Exploring the Affordances of VR for Musical Interaction Design with VIMEs

Anıl Çamcı
University of Michigan
acamci@umich.edu

Matias A. Vilaplana Stark
University of Michigan
matvilap@umich.edu

Lusi Wang
University of Michigan
ruthlw@umich.edu

ABSTRACT

As virtual reality (VR) continues to gain prominence as a medium for artistic expression, a growing number of projects explore the use of VR for musical interaction design. In this paper, we discuss the concept of VIMEs (Virtual Interfaces for Musical Expression) through four case studies that explore different aspects of musical interactions in virtual environments. We then describe a user study designed to evaluate these VIMEs in terms of various usability considerations, such as immersion, perception of control, learnability and physical effort. We offer the results of the study, articulating the relationship between the design of a VIME and the various performance behaviors observed among its users. Finally, we discuss how these results, combined with recent developments in VR technology, can inform the design of new VIMEs.

Author Keywords

Virtual reality, interaction design, musical expression, immersive media

CCS Concepts

•**Human-centered computing** → **Virtual reality;** *Human computer interaction (HCI);* •**Applied computing** → **Sound and music computing;**

1. INTRODUCTION

The multimodal and interactive experiences enabled by the modern VR platforms constitute a new playground for the digital music instrument designer. Over the past few decades, developments in physical computing and fabrication have facilitated the design of many new interfaces for musical expression. Today, VR is rapidly becoming a new frontier in this area, enabling artists and researchers to explore the novel affordances of immersive systems for the design of digital musical interfaces.

In this paper, we discuss the concept of virtual interfaces for musical expression (VIMEs) in terms of the unique opportunities and challenges that they bring to musical interaction design. Building upon existing studies on virtual instruments, we discuss musical interaction design considerations in VR based on case studies. We describe four VIMEs with details pertaining to their conceptual design, technical implementation and usability. We then discuss the details and results of a comprehensive user study that investigates how the different design strategies adopted in these four VIMEs affect the users' behaviors in VR. Finally, we evaluate how the results of this study can contribute to the design of new VIMEs, informing the work of researchers and performers alike.

Licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). Copyright remains with the author(s).

NIME'20, July 21-25, 2020, Royal Birmingham Conservatoire, Birmingham City University, Birmingham, United Kingdom.

2. RELATED WORK

Previous projects have harnessed the potential of virtual environments for musical expression long before the recent wave of VR platforms. These earlier works adopt a variety of approaches to musical interaction design, such as the virtualization of existing instruments in CAVE-like environments [11] and the use of gamification techniques for collaborative music production [3].

For instance, Berthaut et al. explore bidirectional audio-visual mappings in a virtual space where graphical changes are mirrored in sound while musical events are visually displayed in the form of 3D widgets that the performer can control with a hand-held device [2].

More recent projects have incorporated modern VR platforms, game engines, and motion tracking technologies into the design of virtual musical instruments. For instance, *Wedge* [12] and *Chromachord* [9] both make use of a Leap Motion depth camera strapped to an Oculus Rift headset to allow the performer to use their hands to interact with virtual objects. Based on a similar hardware configuration, the VR performance *Carillon* takes place in a game-like VR that is conceived as a stage for a carillon performance where remote participants interact with the virtual instrument over the network [10].

Most recently, Palumbo et al. created a VR instrument modeled after modular synthesizers. By allowing users to design modular synthesis patches in virtual space, this instrument facilitates an audio programming workflow that is reminiscent of patch design in Max [13]. In a similar project called MuX, Andersson et al. offer a sandboxed audio programming environment that allows the user to build interfaces from modules that are based on physical models [1]. In a more commercial project, *EXA* gives the user more flexibility with abstract shapes that can be fashioned into custom interfaces that can be configured with instrument sounds and various musical scales [8].

Ranging from collaborative experiences to design tools and performance systems informed by existing musical traditions, these projects exemplify the breadth of musical interfaces that can be implemented in VR. With the case studies discussed in this paper, which share conceptual and technical similarities with some of these existing projects, we hope to further our understanding of user experiences with virtual interfaces for musical expression.

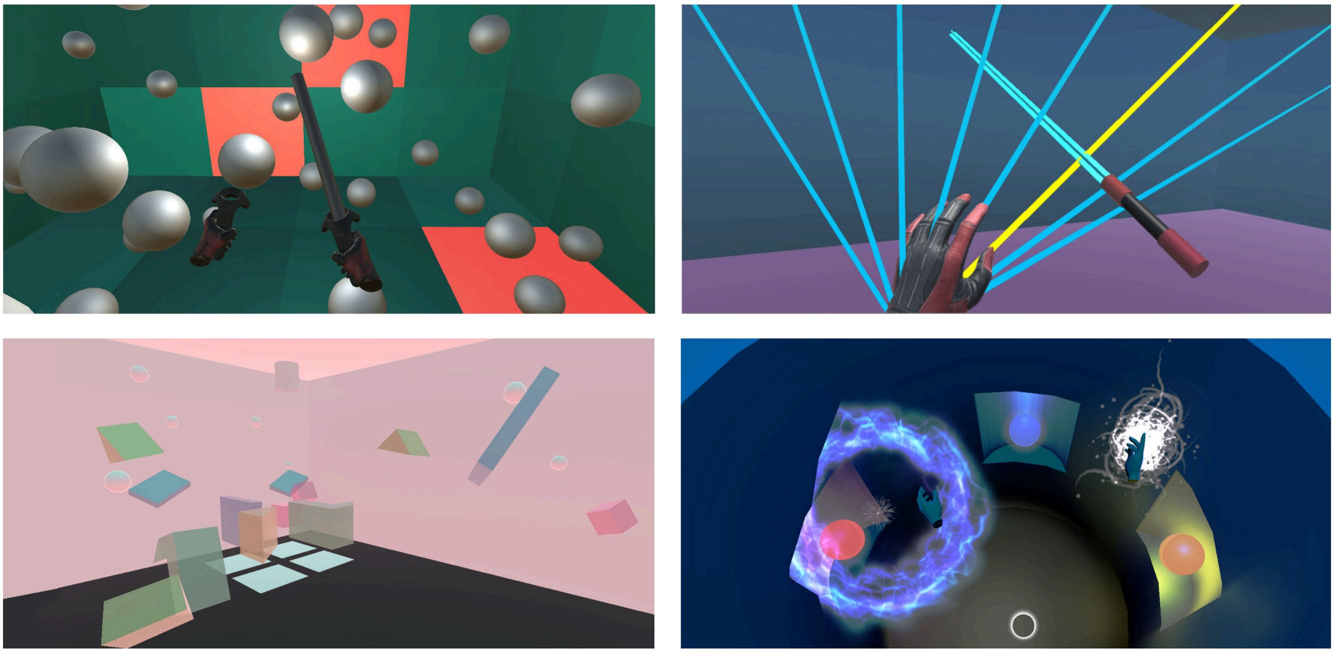

Figure 1: *Ball Pit* (top left), *Laser Harp* (top right), *Marbells* (bottom left), *ORBit* (bottom right).

3. VIMES

Digital musical instruments can sever the tie between a physical expression and its acoustic outcome. Although this may disrupt the action-perception feedback loops inherent to musical performance, the computational mapping between performance input and auditory output can lead to otherwise impossible associations between movement and sound. VIMEs extend this evolution by employing designs that further liberate a musical interface from physical constraints. Exploiting the audiovisual mechanics of an arbitrarily defined virtual environment, and utilizing room-scale motion tracking, VIMEs can enable musical interactions that traverse a line between real and imaginary spaces.

NIME studies often combine software design, signal processing and physical computing to conceive novel tools for musical creativity. If VR experiences can be characterized as those where the user's presence is mapped onto a computationally-defined space, it can be argued that many NIMEs already exhibit VR-like qualities through their use of motion tracking and spatial audio. However, many digital musical instruments have also made more explicit use of VR as discussed in the previous section. In Serafin et al.'s design principles for virtual reality musical instruments (VRMIs), 3 layers for the evaluation of such instruments are proposed: an interaction design layer that is concerned with the mapping between the input and the output modalities; a VR-specific layer that is focused on the intrinsic qualities of the virtual experience; and a third layer that deals with the overall qualities and the goals of the interactions implemented with a VRMI [14].

These layers touch upon issues relating to NIME design on one end and VR content creation on the other; the area between these two poles can be considered as the main space for the development of VIMEs, which can be conceived as musical systems that extend beyond instruments. A VIME, for instance, can be formulated as a virtual sound installation or an interactive experience. Mirroring the broad scope of applications in NIME research, a VIME can range from a hand-held virtual object to an infinite environment, the shared quality between the two being their facilitation of musical expression.

3.1 Implementation

The VIMEs discussed in this paper are designed for room-scale virtual experiences using the Oculus Rift S and the HTC Vive platforms. While some of the designs rely on physical movement, others exploit virtual navigation within environments that extend beyond the confines of the tracked space. Although inside-out tracking systems, such as the one used in the Oculus Rift S, that rely on plane detection are suitable for arbitrarily defined play areas, outside-in tracking, which the HTC Vive employs, offers the advantage of constant controller tracking even when the controllers are outside of the user's field of view. The choice of hardware platform for each design was accordingly informed by the tracking capabilities and hardware controller layouts of the individual platforms. All four VIMEs were designed using Unity 3D. Google's Resonance Audio API was used for the binaural spatialization of the sounds.

3.2 Case Studies

To better understand the affordances of VR for musical interaction design, we developed four VIMEs as case studies. While each VIME explored a different facet of interaction design for VR, they were all informed by the following design guidelines:

- The interface should only be possible to implement in VR, exploiting interactive, virtual-physical, acoustical, and visual affordances of VR.
- The interface can be governed by existing musical traditions, mental models and learned sensorimotor behavior or deviate from these entirely in the form of a novel interface.
- The scale of the interface can range from that of an instrument to that of an arbitrarily sized environment conceived as a musical system.

3.2.1 Ball Pit

Ball Pit is designed for the HTC Vive as a musical interface at the scale of an environment. As seen in the top-left panel in Fig. 1, the user is situated in an enclosed space with tiled walls, where they can move around physically in room scale. Using the left-hand controller, the user can spawn an arbitrary number of balls in three different sizes. Once thrown into the room, the balls move without being affected by gravity. The user can grab the balls in mid-air or strike them with a virtual baton mapped to the right-hand controller. The user can also remove the most recently spawned ball or destroy all balls at once.

Each tile displays a different shade of green and is tuned to a unique pitch from the chromatic scale. When a tile is struck by a ball, it lights up in red and emits a sound pulsaret reminiscent of the bouncing of a ping-pong ball. While the pitch of the sound is based on the tuning of a tile, its octave is determined by the size of the ball (i.e. smaller balls produce higher-octave sounds).

3.2.2 *Laser Harp*

Designed for the HTC Vive, *Laser Harp* re-imagines the traditional string instrument as a room-scale interface that stretches between the performer and the virtual environment. Eight strings extend between the user's left hand and a bridge affixed to the ceiling as seen in the top-right panel in Fig. 1. A virtual bow mapped to the right-hand controller allows the user to strum and bow the strings.

The sounds are synthesized using the Karplus-Strong algorithm to achieve plucked and sustained string sounds with velocity sensitivity. The touchpad on the left-hand controller can be used to scroll through four different musical scales, namely major, minor, chromatic and whole-tone scales. Moving the left hand, the user can stretch the strings to tune them. The rotation of the left hand controls the cut-off frequency of a low pass filter, allowing the user to alter the timbre of the instrument.

3.2.3 *Marbells*

Marbells is designed for the Oculus Rift S in the scale of an environment wherein the user can design a musical causality system by placing resonant blocks in the pathway of marbles that are ejected from four pipes in each corner of a virtual room as seen in the bottom left panel in Fig. 1. A bell-like sound is heard when a marble comes in contact with a block. The marbles adhere to traditional physics as they bounce off of the blocks and disappear once they reach the floor. The virtual environment extends beyond room scale to allow for elaborate causality systems; while the user can move physically within the confines of the play area, they can also use the thumbsticks on the Touch Controller to navigate beyond this area.

The user can spawn an arbitrary number of blocks from a selection of four different types, each with different timbral characteristics. This is achieved by grabbing a block from the pedestals located in the middle of the room; a new block is spawned every time an existing block is grabbed. The blocks can be arbitrarily resized and positioned in space where they suspend in midair when released. The user can point to pipes to change the clock division of the rate at which they emit balls and point to blocks to change their tuning on a pentatonic scale, creating diverse polyphonic and polyrhythmic structures.

3.2.4 *ORBit*

Designed for the Oculus Rift S, *ORBit* encapsulates the performer in a cylindrical structure, which serves as a gallery of virtual musical objects as seen in the bottom-right block in Fig. 1. Among these objects are 4 types of orbs that emit sounds and 2 auras that function as audio effects. The user can grab these objects and move them anywhere within their reach. Neither gravity nor ballistic forces affect the objects, which remain suspended in air when released.

Once removed from its compartment, an orb gives out a droning synthesized sound that is unique to it. The orb's position on the vertical axis controls the pitch of the drone that is spatially affixed to it. The two aura objects, which offer low-pass filtering and distortion effects, adhere to a similar interaction model. An aura's proximity to an orb controls the strength of its effect on the orb's sound.

4. EVALUATION

We conducted a user study to understand how the different design considerations that govern these four VIMEs affect the users' experience in terms of control, physical effort and immersion among other factors.

4.1 Participants

20 users between the ages of 20 and 35 participated in the study. 7 users reported having received college education in music. Among these were performers, music tutors, and music technology students. All remaining users reported having a basic level of experience with music through music lessons or participation in bands. 13 users reported prior experience with VR, with 7 of them identifying themselves as VR content creators. 18 participants reported prior experience with gaming.

4.2 Procedure

The study sessions are carried out with one user and two testers at a time. Each user performs with all four VIMEs in a shuffled sequence to avoid order-bias effects across participants. The names of the instruments are not disclosed to the user. Each session lasts for approximately one hour.

At the beginning of each session, Tester A describes the study design and asks the user to fill out a demographics survey. After Tester B explains the basic functions and controls of the first VIME, the participant is asked to perform with this interface in the form of an open-ended exploration. Various types of data pertaining to user behavior is logged during the performance. The participant is then asked to respond to a survey followed by a five-question interview carried out by Tester A. The above steps are repeated with the remaining VIMEs. Once all four performances are completed, the user fills out an exit survey consisting of six forced-ranking questions. This is followed by a final interview question regarding the criteria used by the participant to rank the different designs.

The per-VIME surveys include the same set of Likert-scale questions about the following criteria: intuitiveness of the interface, ease of learning how to use the VIME, the extent to which the user felt in control of the interface, and whether the user would be able to replicate their performance with the VIME. Additionally, the user is asked to estimate the time they spent playing with the VIME and whether the interface reminded them of any other object, concept or experience.

In the exit survey, the user is asked to force-rank the VIMEs on the following scales: From the easiest to control to hardest to control, from the one that demanded the most physical effort to one that demanded the least, from the most instrument-like to most game-like, from the most traditional to least traditional, and from the most expressive to least expressive. We describe the expressiveness of a VIME as the extent to which the VIME allowed the user to actualize an idea, musical or otherwise. In the final question of the exit survey, the user is asked to rank the four VIMEs from the one they liked the best to the one they liked the least. In a follow up interview question, the users were asked to expand upon this ranking.

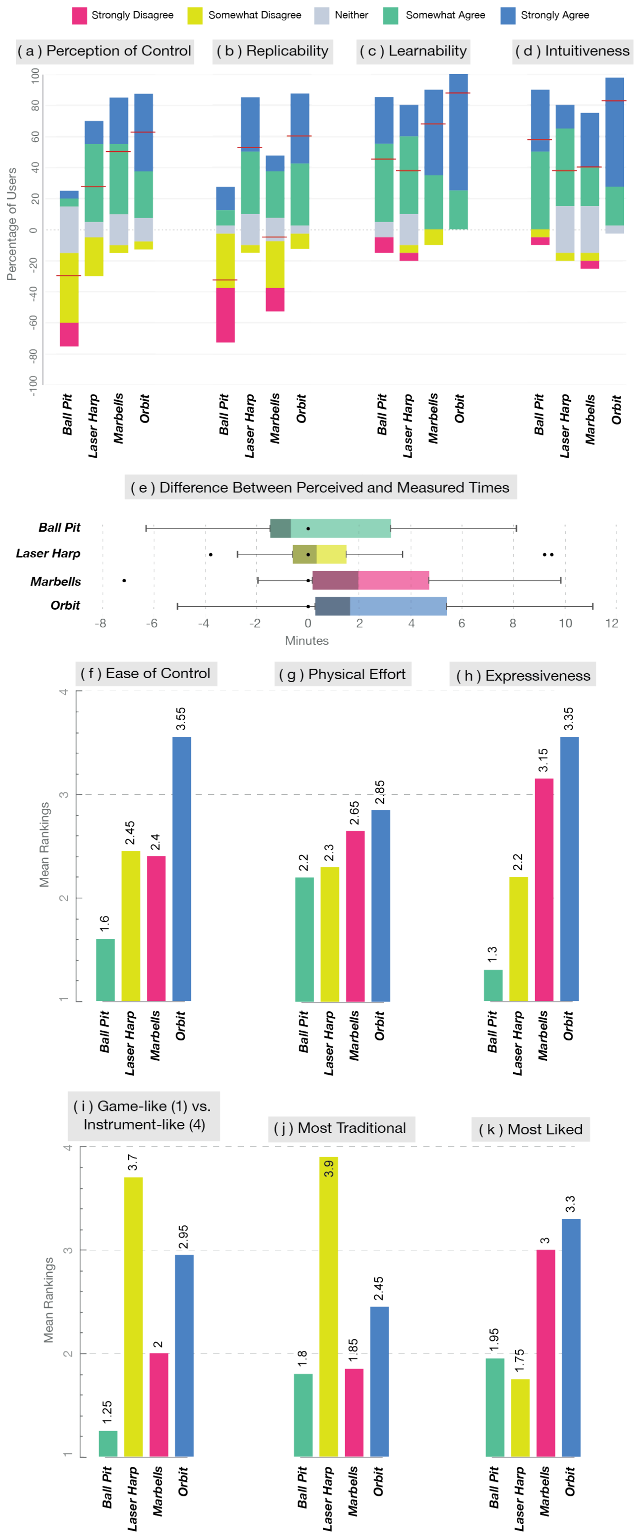

Figure 2: Results of the per-VIME and exit surveys. (a) "I felt in control of this VIME," (b) "The interface was intuitive," (c) "It was easy to learn how to play," (d) "I can replicate my performance," (e) average differences between perceived and actual performance times per VIME. The exit survey results show the average rankings for the four VIMEs: (f) easiest to control, (g) the amount of physical effort required, (h) most expressive, (i) most traditional, (j) most game-like to most instrument-like, (k) most-liked to least-liked.

5. RESULTS AND DISCUSSION

The average duration of performance with each VIME was as follows: *Ball Pit*: 7.1 minutes (SD: 2.14); *Laser Harp*: 6.64 minutes (SD: 2.51), *Marbells*: 8.23 minutes (SD: 2.11); *ORBit*: 6.59 minutes (SD: 2.72). The data from the per-VIME survey questions can be seen in Figs. 2(a, b, c, d, e) with (e) showing the average difference between the estimated and the actual performance times for each VIME. The results of the exit survey can be seen in Figs. 2(f, g, h, i, j, k).

5.1 Performance Behaviors

In this section, we discuss some of the performance trends we observed among users per each VIME, exploring how our design decisions were reflected in user behaviors.

In *Ball Pit*, although it is possible to adjust the aim and force of throwing to achieve certain pitch and rhythmic qualities, once the space is populated with enough balls, it becomes harder to exert control over the musical outcome. This design decision was reflected in the users' performance on a spectrum from controlled randomness to chaos. Some users spawned a limited number of balls and tried to maintain control over their movement. Conversely, others adopted an approach where they spawned an abundance of balls and struck them randomly as they flew by. Accordingly, in the follow-up interviews, some users expressed having tried to perform musical ideas while others likened their experience to playing a sports game. A transition between these two behaviors was also observed: gradually increasing the number of the balls, some users shifted their involvement from that of an active performer to that of a passive observer.

The audiovisual design of *Laser Harp* was informed by string instruments. Accordingly, most users expressed wanting to perform preconceived musical ideas with this interface. Much like with real instruments, this approach requires a learning process that may not be possible to fulfil within the confines of a user study. In our analysis of the string collision times, we observed a tendency towards playing adjacent strings in rapid successions. We believe this is mainly due to the fact that, unlike real string instruments, *Laser Harp* allows the user to pass the bow through the strings. In a recent study that compares tactile and virtual interactions with VR instruments, a similar user preference towards cutting through the strings is reported [5]. The authors attribute this tendency to the lack of physical anchoring that facilitates extended bowing gestures, and the users' preference towards exploiting the lack of physical limitations in VR.

The mode of interaction in *Marbells* was inspired by the Rube Goldberg Machine, which is a system designed to perform a task through an elaborate—and mostly redundant—chain of events. With most users, we observed an initial rough construction behavior, where they explored the response of the system and the governing rules of physics. This then transitioned into a more deliberate approach, where they fine-tuned the musical causality system. Similar to the behaviors observed with *Ball Pit*, as the users introduced more blocks into the environment, the complexity of the musical output increased with emergent structures that elude predictability. While some users embraced such structures by allocating more time to rough design, others worked with less blocks, focusing on finer adjustments.

The core interaction in *ORBit* adopts the common pitch-to-height mapping by way of allowing users to place an arbitrary number of sound-producing orbs at different heights in a confined space. In response to this design choice, we observed two primary performance behaviors: forming chords and performing phrases. In the first, the users were focused on the harmonic relationship between multiple orbs placed at different elevations. In the second, the users held on to an orb and moved it vertically to perform melodic structures. This behavior was also observed with the effect auras, when users, for instance, used the distortion effect to emphasize different notes in a chord.

5.2 Mental Associations

When dealing with a system, we associate it with external referents to form mental models that inform our understanding of the said system [7]. Accordingly, we explored how mental associations affect a user's approach to different VIMEs. After each performance, we asked users if they associated a VIME with an object, concept or previous experience. Furthermore, given the inherent overlap between VR and gaming with regards to mechanics, rule systems and reliance on similar design tools, we asked users to rank all four instruments on a spectrum from game-like to instrument-like in the exit survey.

8 users likened their experience with *Ball Pit* to playing a pitched percussion instrument due to the sounds generated by the system and the bodily interactions it prompts. 5 users associated these interactions with sports games, such as racquetball and squash. This is reflected in Fig. 2(i) with *Ball Pit* ranked as the most game-like VIME. Some users described the rain-like soundscape of more chaotic scenes as being therapeutic and meditative.

12 users associated *Laser Harp* with string instruments such as harp, violin and cello. Furthermore, Fig. 2(i) shows that *Laser Harp* was ranked as the most instrument-like VIME. We believe this strong association to an existing instrument have prompted users to adopt a more traditional approach with attempts at playing musical phrases akin to those that would be played with a violin.

Many users associated *Marbells* with kinetic sculptures, legos and Rube Goldberg Machines. The latter, which was a point of departure in our design, was used by 5 participants without our prompt. The users expressed that the designs made with this VIME could be presented as installations.

Finally, *ORBit* was most commonly associated with music production tools, digital audio workstations (DAWs) and MIDI interfaces. Furthermore, users made explicit references to concepts like chords, music theory and DJ tools when describing their experience with this VIME.

5.3 Presence

In their analysis of how to gauge presence and involvement in virtual experiences, Witmer and Singer focus on various factors including the user's ability to control events, compelling behaviors exhibited by virtual objects, a convincing sense of mapping across modalities, and losing track of time among others [15]. Although a relationship between perceived performance time and the illusion of presence would require further research, our results do show a positive correlation between VIME preference and the difference between estimated and actual performance times, as seen in Fig. 2(e) and (k), with the assumption that greater differences imply further loss of tracked time.

Most users expressed that the behavior of virtual objects conformed to their expectations with each VIME. Even when a behavior defied the rules of physics, the users were able to adjust to it as long as it was consistent. Users mostly agreed that the spatial mapping across different modalities was convincing in all VIMEs. With *ORBit*, in particular, users reported that the confined space, as seen in the bottom right panel in Fig. 1, felt more immersive than some of the more expansive environments in other VIMEs. Some users expressed that, due to the spatial limits imposed by the encapsulating structure, the orbs were placed at a closer proximity to the user's head; this likely made the binaural effect more prominent than other VIMEs and gave the users a more convincing sense of spatial mapping.

5.4 Control

Dobrian and Koppelman describe control as a precondition for expressive interfaces but add that the mere presence of control does not guarantee expression [6]. To afford control, an interface should be able to map gestures to sonic results with an appropriate amount of detail. Although it is unreasonable to expect first-time users to gain sufficient expertise with a VIME in the context of a user study, we asked our users to rank the four VIMEs in terms of expressiveness based on their limited experience. As mentioned earlier, expressiveness in this context was described loosely as a VIME's ability to help the user actualize an idea, musical or otherwise. With that in mind, we do observe a correlation between the rankings on ease of control and expressiveness as seen in Figs. 2(f) and (h), and the users' evaluation of the extent to which they felt in control of individual VIMEs shown in (a).

With VIMEs that were ranked as most instrument-like, the users evaluated their performances to be mainly replicable. Conversely, the most game-like VIMES, namely *Ball Pit* and *Marbells*, were ranked relatively low in replicability with notable disagreement among users. This disagreement can be explained by the aforementioned variations in the performance behaviors among users ranging from exertion of control over the VIME to influencing an ongoing emergent musical process.

All VIMEs were ranked relatively high in terms of learnability. Moreover, learnability and intuitiveness rankings show strong agreement with the exception of *Marbells*, which was rated highly learnable yet moderately intuitive. With this VIME, we observed that most users focused on the placement of blocks in the scene, while making little use of the tempo and pitch change features. This VIME, on the other hand, was ranked relatively high in expressiveness; being able to make expressive use of a VIME might have given the users the sense that they learned how to use it even though they did not exploit some of its functionality.

While the physical effort rankings seen in Fig. 2(g) did not show significant agreement across users, these give some interesting insights on how physical effort might be perceived in VR. *ORBit* required users to stand in one spot and move virtual objects on the vertical axis. Relying on virtual navigation, *Marbells* demanded a similar type of physical activity for placing blocks in 3D space. On the other hand, *Ball Pit*, which is a room-scale experience that the users often associated with sports games, was ranked lower in terms of physical effort. A possible explanation of this counterintuitive outcome is the users' interpretation of the dexterity involved in placing the virtual objects in *Marbells* and *ORBit* to be more taxing than the free form physical activity performed with *Ball Pit*. Some users also described the virtual navigation in *Marbells* as requiring effort; while this did not entail exertion of physical energy, the mental taxation might have contributed to a perceived sense of effort. Furthermore, in the aforementioned behavior of performing phrases with *ORBit*, some users acted out dance-like moves and associated this experience with dancing, which might explain a level of perceived physical effort. Finally, another possible explanation is that the reported extent of external activity in *Ball Pit* (i.e. balls flying by) might have dwarfed the users' own sense of proprioceptive involvement.

6. CONCLUSIONS

VR offers distinct possibilities for musical interaction design, particularly with regards to the liberation of the interface from physical constraints and the ability to map user input to audible outcome using the arbitrarily defined mechanics that underlie an immersive sonic environment. Given the users' ability to adjust to diverse virtual interaction schemes implemented in our case studies, we hope these findings will encourage designers to further exploit the affordances of VR with regards to physics and mechanics.

We observed a significant effect of mental models on how users approached each VIME. Interfaces that were viewed more as a game were evaluated by a different set of standards then those that were viewed as an instrument. We believe that through interaction mechanics that strike a balance between these two paradigms, it might be possible to implement VIMEs that serve more diverse creative goals. Given the flexibility of VR in terms of visual and mechanical design, it can serve as a platform not only for creative expression in and of itself but also for prototyping NIME ideas prior to physical implementation.

While all four VIMEs supported room-scale motion tracking, the virtual architecture of a VIME significantly affected how users utilized the tracked space. With *ORBit*, in particular, confining the users in a narrow structure have prompted them to place objects in close proximity and gain a more prominent sense of audio spatialization as a result.

Although the relationship between games and music has been extensively studied [4], we believe the close collaboration between VR and gaming offers new opportunities for VIME design in terms of utilizing gamification in musical systems. When asked about how they viewed these VIMEs being used, most users referred to training, and music education applications. VR is indeed being increasingly used for educational purposes across many disciplines including music. We believe that VIMEs can support diverse learning experiences at the intersection of music education and gaming, alongside their more immediate applications in music production and audiovisual performance.

With VR controllers, which the average user does not have extensive experience with, we found it important to implement tiered control schemes that can support expressive behavior supplemented by more sophisticated interactions that enable extended functionality. With *Marbells*, for instance, the users were able to achieve desirable results with low-level interactions like grabbing and moving, while other adjustments were available to them through the use of the buttons on the controller.

The only instance of a user expressing discomfort was when a user with no prior experience with VR reported feeling motion sickness while using the virtual navigation in *Marbells*. Although this interaction was implemented using Oculus' recommended set of guidelines, this occurrence should prompt the VIME designer to consider different levels of VR experience when designing a system.

In the near future, we hope to explore the effects of VIME design on performance in collaborative contexts. For instance, constructing a musical causality system in *Marbells* with multiple users can reveal new performance dynamics and create opportunities for networked performances. We also hope to further our investigation into the perception of physical effort in VR as this can have significant implications for not only musical interaction design but also other forms of kinesiological research in VR.

7. ACKNOWLEDGMENTS

We would like to thank the following people who contributed to the design of the four VIMEs featured in this paper: Fee Christoph, Timothy Everett, Jessica Glynn, Utku Gürkan, Oren Levin, Frannie Miller, Rishane Oak, Benjamin Roberts, Ayal Subar, Aaron Willete.

8. REFERENCES

[1] N. Andersson, C. Erkut, and S. Serafin. Immersive audio programming in a virtual reality sandbox. In *2019 AES International Conference on Immersive and Interactive Audio*, March 2019.

[2] F. Berthaut, M. Desainte-Catherine, and M. Hachet. Interacting with 3d reactive widgets for musical performance. *Journal of New Music Research*, 40(3):253–263, 2011.

[3] J. N. Bott, J. G. Crowley, and J. J. LaViola Jr. Exploring 3d gestural interfaces for music creation in video games. In *Proceedings of the 4th International Conference on Foundations of Digital Games*, pages 18–25. ACM, 2009.

[4] A. Çamcı, C. Çakmak, and A. G. Forbes. Applying game mechanics to networked music HCI applications. In S. Holland, T. Mudd, A. McPherson, and M. Wanderlay, editors, *New Directions in Music and Human-Computer Interaction*, pages 223–241. Springer, 2019.

[5] A. Çamcı and J. Granzow. Hyperreal instruments: Bridging VR and fabrication to facilitate new forms of musical expression. *Leonardo Music Journal*, 29:14–18, December 2019.

[6] C. Dobrian and D. Koppelman. The "e" in nime: Musical expression with new computer interfaces. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, page 277–282, 2006.

[7] J. K. Doyle and D. N. Ford. Mental models concepts revisited: some clarifications and a reply to lane. *System Dynamics Review*, 15(4):411, Winter 1999.

[8] EXA: The Infinite Instrument. `https://store.steampowered.com/app/606920/EXA_The_Infinite_Instrument/`. Accessed 27 Jan 2020.

[9] J. Fillwalk. Chromachord: A virtual musical instrument. In *2015 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 201–202, 2015.

[10] R. Hamilton and C. Platz. Gesture-based collaborative virtual reality performance in carillon. In *Proceedings of the International Computer Music Conference 2016*, pages 337–340, 2016.

[11] T. Mäki-Patola, J. Laitinen, A. Kanerva, and T. Takala. Experiments with virtual reality instruments. In *Proceedings of the Conference on New interfaces for Musical Expression*, pages 11–16, 2005.

[12] A. G. Moore, M. J. Howell, A. W. Stiles, N. S. Herrera, and R. P. McMahan. Wedge: A musical interface for building and playing composition-appropriate immersive environments. In *2015 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 205–206, 2015.

[13] M. Palumbo, A. Zonta, and G. Wakefield. Modular reality: Analogues of patching in immersive space. *Journal of New Music Research*, 49(1):8–23, 2020.

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

[15] B. G. Witmer and M. J. Singer. Measuring presence in virtual environments: A presence questionnaire. *Presence*, 7(3):225–240, 1998.