the lead vocal suggestiveness both grounded and guided the narrative. The binaural soundbed of keys, pads, strings and backing vocals positioned the listener at the center of the musical work. Additional binauralization of a lead synth and flugelhorn guided the experiencer to visual cues. Lastly, the object-oriented sounds of birds and user-initiated paintbrush tones reinforced the experiencer's narrative-implied gaze and engagement.

With a stronger understanding of the spatialization process and affordances of virtual reality, the analysis of prior VR music videos offered by hip-hop artists Run The Jewels and Travis Scott, guided by single non-spatialized audio files, should make that much more sense. Atherton and Wang review similar audiovisual-centric VR music videos "Fantasynth: Chez Nous" and "Surge", stating that "unfortunately, [they] both spatialize audio in the stereo field without matching the location to corresponding animations in the scene; not only is this potentially disorienting, it misses a low-hanging fruit that could have facilitated stronger Immersion" [24].

Beyond spatialization, it should also be noted that audio immersivity can expand beyond how it might sonically envelop the user, but in how it can command narrative. As discussed earlier, challenges in virtual reality arise when immersion and interactivity attempt to coexist alongside a meaningful narrative.

> **Principle 1.3: Audio should be interactive.** In a virtual reality with audio as a first-class modality, users can meaningfully affect audio; their actions have a significant musical or sonic consequence. Such a reality can create opportunities for users to exercise creative agency through musical expression. [24]

As Atherton and Wang expressed above in philosophies of design, well-designed VRMIs offer a balance of immersivity and interactivity of audio through user gestures. The following VRMI case studies offer a brief but diverse palette of how existing developers are approaching multimodal VRMI application, specifically in the audiovisual and audio-haptic space. These VRMIs exist in a vacuum that does not concern narrative, however the goal for the music artist is to extract approaches that may relate to their own narrative-based music-centered VR works.

Figure 7-12: Images of the "Synesthsizer" (left) in which RGB values of pixels fro images are converted to melodies, and the "Synesthesia Laboratory" (right), in which audio signals are transformed into visual shapes: in this instance, representation of the word.

In "Synesthsizer: Physical Modelling and Machine Learning for a Color-Based Synthesizer in Virtual Reality", developer Giovanni Santini explores the subjective subset of the synesthetic condition, chromesthesia, a sensory overlap condition in which upon the trigger of sensory stimuli, one "links sight to hearing, producing color in response to sound" [20]. With the Synesthesizer, the author uses this concept to create a VRMI that converts RGB values of pixels from virtual images into melodies…based on a system of rules…linking color to pitch" [20]. In "Towards a Synesthesia Laboratory: Real-time Localization and Visualization of a Sound Source

for Virtual Reality Applications", author Ahmet Kose applies a series of "recently developed methods for detecting sound source in a three-dimensional space around the listener [and then] transforms the audio signal into a series of visual shapes, such that the size of each shape is determined by the loudness of the sound source, and its color is determined by the dominant spectral component of the sound" [1]. Amplitude and frequency index data are analyzed in real-time to create visual representations of the evolving sound source via size and color. The goal of this project is to present a synesthetic experience within VR to a participant without the condition. Though the synesthetic experience is highly subjective, the purpose of a project (and accompanying paper) like this is the deep explanation of approaches to reason and execution. For instance, the author's instruction of approach to real-time signal extraction by way of the Mel-Frequency Cepstral Coefficient (MFCC) method offers artists the tools (and equations) to map both determinant and generative elements of music to visual properties in a foundationally meaningful way. Though Kose's mapping of sounds to shapes may seem artistically crude at a superficial level, it is best to view this as a model for artists to develop their virtual musical worlds in imaginative ways.

For instance, a notable example of effective interactivity within an immersive narrative-based musical VR piece is "12 Sentiments for VR", in which the virtual world itself and elements within react musically in real-time to user gestures, just as shapes from the above "Synesthesia Laboratory" were created from real-time sound-generation. In "12 Sentiments" second act, "users collect sunlight with their hands, represented as leaves, and a chord swells up as the leaf absorbs sunlight". The longer the user collects sunlight, the more the chord and timbre grow in complexity. In the fourth act, the user embodies the wind with clouds as hands. When their cloud-hands move, each triggers a note. Other hand gestures increase pitch and arpeggiation. Essentially "the user can meaningfully affect the virtual world's music without the experience being framed as a musical instrument" [24]. Each of these acts offers the user a different engagement and control

over virtual musical elements, but unlike VRMIs, all exist within an overarching narrative; there is distinct purpose in how the user interacts and informs each act.

One specific aspect of Principle 1.3 is audio-driven alignment:

> **Audio-driven architecture.** Use audio-driven architecture when sensible (e.g. informing graphics from audio timing). Audio signals and strong audio-based timing can be used to drive other modalities like visuals and physics simulation when it makes sense to do so to preserve musicality. [24]

In our examination of musical narrative functionality, author Johnny Wingstedt discussed the idea that multimodality is a principle of musical composition (i.e. combining instruments to create varying timbral effects or the building of complex polyrhythmic percussive layers) so to extend this idea out to other forms of media is quite natural. VR is a great connector to this idea because it offers the unique ability to cohesively translate an idea across modalities. When that idea is prioritized within the music, the chance to secondarily affect visual, physics-based, or haptic behaviors mapped directly from musical and audio properties lends to an organically immersive experience that acts as a natural extension of musical narrative functionality and multimodality. It is one of the most exciting aspects of developing in virtual reality and it can be explored through both generative and fixed media audio.

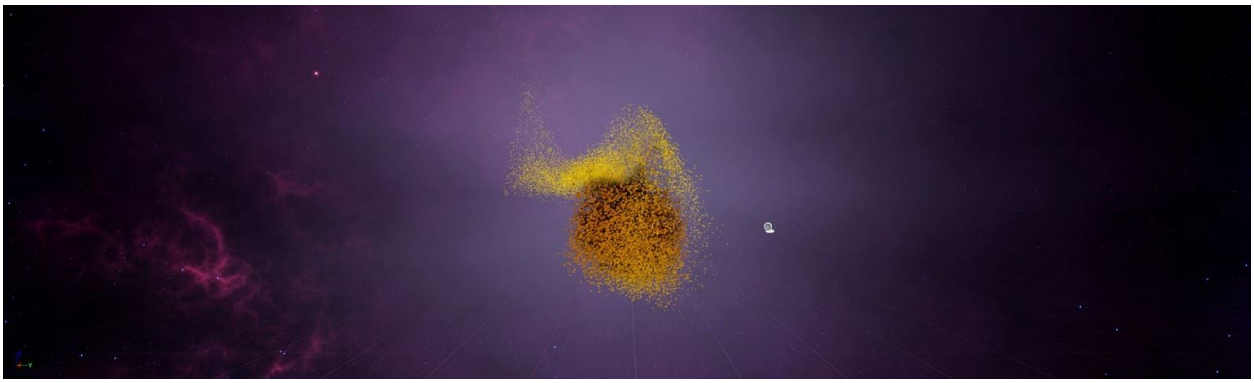


Figure 7-13: A particle system with parameters mapped to the amplitude, frequency index and frequency values of a flugelhorn recording, as well as a lead and backing synth solo were analyzed to gather real-time amplitude, frequency index and frequency value data.

For instance, in "Pigments of Imagination", a series of parameters for each instrument and sound element were mapped to material, textural and behavior properties of visual assets and systems in the VR world, allowing for both manual automation and real-time reactivity [62]. For example, using Unreal Engine's *Audio Synesthesia* non-real-time analysis assets [6], a solo flugelhorn was analyzed to discretely gather non-real-time amplitude, frequency index and frequency value data that was then mapped to a series of physics-based parameters in an emergent Niagara (particle) system. This audio-reactive particle system was then spatially mapped to where the flugelhorn existed within the binauralized 7.1 soundbed, adding a key visual cue to existing spatialization. This visual cue and spatial location allow the flugelhorn to function as a guide through the overall narrative of the VR work, reinforcing the idea that, while audiovisual reactivity can be novel on its own, meaningful design is informed by narrative. It should also be noted that in the VR music video space, there is room for fixed media and generative audio to co-exist. Whereas the audio files in "Pigments" are pre-analyzed, future practices aim to graduate from the Audio Synesthesia plugin and allow for real-time amplitude and frequency analysis of procedural (generative) sounds as well using similar Fourier Transform technique as discussed above in the Synesthesia Laboratory's application of the MFCC Method.

Lens 2: Design to the Medium

What made me initially connect to virtual reality as an experiencer was its ability to recreate imagination in a meaningful and immersive way. Whereas 360 video and other hyper-realistic approaches to virtual reality have their own uses, the magic in VR is in its potential to explore beyond what is possible in the physical world. In "Pigments of Imagination", this idea was explored by beginning the user's journey as a "passive experience on a two-dimensional screen within the virtual space which then smoothly unfolds into three dimensions as a means to reinforce the transition into a continuously evolving, surreal environment" [62]. Atherton and Wang reinforce this sentiment by recommending that artists design to the unique affordances of the medium. The first principle in this lens states:

> **Don't Port:** avoid copying something from an old medium to a new medium without considering how the two are different and what the new medium offers

One reason virtual reality has yet to establish a recognized narrative structure is because similar forms of established media, like film and video games, exist in different formats, often two-dimensional. The idea, it could be argued, is to determine the key emotional components and narrative functions from other media and reestablish them given the rules, behaviors, and affordances of virtual reality. Some key aspects to reimagine when approaching the surreal nature of VR include:

Virtual Physics: Virtual physics offers the opportunity to abstract fundamental rules of the physical world. It is completely plausible to build a real-world-physics-based system in VR that mirrors a real-world experience. The fun comes in abstracting these existing systems and exploring the results, something only possible in a virtual world. For instance, in "Pigments of Imagination", all interactive objects in the child's room maintain a certain degree of real-world physics (gravity, collision and haptic sensations), however the paintbrush, key to enriching

immersion and narrative, maintains collision and haptic feedback but no gravity. This lends to an intended surrealism of the work but also focuses the experiencer on the item most worth their attention.

> **Scale.** VR, unlike physical reality, has few restrictions on scale. This opens up the possibility of drastically differing scales, as well as objects that change size in real time. [24]
>
> **Time.** With the use of virtual physics engines also comes the ability to manipulate virtual time. [24]

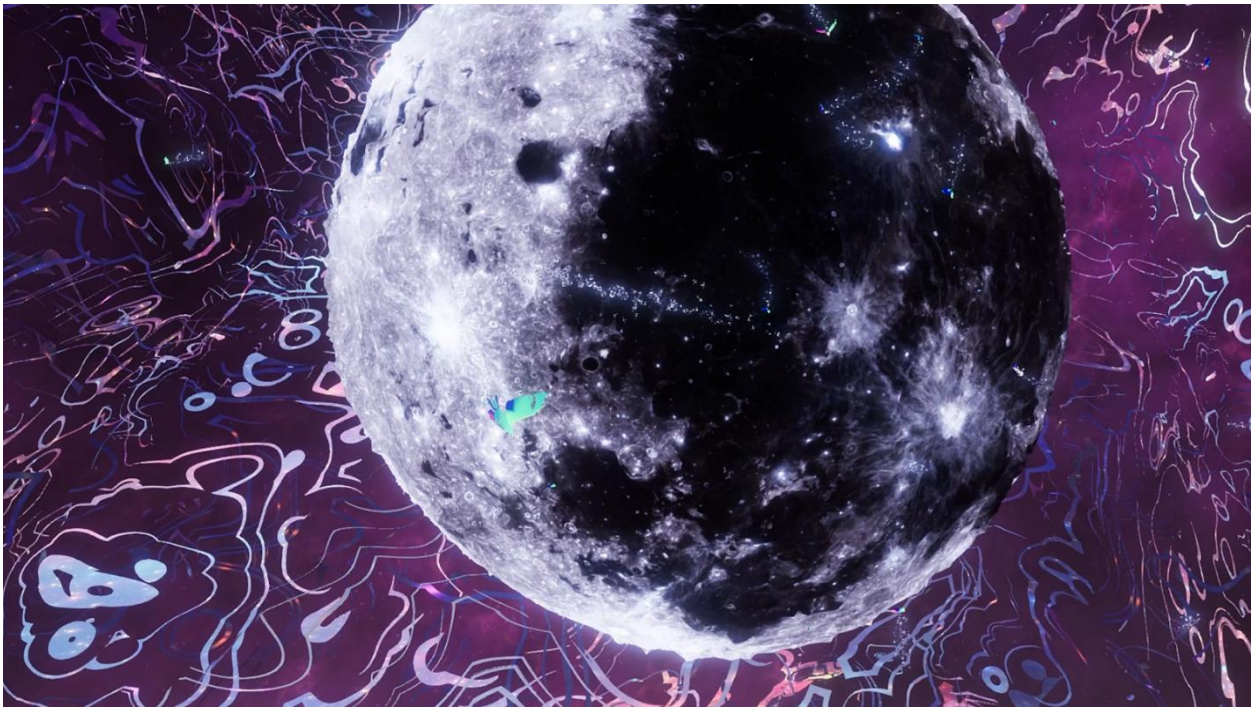


Figure 7-14: An image of the approaching moon in “Pigments of Imagination”

In “Pigments of Imagination”, we explored a collaboration between scale and time. To achieve the experience of the user traveling across the sky to an augmented moon at a fixed time rate guided by a portion of song length, the position of the user, and size and coordinates of the moon had to cohesively and continuously shift within this fixed timeframe to give the illusion of a slow ascension toward an object that at start appeared small but upon completion of ascension

appeared massive. Conversely, time was also altered within the fixed composition, which still held enough elasticity to expand or collapse measures in certain movements in accordance with virtual reality needs, specifically during the passive portions.

"12 Sentiments for VR" exemplifies time alteration between two similar acts to alter the emotional response of the experiencers. As discussed above, "Movement 4 has users embody the wind to blow small seedlings skyward; in movement 8, users try to do the same thing, but the seedlings are no longer buoyant enough to fly, and they sink back to the ground. If the seedlings were to fall at the same speed in movement 8 as they did in movement 4, the emotional effect would be comical, rather than melancholy as the accompanying score aimed. Therefore, the author slowed the time of the seedlings' drop, such that they remained suspended in air momentarily. This slowing more accurately matched the intended emotion of the movement [24].

Lens 3: Designing for the Body

> **Tactility Matters -** The physical response of playing an instrument is an essential part of developing musical skill on many instruments. [24]

Though "Pigments of Imagination" offers haptic feedback when users pick up instruments and responsive musical tones when they use the paintbrush, there are few other responses to the existing musical assets in the virtual room. This is a feature we would like to explore for future works. For instance, in the *generative virtual composition* "Crystal Vibes", the authors create an "artistic piece that combines spatial audio, endless fractal geometries, full-body haptic vibration feedback and sound-to-color visualization based on empirical biological functions related to human auditory and visual perception" (Outrum, Kinishi). Specifically, the work reaches beyond just audiovisual and spatial cues as discussed above, and maps for haptic feedback via a "synesthesia suit" with 26 vibrating actuators located on the head, arms, body and legs" which provide layers of vibration throughout the body "that allow the user to experience events, sounds and touch within virtual environments, engaging an extra dimension of the user's sensory

experience for greater emersion and synesthesia-like crossed-sense modalities"(Outrum, Kinishi).

> **Movement Matters** - Part of reality is the ability to move through it. Many virtual realities feature spaces that are much larger than the physical space the user is playing from. To compensate, each experience chooses a movement paradigm – a way of interacting with the world that causes the representation of the user's body to move through it…Thoughtfully designed movement paradigms have the opportunity to contribute to the aesthetic unity of the experience. [24]

How a user moves through a virtual space ultimately defines the aesthetic design of that space through the experience of the user. Therefore, it makes sense, specifically in narrative-based experiences, to subtly design the user's experience through the space in a way that aligns with the creator's original intentions for the experience. One notable concern in designing for movement is to empathize with the embodied position of the user and be wary of developing movement cues accordingly. Failure to do so can create cybersickness, a "visually induced motion sickness resulting from immersion in a computer-generated virtual world" [27].

As related to "Pigments of Imagination", the user transitions from full walkability on a flat surface in line with their height and natural perception of a space; its physics, gravity and size. However, at a certain point the user involuntarily ascends toward the moon, instantly shifting their expectations of the world's 'rules' and if not handled carefully, their sense of embodiment and presence in the virtual space. There is an extraordinarily strong chance that if not handled delicately, this pivot, necessary for the narrative, could disorient the user out of the experience completely. With this in mind we mapped a nonlinear time rate of ascension such that the user barely realized they were moving beyond their own will. As they prepare to transcend through the room's ceiling into the night sky above, the user is lifted slowly and held in place softly as to allow a moment to soak in through sight, sound and the 'feeling' of low-stakes (close to the floor) levitation; that they still have control over their virtual body, however their 'body' will move

through the virtual space guided by the narrative (akin to getting in a car or on a roller-coaster). The music composition recognized this pivot as well, moving from a busier sound-world involving narrative-informing lyrics to a more inviting musical space that aimed to offer a sonic 'safety-net' for the potential disorientation. We felt it was very important to give the user this opportunity to reconcile this shift; to produce in them not anxiety at a loss of agency, but a relief that they can relax, let the experience unfold around them and become more focused on the world, the narrative and the movements of 'narrative guides' strategically and subtly placed throughout the experience.

Lens 4: Doing vs Being

Atherton and Wang define *doing* as intentional action to achieve a goal. By contrast, they define *being* as the way one simply exists within the virtual world.

> We posit that so far, VR experiences have not given enough attention to being.VR often borrows heavily from traditional video games, which tend to focus on doing: moving, attacking enemies, achieving goals. However, as an immersive medium, VR also has powerful affordances for being: existing in an environment, absorbing the sights and sounds and finding one's place in the virtual world. This kind of being invites introspection and reflection, calm and the processing and development of emotions. [24]

Principle 4.1: "Design to balance doing (action) and being (reflection)"

We can also think of doing versus being as interactivity versus immersion. For music artists, virtual reality is a medium that has the potential to invite presence through immersion, to invite being versus doing. This reinforces the idea that music-centered VR works (i.e. VR music videos) that utilize both mediums to present the emotional and narrative cues of the music have been offered, unlike the mobile devices and laptop that inspire various strains of non-musical active engagement), an environment suited well for being rather than doing. In fact, in pieces like "Pigments of Imagination" and "12 Sentiments for VR", the user is purposefully 'not' doing. The

creator is in greater control of the user's environment and journey, allowing them to 'be' rather than 'do'. There are no real ‘goals’, they are allowed and encouraged to look around and listen and engage without definitively altering their journey.

Again, “12 Sentiments for VR” exemplifies the balance of doing vs. being, specifically in a narrative framework where the trade-off between immersion, interaction and narrative are most challenging. This works because “12 Sentiments” employs the Bardic narrative philosophy proposed by Blom & Beckhaus: create an overarching storyline and build “local pieces” within [31]. “12 Sentiments” exists in 12 scenes. The user is allowed to “be” (immersion/presence) in each scene and “do” (interact) as they’d like. The narrative “predominantly advances between movements rather than during them [24]. The downside, Atherton states, is that “the user is unable to affect the broader narrative; the meaningful actions available to them instead affect the development of the music itself” [24]. This exemplifies the challenges of the three-way trade-off between immersion, interactivity, and narrative: there will always be something lost when seeking more from the other two. This is why the *Communication Diagram* is so crucial: when these critical moments of uncertainty occur (and they will repeatedly), the artist can always travel back to the emotional center to help inform their decision in these trade-offs. It is in these thoughtful choices where not only the backbone of the work takes shape, but also the artist’s understanding of their reasoning for making it.

> **Doing and being for narrative-aesthetic coherence.** The user progresses through 12 different scenes and has the opportunity to be in each, enabling them to experience a coherent immersive and interactive world. Here, the narrative predominantly advances between movements rather than during them; the end result is that the user experiences many immersive individual scenes that punctuate a broader narrative. [24]

This harkens back to Blom and Beckhaus’s research on the Bard’s use of ‘local pieces’, in which overall narrative architecture is maintained while intermediary elements can be localized

to each storyteller. This approach is, speculatively speaking, the direction I intend to take with my next VR experience, “Do You Smile In Your Dreams?”, a forked-path binary-tree narrative that allows more interaction within each scene while maintaining a narrative arc that contains the more intentional emotional components [31].

> **Embodiment**. Doing and being are both intricately linked with embodiment and presence. When a user is being, they are settling into the environment and feeling like a part of it, which helps them connect with their virtual body if it is represented thematically appropriately (see Principle 6.1).When a user is doing, they take actions ‘as’ the virtual body, further reinforcing the link. Both are necessary to help a user connect fully with their virtual body. [24]

> **Gaze:** Something as commonplace as the direction of the user’s gaze can be used as an effective mediator between doing and being…While looking up is associated with awe and admiration, looking back is often associated with nostalgia and regret. We look back physically at places we have left behind, and metaphorically at moments in time in the past. [24]

“Pigments of Imagination” uses key diegetic elements of the world as subtle guides, such as birds, a paintbrush, and the moon, to direct the user's attention toward relevant narrative moments as well as offering overarching views to be absorbed by the user to reinforce foundational emotional messaging. It should be noted that these gazes are not possible in any other medium. A two-dimensional video would not allow a user to make the decision to look up in awe and then look down nostalgically or regretfully. This is an affordance of VR that allows for a magnification of that moment within the original composition. In turn, the musical and lyrical components of the song at that moment intentionally score these conflicting emotions brought on by the user's decision to both look up at a moon and look down at their room. As stand-alone music, this section of the composition may elicit similar emotions to a subjective listener, however when paired with a VR experience that guides the user to make a conscious decision to move their gaze in such a way to organically manifest a complex set of emotions, the potential in the music elevates to reinforce realizations made naturally by the user. This is powerful because in allowing the user to make these subtle but meaningful decisions, the emotional resonance of the piece

remains pure, unobstructed by the creator's manipulations of the user's emotions, something only possible because of the 'choice' that VR gives the user, even in a loosely interactive, fixed-media music-centered experience.

Lens 5: Design for Play

> Designed well, virtual reality can fulfill some of the preconditions for play, such as making users feel protected and free and giving them ways to express themselves. In this lens, we examine play as a simultaneous expression of doing and being. [24]

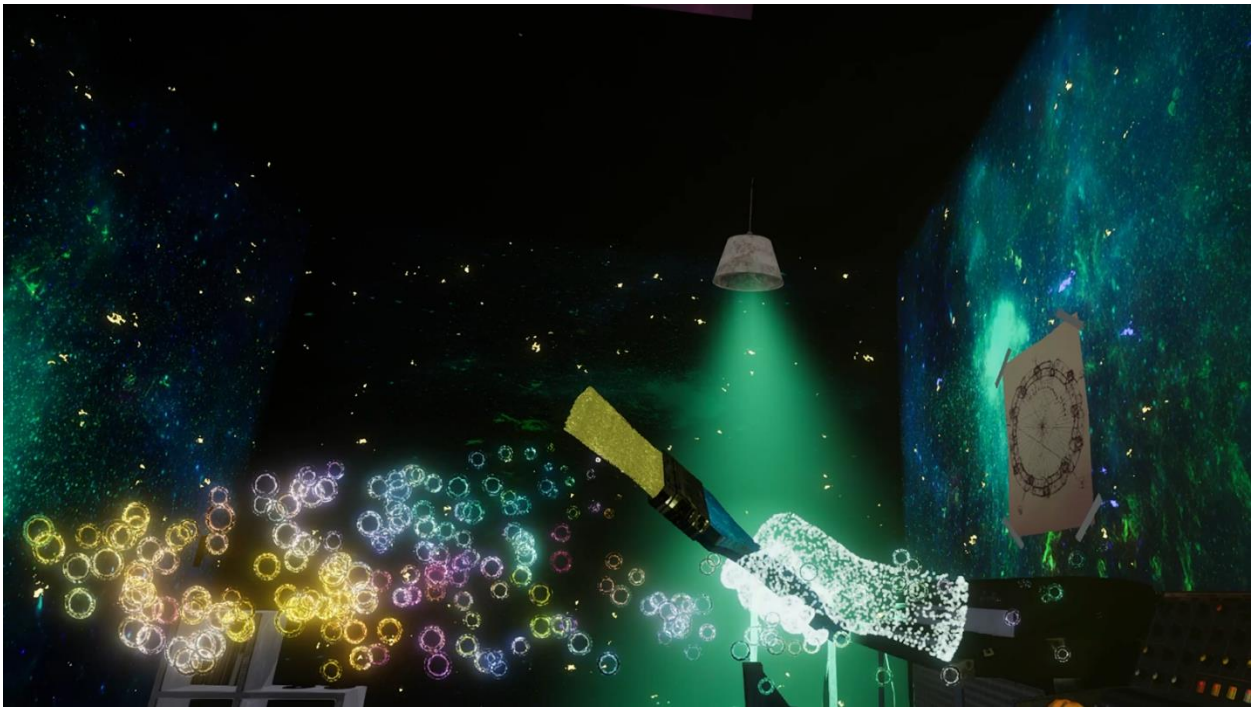


Figure 7-15: In “Pigments of Imagination”, a paintbrush is designed to let the user interact with the world and express themselves without sacrificing narrative structure.

In “Pigments of Imagination, user interaction, specifically toward play, is strongly invited, particularly before the narrative journey begins. As shared above, users can pick up most items (instruments, records, chairs and paint cans) found in the opening scene bedroom. They can hold them or toss them around. A paintbrush appears on a desk under a small spotlight, hovering and

rotating above the desk's surface. The implication, in comparison to other items in the room, is that this item is meant to be engaged with. Users can grab the paintbrush with either hand. Naturally, intuition is to wave it around, at which point tactile vibrations and the emergence of colorful paint particles spawn from the paintbrush [46]. Additionally, musical sounds launch directly from the paintbrush via object-based spatialization at varying tones depending on the speed of the brush's movements, giving the user a slight impression that they are creating musical elements within the world as well. The paintbrush can be tossed around the room, at which point it defies 'real' physics slowly bouncing off walls, floor and ceiling, spinning and spawning various musical tones that, to the user's ear, move spatially in alignment with the coordinate and movement of the paintbrush. As the user ascends toward the night sky, the paintbrush naturally disintegrates, allowing the user reclamation of two free hands. Though the user is confined to the journey toward the moon (a predetermined ascension), with their hands they can 'paint the sky' just by waving their arms and hands in any direction toward the surrounding sky. In response to this gesture, streaks of paint scatter the sky specific to the direction and gesture they waved to. This allows for a certain level of 'play', purposefully back-seated to 'movement guides' (birds, the moon, particle systems) designed to push the user to experience an awareness of their journey toward the moon. This subtle choice is made because the user's awareness of the moon and their journey to it is more important for the narrative and therefore more in service to the emotional truth. The option for the user to 'paint the sky' reinforces the surrealism of the world, its immersive traits and therefore the user's presence, however not at the expense of narrative and emotional resonance.

Chapter 7, in portion, is a reprint of the material as it appears in *In Proceedings of the 29th ACM Symposium on Virtual Reality Software and Technology (VRST '23)*. Association for Computing Machinery, New York, NY, USA, Article 71, 1–2. The thesis author was a co-investigator and co-author alongside Eito Murakami.

Chapter 8 : Conclusion

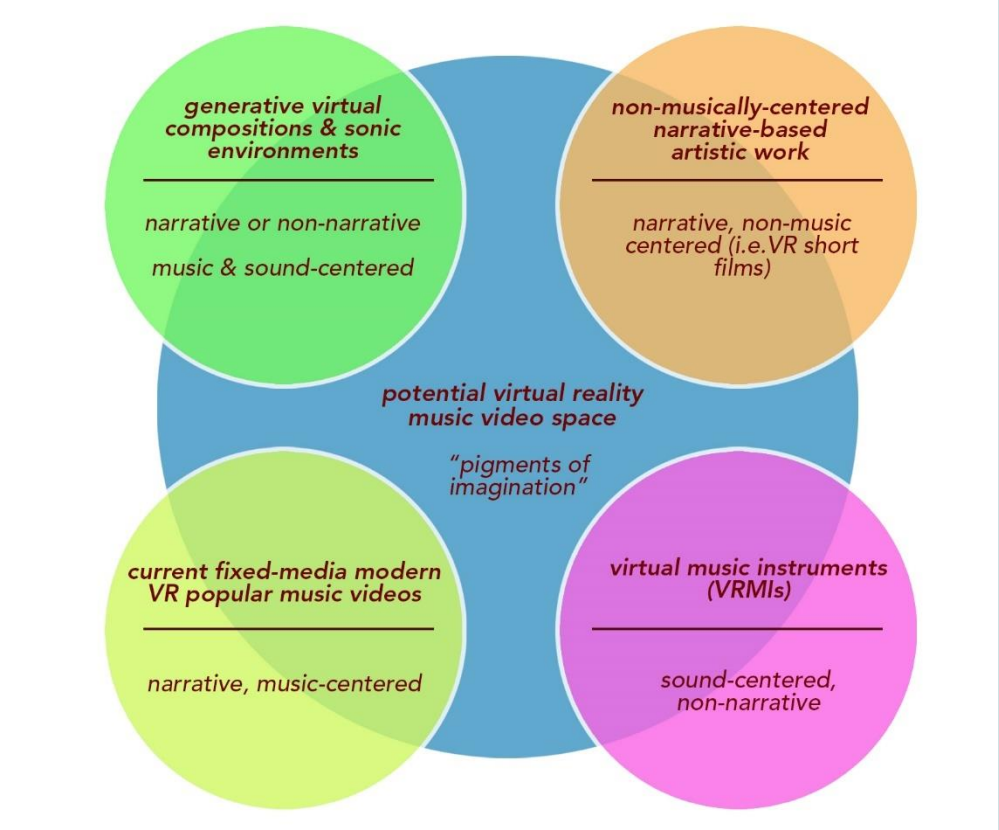


Figure 8- 1: Diagram of a reimagined "Virtual Reality Music Video" field informed by four surrounding VR music-incorporated VR fields.

What I have offered in this thesis is a route for multidisciplinary music artists to actively engage their listeners in a time where the listening experience can often feel deprioritized to other modalities. Consequently, this document revolved around methods of artistic practice wherein the artist can center the emotion of their work within music and have that emotion resonate via narrative through the adjacent medium of virtual reality. In doing so, this thesis diagrammed the ways in which emotion, narrative(s) and collaborative mediums iteratively interface with each other to achieve a prioritized emotional resonance from the artist to the experiencer. Conventions of each medium (music and VR) were examined in such a way as to illustrate how they might fundamentally change when friction between them occurs, thus creating a set of altered design

principles for an emergent fixed-media music-centered field within VR. Specifically, I narrowed this collaboration of mediums to the field of “VR music videos” and provided case studies from within the field and adjacent VR-related fields that might better inform artistic practice and maximize results. One of these case studies was the author’s work “Pigments of Imagination", which offered examples of these altered principles in action, but also areas that could be further explored within the work or future works.

Though “Pigments of Imagination” was quite successful both in its presentation of the author’s initial vision and participant reception thus far, there is room for improvement. For instance, when the participant appears in the room, they are surrounded by a bed, desk, chair, and instruments among other bedroom items. User intuition is typically to explore the level of interactivity in the room, specifically with the instruments. Users attempt to play an electric piano and synthesizer, however neither is interactive to this level. Though it was our intention to do so, we felt the time and effort to turn these into interactive VRMIs was not worth their contribution to the narrative; they were more so created to act as decoration with light interactivity. However, upon watching various participants attempt to play the instruments, we underestimated how this increased level of interactivity would add to immersion without taking away from narrative. Future projects will take this level of interactivity under consideration. Also, the tonal pads triggered by the user’s swipe of the paintbrush are not as musically additive as they could be which means they do not add to the composition in as meaningful a way. Future projects may approach these interactive elements with generative sound created within the DSP engine rather than pre-existing sample clips.

Audiovisual reactivity may be the largest area for future research and application. Since the creation of “Pigments”, I have begun formally researching and developing VR sub-projects that study audiovisual mapping in a more deliberate way. For instance, inspired by the synesthetic

VRMIs discussed in this document, current work involves applying the results of real-time STFT analysis and synthesis techniques toward particle system parameters in Unreal Engine. Future work also seeks to develop and then graduate music-centered narrative-based projects from the three-dimensional space of VR into communal dome settings. In this graduation, a completely new set of questions about immersion, interactivity and narrative dynamics arise. However, in this thesis's guiding principles, a path forward seems quite possible.

Challenges aside, we feel "Pigments of Imagination" met our initial goals and exceeded numerous expectations. As with any creative endeavor in which the artist learns and applies in real-time, a point comes where they must decide when to continue editing their current work or put that knowledge toward a new work. In the case of "Pigments", the goal (and moral of the work) was always to celebrate the process. With this paper, that process has been thoroughly analyzed and mined for lessons that can be applied in future work, therefore the project is complete. Though we have presented the work publicly in a few conference and festival settings, the next challenge becomes seeking out a wider public forum: galleries, immersive spaces, festivals and ultimately distribution to a general marketplace where the very music artists this thesis targets can experience the work in a tangible way. One thing about the uniqueness of a VR experience is how excruciatingly difficult it is to translate it in any other way but to have someone just experience it themselves. One prospect is to expand the spherical singular experience of Vr into a communal dome-setting (similar to Lanier's definition of the CAVE in VR system classifications), something I am excited to pursue both at a design and presentation level with future works.

15 years ago, I wrote a series of articles in a local Cincinnati paper as a struggling independent artist offering advice for other struggling independent artists. And I closed one article out analogizing these monumental goals we set for ourselves to the image of a beautiful glowing moon in the distance, which from afar looks stunning. And how early on our artistic

journeys we stare at that moon and dream about all of the potential fulfillment that it will offer us, and we hold onto it for dear life because it's all we have. And we begin our trek fueled by the promises of what the glow of that moon will offer us. So much so that we continue along an often lonely, thrilling, sad, adventurous, and disillusioning path toward this grand goal. But something happens along the way. The closer we get to this moon, its ephemeral glow, once catalyzed by the bliss of ignorance, is slowly replaced by very real and jagged craters, rocks, and dust. The glow that guided us here, is gone. We reach this goal, and it looks nothing like we had imagined. And it is quite heartbreaking, because the promise of what we think this goal will bring us is what fueled us through the loneliness, the sadness, the challenges, and the burnouts. Even more precarious, the narrow lens required of focusing so specifically on a single goal can blind you to the moments and people in life who may not be there with you to see that goal complete. This insight is what led to the creation of "Pigments of Imagination", both as a standalone song and full virtual reality experience.

I pursued this project because before starting it, I had no idea how to create it and that was exciting. In virtual reality, I found a way to three-dimensionalize my own imagination using a medium and tool that I had no prior knowledge of, a platform that might more properly extend my artistic intent to the experiencer. In the process of learning how to use this tool, the parameters of my imagination expanded, and it occurred to me: throughout the process of creating this work, I now know more than I could imagine at the outset. Ideas that were so big slowly became small: growth in real-time. This is a deeply empowering and exciting revelation that I wish to pass on to anyone else, specifically artists looking to push their craft forward in innovative directions. The process of turning big to small, of quantifying your imagination and sharing it; that you will one day know more than you are capable of imagining right now, and through your curiosity to expand, I will learn and imagine in even greater and deeper ways. This process is so special, so

powerful a mechanism for all of us as humans to undertake and pay forward. My hope is that this thesis and this project contributes toward that cycle.

References

1. Ahmet Kose, Aleksei Tepljakov, Sergei Astapov, Dirk Draheim, Eduard Petlenkov & Kristina Vassiljeva. 2018. Towards a Synesthesia Laboratory: Real-time Localization and Visualization of a Sound Source for Virtual Reality Applications. *Journal of Communications Software and Systems*. 14. 10.24138/jcomss.v14i1.410.

2. Angelique Scharine, Kara Cave & Tomasz Letowski. 2009. Auditory Perception and Cognitive Performance. 10.13140/2.1.3160.1925.

3. Anıl Çamcı. 2019. Exploring the Effects of Diegetic and Non-diegetic Audiovisual Cues on Decision-making in Virtual Reality. In Proceedings of the 16th Sound & Music Computing Conference SMC 2019 (Málaga, Spain). Sound and Music Computing Network, 195–201. https://doi.org/10.5281/zenodo.3249291

4. Anna Clarke. 2020. *Supporting Pro-Amateur Composers Using Digital Audio Workstations*. Ph.D. Dissertation. University of Nottingham. Nottingham, England.

5. Anthony Newcomb. 1987. Schumann and Late Eighteenth-Century Narrative Strategies. *19th-Century Music*, vol. 11, no. 2, 1987, pp. 164–74. *JSTOR*, https://doi.org/10.2307/746729. Accessed 16 May 2024.

6. Audio Synesthesia & Unreal Engine. 2022. *Working with Audio Synesthesia*. Retrieved May 8, 2024 from https://docs.unrealengine.com/4.27/en-US/WorkingWithAudio/Synesthesia/

7. Bosun Xie. 2013. Head-Related Transfer Function and Virtual Auditory Display (2nd Edition) - 1.1 Spatial Coordinate Systems (pp. 1). J. Ross Publishing, Inc. Retrieved from https://app.knovel.com/hotlink/pdf/id:kt00U8SF91/head-related-transfer/spatial-coordinate-systems

8. Brian Solomon. 2014. Facebook Buys Oculus, Virtual Reality Gaming Startup, for $2 Billion (April 2014). Retrieved June 9, 2024 from https://www.forbes.com/sites/briansolomon/2014/03/25/facebook-buys-oculus-virtual-reality-gaming-startup-for-2-billion/

9. Commercial Acoustics. 2020. How Sound Travels (Dec 2020). Retrieved May 16, 2024 from https://commercial-acoustics.com/how-sound-travels/

10. D. J Chalmers. 2017. The virtual and the Real. *Disputatio, 9(46),* 309–352.

11. D. Marini, R. Folgieri, D. Gadia, & A. Rizzi. 2012. Virtual reality as a Communication Process. *Virtual Reality*, (16), 233-241. Retrieved from DOI 10.1007/s10055-011-0200-3

12. David Hesmondhalgh. (2022). Streaming's Effects on Music Culture: Old Anxieties and New Simplifications. *Cultural Sociology, 16(1),* 3-24. https://doi.org/10.1177/17499755211019974

13. Devin Dolan & Michael Perets. 2015. Redefining the Axiom of Story: The VR and 360 Video Complex. Medium (December 20, 2015). Retrieved 1 April 2024 from https://medium.com/@devon.michael/redefining-the-axiom-of-story-the-vr-and-360-video-complex-bee3c20d69df.

14. Dolby Games. 2023. *Dolby Atmos and Unreal Engine*. Retrieved June 28, 2023 from https://games.dolby.com/atmos/dolby-atmos-for-unreal-engine/

15. Eduard Hanslick. 1955. Om det sköna i musiken. Uppsala: Schismens skriftserie 1, (orig. Vom Musikalisch-Schönen, 1854)

16. Francis Hamit. 1993. *Virtual Reality and the Exploration of Cyberspace*. Sams Publishing.

17. Francis Rumsey & Tim McCormick. 2014. Sound and Recording: Applications and Theory (7th ed.). Routledge. https://doi.org/10.4324/9780203756232

18. Fred Maus. 1991. Music as Narrative. *Indiana Theory Rev. 12.*

19. George Waddell & Aaron Williamon. 2017. Eye of the Beholder: Stage Entrance Behavior and Facial Expression Affect Continuous Quality Ratings in Music Performance. *Frontiers in Psychology* 8: 513.

20. Giovanni Santini. 2019. Synesthesizer: Physical Modelling and Machine Learning for a Color-Based Synthesizer in Virtual Reality. *Montiel, M., Gomez-Martin, F., Agustín-*

Aquino, O.A. (eds) Mathematics and Computation in Music. MCM 2019. Lecture Notes in Computer Science(), vol 11502. Springer, Cham. https://doi.org/10.1007/978-3-030-21392-3_18

21. Grahame Weinbren. 1997. The Digital Revolution is a Revolution of Random Access. *Telepolis: Magazin der Netkultur*, 17 February 1997 <http://www.heise.de/tp/english/special/film/6113/2.html>; accessed March 18 2024

22. Holly Rogers. 2014. *Music and Sound in Documentary Film*. Routledge.

23. International Society for Presence Research. 2000. The Concept of Presence: Explication Statement. Retrieved April 6, 2024, from http://ispr.info/

24. Jack Atherton, & Ge Wang. 2020. Doing vs. Being: A philosophy of design for artful VR. *Journal of New Music Research*, *49*(1), 35–59. https://doi.org/10.1080/09298215.2019.1705862

25. Janet Murray. 2016. Not a Film and Not an Empathy Machine. *Immersive News*. Retrieved March 2024 from https://immerse.news/not-a-film-and-not-an- empathy-machine-48b63b0eda93

26. Jaron Lanier. 2017. *Dawn of the new everything: A journey through virtual reality*. London: The Bodley Head.

27. Jason Jerald. 2015. The VR Book: Human-Centered Design for Virtual Reality. *Association for Computing Machinery and Morgan*; Claypool.

28. Jesse Damiani. 2022. *The Essential VR Music Videos of 2018*. Retrieved June 28, 2023 from https://www.forbes.com/sites/jessedamiani/2018/12/28/the-essential-vr-music-videos-of-2018

29. John Bucher. 2017. Storytelling for Virtual Reality: Methods and Principles for Crafting Immersive Narratives (1st ed.). Routledge. https://doi.org/10.4324/9781315210308

30. Johnny Wingstedt. 2005. Narrative Music Towards an Understanding of Musical Narrative Functions in Multimedia.

31. Kristopher Blom & Steffi Beckhaus. 2005. *Emotional storytelling*. University of Hamburg, Germany.

32. Liang Sun, Xuan Zhong & William Yost. 2015. Dynamic Binaural Sound Source Localization with Interaural Time Difference Cues: Artificial Listeners. *Journal of the Acoustical Society of America*. 137. 2226-2226. 10.1121/1.4920112.

33. Liz Pelly. 2018. Streambait Pop. *The Baffler* (Dec 2018). Retrieved May 9 2024 from https://thebaffler.com/downstream/streambait-pop-pelly

34. Louis Pisha, Siddharth Atre, John Burnett, & Shahrokh Yadegari. 2020. Approximate diffraction modeling for real-time sound propagation simulation. *The Journal of the Acoustical Society of America*, *148*(4), 1922. https://doi.org/10.1121/10.0002115

35. M. Fernanda Astiz. 2020. Storytelling in the Higher Education Classroom: Why it Matters. *Coll. Teach.* 68, 187–188. doi:10.1080/87567555.2020.1785382

36. M.C. Green, T.C. Brock & G.F. Kaufman, G.F. 2004. Understanding Media Enjoyment: The Role of Transportation Into Narrative Worlds. *Communication Theory*, 14(4), 311-327. Retrieved from https://doi.org/10.1111/j.1468- 2885.2004.tb00317.x.

37. Marc Ernst. 2004. Merging the senses into a Robust Percept. *Trends Cogn Sci*. 2004 Apr;8(4):162-9. doi: 10.1016/j.tics.2004.02.002. PMID: 15050512.

38. Marie-Laure Ryan. 1999. Immersion vs. interactivity: Virtual reality and literary theory. *SubStance*, 28(2), 110–137.

39. Marko Ciciliani. 2019. Virtual 3D environments as composition and performance spaces. *Journal of New Music Research* 49, 1 (December 2019), 104–113. https://doi.org/10.1080/09298215.2019.1703013

40. Mehmet Akten. VIS143. Class Lecture. Topic “Virtual Environments”. University of California, San Diego. October 7, 2021.

41. Mel Slater., and Sylvia Wilbur. 1997. A Framework for Immersive Virtual Environments (FIVE): Speculation on the Role of Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, *6*(6).

42. Michael Betancourt. 2013. The History of Motion Graphics : From Avant-Garde to Industry in the United States . United States: Wildside Press.

43. Michael Ogden. 2019. The Next Innovation in Immersive [Actuality] Media Isn't Technology- It's Storytelling. *The Asian Conference on Media, Communication & Film 2019 Official Conference Proceedings*.

44. Michael Schutz. 2008. Seeing Music? What Musicians Need to Know about Vision. *Empirical Musicology Review* 3: 83–108.

45. Michael Vallance & Phillip Towndrow. 2022. Perspective: Narrative Storyliving in Virtual Reality Design. *Frontiers in Virtual Reality*. 3. 779148. 10.3389/frvir.2022.779148.

46. Miguel Melo, Guilherme Gonçalves, Pedro Monteiro, Hugo Coelho, José Vasconcelos-Raposo, and Maximino Bessai. 2022. Do Multisensory Stimuli Benefit the Virtual Reality Experience? A Systematic Review. *IEEE Transactions on Visualization and Computer Graphics* 28, 2 (February 2022), 1425–1442. https://doi.org/10.1109/TVCG.2020.3010088

47. Mike D'Errico. 2022. *Push: Software Design and the Cultural Politics of Music Production*. Oxford University Press. https://global.oup.com/ushe/product/push-9780190943318?cc=us&lang=en&

48. Mike Senior. 2018. *Mixing Secrets for the Small Studio (2nd ed.)*. Routledge. https://doi.org/10.4324/9781315150017

49. Paul Milgram & Fumio Kishino. 1994. Taxonomy of Mixed Reality Visual Displays. *IEICE Transactions on Information and Systems*, *E77-D*(12), 1321–1329.

50. Pola Weiß. 2019. The Key: Straddling Mystery and Harsh Reality (October 2019). Retrieved May 20, 2024, from https://vrgeschichten.de/en/the-key-straddling-mystery-and-harsh-reality

51. Richard Cytowik. 2018. *Synesthesia*. The MIT Press. https://doi.org/10.7551/mitpress/10568.001.0001

52. Rob Hamilton. 2019. Coretet: A 21st Century Virtual Reality Musical Instrument for Solo and Networked Ensemble Performance. *IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, Osaka, Japan, 2019, pp. 1305-1306, doi: 10.1109/VR.2019.8797825.

53. Roland Barthes, R. 1981. *Camera lucida: Reflections on photography*. London: Vintage Classics.

54. S. Benford, J. Bowers, L.E. Fahlén, C. Greenhalgh & D. Snowdon. 1997. Embodiments, Avatars, Clones and Agents for Multi-User, Multi-Sensory Virtual Worlds. *Multimedia Systems*, 5(2), 93–104.

55. S.V. 2024. Oxford English Dictionary. "emotion (n.)," https://doi.org/10.1093/OED/2262908073.

56. Sachin Rekhi. 2017. Interaction Design (January 2017). Retrieved June 9, 2024 from https://medium.com/@sachinrekhi/don-normans-principles-of-interaction-design-51025a2c0f33

57. Shahrokh Yadegari, John Burnett, Eito Murakami, Louis Pisha, Francesca Talenti, Juliette Regimbal, and Yongjae Yoo. 2022. Becoming: An Interactive Musical Journey in VR. *In ACM SIGGRAPH 2022 Immersive Pavilion (Vancouver, BC, Canada) (SIGGRAPH '22)*. Association for Computing Machinery, New York, NY, USA, Article 2, 2 pages. https://doi.org/10.1145/3532834.3536209

58. Steven R. Holtzman. 1994. *Digital Mantras: The Languages of Abstract and Virtual Worlds*. MIT Press (MA).

59. T Venkatesan & QJ Wang. 2023. Feeling Connected: The Role of Haptic Feedback in VR Concerts and the Impact of Haptic Music Players on the Music Listening Experience. *Arts*. 2023; 12(4):148. https://doi.org/10.3390/arts12040148

60. Tilmann Haberman. 2019. Emotion and Narrative: Perspectives in Autobiographical Storytelling. Cambridge University Press. 10.1017/9781139424615.

61. Tim Robley. 2016. Notes on Blindness is One of the Most Eye-Opening Documentaries You'll See All Year - Review. The Telegraph (June 30, 2016). Retrieved May 12, 2024, from https://www.telegraph.co.uk/films/2016/06/30/notes-on-blindness-is-one-of-the-most-eye-opening-documentaries/

62. Timothy Gmeiner & Eito Murakami. 2023. Pigments of Imagination: An Interactive Virtual Reality Composition. *In Proceedings of the 29th ACM Symposium on Virtual Reality Software and Technology (VRST '23)*. Association for Computing Machinery, New York, NY, USA, Article 71, 1–2. https://doi.org/10.1145/3611659.3617201

63. Trevor Wishart. 1996. *On Sonic Art*, 2nd edition. Ed. Simon Emerson. New York: Routledge.

64. Tricia Rose. 1994. *Black Noise : Rap Music and Black Culture in Contemporary America*. Hanover, NH :University Press of New England, 1994.

65. Vijay Iyer. 2004. *Exploding the Narrative in Jazz Improvisation: The New Jazz Studies*. 10.7312/omea12350-020.

66. "virtual reality." 2024. Mirriam-Webster.com. Retrieved June 10, 2024 from https://www.merriam-webster.com/dictionary/virtual%20reality

67. Zach Bresler and Stan Hawkins. 2022. 'A Swarm of Sound' Audiovisual Immersion in Björk's VR Video 'Family'. *Music, Sound, and the Moving Image* 16, 1 (Spring 2022), 29–52. https://muse.jhu.edu/article/860469

68. Zoe Camp. 2016. Run the Jewels Launch Virtual Reality "The Crown" Video. Pitchfork (March 10 2016). Retrieved May 9, 2024, from https://pitchfork.com/news/63961-run-the-jewels-launch-virtual-reality-crown-video/