JOURNAL OF NEW MUSIC RESEARCH
https://doi.org/10.1080/09298215.2019.1705862

Doing vs. Being: A philosophy of design for artful VR

Jack Atherton and Ge Wang

Center for Computer Research in Music and Acoustics, Stanford University, Stanford, CA, USA

ABSTRACT
We present a practical philosophy for how to design artfully for virtual reality. Framing this design philosophy is the balance of *doing vs. being* as a central duality in all design for VR. The values underlying this way of thinking draw from a diverse set of interdisciplinary perspectives and from our own practice in computer music design. We develop a set of design principles, through which we critically analyse several popular musical VR experiences as well as some of our own creations. These include a tool for audio programming in VR, a set of design etudes on musical instruments and a multi-movement artistic work. Through these principles and case studies, we hope to provide lenses – both critical and practical – to inform the design of artful VR.

ARTICLE HISTORY
Received 26 March 2019
Accepted 5 December 2019

KEYWORDS
Virtual reality; artful design; musical interaction; design philosophy; aesthetics; human–computer interaction

1. Introduction

What does it mean for virtual reality to be *artful*? How do we design *well* for virtual reality? As the medium of VR enters a phase of widespread commercial adoption, these are central questions that designers of VR experiences grapple with. Their answers lie in a critical examination of values, an evolving notion of craft for VR design and an exploration of its ethical and humanistic implications.

Our work explores these questions through the design of musical experiences. We chose to focus on music because of its power to create rich human experiences. There is a universality to music. It requires both activity and reflection. Music allows us to play, to experience and bring about emotions, to bond with one another. Music lends realities a fullness and helps make an argument for their existence. But also, musical experiences make for telling case studies because their design requires richness and nuance. To craft expressive musical experiences in VR requires not only sound, but also visuals and interaction. The conjunction of these aspects motivates one of our central questions: how to build total experiences in VR.

The design practice presented in this paper is grounded in many perspectives on the medium of VR. They outline issues of immersion, interaction, telepresence and embodiment as core elements of virtual reality. Each perspective also contributes its own notion of what the medium should support, ranging from narrative and history to rich social interaction. At the same time, we ground our notions of the craft and ethics of design in prior design practice; in particular, virtual reality musical instruments (VRMIs) and the philosophy of Artful Design (Wang, 2018).

This article contributes a practical philosophy for the design of artful VR. It consists of a set of lenses and principles for designing well in VR, covering both the craft and the humanistic dimensions of the medium. It articulates strategies and virtues for creating total worlds. Among these perspectives is the use of audio as a first-class modality alongside visuals and interaction. We also introduce a new lens for designers of artful VR: the balance between *doing* and *being* as a central duality in all design for VR. Collectively, these lenses constitute a philosophy of design: they capture how we have thought about designing for VR in the context of both building experiences and reflecting on what we have built. As a *practical* philosophy, its theories are meant to be put into practice by real designers building real-world systems.

The rest of this article will proceed as follows: first, we enumerate the lenses and design principles of our philosophy in order to give the reader an idea of the scope of the article. To motivate these, we consult perspectives from several disciplines to formalise a definition of what constitutes a virtual reality. We then endeavour to illuminate our design principles through a number of case studies. Overall, we argue that artful VR design expresses human

CONTACT Jack Atherton lja@ccrma.stanford.edu Center for Computer Research in Music and Acoustics, Stanford University, Stanford, CA, USA

values, makes use of the core nature of the medium, and articulates a balance between doing and being.

Lens 1: Don't forget about audio

- Principle 1.1: Audio should be dynamically generated.
- Principle 1.2: Audio should be immersive.
- Principle 1.3: Audio should be interactive.

Lens 2: Designing to the medium

- Principle 2.1: Don't Port (Corollary): Make things that would be impossible in the physical world.

Lens 3: Doing vs. Being

- Principle 3.1: Design to balance *doing* (action) and *being* (reflection).
- Principle 3.2: Look up! Use gaze to modulate between *doing* and *being*.

Lens 4: Interaction

- Principle 4.1: Drive interaction design with aesthetics.
- Principle 4.2: Multimodality is a virtue.
- Principle 4.3: Make space for *being* alongside *doing* in interaction.

Lens 5: Immersion

- Principle 5.1: Create worlds that enhance *doing* and *being* through animus.
- Principle 5.2: Balance stylisation and realism.

Lens 6: Designing for the body

- Principle 6.1: Design for virtual embodiment.
- Principle 6.2: The body is an implicit medium where *being* supports *doing*.
- Principle 6.3: Movement matters.

Lens 7: Designing for play

- Principle 7.1: Play is both an activity and a state: a synthesis of *doing* and *being*.

Lens 8: Designing for social

- Principle 8.1: Replicate baseline social interactions; redesign the rest.
- Principle 8.2: Support many kinds of social engagement.
- Principle 8.3: Design for social *doing* and social *being*.

2. Reality with a capital 'R'

We seek to build total experiences that wholly transport the user to another reality (or give the perception of doing so). This goal necessitates a definition of what constitutes a 'reality'. A broad range of disciplines (Chalmers, 2017; Pimentel & Teixeira, 1993; Ryan, 1999) define the baseline requirements for virtual reality as:

- virtual – to be a virtual reality, the reality must be simulated (e.g. computer-generated).
- immersive – to be a virtual reality, the reality must give its experiencers the sensation of being surrounded by a world.[1]
- interactive – to be a virtual reality, the reality must allow its experiencers to affect the reality in some meaningful way.

With the goal of creating total, all-encompassing experiences in mind, one aspect of designing well for VR is thus to align the designed experience with this definition of virtual reality.

Experiences that satisfy these conditions bring about a number of powerful effects. The nuanced combination of immersion and interaction leads to an effect called *telepresence* (or *presence*), which is the effect of feeling more present in the virtual environment than in physical reality (Ryan, 1999; Steuer, 1992).

Presence brings with it the effect of *(virtual) embodiment*, the ability for a user to feel and act as though a virtual body is their own. This 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 (Benford, Bowers, Fahlén, Greenhalgh, & Snowdon, 1997; Groom, Bailenson, & Nass, 2009; Schultze, 2010).

2.1. Terminology

Throughout this article, we will use the term *VR* to refer to the hardware systems for delivering immersive experiences and to refer to the immersive experiences themselves. We will use the term *virtual reality* to refer to a VR experience that specifically satisfies the definition presented above.

3. Literature review

Beyond this definition of virtual reality, what other perspectives should we use to inform the design of our VR

[1] Immersion has also been defined in a bottom-up sense as technology that 'delivers the ability to perceive through natural sensorimotor contingencies', i.e. technology that delivers stimuli to be used by our perception systems, such as visuals, sound, touch (Slater & Sanchez-Vives, 2016).

experiences? Here follow ideas from several disciplines that ground our design philosophy.

3.1. Philosophy

3.1.1. Virtual realism

David Chalmers' essay 'The Virtual and the Real' outlines a position called *virtual realism*, which posits that virtual realities are genuine realities; that virtual objects are real objects; that what goes on in virtual reality is truly real; that virtual experiences can be as valuable as non-virtual experiences; and indeed that 'life in a rich VR is roughly as valuable as ordinary non-virtual life' (2017). In this view, virtual embodiment (the perception of inhabiting a virtual avatar) is not a fiction and VR allows us to actually do things rather than simply having the flavour of doing things.

Chalmers identifies a number of aspects of life that cannot be experienced in current VR at the same level of richness as in physical reality. Some of these can be rectified by future advances in technology, while others might be impossible to achieve in the medium. The former aspects include relationships, embodiment and quality (e.g. visual resolution, number of modalities), while the latter ones include rich histories, naturalness and birth and death.

In turn, we might also further reflect on what we consider important in music-making: emotional connection, expression and regulation; social bonding and community; virtuosic performance; storytelling; aesthetic self-fashioning; reflections of our humanity. How well might each of these aspects be supported by the medium of virtual reality in the future? Such considerations can range from a high level (e.g. how to use the medium to support the community functioning of a group of music-makers) to a low level (e.g. how developments[2] in haptic feedback and tangibility can make virtual instruments feel more 'real' and possibly more virtuosically playable).

3.1.2. Human flourishing

Eudaimonia or human flourishing is an ancient concept that refers to living well in a deep, rich, meaningful way. One of the more recent notions of this concept is the Capability Approach created by Amartya Sen, which suggests that rather than prescribing certain actions or states of being that everyone must fulfil in order to flourish, we should instead prescribe capabilities, which are actions or states of being that everyone must have the *ability* to fulfil *should they so desire* (1999). Martha Nussbaum has created one of the most canonical, widely accepted lists of capabilities (2007). David Hesmondhalgh has commented on how music can play a central role in fulfiling some of these capabilities (2013). In particular:

4. Senses, Imagination, and Thought: the ability to experience and produce musical works (and more).
5. Emotions: the ability to experience emotions, especially emotions directed toward others (such as longing, grieving, compassion, gratitude, justified anger).

7a. Affiliation: the ability to engage in social interaction at a variety of levels, from small communities to society and to empathise with others.

9. Play: the ability 'to laugh, to play, to enjoy recreational activities'.

As we develop VR further and start to weave it into the fabric of society, we should keep these capabilities in mind as guiding forces on the moral-ethical level of design. We should design VR experiences where people can express themselves musically, rather than just experiencing what others have created (capability 4). Our designs should keep in mind empathy (capability 7a) and the healthy development and processing of self- and other-directed emotions (capability 5). Our designs should preserve and enhance our users' ability to engage socially with a community and with society at large, rather than degrading this ability as current 'social networks' arguably do (capability 7a). Our designs should foster and encourage play through a combination of feelings of safety, whimsicality and opportunities for creative self-expression (capability 9).

3.2. Literature

In her 1999 review of the field, 'Immersion vs. Interactivity: Virtual Reality and Literary Theory', Marie-Laure Ryan explores problems and prospects of virtual reality as it relates to literature. Ryan proposes that in literature, a trade-off between immersion and interactivity comes about when a *narrative* is introduced. Ryan notes that in textual fiction, the more interactive the text becomes, the less immersive it becomes. A common example is a choose-your-own-adventure book that ask readers to choose actions at key moments and flip to other sections of the book. The action of doing this takes the user out of the world of the adventure and reminds them that they are experiencing it through a medium; the flipping of the pages is an action that they take in the physical world which may or may not overshadow the action they are virtually taking by choosing the next page. It reminds readers that the world was *created*, that it doesn't just stand on its own, and that it didn't come about from nothing.

[2] For example, see Berdahl and Huber (2015).

Ryan thus suggests that a major problem to be solved is how to maintain *narrative coherence* and *aesthetic value* while still allowing the experiencer to affect the reality. One important subcomponent of this is to make sure that the experiencer can reasonably anticipate what will happen as a result of their actions in the reality. A larger 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). Ryan points out that experiences that don't rely on narrative much or at all can achieve the best fusion of immersion and interactivity. The examples she provides are flight simulators, sports games, visual displays and (perhaps most auspicious for creators of musical experiences) combinations of visuals, sound and dance.

3.3. Communication

Jeremy Bailenson's book *Experience On Demand* (2018) covers many important affordances of VR, some of which he came upon through directing the Virtual Human Interaction Lab at Stanford University. The book discusses how experiencing alternative, physically impossible realities can be exhilarating and even have profound impacts on how we live our day-to-day lives in physical reality. For example, chapter 6 details promising VR work in reducing pain, such as in patients who are recovering from serious burns.

Despite these compelling effects, Bailenson believes that the most important part of virtual reality will be its social dimension. More than photorealistic environments and avatars, he believes that what makes a virtual environment feel real is the community of people you can interact with while in that reality. It is also not sufficient for a group of people to just show up together in a virtual environment; we need the design of social virtual environments to replicate the necessary parts of physical interaction for the experience to be satisfying and believable. (For example, video conferencing with webcams does not preserve social cues of eye contact because if we direct our gaze to our screen, our webcam will display us looking down.)

Most VR right now focuses on single-person experiences, with the exception of fledgling social networks. Even these experiences have a long way to go toward a satisfyingly rich vocabulary of virtual social interactions.

3.4. Ethnomusicology

Andrew Killick's treatise on *holicipation* outlines it as the under-studied phenomenon of solitary music-making for personal satisfaction (2006). People who engage in holicipation make music for the sheer enjoyment of it, rather than because they are looking for admiration or practicing for a performance. Killick asserts that holicipation is 'one of the most widespread forms of musical activity' and outlines how it is primarily ignored by ethnomusicology, the psychology of music, sociological study of music in everyday life and more fields.

The idea of making music as an end in itself for personal satisfaction is intricately tied up in the notion of human flourishing. Holicipation can be used to process one's emotions, to play and of course to express oneself artistically.

Holicipation is an important concept for VR, since it is a medium that by default involves solitary experience. As we learn more about holicipation, we should pay attention to how these findings can inform design practices for VR experiences that support play and human flourishing through making music on an individual level.

3.5. Art

Hakim Bey's notion of *immediatism* is art-making that is done for immediate pleasure and for the small localised social context surrounding the artist (1994). It has significant overlap with Rich Gold's notion of folk art (2007) and could be considered a version of holicipation that instead involves more than one person at once. There are a number of compelling assertions in Bey's manifesto, which assert the effects of immersion, capitalism and other forces on the mediation of experience, but perhaps the most relevant to us are:

vi. 'Real art is play, and play is one of the most immediate of all experiences'.
xi. Immediatism involves two or more people engaged in creative play, by and for (only) themselves.
xii. In immediatism, all spectators must be performers.

This lens would suggest that we must encourage our users to express themselves artistically through play, not only in isolation (which is so easy in VR), but socially, in small yet meaningful communities.

3.6. Design

3.6.1. VMIs and VRMIs

Virtual musical instruments (VMIs) are relevant to discussions of making music in VR. VMIs are gestural interfaces that use movement tracking to represent a human body part virtually. This is often a hand with 6 degrees of freedom (i.e. 3D position and rotation) as in a VR controller (Mulder, 1998). Central questions of the field of VMIs involve musical expression, mental and physical effort, and complex continuous gestural control. The most relevant question to interaction design for music in VR, however, is that of kinaesthetic feedback; for example, Mulder identified the necessity for such feedback in order to play regular drum rhythms with skill (1994). There has been some work in creating haptic hand controllers that could be used in VR. For example, the Haptic Hand is a prototype of a musical interface that places human fingers on plastic keys (Berdahl & Huber, 2015). The Dexmo is a commercial exoskeleton glove approaching mass-production that applies resistive force to each finger individually to give the wearer the sensation of touching a virtual object (Dexta Robotics, 2019).

VMIs can sometimes involve visuals (Mulder, Fels, & Mase, 1999), but those involving an immersive visualisation in a headset are better described as VRMIs (virtual reality musical instruments) (Serafin, Erkut, Kojs, Nilsson, & Nordahl, 2016). In a 2016 survey, Serafin et al. outline the need for more works for VR that include audio in a primary role and present a few preliminary design principles for doing so. They also discuss Rob Hamilton's perspective of virtual environments as instruments with gesture decoupled from sound (2009).

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.

3.6.2. Artful design

Artful Design is a philosophy of how to design well that originates in computer music but can be used as a lens for most design projects (Wang, 2018). It focuses on how to shape technology with craft, ethics and aesthetics. Such a philosophy is especially helpful when grappling with questions such as 'Why do we design?' and 'What are we designing for?' in the face of a highly commercial wave of technology that carries with it assumptions that are not always aligned with values of music-making. Artful Design encourages its practitioners to put the ethics of how they will live alongside their creations at the bedrock of their work, rather than as an ethical 'leash' to restrain the final product. It encourages them to design not (only) from specific user 'needs' but from the human values that underlie those needs. It does so with a treatment of aesthetics that include not only practical design questions and 'deficiency needs', but also hidden dimensions such as emotional, social and moral aesthetics and 'growth needs'.

Artful Design connects particularly well to the notion of human flourishing. It uses the earlier Greek term *eudaimonia* to refer to a life in which one thrives, with a particular focus on meeting growth needs that enable us to strive for our full potential and self-actualise. These include our need for belonging (social) and our need to appreciate harmony and beauty (aesthetics). Artful Design reminds us that the way we shape technology touches people, altering their lives and their happiness. Our choices in design should therefore be bound to the same standard of ethics as we hold ourselves to in everyday life. We should design to help one another flourish.

4. Case studies

To expound our design principles, we will explore a number of case studies, referenced when appropriate throughout the article. These include design critiques of our own work as well as some of the most popular commercial musical VR experiences of 2018. The latter commercial experiences do not necessarily represent the state of the art in terms of exploration of the possibilities of music in VR, though some do get far. They are, however, representative of the first broad commercial wave of VR technology, released at a time when design practices continue to evolve and included here to help illuminate how our design philosophy might inform broader trends in commercial VR. Here follows an overview of the experiences discussed, shown also in Figure 1.

Fantasynth: Chez Nous is a 7-minute journey through a land of lights and angular geometry that serves as an animation or visualisation of a soundtrack (HelloEnjoy & N'To, 2017). Users can turn their head to look at the mechanised and dangerous world around them but cannot do anything to change their movement (after a couple minutes of stationary experience, they are moved forward at a constant rate) or to affect the playback of the music (it is a static audio file that plays back linearly). This piece does not include a strong narrative, focusing instead on showing fantastical scenery.

Surge is a 4-minute audiovisual journey of small cubes coming to life to form colossal beings (van Meerten, 2016). Like *Fantasynth*, users in *Surge* can turn their head to look at the world of blocks around them, but cannot move (the world moves past them in little jumps timed to the beat of the music) and cannot affect the music. This piece has more of a narrative, depicting the origins of a technological life form.

Playthings: VR Music Vacation is a set of playful and colourful musical instruments made out of food, along with a brief rhythm game (Brower, 2016). Users progress through stages learning to use each of the tools of the game (for hitting close instruments, hitting far away instruments and moving instruments) and getting used to the various instruments themselves (e.g. jelly bean chimes, hot dog marimbas). In the final stage, the mode switches from exploratory free play to a rhythm game

Figure 1. The case studies. Left-to-right, top row: *Fantasynth: Chez Nous*. *Playthings: VR Music Vacation*. *TheWaveVR*. *The Lune Rouge Experience*. Middle row: *Aktual*. *VRAPL*. *Canyon Drum*. *Wheebox*. Bottom row: *Surge*. *Twist Flute*. *Shred Head*. *12 Sentiments for VR*.

where food instruments move toward the user, who must try to hit them with correct timing in order to play along to a prerecorded song.

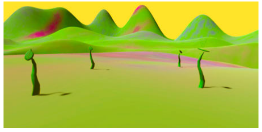
TheWaveVR is a social VR platform for musical VR experiences based on commercial music (TheWaveVR, 2017). Its default areas are set in a series of dimly lit caves, allowing users to customise their body amongst several abstract choices and to navigate to a series of musical experiences contained within a series of arcade-like facades. The platform includes experience points for levelling up and in-game currency for purchasing visual novelty items such as a power-up that allows one to generate butterflies from one's hands. Both experience points and in-game currency are earned by visiting musical experiences.

Aktual is a TheWaveVR experience featuring an iridescent stage with trippy and sometimes grotesque visuals for Ash Koosha's audiovisual show of the same name (Koosha & Strangeloop Sab, 2017). Users in this experience can teleport around the small stage, look up at a skeletal DJ who is 'playing' the music, look at the environmental visuals and pick up an occasional object that appears. Once such object is an abstract ball of scribbles that thrashes around blocking the user's vision when they move it. There is no way for the user to affect the music or the broader environmental visuals, which play back linearly.

The Lune Rouge Experience is a TheWaveVR experience featuring a colourful stage and a set of 4 musical toy regions that let users play with and 'remix' TOKiMONSTA's album *Lune Rouge* (Lee & Strangeloop, 2017). The central stage area plays through the album linearly. However, users can teleport across the virtual environment to four different areas to play with musical toys that are based on individual instruments or voices from four of the songs from the album. For example, one toy allows users to pick up spheres that play a loop of an instrument and move the spheres in and out of objects that apply different audio effects, including filters, echoes and reverb. All of the toys require more than one person present to play all of the parts simultaneously.

VRAPL (VR Audio Programming Language) is our prototype of audio-based live programming in VR. Users create sculptures out of 3D programming blocks that can generate audio and control the physics of the world with audio signals. The blocks include audio-recording comments, function blocks that the user can teleport inside of, and oscillators that can change their frequency when they are made smaller or larger, among many others.

The following four experiences are design etudes (small VR experiences intended to explore a specific concept). *Canyon Drum* is an exploration of many ways to play a drum that is very large and very far away in VR. The goal was to find a method of interaction and tactile response that could allow amateur users to play the drum with a similar virtuosity as they could a physical drum. *Wheebox* is a toy where users launch projectiles that both have fun and get nervous. The goal was to explore *animus*, or the feeling that the environment is alive and full of creatures that have their own feelings, desires and goals beyond responding to the user. *Shred Head* is a set of two objects that can be played either as wind chimes with a virtual hammer or by placing the objects on one's head and swinging them around as if they are wigs. The



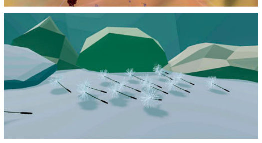
Movement 1: timid exploration to exhilaration. A young vine (the user) in a community of vines learns to grow by looking up at the sun and sprouts leaves.

Movement 2 : excitement, longing. The vine learns to catch sunlight on its leaves.

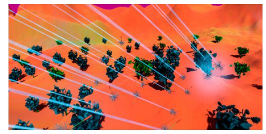
Movement 3: intensity, weight and weightlessness, awe. The vine uses flowers to sprout and launch seedlings into the air.

Movement 4: lightness, flowing. The wind (the user) helps the newly sprouted seedlings learn to fly upward in search of a new home.

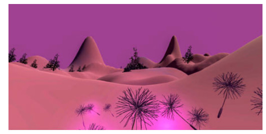
Movement 5: unsure, mournful, peaceful. The seedlings (the user) fly on a jetstream across the landscape, looking forward to a new home and looking back on seedlings that fall behind to be lost forever.



Movement 6: bleak, serene. Rocks in a barren landscape (the user) look on as the seedlings are lost and stuck, but find their way when they are lifted by a strange, magical wind that is called by the user.

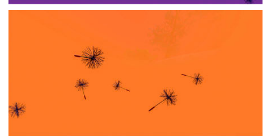


Movement 7: resolved, strong. The seedlings (the user) use the wind to control how they move across the landscape and find a new home.



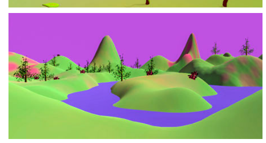
Movement 8: melancholy, weak. The wind (the user) can no longer help the seedlings along, which have grown heavy and rest on dusty ground.

Movement 9: calm, stasis. The ground (the user) absorbs drops of rain to nourish the now half-buried seedlings.



Movement 10: frenetic. The seedlings (the user) vibrate excitedly and with some confusion; something big is coming.

Movement 11: sublime. The seedlings sprout.



Movement 12: joyful, fulfilled. The user grows upward from the perspective of a sproutling to gain a view of the entire world.

Figure 2. The aesthetics that motivated the design of the movements of *12 Sentiments for VR*.

goal was to explore unexpected interactions and whimsy. *Twist Flute* is a virtual flute that users play by blowing into their headset; hand distance controls pitch and hand twisting controls timbre. The goals were to make a physically impossible version of a physical instrument and to use multimodal input (i.e. not just the controllers, but also the breath).

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.

Each of the movements is outlined in Figure 2. Here follow brief descriptions of the movements: the user's point of view, the environment and the musical actions available to the user.

1. The user is a young vine surrounded by similar vines who are animated to look up. When the user looks up, they catch sunlight and grow. Depending on how far up the user looks, different chords play.
2. The user has grown vines and leaves (their hands). They are surrounded by similar vines attempting to catch rays of sun on their leaves. The chord used and the quality of the timbre are controlled by the user according to which leaf they use and how long they hold it in each sunbeam. As the user catches more sunlight, the music becomes more developed, increasing in complexity and pitch height.
3. The user now has flowers for their hands. When they squeeze a flower, a chord repeats at at increasing tempo and seedlings spawn on the surface of the flower; when they release it, the seedlings jump into the air and play a high arpeggio with a decreasing tempo. The user controls the chord played (which hand they use) and the length of the arpeggio (how long they squeeze the flower).
4. The user is no longer a plant, instead embodying the wind with two clouds for hands. The shell of the flower plant is laying on the ground and the air is full of seedlings that jump upward one by one, slowly playing an arpeggio. When the user moves their clouds, two different notes play (one per hand); when they squeeze one of their clouds, a high-pitched chord fades in as well; and when they swing their cloud and stop squeezing, the cloud of seedlings is blown in the direction of their swing and plays a rapid arpeggio. The length of the arpeggio and the number of seedlings affected is controlled by how long the user squeezes.

5. The user follows the cloud of seedlings as they travel across the landscape in a jetstream. They have similar controls to movement 4, but cannot affect the general direction of the seedlings. Additionally, the user's gaze direction affects the mood of the music; when they look forward onto the sunny side of the landscape, the mood feels pastoral and calm, whereas when they look backward toward the shadowed side of the landscape to view seedlings that are accidentally left behind, the chords used are changed so that the mood feels somber and melancholy.
6. The user is a rock at the bottom of a barren valley where the seedlings have become stuck, rendered flightless. Gusts of wind sonified in musical chords blow the seedlings around; the seedlings emit detuned pitched percussive sounds when they collide with one another or the ground. The first set of wind chords are a short cycle crafted to evoke bleakness. When the user looks up for long enough, a cloud of glowing particles descends and transforms the wind chords into a longer, more serene-sounding cycle. With this change, the landscape begins continually and subtly changing colours, and the seedlings are slowly and shakily lifted out of the valley, falling back down many times during the process.
7. The user embodies the wind as in movement 5, but now each gesture controls the direction and speed of travel of the seedlings. The user's hand movements and squeezes still control the music; the speed of their gestures now also subtly decreases or increases the tempo. At higher tempos the textural density of the music is increased.
8. The user embodies the wind and has roughly the same musical controls as in movement 4, but now the seedlings cannot float and are briefly suspended in the air by a slowing of virtual time before falling back to the ground, occasionally becoming buried.
9. The user embodies the ground and has a split perspective; they can see the landscape being rained on above ground and the seedlings buried below the surface of the ground. The more they raise their arms, the more rain they capture, increasing the brightness of the music's timbre and the frequency of the rain arpeggio.
10. The user is now fully underground, seeing the seedlings vibrate as they prepare to grow. They control the rhythm, tempo and overall advancement of the music by squeezing and releasing their hands to start and stop some of the streams of notes.
11. The user is quickly moved above ground to see the sudden sprouting of the seedlings. They have the same musical controls as in movement 10, though this movement is shorter.
12. The user experiences a continual change in perspective whenever they look up in this movement. They begin with the perspective of a tiny being at the base of the new sproutlings and grow as they look up, quickly eclipsing the sproutlings and growing into a cloud of floating seedlings. These seedlings begin swirling and jumping around the user, playing individual arpeggio notes once they are activated by the user's growth, as the user grows further to become the size of the landscape itself. At this point, wind chords similar in timbre to those of movement 6 come to envelop the entire landscape, and eventually the movement ends.

5. Lenses and design principles

The process of design involves a series of choices, and a thoughtfully designed artefact reflects the values underlying those choices. The design principles presented below are not prescriptions for what is 'good' design, but instead are propositions, contextual ways of thinking about design choices.

In the pursuit of exploring the nuances of these design principles, we have applied them in design critiques of the artefacts described above. We are not concerned with the original intentions of the designers of these artefacts; indeed, some of these artefacts were created by designers whose original intentions it is impossible to know. Instead, the design critiques 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?'

Lens 1: Don't forget about audio

As stated in Section 2, one of our central goals is to build total experiences. Many of these experiences combine the musical, the visual and the interactive into a cohesive whole. For these experiences to present their own realities, audio must play as much of a central role as visuals and interaction. Since the modality of audio is often neglected by work outside the field of music technology, a focus on audio as a first-class modality is essential in order to achieve the balance of interaction, audio and visuals that is necessary to create artful experiences.

Our notion of how to treat audio in VR is motivated by the above definition of a virtual reality. A virtual reality with audio as a first-class modality prioritises audio along each of these dimensions:

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

Principle 1.1: Audio should be dynamically generated.

On the surface, this property is easy to achieve. Any experience with sound delivered via a computer other than the direct, unmodified playback of static, pre-recorded audio files will have dynamically generated audio. Here are two more nuanced interpretations of this principle:

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 (e.g. *Fantasynth* and *Surge*). This is not to say that the playback of audio files cannot be responsive; for example, an audio file that is pitched up or down according to the user's actions (such as the intensity of a throwing gesture) could be considered responsive, as could a scene of audio files that is customised with a live recording of the user's voice.

Alignment to medium. Dynamically generated audio can benefit from some of the unique affordances of the medium. In VR, audio can be synthesised 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 (Cook, 2002).

The converse of this principle is worth considering as well: aspects of an experience that are unique to the medium can often benefit from using the modality of audio. *VRAPL* uses audio for the design of its comments. Since VR does not easily afford typing, users leave audio recordings rather than textual notes near their programs (Figure 3).

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).

Figure 3. Audio comments in *VRAPL* allow users to leave verbal descriptions to be played back later. This comment will play back a recording that says, 'this is a sine wave with frequency 337 Hz'.

Spatialisation. Sound spatialisation is an essential and surprisingly oft-overlooked affordance of VR. Surrounding the user with an environment full of objects that make localised sounds will help give the sense of being in a living world. Common sense can go a long way here; if audio and visuals suggest that a sound emanates from an object, the sound will contribute to the experience if it is spatialised to that object's location. Matching the audio location to the visual location avoids spatial confusion. Unfortunately, *Fantasynth* and *Surge* both spatialise 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.

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.[3]

Multiple ways to interact. One more obvious way to enable users to meaningfully affect audio is to create virtual musical toys and instruments, as in *Lune Rouge* and *Playthings*. We can also work with a broader notion of musical interaction in VR that is not restricted to notions of 'instruments' and 'interfaces'. For example, the second movement of *12 Sentiments* enables users to create music through interaction with the environment. Here, users collect sunlight with their hands, represented as leaves; a chord swells up as the leaf absorbs sunlight (Figure 4).

[3] Related lens: Nussbaum's Capability 4: Senses, Imagination, and Thought (2007).

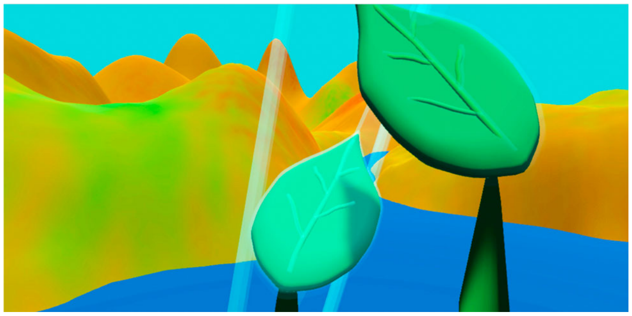


Figure 4. In movement 2 of *12 Sentiments*, the user affects the music by collecting sunlight with their leaves (hands). (See a video: https://ccrma.stanford.edu/~lja/dvb/leaf-hands.)

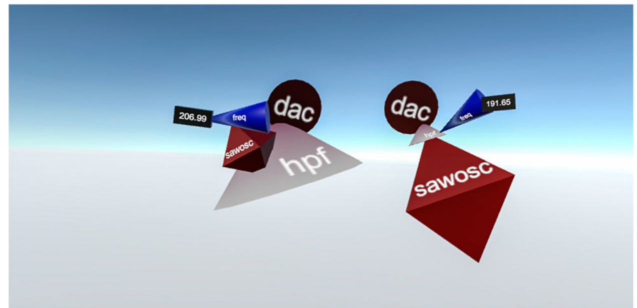


Figure 5. Flexible sculptural assembly allows users to express the same program in different ways, according to their intention and personal aesthetic. Here are two identical programs.

The user can advance the narrative by collecting sunlight with either hand, but each hand has its own chord with its own musical consequences. Here, the user can meaningfully affect the virtual world's music without the experience being framed as a musical instrument.

Alignment to medium. Sonic interactions work well when they are aligned with the basic affordances of the technology. Commercial VR controllers afford pushing buttons, squeezing, using a two-dimensional touchpad, and relatively precise hand position and velocity information. They do not afford the dexterity, touch sensitivity, or touch accuracy that playing musical instruments in physical reality often requires. Align users' sonic and musical interactions with what can be implemented well in the medium (see also Principle 2.1). Alternatively, augment the medium to include affordances necessary for desired sonic interactions, such as in the field of VMIs (Berdahl & Huber, 2015).

For example, consider different treatments of rhythm. In order to create rhythms in *Playthings*, one must individually play each note with correct timing; the system does nothing to help align user actions to a time grid, making it difficult to attain precisely timed percussive music. By contrast, the musical experiences in *Lune Rouge* and *12 Sentiments for VR* leave rhythm fixed at a local level and give the user more broad controls over volume, tempo and digital effects. We believe that these higher-level controls are better aligned to the broad gestural affordances (and the aforementioned lack of affordances for virtuosity) of current commercial VR controllers.

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. This notion is explored further in Principle 4.2.

Lens 2: Designing to the medium

Having discussed the modality of audio, let us now turn to the overall experience as afforded by the medium. With any new medium, there is an opportunity to make the most of its new affordances, while borrowing as appropriate from existing media. Design to the medium; use it to accomplish feats that would be impossible otherwise.

Don't port

A common design principle for new media is *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. For example, *VRAPL* could be considered an evolution of block-based programming. The core insight that *VRAPL* builds upon originates from the eleVR research group, who realised that the rigid grid ordering used in tangible and browser-based block-based programming does not need to be ported to block-based sculptural programming in VR (Hart & Eifler, 2017a). Instead, programming blocks can be inserted into sculptures at any angle, allowing the user to consider aesthetics when choosing the structure that makes the most sense for communicating their ideas (Figure 5).

Principle 2.1: Don't Port (Corollary): Make things that would be impossible in the physical world.

If porting does not result in fruitful design decisions, then what is left is to take advantage of the affordances

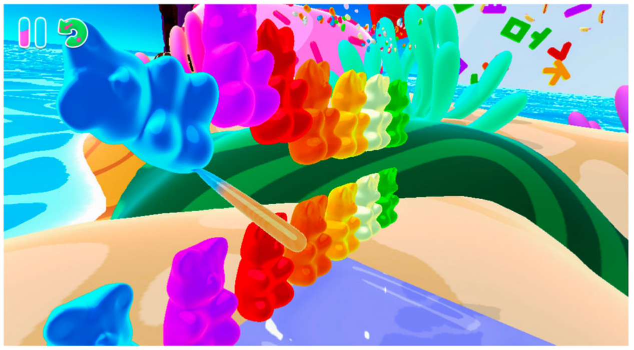


Figure 6. *Playthings* uses springy virtual physics that deviate from our expectations from physical reality.

of the medium to do things that would not be possible without it.[4]

Virtual physics. Virtual physics engines offer opportunities for non-realistic and unconventional uses. Instruments in *Playthings* spring back and forth in place after they are hit; this adds whimsy to the interaction that couldn't easily exist in the physical world (Figure 6).

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. For example, *VRAPL* allows users to change the size of any programming block or virtual world object at any time. This extra dimension of information enables novel ways of controlling programming blocks, such as lowering the frequency of an oscillator by making the block bigger (Figure 7). It also enables whimsical, physically impossible interactions, such as creating a giant bongo bigger than a house that can be thrown around as if it weighs nothing. In the spirit of audio as a first-class modality, size information can be used as an input to audio programs, such as changing the pitch of the bongo when it grows in size.

User perspective. Another oft-overlooked affordance of VR is to change the size of the perspective of the user at will (Hart & Eifler, 2017b). *VRAPL* function blocks use the metaphor of looking inside something to see how it works. Functions are tiny rooms with windows looking in; when a user sticks their head through the window, they are transformed to be a tiny programmer operating in the room as if it were a normal size (Figure 8).

Time. With the use of virtual physics engines also comes the ability to manipulate virtual time. *12 Sentiments for VR* uses this affordance in movement 8. 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. When happening in (virtual) real time, this physics simulation is quick enough to be comical, which doesn't align with the intended emotional aesthetic of 'melancholy; weak' for movement 8. In response, we added a slowing in time that briefly suspends the seedlings at the top of their arc every time the user blows a gust of wind (Figure 9). Informal comments from users who were shown both versions suggested that the use of time-slowing vastly improved the movement's alignment with its intended aesthetic.

Lens 3: Doing vs. Being

The discussion of what is unique to the medium of VR brings us to our formalisation of the concept of *doing vs. being*. The balance between these two states is a central lens in the design of artful virtual reality. We define *doing* as taking action with a purpose; intentionally acting to achieve an intended outcome. By contrast, we define *being* as the manner in which we inhabit the world around us. Users of VR spend their time in either or both of these two modes.

One useful lens for understanding doing vs. being is *Artful Design* Principle 1.5, 'Design is Means vs. Ends' (Wang, 2018). Doing is more closely aligned with means-to-an-end, 'that which serves an external purpose, use, or function', while being is more closely aligned with ends-in-themselves, 'something good in itself, whose value lies primarily in its intrinsic worth'. Being also aligns with a definition of *form*, 'how a thing *is*' (Wang, 2018, Principle 1.6): being is how we exist and fit into the 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.[5]

Doing and being have analogues in interaction and immersion. Doing is aligned with purposeful *action*, taking the form of exploration or control. Being has to do with inhabiting the world in which one is immersed. Doing is outward action that affects the surrounding world, while being arises from taking the time and space

[4] Related principles: Wang (2018, Principle 4.5): 'Design things with a computer that would not be possible without!'; Serafin et al. (2016, Principle 5): 'Consider Both Natural and "Magical" Interaction'.

[5] Related lens: Nussbaum's Capability 5: Emotions (2007).

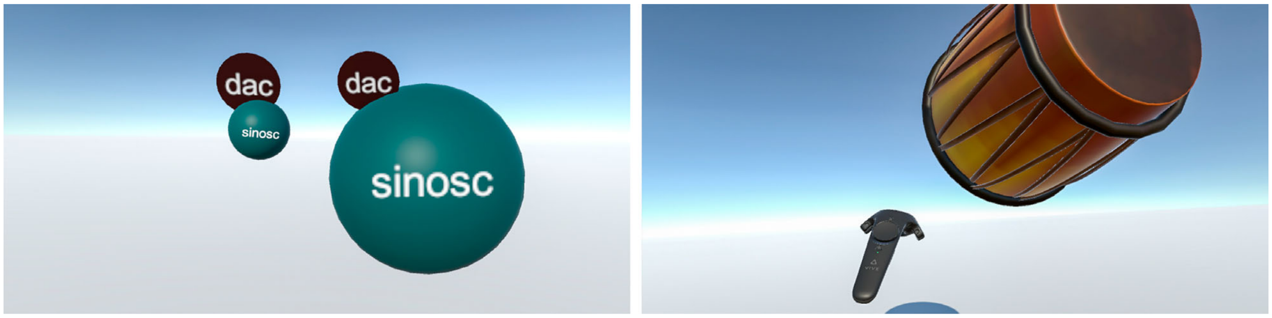


Figure 7. Size can be changed on a whim in *VRAPL*. Users can lower the frequency of a sine wave by making it bigger (left); they can just as easily create a bongo as large as a house and throw it as if it's as light as a feather (right).

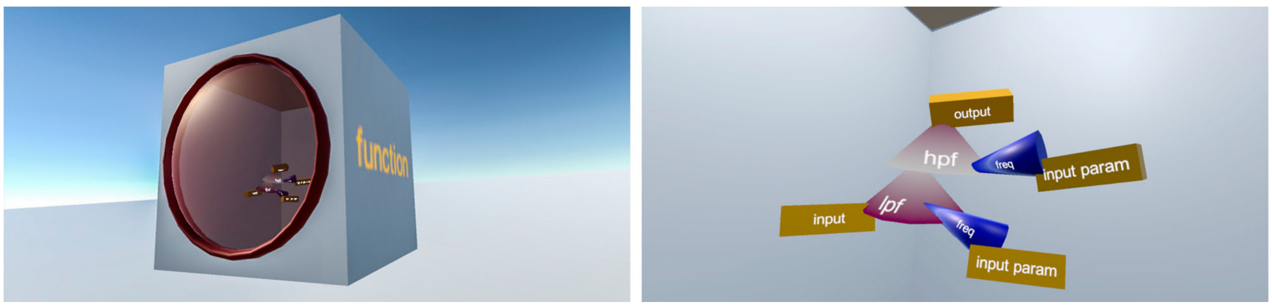


Figure 8. *VRAPL* makes use of varied user perspectives, allowing users to shrink down to modify functions.

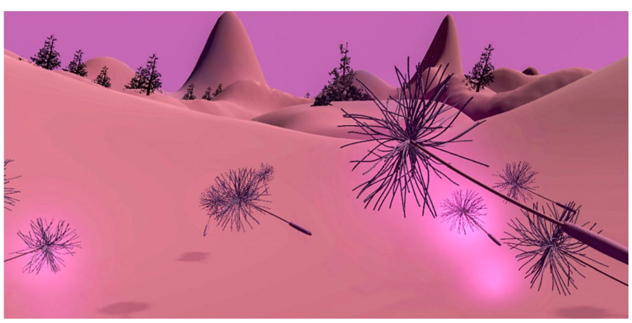


Figure 9. Time-slowing in movement 8 of *12 Sentiments* gives the environment the gravitas necessary for achieving its intended aesthetic: 'melancholy, weak'. Compare a version with time slowing to one without it: https://ccrma.stanford.edu/ ~ lja/dvb/slowing.

for stillness and inward *reflection*. A key difference here is that doing and being are human states, whereas interactivity and immersiveness are properties of a particular experience. Doing and being require interacting or being immersed with *intentionality* and *active engagement* from the user.

Principle 3.1: Design to balance doing (action) and being (reflection).

Doing vs. being is an essential duality in the design of virtual realities. More so than other media, virtual reality requires thoughtful navigation between interaction and immersion, action and reflection, and doing and being in order to achieve a total experience.

Doing and being for narrative-aesthetic coherence. *12 Sentiments for VR* uses the framework of doing vs. being to attempt to overcome the three-way trade-off between interaction, immersion and narrative mentioned by Ryan (1999). 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. The downside of this approach 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.

An experience need not be equally interactive in all its stages, however; situations where the user's agency to affect the world is reduced can afford moments of more intense *being* that also allow a narrative to advance. In movement 5 of *12 Sentiments*, the user's ability to control the movement of the seedlings that are caught in a jetstream is purposefully limited, and in movement 6, they cannot control any aspect of the virtual world with their hands at all. While the user can still affect the music and visuals by turning their head, their agency in these movements is greatly reduced compared to many of the

Figure 10. *12 Sentiments* uses looking up to encourage users to enter the mode of *being*.

other movements. This purposeful reduction in agency aligns with and makes more space for the user to *be* in the movements' intended aesthetics 'unsure' and 'bleak'. The user's limited capability to interact during these movements also allows the narrative to progress through its midpoint climax (the seedlings losing their way, becoming stuck in a rocky valley and being freed by the wind) in an *onscreen*, rather than offscreen, way. This shift in the balance between doing and being thus allows *12 Sentiments* to maintain narrative-aesthetic coherence in these climactic moments.

Music. Doing and being are both essential parts of making music. Making music involves action: playing notes and rhythms, arranging, dancing. Making music also involves being: reflecting on what is made, listening to understand deeply and proactively embodying the music. When designing virtual realities that treat audio as a first-class modality, consider enabling users to make music through both doing and being.

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.

Principle 3.2: Look up! Use gaze to modulate between doing and being.

Something as commonplace as the direction of the user's gaze can be used as an effective mediator between doing and being. For example, one gaze direction that *12 Sentiments for VR* makes extensive use of is looking *up*. In physical reality, looking up is almost always associated with *being*. When we look up, we are admiring the sky, or an impressive mountain, or some grand architecture; or, we are simply daydreaming. We are almost never *doing*, trying to accomplish something or get somewhere. Because of this, each movement of *12 Sentiments* features a sky 'fascinator' with an ever-changing aurora. Additionally, the first, midpoint and final movements require the user to look up to progress through the scene. In the first movement, the user looks up to grow as a small sproutling. In the sixth movement, the user looks up to call upon gusts of wind to free a group of seedlings that are trapped in a barren rocky valley (Figure 10). In the twelfth movement, the user looks up to change their perspective and grow to watch over their valley. Each of these interactions encourages the user to appreciate the environment around them and spend time *being*.

The user's gaze can be used to navigate between modes of doing and being that are linked with more specific emotional aesthetics. 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. The fifth movement of *12 Sentiments* makes use of this metaphor to reinforce its intended aesthetic, 'unsure, mournful, peaceful'. When the user looks forward, they see a cloud of seedlings jumping forward in a gust of wind, happily moving across a landscape, and hear a major chord that reinforces the feeling of untroubled progress. This is a state of doing. However, occasionally a seedling mis-jumps and falls behind; this encourages the user to turn backward and see the seedling tumble behind them. When they do, they see that the seedling is lost to its group forever as it hovers in place where it was lost; they also see the landscape they have left behind in deep shadow; and they hear a minor chord play. Here, there is nothing to do; instead, the experience works to allow the user to inhabit a state of mourning and loss. As the scene progresses, the group of seedlings gradually thins out and

Figure 11. Looking back helps users navigate between a peaceful *doing* and an unsure, mournful *being*.

Figure 12. Very different vibration patterns foster different aesthetics in movements 3 and 4 of *12 Sentiments*.

becomes less confident (the tempo decreases and some notes of their arpeggio are missing). This alternation between feelings of progress at moving forward and feelings of regret at leaving seedlings behind, an alternation between doing and being, produces the intended paradoxical aesthetic of feeling unsure, mournful and peaceful all together (Figure 11).

Other metaphors exist too: for example, looking *around* is linked with confusion and being overwhelmed; looking to the *side* for extended periods of time is linked with boredom (e.g. staring out the window while working). Designers can take advantage of the resulting effects whenever encouraging users to look in a direction other than forward.

Lens 4: Interaction

In our definition from Section 2, a major requirement for an experience to be a reality is that it should be interactive; realities are full-blown worlds that users can not only inhabit but affect meaningfully. A reality's interactions define the 'rules' and constraints of the world, and need to be designed from the beginning of the reality's construction. There are no one-size-fits-all solutions to interaction design; the right interaction is contextual to time, place and individual.

Let us now consider interaction through the lens of doing vs. being. Nuanced interaction allows users to engage in the mode of doing while they take actions but also leaves space for users to engage in being when the conditions are right.

Principle 4.1: Drive interaction design with aesthetics.

Interaction cannot help but come with an experiential dimension. Designing backward from the feeling that the user is intended to have can help align user interactions with the experience's intended aesthetic.[6]

Subtle details in interaction design reinforce aesthetics. Many movements of *12 Sentiments for VR* use the controller's vibration as a haptic feedback mechanism. The strength and pattern of the vibration is carefully chosen to match the movement's aesthetic. For example, movement 3, 'intensity, weight and weightlessness; awe', features a strong vibration with a quick repeating pattern that cuts in and out abruptly, whereas movement 4, 'lightness, flowing', features a weak vibration pattern that fades in slowly from barely noticeable to only slightly noticeable (Figure 12). These subtle details contribute to a kind of aesthetic unity in each of the movements.

Principle 4.2: Multimodality is a virtue.

Humans interact with the physical world in multiple modalities (e.g. sight, sound, touch, breath). It would make sense to take advantage of this breadth in designing interaction in virtual reality.[7] For example, our *Twist*

[6] Related principles: Wang (2018, Principles 4.7, 5.2): 'Aesthetics is not a passive thing, but an active agent of design'; 'There is an aesthetic to interaction'.
[7] Related principles: Wang (2018, Principle 3.1): 'Design sound, graphics, and interaction together'; Serafin et al. (2016, Principle 1): 'Design for Feedback and Mapping'.

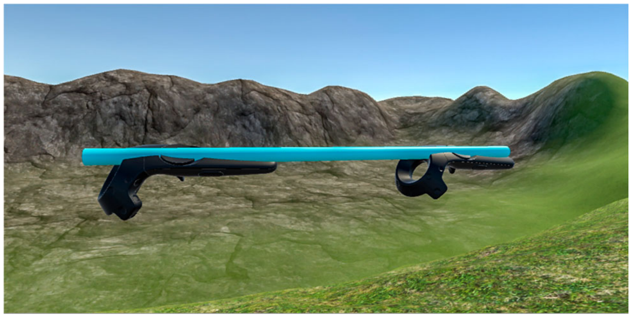


Figure 13. The *Twist Flute* is activated by breath. (See a video: https://ccrma.stanford.edu/ ~ lja/dvb/flute.)

Flute instrument is activated by the user's breath via a microphone located near the user's nose (Figure 13). VR allows us to seek unconventional, even whimsical uses of available technologies beyond the screen, headphones and hand controllers.

Multimodal interaction feedback. Multimodal interaction feedback can create fuller experiences. When users collect sunlight with their leaves in the second movement of *12 Sentiments*, a single chord swells to be louder, the sunlight surrounding the leaf becomes more opaque, and the vibration haptic feedback on the controller becomes stronger. These three modes work together to make the musical gesture of a swell more satisfying and whole, and to make the experience more immersive (Figure 4).

Audio-driven architecture. In audio-driven systems, audio-based timing controls graphics and interaction to achieve a robust control over time that is not afforded by other timing mechanisms such as the system clock (Atherton & Wang, 2018). For example, audio-driven architecture can be used when precise rhythms are necessary. In movements 4, 5, 7 and 8 of *12 Sentiments*, seedlings jump forward or up in a regular rhythm, creating an arpeggio (Figure 14). This timing is generated from the audio; the audio thread tells the visual thread that a note has happened and a seedling should now jump, and not vice versa. This architecture avoids jitter in musical timing that would distract the user and make the experience of the music unsatisfying. Similarly, the 'control object' block in *VRAPL* enables users to control the physics of virtual objects using audio signals (Figure 15). This enables the control of virtual physics and visuals via exacting audio timing. This cross-modal affordance can be used to play a virtual bongo with a regular repeating rhythm, among countless other uses.

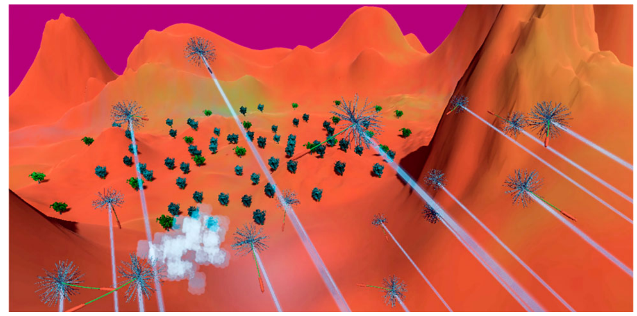


Figure 14. Movement 7 of *12 Sentiments* uses audio-driven timing to control the jumping seedling arpeggio. (See a video: https://ccrma.stanford.edu/ ~ lja/dvb/timing.)

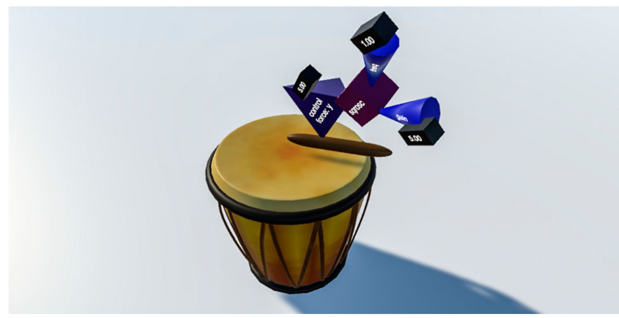


Figure 15. The 'control object' block in *VRAPL* enables users to control the physics of virtual objects using audio signals. (See a video: https://ccrma.stanford.edu/ ~ lja/dvb/physics.)

Principle 4.3: Make space for being alongside doing in interaction.

Interaction is often aligned with doing, but some interactions support being. For example, the second movement of *12 Sentiments for VR* has users learn to collect sunlight by putting their hand, represented as a leaf, into rays of sunlight that slowly fade into existence (Figure 4). This interaction occurs on a slow time scale that enables the user to sink into and appreciate the resulting swell in visuals, sound and vibration; it makes space for the user to enter a mode of being. Making space for being in interaction is one way to achieve a synthesis of doing and being.

Lens 5: Immersion

According to our definition, another major requirement for an experience to be a reality is that it should be immersive; that is, users should have the sensation of being surrounded by a coherent, consistent and *alive* world so that they can believe strongly enough they are in a new reality that they temporarily forget about their physical reality (i.e. feel *presence*). Through the lens of doing vs. being, an immersive reality allows users to *be* in that world, to inhabit it, to feel like a part of it.

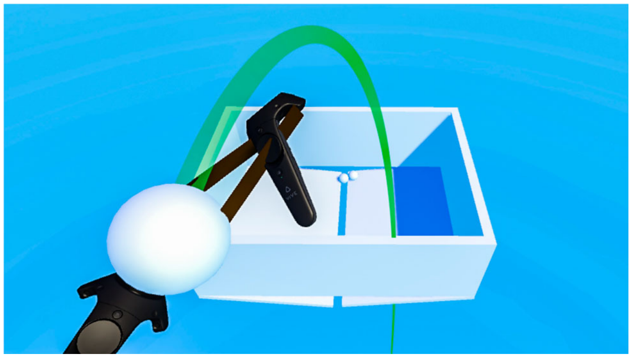


Figure 16. The projectiles of *Wheebox* have animus in the way they react to being thrown ('whee!') and to being stuck (wiggling and chattering). (See a video: https://ccrma.stanford.edu/ ~ lja/dvb/animus.)

Principle 5.1: Create worlds that enhance doing and being through animus.

We once had a conversation with Rebecca Fiebrink where we discussed the notion of *animus*, or the sense that the creatures in a virtual world have their own needs, desires and emotions, and will act accordingly, not always attending to the whims of the user. Because animus is a property borrowed from physical reality, virtual realities with animus can feel much more richly immersive than those without.

Even subtle creature actions can go a long way toward giving a reality a sense of animus. In our design etude *Wheebox*, users can fire spherical projectiles with a slingshot interaction mechanic. The projectiles exclaim 'whee!' when they are fired. Nearby is a large box that the user can fire projectiles into. At first, it seems that the projectiles go exactly where they are fired. As the number of projectiles in the box increases, however, the projectiles start to wiggle a little and chatter nervously amongst themselves (implemented with granular synthesis on the 'whee!' sound file). At some point, the bottom of the box drops out and the projectiles cry out in shock before falling to meet their doom. The projectiles' movements and chatter sounds are subtle, but they suggest that the projectiles are more than simple ammunition (Figure 16).

Animus and interaction. Though it is a property that primarily enhances immersion and *being*, animus without interactivity can lose its effect somewhat if the user ever tries to interact with the creatures. Midway through *Surge*, building blocks start stumbling out of the floor, occasionally building up enough to form walking legs before collapsing back to the floor (Figure 17). This is a moment of captivating animus, but building blocks suggest interaction, and yet the user cannot grab any of the blocks or help or hinder this newly forming creature (i.e. engage in *doing*) in any way.

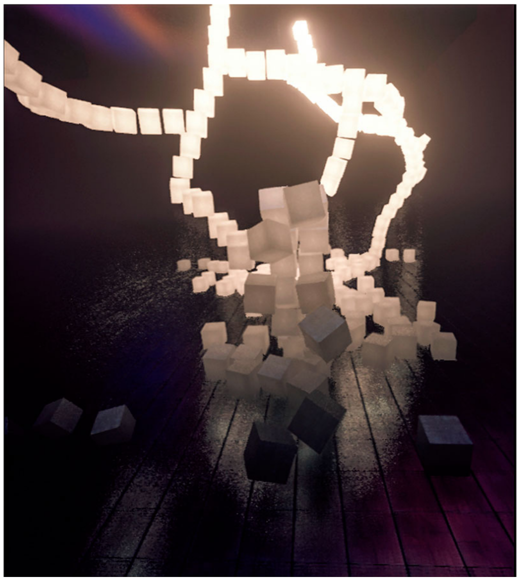


Figure 17. The user cannot interact with the building blocks that stumble up out of the ground to form legs in *Surge*.

Principle 5.2: Balance stylisation and realism.

In computer graphics, something is *stylised* when it is rendered non-realistically with intentional, consistent idiosyncrasies. Stylisation can apply to *form*: for example, a ray of sunlight represented as a rotating prism. Stylisation can also apply to *function*: for example, virtual physics rules that allow a ball to bounce indefinitely. (Note that although the notions of stylisation and realism are often applied to visuals, they are relevant in other modalities as well.)

Stylisation is noticeably different from a failure to achieve realism. Commercial graphics technology tends toward realism; we create graphics cards with more and more resources in order to run algorithms that better approximate the visuals of physical reality (Glassner, 1989). However, VR will probably never fully replicate physical reality; parts of it will always be an approximation (Lanier, 2017). Stylisation helps users suspend their disbelief to achieve presence (Ryan, 1999). These attributes argue for a balance between representing aspects of the virtual world in a realistic way and giving the virtual environment deep richness with judicious stylisation.

The lens of stylisation can also be applied to the modality of audio. Much work is devoted to the development of physically motivated audio spatialisation techniques like HRTFs, which often improve performance in laboratory sound localisation tasks over more stylised techniques like panning (Larsen, Lauritsen, Larsen, Pilgaard, & Madsen, 2013). It is possible, however, that such advanced techniques are not necessary for users of virtual

Figure 18. Stylised vine connections (left) prevent the jarring effect of seeing a disconnected vine (right).

realities to suspend their disbelief; thus, such techniques may not be necessary in cases where rapid accuracy in sound localisation is not critical. Similarly, in the case of physical modelling synthesis, it may be more important for the audio to sound *good* than for it to sound *accurate* to physical reality.

Stylisation, expressiveness, immersion. Stylisation can help create expressive and immersive interactions when available tools and techniques make realism difficult. An attempt at realism that fails the user's preconceptions from physical reality can be anti-immersive, causing the user to notice the limits of the medium and remember that they are having a mediated experience. Stylisation can signal to a user that the world is not attempting realism, so that they appreciate the world's style rather than becoming frustrated or confused at perceived breaks in realism. For example, the second movement of *12 Sentiments for VR* places the user in the body of a vine plant. Two vine offshoots connect to the user's hands, which are represented as leaves. At first, we tried to fully connect the vine to the user's hand, which is difficult because the user can rotate their hand at any angle. We tried a few heuristics for connection that looked good in most positions, but could not possibly look realistic in all orientations. We remedied this by creating a new version of the vine where the end of the vine comes to a point, and the leaf simply hovers nearby (Figure 18). This stylisation avoids the distraction of breaks in realism and often delights users at the magic of having a floating leaf hand.

Aesthetic-driven stylisation. Stylisation contributes to experiential unity when it aligns with an experience's intended aesthetic. In the ninth movement of *12 Sentiments*, the user embodies a section of the earth in which many seedlings are buried; they look to the sky and catch raindrops with the intended aesthetic of 'calm, stasis'. Our original vision for the scene involved explorations in time slowing (as in movement 8; see Principle 2.1), wherein users would reach out and put their hand through large, slow-falling raindrops that would jiggle slightly in response to being touched. Testing this interaction revealed it was far too playful and active for the intended aesthetic of the scene. We replaced this representation of rain with a particle system of many smaller droplets. This representation was a subtler, more realistic depiction of rain, but also ultimately served the scene better.

Lens 6: Designing for the body.

Having elaborated on doing vs. being through interaction and immersion, let us now consider how it relates to three other key aspects of the medium of VR: the body, play and the social dimension.

Principle 6.1: Design for virtual embodiment.

Embodiment is a large part of how the user 'fits into' the world, which they do through doing and being in their virtual body. The user's ability to *be* is enhanced when the representation of their body thematically coheres to the reality in which they are immersed (e.g. through a virtual avatar). When appropriate, represent the user's hands, possibly their head, any other aspects of their body that are tracked, and anything else that would implicitly exist between those points (Figure 19).[8] It may be beneficial to represent the user's body: for example, in movements 1–3 of *12 Sentiments for VR*, the user's hands are leaves or flowers, and a vine is rendered to represent their body and arms, up toward their neck. (It is implied that the user's head is a leaf as well, since they are surrounded by other plants that have leaves for heads.) However, providing a virtual body for the user is not always thematically

[8] Related principle: Serafin et al. (2016, Principle 8): 'Represent the Player's Body'.

Figure 19. In movement 2 of *12 Sentiments*, the user is a plant, and their hands are represented as solid leaves; in movement 4, they are the wind, and their hands are represented as amorphous clouds; and in movement 6, they are the earth, and their hands are not represented.

appropriate. In movements 4, 5, 7 and 8 of *12 Sentiments*, the user embodies the wind. Their hands are rendered as swirling clouds, but the rest of their body is invisible. In movements 6 and 10, the user does not need to take any actions that necessitate putting their hand in a specific place, and they embody the earth itself, so no part of their body is rendered virtually.

Principle 6.2: The body is an implicit medium where being supports doing.

In addition to merely representing the body to facilitate the effect of virtual embodiment, VR allows the designer to treat the body as a medium for aesthetic and artistic effect.[9]

The body is a locus where doing and being feed back on each other. People are *doing* when they take actions, but to do so they rely on properties of the sensation of *being* in their body: spatial awareness, gaze, proprioception and tactility all contribute to our sense of being and support our ability to *do* with our body.

Spatial awareness. Humans have finely tuned capabilities of spatial awareness and spatial memory; we are proficient at remembering familiar routes and the locations of objects we have used in physical reality. These abilities can be transferred to virtual reality. *VRAPL* allows users to surround themselves with programs and unused blocks to 'tinker' with and swap out subprograms. This presents a different paradigm than in text-based programming, where this act of swapping usually requires careful copying and pasting. Furthermore, the user's programs are also recreated every time they enter the experience, allowing them to use their spatial memory to find and navigate through old programs easily, contrasting with the more linear approach of text-based programming.

Gaze direction. As discussed above in Principle 3.2, the direction of the user's gaze can be used to navigate between doing, being and countless other nuanced emotional aesthetics.

Virtual proprioception. It can be empowering to users to be able to predict how their own body will be rendered in the virtual world. Consider a social musical VR experience: dancing alone in *Lune Rouge* is enjoyable, but when multiple avatars inhabit the same space, users may begin to think consciously about how they will be perceived. Since only three points of the user's body are tracked (hands and head), users' nuanced movements in physical reality may not be directly reflected in VR. It is possible that a user who is not confident about how their virtual body is represented would be less likely to want to continue dancing than a user who has a firm understanding of how their movements will translate to their virtual body (Figure 20).

[9] Related principle: Wang (2018, Principle 5.4): 'Bodies matter'.

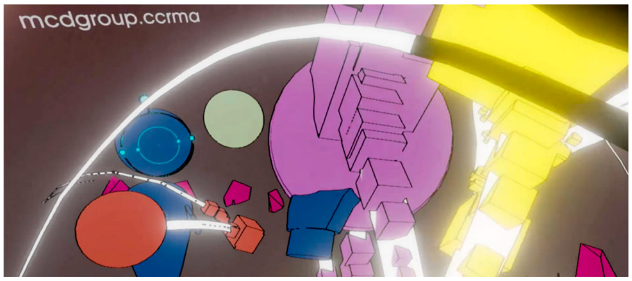


Figure 20. Though it shows rough approximations of other users' bodies, *Lune Rouge* does not give users a clear understanding of their own virtual body's appearance.

Tactility matters. The physical response of playing an instrument is an essential part of developing musical skill on many instruments. The field of VMIs can be a source of inspiration when bringing physical musical interactions to VR with appropriate nuance. Sometimes, it may be helpful to design 'inside-out' by working backwards from the available mechanisms of the underlying system, rather than shoehorning existing physical interactions into these mechanisms (Wang, 2018, Principle 2.2). However, this is not always possible when the design calls for a very specific interaction, such as playing a drum.

Our design etude *Canyon Drum* explored many different interaction techniques for playing a physically impossible drum that is very large and very far away from the user's virtual location (Figure 21). Most used the VR controller's haptic vibration feedback.

- Non-linear arm extension (see Poupyrev, Billinghurst, Weghorst, & Ichikawa, 1996) proved near-impossible to use; we believe it is better suited to grabbing items that are just out of reach at a medium distance.
- Enabling users to hit the drum with a laser pointer, with tactile feedback from the click of the controller's touchpad, afforded timing flexibility and flexibility on where the drum is hit, but no control over dynamics.
- Air drumming (see Dahl, 2014) with vibrational haptic feedback was somewhat satisfying, but had many false positives and negatives and did not allow users to play very quickly due to a necessary debounce time.
- Air drumming to control large mallets suspended above the drum with vibrational haptic feedback was more successful than air drumming due to the addition of visual feedback. However, this technique did not allow flexibility of where the drum is hit and was somewhat unwieldy for playing very fast rhythms.
- Hitting a surrogate surface in front of the user that mimicked the surface of the drum and gave the user vibrational haptic feedback afforded flexibility in timing, dynamics and drumhead position, but was still somewhat unnatural without the more forceful tactile feedback of physically hitting an object in physical reality.

Our conclusion based on all these attempts is that we have not found a satisfying interaction design for the problem of hitting a drum that is very far away; we have moved on from this problem. It may be that musical interactions that rely on hitting something require a tactile response that uses more advanced or specific technology than the buttons and vibration available in the commercial Vive controllers. For example, a controller could be designed that has the power to stop its own movement as if it had hit an object. Until more advanced technology is available, it may be more fruitful to align musical interactions in VR with the level of tactility afforded by the technology currently available.

Principle 6.3: 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. In choosing a movement paradigm, designers are implicitly choosing the aesthetics, affordances and physiological effects that are brought about by that paradigm.

Aesthetic-driven movement design. Thoughtfully designed movement paradigms have the opportunity to contribute to the aesthetic unity of the experience. The fifth movement of *12 Sentiments* features a group of seedlings travelling across a vast landscape. We spent a while designing various movement paradigms for travelling vast distances in the air. One that was more technically successful moved the user's point of view in the same direction as they moved in the real world, but amplified the distance much further. The problem with this movement paradigm was that it made the user feel incredibly powerful, as if they were looking top-down on the world from a god's perspective; this did not align with the intended aesthetic of 'unsure, mournful, peaceful' (Figure 22). We realised that a movement paradigm where the user cannot control their general direction would much more eloquently align with the intended aesthetic. In the final version of the movement paradigm, the user follows the seedlings across the landscape and can do nothing to affect where they go or to prevent the loss of the seedlings that fall behind them. This movement paradigm uses the lens of *being* for intentional aesthetic effect.

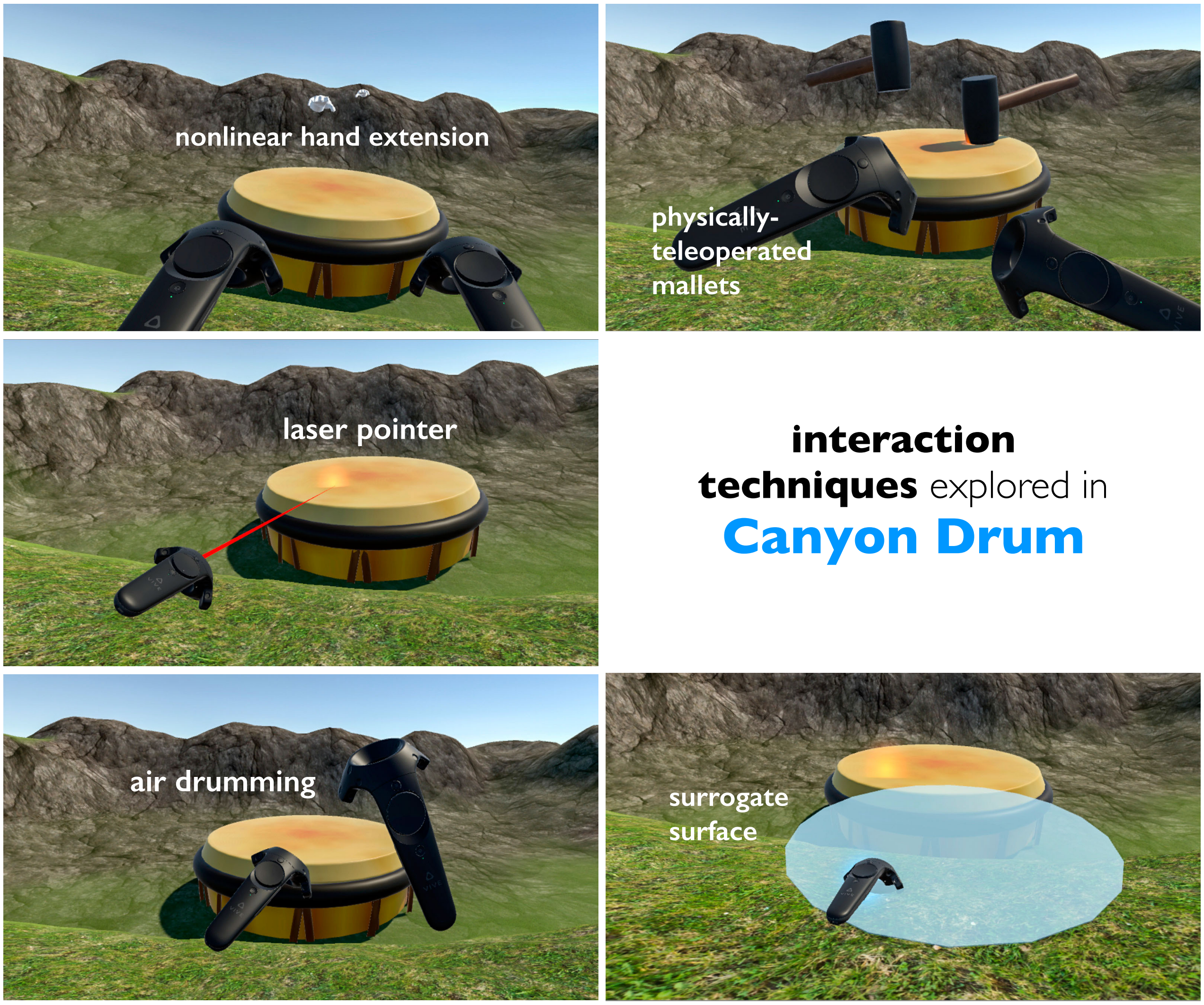


Figure 21. The interaction techniques we explored in *Canyon Drum*.

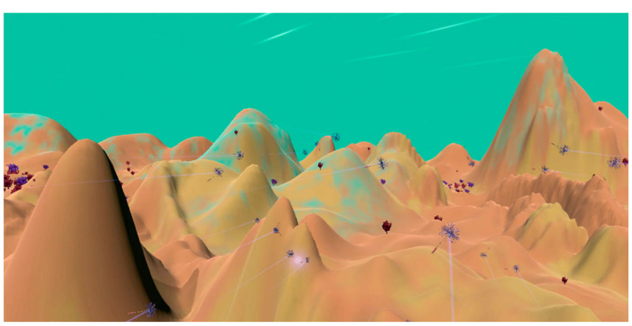


Figure 22. The initial active movement paradigm created for movement 5 did not fit the aesthetic of 'unsure, mournful, peaceful'. The next iteration largely removed users' agency to choose where they move in order to better align with this aesthetic. (See a video: https://ccrma.stanford.edu/ ~ lja/dvb/movement.)

Movement constraints and contextual social practices. Added social constraints on movement may help users feel more comfortable. While teleportation is useful for moving around in *TheWaveVR*, it is possible for a user to accidentally teleport so that they are standing in the same space as another user. In our experience, accidentally 'landing in' another user is jarring as there are many social implications to standing very close to another person in physical reality. Consider limiting exactly where users can move so that their virtual body does not occupy space already used by another virtual object or virtual body.

Teleportation and thematic coherence. Teleportation is a speedy movement paradigm that is used frequently in VR because it is well-aligned to the affordances and constraints of the medium, allowing users to explore a large virtual space from within a smaller physical space. Since this movement paradigm is so widespread, it deserves particular attention here. Consider implementing teleportation when thematically appropriate and avoiding its use otherwise. *TheWaveVR* and *VRAPL* depict futuristic technological environments, and both use teleportation.

12 Sentiments does not use teleportation because a major focus of the experience is allowing users to embody and *be* in various virtual beings that cannot teleport, such as plants and rocks.

Cybersickness. Differences in movement between the virtual world (visuals) and the physical world (user's vestibular system) may make users feel sick (Davis, Nesbitt, & Nalivaiko, 2014).[10] The primary culprit here is acceleration that does not match the user's acceleration. In the middle section of *Surge*, the floor moves a small amount below the user on every beat of the song, accelerating and decelerating quickly until it has moved the distance of one tile away from the user. This lurching can be disorienting. Conversely, *Fantasynth* presents an alternative movement paradigm that avoids cybersickness, moving the user forward at a constant velocity and representing the regular beat of the music with other visual cues.

Lens 7: Designing for play

VR is particularly well-suited for play. As *Artful Design*, Principle 6.9 states, 'Play is free, voluntary, uncertain, unproductive by choice; it occurs in a separate space, isolated and protected from the rest of life' (Wang, 2018). VR is an isolating medium that removes a user from physical reality to a separate, virtual reality. Designed well, virtual reality can fulfil some of the preconditions for play, such as making users feel protected and free and giving them ways to express themselves.[11] In this lens, we examine play as a simultaneous expression of doing and being.

Principle 7.1: Play is both an activity and a state: a synthesis of doing and being.

Play helps people reach a synthesis of doing and being. When someone is playing, they are *doing*, taking actions, but they are also taking those actions to achieve and maintain a state of *being* that is restorative and protected from the rest of life. This is similar to the state of *flow*, where someone feels an energised focus on a task, which requires *doing* to negotiate challenge as well as *being* in a state of engagement (Csikszentmihalyi, 2009); for example, the virtuosic playing of a piece of music.

A synthesis of doing and being requires skilful navigation of the craft of designing for play. This includes nuanced use of multimodality, designing for bodies and movement, designing to the medium and the fashioning of unexpected interactions.

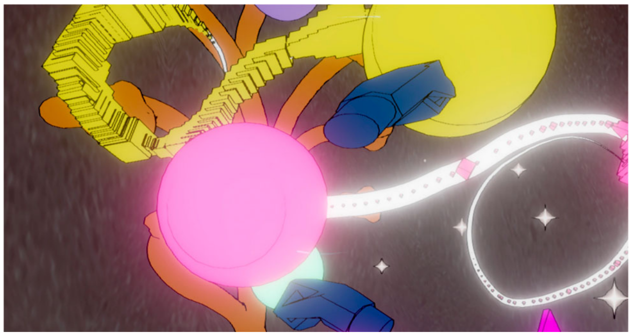


Figure 23. Having to move to generate sound in *Lune Rouge* can encourage users to dance.

Multimodality. Multimodal design for play ensures that signals to engage in play are present throughout the experience. The colours of *Playthings* are bright and saturated; the sound files are cute and whimsical. The instruments are easy to play and made of food, and both the skill floor and the skill ceiling are relatively low. All this communicates cohesively that the music made in Playthings is not so serious, lowering users' inhibitions and encouraging them to play.

Bodies and movement. Users can be encouraged to play through movement. For many people, making music is intricately linked with dancing, and dancing is easier in a less inhibited, playful state. The musical toys that use parts from TOKiMONSTA's music in *Lune Rouge* often require users to manipulate an object that only makes sound while it is being moved. Since the resulting sound has a groove, there is no way to make an unpleasant sound; this lowers users' inhibitions so they feel they can dance. Since the user is already moving in order to produce the sound, it is easy to shift into dancing with the instrument (Figure 23). The user's movement and the generated music together form the cue to start dancing and engage in a playful synthesis of doing and being.

Physical impossibilities. VR's affordances for physical impossibilities can allow users to be silly. In informal laboratory settings, we have observed many users of *VRAPL* derive much satisfaction from scaling up a bongo so that it is the size of a house, then throwing it around as if it were as light as a feather (Figure 7).

Unexpected interactions. Interactions (musical and otherwise) in unexpected places lets users engage in discovery. Our design etude *Shred Head* features two dangling wind chimes with many segments that swing back and

[10] Related principle: Serafin et al. (2016, Principle 3): 'Prevent Cybersickness'.
[11] Related principle: Wang (2018, Principle 6.12): 'Tekhné: Art and Expression as Play'; related lens: Nussbaum's Capability 9: Play (2007); for solo musical experiences, related lens: holicipation (Killick, 2006).

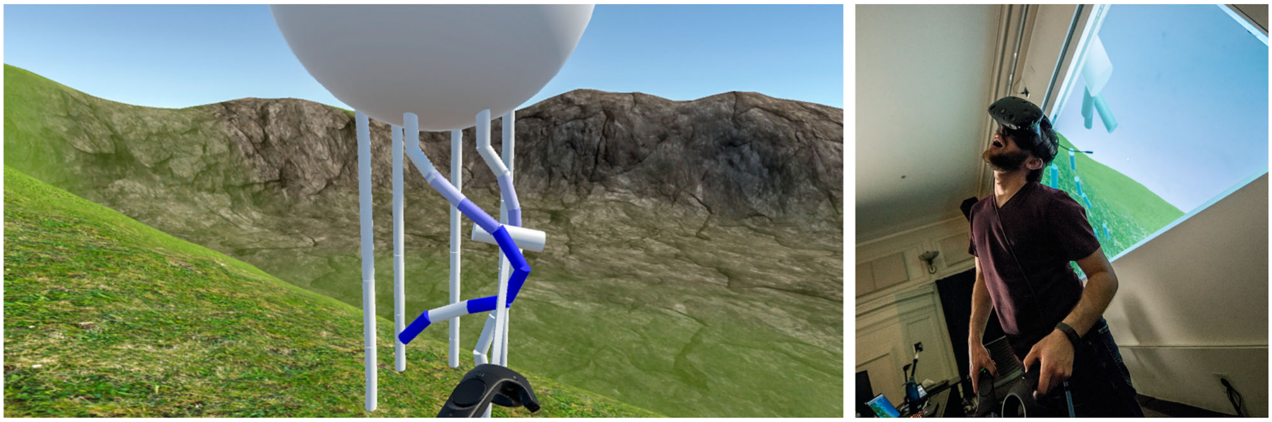


Figure 24. To make sound in *Shred Head*, users can stroke wind chimes or wear them as a wig.

Figure 25. The social dimension of *Aktual* (left) mimics a nightclub in the physical world. By contrast, *Lune Rouge* (right) adds several musical toys that let users remix the album's music, only steps away from the central stage area (visible here through the gap in the far wall).

forth. Users can stroke the wind chimes with a tiny virtual hammer, but one less obvious interaction we found to be more satisfying and whimsical is to place the wind chimes on one's head. In this case, the chimes become a musical wig that is activated by the user swinging their head around (Figure 24).

Lens 8: Designing for social

Since VR is by default an isolating medium, design for social interaction comes much less naturally than design for play. Yet, social interactions are central to human existence. Much of music is making music with other people; designing virtual realities with social interaction in mind can support this end.[12] Some relevant lenses here include immediatism and folk art, which encourage us to be mindful of the duality between design for individuals and groups (Gold, 2007) and to find ways to encourage people to play socially in small yet meaningful communities (Bey, 1994).

Principle 8.1: Replicate baseline social interactions; redesign the rest.

Bailenson's lens suggests that we should replicate a baseline set of social interactions from physical reality in order to create rich, satisfying, believable virtual communities (2018). In accordance with Principle 2.1, however, well-designed virtual realities make specific use of the medium to accomplish feats that would not be possible outside it. *Aktual* has an otherworldly, grayscale, iridescent, creepy aesthetic, which would be difficult to accomplish in physical reality. However, the social dimension of this experience is that a virtual being stands behind an elevated platform with turntables; audience members stand on a dance floor listening. While this experience couldn't quite exist in physical reality, it might have been more compelling if there were something about the social dimension that differed from visiting a nightclub (Figure 25). In contrast, *Lune Rouge* also has a central dance floor that plays back the titular album on virtual speakers, but it adds to this several musical toys that allow small groups of people to make their own music out of earshot of the dance floor.

[12] Related principle: Serafin et al. (2016, Principle 9): 'Make the Experience Social'.

Figure 26. *Lune Rouge* supports music-making with small groups and large crowds.

Principle 8.2: Support many kinds of social engagement.

Different users may have different social needs. Designed experiences can support many of these needs through a diversity of social interactions.

Familiarity. Social engagement can be supported on various levels of familiarity.[13] *The Wave VR* platform features 'waves', which are akin to raves and allow many strangers to come together to experience music collectively. The platform also allows users to host small parties in their home caves, enabling a more intimate experience without strangers.

Scale. Designed experiences can support various levels of social scale. In *Lune Rouge*, all the musical toys require multiple people to fully use. There are two toys with four musical spheres, supporting 2–4 users. One of the toys features three consoles, each supporting at least one person. The final toy, however, features a ring of slowly falling musical objects with two musical spheres in the centre. This layout suggests a single musical performer in the centre, with many onlookers assisting them by playing the occasional sphere sound effect along the outside (Figure 26). This toy is compelling because it uses the affordances of VR to transform a many-person social scenario from physical reality (that of a dance ring in a club) into a magical musical interaction. It also combines users who are *doing* and users who are *being* into the same social, musical context.

Principle 8.3: Design for social doing and social being.

Like play, socialisation is an activity (*doing*) that creates its own space for *being*. Social doing is engaging in shared activities; social being lies in the feelings of togetherness, camaraderie and community that result.

[13] Related lenses: Nussbaum's Capability 7a: Affiliation (2007); Artful Design Model 7.2, 'Rings of Familiarity in Social Design' (Wang, 2018).

Music-making. Making music is often a collaborative process. Playing with other people adds a richness and complexity to the activity beyond playing alone. Engaging in music collectively also naturally supports both doing and being: doing through taking musical actions and being through listening, immersing oneself in the music and preparing for one's next contribution. For example, *Lune Rouge* supports being via listening to music in a group on a dance floor; it also supports doing via actively playing with musical toys. Design for social music-making.

What supports social doing? Social doing is supported by shared activities. Especially in VR, the richest forms of these activities involve the body and embodiment: music-making, dancing, movement in general. Social activities can benefit from structure: predetermined structure, as in goal-oriented games, or on-the-fly structure, as in undirected play. The most basic requirement for social doing is co-location: supporting the existence of multiple people in one space, all with the ability to perceive each other richly with low latency. This property is not yet trivial but will likely become so with the development of new libraries. Social doing requires the ability to coordinate and collaborate; participants that are engaged in an activity that is meant for social doing can benefit from the ability to clearly perceive each other's actions.

What supports social being? Social being is supported by a sense that one *belongs* as part of a group. This feeling can be strengthened by the experience of a shared world with intentional architecture. For example, consider the dance ring toy of *Lune Rouge*: a ring of slow-falling musical objects surrounding a more expressive toy for one or two people encourages the participants to engage and be absorbed in a leader–follower group musical activity (Figure 27). More fundamentally, social being is supported by the experience of a coherent world. People can engage more deeply with each other if they have the sense

Figure 27. The architecture of *Lune Rouge* supports social being.

that they share a fully realised reality: that they are experiencing what the other is experiencing, and that they can affect the world together.

6. Concluding remarks: what makes 'good' design for VR?

How do we design well for VR? What does it mean for VR to be artful? Good design for VR expresses human values. These values can originate from the above design principles as well as other deeply held perspectives. Good design for VR uses the core properties of the medium – isolation, interaction (agency), immersion, presence, embodiment, perspective – in the pursuit of greater human goods. Consider:

Access. Good design for VR promotes access. Use the medium to make things accessible to people who otherwise wouldn't have access. Do this in ways that are at least as nuanced and rich as in physical reality.

Craft. Design to the medium. For example, Artful Design proposes a 'sanity check': 'Does the end product justify the technology? Does it do at least one thing that can be achieved by no other means? Does the design use the medium to support the right interplay between technology and humans?' (Wang, 2018, p. 252).

Total systems. Speak to all our senses with equal consideration. Find balance between audio, visuals and interaction. Mix realistic and fantastical elements. Find balance between interaction and immersion in service of coherent worlds. Support play for individuals and groups. Create total experiences that wholly transport the user to another reality.

Doing and being. Achieve a radical synthesis of doing and being. Give people the means to reflect as well as the ability to take intentional action. Use doing-and-being as a tool for thinking about embodiment; well-embodied interaction that speaks to all our senses feeds back on our sense of presence, on our ability to exist intentionally in a space. Through the artful combination of the modes of doing and being, we can unlock a greater understanding of the medium of virtual reality, and perhaps even of our experience outside of it as humans in a world.

Human flourishing. Support expression through art, emotional development, social connection through community and play. Encourage holicipation and folk art. Provide opportunities for living rich, full lives.

Artful VR achieves balance between these aims, making good use of the medium to craft total experiences that support human flourishing for all.

The principles presented here are not meant to be exhaustive. They leave room for future practice to explore new ways of thinking, new principles, different values; new insights with which to consider both the craft and the meaning of design.

We see a few open questions for the medium. The exploration of isolation and connection: what is the nature of connection in a medium designed to isolate, and how can we achieve the richness of connection that is already possible in physical reality? In music: what is the nature of musical performance in VR? What new musical interactions, instruments, performance practices are possible? How do we find and support the right interplay between this technology and humanity? The future is exciting as we begin to answer for ourselves what are the philosophical, artistic and humanist implications of virtual reality.

Acknowledgements

Thanks to the three anonymous reviewers for their thoughtful feedback. Thanks to Rebecca Fiebrink for her conversations on the concept of animus. Thanks to Nicola Dibben for introducing us to *Why Music Matters* and *holicipation*, and for playing social VR experiences critically with us. Thanks to the members of the CCRMA VR/AR Design Lab – Kunwoo Kim, Charles Foster, and Marina Cottrell – for their valuable critiques. Thanks to Anıl Çamcı and Rob Hamilton for organising the Audio-first VR Workshop at NIME 2018.

Funding

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1656518.

ORCID

Jack Atherton http://orcid.org/0000-0002-6452-5020
Ge Wang http://orcid.org/0000-0002-5170-5460

References

Atherton J., & Wang G. (2018). Chunity: Integrated audiovisual programming in unity. *Proceedings of the 2018 conference on new interfaces for musical expression*. Blacksburg, VA.

Bailenson, J. (2018). *Experience on demand: What virtual reality is, how it works, and what it can do*. New York, NY: W. W. Norton & Company.

Benford S., Bowers J., Fahlén L. E., Greenhalgh C., & Snowdon D. (1997). Embodiments, avatars, clones and agents for multi-user, multi-sensory virtual worlds. *Multimedia Systems*, 5(2), 93–104.

Berdahl E., & Huber D. (2015). The haptic hand. *Proceedings of the 2015 conference on new interfaces for musical expression*. Baton Rouge, LA.

Bey, H. (1994). *Immediatism*. Edinburgh: AK Press.

Brower G. M. (2016). *Playthings: VR music vacation*. Always & Forever Computer Entertainment.

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

Cook, P. R. (2002). *Real sound synthesis for interactive applications*. Boca Raton, FL: AK Peters/CRC Press.

Csikszentmihalyi, M. (2009). *Flow: The psychology of optimal experience*. New York, NY: Harper Collins.

Dahl L. (2014). Triggering sounds from discrete air gestures: What movement feature has the best timing? *Proceedings of the 2014 conference on new interfaces for musical expression*. London, UK.

Davis S., Nesbitt K., & Nalivaiko E. (2014). A systematic review of cybersickness. *Proceedings of the 2014 conference on interactive entertainment* (pp. 1–9). Newcastle, NSW, Australia.

Dexta Robotics (2019). *Dexta robotics – touch the untouchable*. Retrieved from https://www.dextarobotics.com/en-us

Glassner, A. S. (1989). *An introduction to ray tracing*. San Diego, CA: Academic Press.

Gold, R. (2007). *The plenitude: Creativity, innovation, and making stuff*. Cambridge, MA: MIT Press.

Groom V., Bailenson J., & Nass C. (2009). The influence of racial embodiment on racial bias in immersive virtual environments. *Social Influence*, 4(3), 231–248. doi:10.1080/15534510802643750

Hamilton R. (2009). Building interactive networked musical environments using q3osc. *Audio engineering society conference: 35th international conference: Audio for games*. London, UK.

Hart V., & Eifler M. (2017a). *Expressive programming in VR | elevr*. Retrieved from http://elevr.com/expressive-programming-in-vr/

Hart V., & Eifler M. (2017b). *The three scales | elevr*. Retrieved from http://elevr.com/the-three-scales/

HelloEnjoy & N'To (2017). *Fantasynth: Chez Nous*. HelloEnjoy.

Hesmondhalgh, D. (2013). *Why music matters*. Chichester: John Wiley & Sons.

Killick A. (2006). Holicipation: Prolegomenon to an ethnography of solitary music-making. *Ethnomusicology Forum*, 15(2), 273–299.

Koosha A., & Strangeloop Sab H. (2017). *Aktual*. TheWaveVR.

Lanier, J. (2017). *Dawn of the new everything: A journey through virtual reality*. New York, NY: Henry Holt & Co.

Larsen C. H., Lauritsen D. S., Larsen J. J., Pilgaard M., & Madsen J. B. (2013). Differences in human audio localization performance between a HRTF- and a Non-HRTF audio system. *Proceedings of the 8th audio mostly conference*. Piteå, Sweden.

Lee J. & Strangeloop (2017). *The Lune Rouge experience*. TheWaveVR.

Mulder A. (1994). Virtual musical instruments: Accessing the sound synthesis universe as a performer. *Proceedings of the First Brazilian symposium on computer music*. Belo Horizonte, MG, Brazil.

Mulder A. (1998). *Design of virtual three-dimensional instruments for sound control* (Unpublished doctoral dissertation). Simon Fraser University, BC, Canada.

Mulder A., Fels S., & Mase K. (1999). Design of virtual 3D instruments for musical interaction. *Graphics Interface*, 99, 76–83.

Nussbaum M. (2007). Human rights and human capabilities twentieth anniversary reflections. *Harvard Human Rights Journal*, 20, 21–24.

Pimentel, K., & Teixeira, K. (1993). *Virtual reality: Through the new looking glass*. New York, NY: Intel/Windcrest.

Poupyrev I., Billinghurst M., Weghorst S., & Ichikawa T. (1996). The Go-go interaction technique: Non-linear mapping for direct manipulation in VR. *Proceedings of the 9th annual ACM symposium on user interface software and technology*. Seattle, WA.

Ryan M. L. (1999). Immersion vs. interactivity: Virtual reality and literary theory. *SubStance*, 28(2), 110–137.

Schultze U. (2010). Embodiment and presence in virtual worlds: A review. *Journal of Information Technology*, 25(4), 434–449.

Sen, A. (1999). *Commodities and capabilities*. Oxford: Oxford University Press.

Serafin S., Erkut C., Kojs J., Nilsson N., & Nordahl R. (2016). Virtual reality musical instruments: State of the art, design principles, and future directions. *Computer Music Journal*, 40(3), 22–40.

Slater, M., & Sanchez-Vives, M. V. (2016). Enhancing our lives with immersive virtual reality. *Frontiers in Robotics and AI*, *3*, 1–47.

Steuer J. (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of Communication*, 42(4), 73–93.

TheWaveVR (2017). TheWaveVR.

van Meerten A. (2016). *Surge*. House of Secrets.

Wang, G. (2018). *Artful design: Technology in search of the sublime*. Stanford, CA: Stanford University Press.