Temporality Across Three Media: Inner Transmissions

Julia J. Mills
CCRMA, Stanford University
jjmills@ccrma.stanford.edu

ABSTRACT

With time inextricable from music, as well as theatre and film, creating effective audiovisual works in digital media relies on an understanding of the medium's effect on the work's temporality. There is a need to analyze the interaction between virtual environments and time perception within a musical context. This paper discusses the sound installation *Inner Transmissions* as a case study for examining this issue. The work compared three media: physical space using radio transmission, web-based 360° photo and video, and headset virtual reality (VR). The core concept of the installation remained the same across formats, designed to allow the issue of temporality to be addressed explicitly. Listener feedback suggests expectations, physical interactivity, and environmental movement are important factors in determining how temporality varies between media, posing as affordances for some and limitations for others. The outcomes of this case study provide an example of how temporality functions in musical works created for virtual environments, serving as a starting point for future research in this area.

1. INTRODUCTION

Music is inherently a temporal art form. As VR emerges as a medium for audiovisual works and sound installations specifically, a clear understanding of its effect on temporality is needed. Temporality in this paper refers to both a listener's perception of time while experiencing a work and the consequences for how they then engage with it: passively or actively, and for how long. An artist's intentions for this aspect of the listener experience can often be distorted by their choice of medium, with some posing many more obstacles to artistic realization than others. While some research has been done on time perception in VR environments, no work has yet analyzed the role of sound in the users' temporal distortions, particularly not in an artistic context.

Grayson Mullen and Nicolas Davidenko (2021) recently documented the effect of time compression[1] when comparing conventional monitor (CM) and VR head-mounted displays (HMDs) [1]. However, this study, as with others investigating the relationship between

cognitive load and time perception in VR [2, 3], focused on gaming contexts. There are no studies that look at the combined perceptual, and specifically artistic, effects of music within VR environments. Music in CM videogames has been shown to affect immersion [4], but the link to VR is missing and, moreover, would not approach the issue from a compositional perspective.

Serving as a case study, comparing the same work across media allows us to analyze the affordances and limitations of how each deals with temporality. This work, *Inner Transmissions*,[2] was created as a physical installation at Stanford University's Papua New Guinea Sculpture Garden, a web-based installation using 360° photo and video, and an installation for VR headset.[3] While the implementations of each installation were quite different, the core content remained consistent across formats. Focusing on the goal of eliciting the same overall feelings from listeners allowed the singular work to be compared across media. Using the physical installation as a baseline, the web-based and VR installations offer opportunities to isolate how temporality functions in digital, musical contexts. There are myriad other implications from the differences in user interfaces and experiences, but narrowing in on how temporality was specifically affected provides windows into many secondary aspects of sound installation composition. The goal of this paper is to highlight the effects to temporality from changes in medium with the same overall listener feelings from each installation.

2. INNER TRANSMISSIONS

Spatialized around each installation space, six compositions were paired with six sculptures in the Papua New Guinea Sculpture Garden. Each corresponded to a different imagined universe: "A Melancholic World" paired with *The Thinker (Yerakdu)*; "A Many-Voiced World" with *The Gates of Hell/Opawe & Namawe*; "The World of Rising Suns" with *Kwoma Spirit House Posts*; "A Hypnagogic World" with *Kawatukit*; "The World of Power" with *Kura*; and "An Angelic World" with *Wanmi and Saw Hokwa* (see Fig. 1). As they moved away from one sculpture and towards another, a visitor would hear radio static gradually overwhelm the composition they were just listening to, and the new sculpture's piece begin to fade in.

Both the compositions and their method of transmission in each medium focused on radio broadcasts as the core

---

[1] Meaning a larger real duration is perceived as a shorter experience.

[2] Virtual installations at https://ccrma.stanford.edu/~jjmills/VFT/

[3] The terms *web-based* and *VR* will be used rather than CM and HMD.

Figure 1. Featured Papua New Guinea Sculpture Garden sculptures; from left to right: *The Thinker (Yerakdu)* by Simon Gambulo Marmos; *The Gates of Hell/Opawe & Namawe* by Marmos and Jo Mare Wakundi; *Kawatukit* by Marmos and Wakundi; *Wanmi and Saw Hokwa* by Naui Saunambui, Yati Latai, Membor Apokiom, et al.; *Kwoma Spirit House Posts* by Apokiom and David Kapa Kaipuk; and *Kura* by Marmos and Wakundi.

aesthetic. The compositions combined air traffic control[4] and Papua New Guinean radio[5] samples with instrumental loops using guitar, violin, kalimba, keyboard, and the programming language ChucK. The samples grounded the compositions in real geographic locations, while their distortion and scattering among the instrumental loops then displaced them from their original place and time.

The pieces were all infinitely looping: listeners could start and stop listening at any point and get a representative sample of the full composition. This further removed any linearity the samples had in their source materials and reinforced the sense that listeners were tuning into forever ongoing transmissions from unseen worlds. Able to freely explore the installation site, the looping of the compositions also perpetuated the sense that the listener is an explorer, peeking fleetingly into worlds that exist on a very different timeframe than themselves. The feeling of searching for something amidst the radio receiver static was mirrored visually by the use of flashlights; the search for a sound source was in turn a search for a sculpture in the dark. This curated sense of agency, striking a tension between the audience's playful exploration and more contemplative listening practice. Each installation visitor was given headphones, so that their experience would remain isolated and personal.

In each installation, listeners were instructed to explore independently. Their location and orientation in the space,

as well as their timing when tuning in to different compositions, changed the final listening experience, which allowed each audience member to curate their own relationship to the work.

2.1 Premises

This core framework set up three main premises that allowed for the explicit handling of temporality: the goal of promoting stillness in listeners, the concept of tuning into parallel worlds, and the emphasis on self-guided exploration of the space. These concepts correspondingly oriented the artist, the listener, and the interaction between them towards temporality as the focus of the installations and their comparison.

2.1.1 Artist Intention

The concept of "promoting stillness" in listeners was the cornerstone of the installation design process. Centering this as the primary goal of each installation laid the foundation for all proceeding design decisions.

2.1.2 Communication with Listeners

To more organically observe how listeners spent their time in the installations, the goal of promoting stillness was not shared, with the exception of in a later focus group. Instead, listeners were primed to think about temporality through the idea that sculptures contained different imagined universes. The program notes stated, "What does it sound like to listen to an entire world? Within each sculpture is another universe, one maybe very similar to our own or maybe drastically different, governed by unfamiliar

---

[4] UNNT (Novosibirsk, Russia), VMMC (Taipa, China), VTBS (Bangkok, Thailand), WBKK (Kota Kinabalu, Malaysia), WBKS (Sandakan, Malaysia), WMKB (Penang, Malaysia), YBCS (Cairns, Australia).

[5] Radio Maria, NBC News PNG, and Nau FM.

Figure 2. Radio transmitters (left) and receivers (right).

forces and energies. *Inner Transmissions* invites listeners into the galaxies hidden within the Papua New Guinea Sculpture Garden." The idea that listeners are tapping into unseen worlds implies there may be multiple planes of time at work. Framing the installation experience in this way opened up listeners to allowing the work to alter their time perception and challenge their sense of pacing in the space.

2.1.3 Listener Agency

To get organic feedback from listeners regarding their experience of each installation's temporality, it was important to allow them to structure their own installation experiences. Encouraged to explore independently, listeners were able to individually determine how they wanted to spend their time in each format of the installation. The compositions had no clear start or end, and so each listener could act on their personal sense of the work's temporality. Whenever they felt they had heard enough, they were free to move elsewhere in the space. Though curated by the artist, the final listening experience is therefore largely up to each visitor and their perception of how time is passing during their experience, yielding more genuine feedback.

3. IMPLEMENTATION

3.1 Physical

The physical installation was implemented using local radio transmission. Multiple radio transmitters, each looping over a different composition, were placed near various sculptures and set to broadcast on the same frequency band. This turned radio receivers into rudimentary proximity sensors; listeners heard whatever transmitter they were closest to. As they explored the space, transitions between compositions could be noisy (pure radio static) or gradual (latching onto multiple transmitters at once and layering their sounds). Since the installation was designed to be experienced at night, flashlights were attached to the radio receivers, linking the search for sound and the search for sculptures (see Fig. 2). Upon arriving to the sculpture garden, visitors were each given a flashlight-receiver, headphones, and a map with program notes.

Figure 3. The view from the web-based installation node for *The Thinker*; the red icon and cone on the embedded map indicate the listener's position and orientation, respectively.

3.2 Web-based

The web-based format of the installation was created using Pano2VR, a program for exporting 360° photo and video projects. Each photo or video was used to create a "node" in the project, an HTML page that allows the visitor to click and drag to explore the image. Videos of each transmitting sculpture were captured, as well as photos of various transition points between them and the entrance to the installation. A flashlight was shone either on the transmitting sculpture or in one of two directions when at a transition point, creating separate "to" and "from" nodes in the installation. An embedded map, the same image as listeners were given in the physical installation, indicated where in the installation the visitor was and what direction they were facing (see Fig. 3). To travel around the site, listeners would click one of the icons positioned around the image and be transported to the new location selected. Audio was spatialized for each node, changing in each ear depending on the direction faced, similar to how the radio receivers' antennae would spatialize sound in the physical space. To authentically incorporate the transmission concept and radio aesthetic to the sound, radio receiver output at each node was recorded, as well as some pure radio static to fill in any gaps in the spatialized audio. Sound sources used Pano2VR's "directional circular" spatialization mode, which diffuses sound as a circular field of a changeable size. In both composition and transition nodes, "filler" static sound sources were placed such that the aggregate sound level was approximately constant as the listener rotated around the space.

3.3 Headset VR

The VR format of the installation was created using Unity and developed for the Oculus Quest 2. The physical site map was used as the basis for the worldbuilding, with trees and non-transmitting sculptures procedurally placed, scaled, and rotated a certain distance from pathways. Clouds were animated to constantly swirl above the
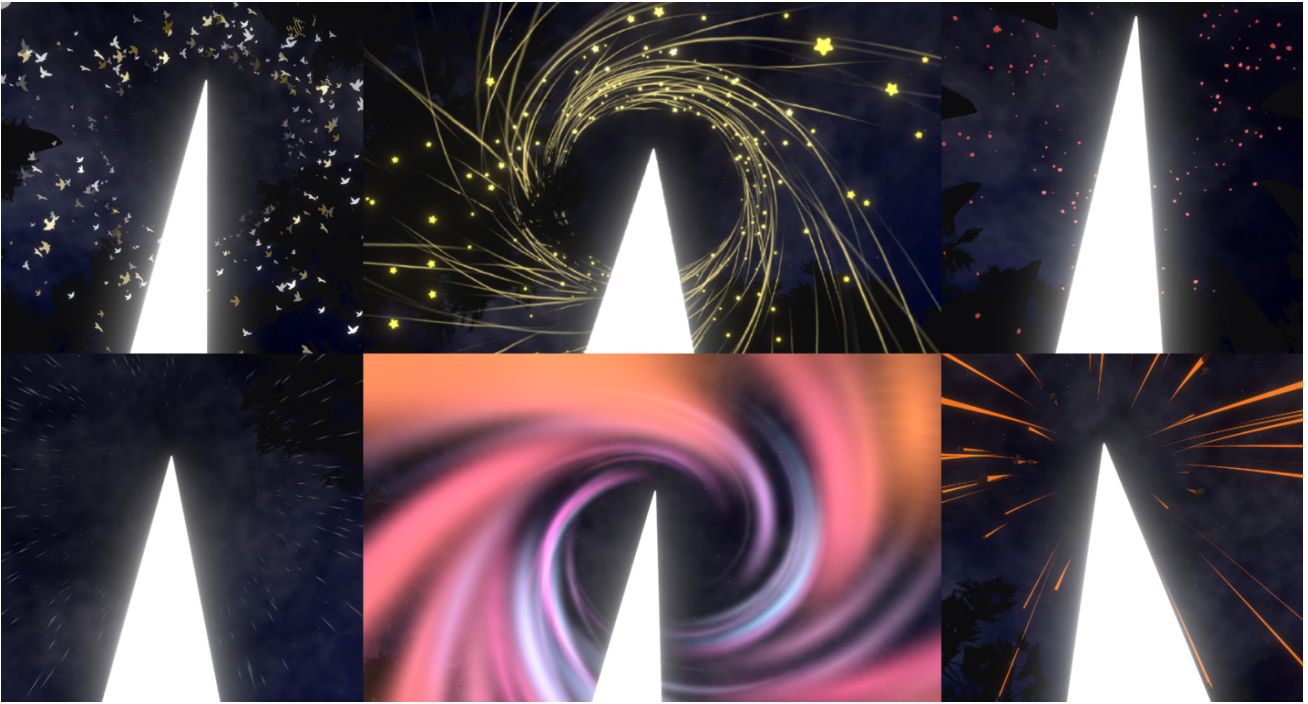
Figure 4. Animations corresponding to each transmitting sculpture; clockwise from top left: "A Many-Voiced World," "A Hypnagogic World," "An Angelic World," "The World of Power," "The World of Rising Suns," "A Melancholic World."

sculpture garden. In part to help with navigability in a largely dark scene, the transmitting sculptures were represented by enormous pillars of light reaching towards the sky. Within the headset, users had modelled hands, one of which held a flashlight. The only controls were to move and rotate, each controlled by a joystick.

Audio was spatialized similarly to the web-based format, in which a mixture of the compositions recorded through radio receivers and radio static was used to build up the soundscape. Both ran on infinite loops. Each composition was located on its pillar's base, while the 24 static sound sources were invisibly placed throughout the scene at ground level. Unity allows audio sources to have individual spatialization curves, so each sound source had a custom loudness vs. distance plot tuned to blend more seamlessly between sources. The sound sources in the scene all used Unity's "3D spatial blend" setting, with the exception of a motion-activated recording of footsteps on dried leaves. The footstep sound was presented as a 2D audio source projecting from the listener headset position.

An additional mechanism was used to reintroduce the thematic content of each composition and further promote stillness in the installation. When standing near a transmitting sculpture, over the course of 30 seconds all other light pillars blinked out, all surrounding foliage and sculptures faded to black, and all radio static faded out. Once the transition was complete, the "inner world" of the sculpture began swirling around the listener and the remaining pillar. The animations corresponded to the compositions as follows: rain for "A Melancholic World," a flock of birds for "A Many-Voiced World," flares of light for "The World of Rising Suns," falling stars for "A Hypnagogic World," meteors for "The World of Power," and cherry blossom

petals for "An Angelic World" (see Fig. 4). Upon the listener moving in any direction, the scene snapped back to how it originally was, and they were once more surrounded by the full sculpture garden with all its radio static.

4. RECEPTION

Installation participants consisted of mostly Stanford students and faculty, with some non-university affiliated visitors from nearby areas. Many of the physical installation visitors had personal ties to the sculpture garden and were present for the Papua New Guinea Sculpture Garden's opening in 1994. All quotes are from a focus group held once after viewing the physical installation and again after viewing all three installations. Other observations of installation visitor behavior will include a mix of those who viewed one, two, or all of the installation formats.

Table 1. Feedback Dimensions

	Physical	Web-based	VR
Expectations	+	-	±
Physical interactivity	+	-	±
Environmental movement	+	±	+

In both visitors' behavior and follow-up conversations, three major themes emerged as dimensions of the installation experience most affected by medium: expectations, physical interactivity, and balanced pacing. For each of these facets, listener feedback indicated to what extent it was an affordance or limitation of the medium. These feedback trends are summarized in Table 1, where "+" indicates a perceived affordance, "-" indicates a perceived

limitation, and "±" indicates a grey area or lack of specific feedback. The following sections provide some examples of listener comments that point to these overall outcomes, as well as their relationship to temporality in each medium.

4.1 Expectations

Visitor expectations surrounding temporality varied widely between media and were frequently mentioned in listener feedback. The physical installation was often compared to visiting a museum, an analogous experience of sitting quietly in front of a piece of visual art. As one visitor said, "I sat down for moments and just thought about things ostensibly unrelated. My favorite thing in museums or installations is when you connect it to whatever you happen to think about. It's a space to do that." The physical installation experience met visitors' temporal expectations, and in this way, expectations can be seen as a potential affordance of dealing with temporality in physical media.

Before viewing the web-based and VR installations, most visitors struggled to imagine or were pessimistic about experiencing them. "It was such a complete experience for me that replicating it would not work. You'd have to do something different… Having the same version in VR, I don't think I'll get the complete experience." There was an expectation that the digital formats would be quite representative, seeking to directly capture what the physical installation experience was like. When reframed as a separate installation experience, one listener said, "[Making] this in VR would be really hard, if not impossible, but it made me wonder what might be achieved in VR that's not worse or any better."

Overall, temporal expectations of the web-based version were quite low, and seeing this medium as a form of documentation rather than a complete experience was a common perspective. One listener described it as "the best you can do without being there or being in VR." Another said, "I think we have to use it knowing that it can't be an experiential substitute because it is so different." Despite the goal of promoting stillness in listeners, it was the installation people spent the least time exploring. When the web-based and VR installations were shown in conjunction, the former often received only a few exploratory clicks before the listener moved on. When asked to explain the short temporal runway the web-based installation had, one visitor phrased it as, "There is no reason to linger in that space – it's the ultimate 'I get it.'" Another commented, "In a way, it's almost expected – this *should* be way more visceral and powerful in person." Visitors did not expect a full experience in the web-based installation but rather a substitute or stand-in for the "real thing." As an independent installation medium, then, this case study indicates that listeners' expectations surrounding temporality may place a strong limitation on this format.

Listeners' temporal expectations in VR came up more with respect to the transition into the "inner world" animations than to the experience as a whole. Though only 30 seconds, people reported feeling that the transition was much longer and frequently questioned if they were following the instructions correctly to trigger it. For comparison, visitors frequently sat in front of sculptures in the physical installation for over five minutes. Listener surprise at the transition length suggests that the time compression documented in VR environments may be limited to gaming settings, or at least function very differently in compositional contexts. As it stands, listener feedback from this case study did not clearly indicate expectations as either an affordance or limitation of this medium.

4.2 Physical interactivity

Comparably prominent in listener feedback was the dimension of physical interactivity in the experience. The relative presence or absence of this aspect largely correlated with the overall perceived effectiveness of each installation medium. In summary, more physical interactivity seemed to act as a motivator in listeners exploring the space, which in turn heightened their sense of temporality in the work.

In the physical installation, interactivity was often mentioned as fundamental to the experience, and visitor behavior was observed as matching this feedback. Listeners frequently stopped to rotate the radio receivers and listen to the effect as they walked around the site. One visitor commented, "I thought I was the only one playing around with it like an antenna, but it seems like a lot of people did. When I arrived [in front of a transmitting sculpture] and the noise cleared up, it was really satisfying." The sense of reward appeared to motivate people to explore the space, and some visitors returned to the physical installation a second time in order to further experiment with the receivers.

Listeners also frequently connected the physicality of the receivers to a sense of independence and play in their exploration of the work. One focus group member said, "A lot of us responded to the idea of a physical device in your hands that was like a divining rod of sound… It feels kind of playful. I'm not completely deterministically going through the motions. I have some agency." This was in contrast with this specific listener's description of the web-based installation as "deterministic." With a core premise of this installation being self-guided exploration, this feedback indicates an effective implementation thanks to the physical interactivity of the installation medium.

The multisensory nature of physical installations also fed into this feedback dimension. When imagining what the virtual experiences would be like, multiple listeners described a list of elements they anticipated would be missing: "For me, the breeze and air, the smell of the nature, and the leaves that I stepped on – they were such good parts of the experience that I feel like if they were absent, [the experience] won't be complete." These aspects of the physical installation were natural affordances of working outdoors, given to the composer by default.

Their absence was a point of negative feedback for the web-based installation, as listeners had anticipated. One said, "I think the physical installation is so much more

dimensional: from the holding of the receiver to the walking, to the leaves, to the breeze, to the sunset, to the tuning in." It is notable that the core mechanism of the physical installation ("the tuning in") was mentioned last in this listener's list. It hints that perhaps the specifics of the work are almost secondary to the overall experience of being in the installation site. Indeed, the correlation between complexity of sensory input and profundity of response in digital experiences has been previously proposed [5]. From a temporal perspective, it seems reasonable to view this increased enjoyment as deepening listeners' connection to the temporality of the work. Exact mechanism aside, the physical interactivity and multisensory elements of the physical installation were clearly perceived as affordances of the medium.

The web-based installation was the least physically interactive format, and as might be expected, visitors reported this as detrimental to their overall experience. Though the spatialization of audio mimicked the responsiveness of the other installation formats, interaction was limited to clicking and dragging using a mouse. One visitor said, "There's not a lot you can do. I cannot go up to a statue and really look at it and see all the grains in the wood." This inflexibility may have been a contributor to the feeling mentioned earlier by visitors that there was "no reason to linger" in this medium.

The listening experience of the compositions themselves was also altered by the lack of physical interactivity. Comparing this installation to the physical version, one visitor said, "The audio feels more disconnected because I'm looking at a photo while listening." When asked if the installation concept was apparent, they said, "I sense that [the music] should have something to do with what I'm looking at," but that the audiovisual connection was otherwise unclear. In contrast, they perceived the radio receivers themselves as responsible for making this connection clear in the physical installation. Thus, the lack of physicality, correlating with a lack of motivation to explore and a lack of conceptual cohesion, was shown to be a potential limitation of this medium.

Perhaps because listeners were focused on the relative novelty of using a VR headset, there was significantly less negative feedback regarding the lack of multisensory experience in the VR installation. The level of physical interactivity in this format was in between that of the physical and web-based versions: listeners were able to see virtual hands mirroring their own movements, one of which held a flashlight. The controllers used the common video game mapping of one joystick for movement and one for rotation, so it is possible that adherence to such a wide-spread convention also contributed to the lack of specific listener feedback on the installation's physical interactivity.

Though less about controller interactivity, the VR installation's triggered "inner world" animations fulfilled a similar function to the radio receivers in the physical work by responding to the listener and giving them a feeling of satisfaction upon "tuning in" to the composition. The reaction of many listeners when animations began was pronounced and extremely positive, particularly the first time,

with some watching motionless and some quickly swiveling around to look. While both installation features are responsive to listener behavior, positive feedback about the physical installation devices and the VR installation animations cannot be substituted for one another as equal forms of physical interactivity. Overall, listener feedback in this case did not clearly articulate the relationship between temporality and physical interactivity in VR as an affordance or limitation.

4.3 Environmental movement

The final dimension highlighted in listener feedback was environmental movement, referring to the background visual elements of an installation site as well as foreground graphics. Listeners did not mention environmental movement in their feedback on the physical installation, but it was indirectly indicated as an affordance during discussions of the web-based format.

In the web-based installation's video nodes, listeners reported and were observed enjoying the small, moving background elements, including the occasional pedestrian or vehicle passing by the edges of the sculpture garden. The clearer the passerby was visible in the video, the larger the response from listeners. A jogger clearly in frame at the start of one node's video was of particular interest, with previous participants eager to point them out to the current listener. In all cases, more movement in the video resulted in longer exploration of the node in question, suggesting the videos' temporality superseded that of the music.

The environmental movement captured in the 360° videos was thus indicated as an affordance to the temporality of the web-based medium. However, it is unclear to what extent credit can be given to the medium itself rather than the physical site being recorded. Though not mentioned explicitly in listener feedback, it is reasonable to infer that these moving background visuals contributed similarly, if unremarkably, to the temporality of the physical installation. Thus, the environmental movement in this case was shown to be an affordance of the physical installation medium.

Animations in the VR installation fulfilled a similar role. Both the clouds and triggered "inner world" animations elicited positive feedback from listeners, with many spending the majority of their time in the installation with their held tilted all the way back to watch. Feedback about the animations was the most explicit in its relationship to the installation's temporality. When asked what aspect of the VR version made them slow down, one focus group member said, "In order to induce stillness, you need movement. Stillness implies the passage of time. In VR, that was the [triggered animations] and the clouds. It's the world kind of falling away," referring to the transition into the animations. The ability to have constantly moving visuals and maintain a dynamic environment was the cornerstone of listeners' perception of temporality in the VR installation. Of all dimensions emergent from listener feedback, perhaps this was the most clearly indicated affordance.

5. CONCLUSION

Inner Transmissions provides a case study for isolating and analyzing the affordances and limitations of working with temporality across different media. Listener feedback highlighted three dimensions of sound installation design particularly impactful on their perception of the work's temporality: their expectations for each medium, its physical interactivity and multisensory elements, and any environmental movement. The physical installation effectively acted as a baseline for listeners, with all three dimensions indicated as affordances to their perception of the medium's temporality. The web-based installation emerged as having the most limitations, with no dimensions shown to be clear affordances. Feedback for the VR installation was largely in between the other two, with the environment's movement and responsiveness emerging as potential affordances.

This work provides a framework for future study in this area, particularly further research of the design dimensions highlighted. They raise questions surrounding listener attention spans for artistic works vs. other content in VR, the role of haptics and other multisensory controls, and the overall level of visual stimulation needed to imbue temporality into a digital environment. While *Inner Transmissions* provides an example of an investigative work in these areas, many more such explorations are needed in order for composers to fully grapple with the digital tools and media swiftly becoming available to them. The issue of temporality examined in this paper is but one of many crucial implications of an artist's choice of medium. Additional comparative works are needed to obtain a more granular understanding of the effects of user interface, specific compositional choices and style, and site-specific phenomena in the context of VR sound installation design.

6. REFERENCES

[1] G. Mullen and N. Davidenko, "Time Compression in Virtual Reality," in *Timing & Time Perception*, 2021. 9(4): p. 377-392. https://doi.org/10.1163/22134468-bja10034

[2] C. Schatzschneider, G. Bruder, and F. Steinicke, "Who turned the clock? Effects of Manipulated Zeitgebers, Cognitive Load and Immersion on Time Estimation," in *IEEE Transactions on Visualization and Comput. Graphics*, 2016. 22(4): p. 1387–95. https://doi.org/10.1109/TVCG.2016.2518137

[3] T. Read, C. A. Sanchez, and R. De Amicis, "Engagement and time perception in virtual reality," in *Proc. of the Human Factors and Ergonomics Soc. Annual Meeting*, 2021. 65(1), p. 913–918. https://doi.org/10.1177/1071181321651337

[4] T. Sanders and P. Cairns, "Time perception, immersion and music in videogames," in *Proc. British Computer Soc. Conf. on Human-Computer Interaction* (BCS-HCI), 2010. p. 160–167. https://doi.org/10.14236/ewic/HCI2010.2

[5] C. Warwick, "Beauty is truth: Multi-sensory input and the challenge of designing aesthetically pleasing digital resources," in *Digital Scholarship in the Humanities*, 2017. 32(suppl_2), p. ii135–ii150, https://doi.org/10.1093/llc/fqx036