# Searching for the Perfect Instrument: Increased Telepresence through Interactive Evolutionary Instrument Design

Dhruv Chauhan
University of Bristol
MVSE School of Engineering
Bristol, UK
dc14690@my.bristol.ac.uk

Peter Bennett
University of Bristol
MVSE School of Engineering
Bristol, UK
peter.bennett@bristol.ac.uk

## ABSTRACT

In this paper, we introduce and explore a novel Virtual Reality musical interaction system (named REVOLVE) that utilises a user-guided evolutionary algorithm to personalise musical instruments to users' individual preferences. REVOLVE is designed towards being an 'endlessly entertaining' experience through the potentially infinite number of sounds that can be produced. Our hypothesis is that using evolutionary algorithms with VR for musical interactions will lead to increased user telepresence. In addition to this, REVOLVE was designed to inform novel research into this unexplored area. Think aloud trials and thematic analysis revealed 5 main themes: control, comparison to the real world, immersion, general usability and limitations, in addition to practical improvements. Overall, it was found that this combination of technologies did improve telepresence levels, proving the original hypothesis correct.

## Author Keywords

Virtual Reality, Evolutionary Algorithms, Adaptable Musical Interface, Qualitative Study

## CCS Concepts

•**Human-centered computing** → **Virtual reality;** *User studies;* •**Applied computing** → **Sound and music computing;**

## 1. INTRODUCTION

Our hypothesis is that creating a novel musical experience as an intersection of three existing fields (Virtual Reality, Evolutionary Algorithms and Musical Interaction), would lead to increased telepresence in users, as opposed to each field in isolation. Although plenty of research already exists into each field, and even between pairs of each, very limited research has been done with all three [2]. The open-endedness and subjectivity of the music domain lends itself to evolutionary algorithms, particularly where a user-centric method is used. Further, Virtual reality offers much potential for new musical expression, demonstrated by the Audio-First Workshop at NIME 2018. This project prototypes an interactive virtual musical system in which a user explores possible musical interactions, and based on their preferences, informs an evolutionary algorithm. We aim to show that the union of these 3 areas increases the users' telepresence with the experience, and to use the resulting analysis to inform further research.



*NIME'19,* June 3-6, 2019, Federal University of Rio Grande do Sul, Porto Alegre, Brazil.

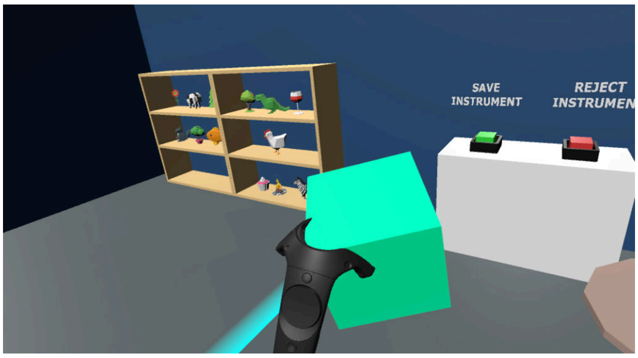


**Figure 1: A snapshot of the final system. The user handles the instrument (green box), and can choose whether to save or reject it by pushing one of the two buttons, influencing the generation of the next instrument.**

In VR literature, Steuer outlines the notion of telepresence: defined as "the notion of being elsewhere" [6]. The final definition he proposes for VR is "simulated environment in which a perceiver experiences telepresence". Steuer considers the various factors that help to create a telepresence, falling under the umbrellas of 'vividness' and 'interactivity'.

The combination of both the fields of VR and musical interaction yields the notion of Virtual Reality Musical Instruments. Jaron Lanier [5] first prototyped an 'abstract virtual world' and 3 VR musical instruments in 1987. These instruments highlight some of the things that are possible with Virtual Reality technology. Since 1987, consumer technology has advanced vastly, and in the time, there have been other explorations between virtual reality and musical instruments. Serafin et al. [5] propose a set of design principals for evaluating Virtual Reality Musical Instruments. We will later use these design principles to evaluate our design.

Finally, we explore existing applications of Genetic Algorithms in the context of music. The work done by Keijzer et al. explores Interactive Evolutionary Algorithms (IEAs) [4], which are named due to their need for a user in the fitness function. Their work is done in the context of learning subjective fitness functions to create 'pleasing' drum patterns. Using a subjective fitness function with evolutionary algorithms, explored in the above papers, as a way for a user to navigate a space of sounds yielded patterns that suited the user more than without a user input. However, personal preference of music and interactions means that what a user deems a successful result may vary dependent on the user and their background.

## 2. INTERACTION WALK-THROUGH

When the player puts on the headset, they see a scene as shown in Figure 1. This includes visualisations of the controllers. A single instrument exists on the table, and the user may interact around and with it as they choose to. When a user passes a controller through the instrument, it changes colour to indicate that the controller is inside. If the user presses the trigger on the back of either controller, while it is held down, the instrument acts as a child of the controller - i.e. it has been grabbed. When the trigger is released the instrument is dropped, but since it acts as a rigid-body it conforms to gravity and maintains any momentum it may have. This allows a user to control the instrument in any way they choose: throwing it up and catching it again, rotating around, waving it around, moving it up and down, holding the object still while physically moving in the space, bouncing it off nearby surfaces, or many other forms of interaction. In addition to this, if the trigger is held before the user inserts the controller into the instrument, and they push against it, the controller acts as a rigid-body, thus allowing a user to push, or nudge the instrument around.

On the left of the user are two buttons, one green, to indicate the saving of the instrument, and one red, which indicates that the user rejects the instrument. These can also be seen in Figure 1. The buttons consist of cubes, that when pressed, trigger the relevant action. The buttons can either be triggered by turning the controller into a rigid-body (by holding the trigger when not in contact with any objects), and pushing downward, or may even be triggered by the instrument itself.

If the save button has been pressed, the genome of the instrument is added to an array of saved genomes. In either case, the current instrument object is destroyed. Now, if the evolutionary limit hasn't been reached, a new random genome is generated and presented to the user, in the same way as above. A new instrument rigid-body object is generated on the table, for the user to interact with. If the evolutionary limit has been reached, the evolutionary process starts. In all cases, text appears for 5 seconds above the buttons, indicating what action has been taken.

The genome is mutated with a crossover algorithm, and the users are presented with an instrument loaded from the next generation. The user may then specify their preferences, and the cycle repeats.

## 3. SYSTEM DESIGN

REVOLVE was designed using Unity, Pure Data (in the form of LibPD), Virtual Reality Toolkit, and tested using a HTC Vive. An image of the system in action can be seen in Figure 1. In addition, a flow diagram of how REVOLVE operates may be seen in Figure 2.

### 3.1 Musical System and Instrument

The musical system is realised as an exposed Subtractive Synthesizer in Pure Data. It works by taking a sawtooth wave and passing it through a VCF (Voltage Controlled Filter), and then a VCA (Voltage Controlled Amplifier). A VCF shapes an audio wave by specifying start, end and Q (resonance at the cut-off frequency) values. VCFs allow for continuous adjustments of values, which is why they were selected. A VCA adjusts the amplitude of the incoming wave: adjusting it with an ADSR (Attack-Decay-Sustain-Release) envelope that shapes the sound. The parameters to these parts of the system would correspond to a specific 'instrument', along with a specific chosen parameter that would be the controlled parameter from Unity.

We chose that the interaction would not be specifically tied to a user interaction or gesture, but rather that the object's properties would inform the output, and the user would have the freedom to interact with the object in a secondary manner — in the way they saw appropriate. A parallel could be drawn here with conventional instruments, whereby they can often be played in multiple ways, but some ways make more sense than others, or rather more users tend to converge on a specific method of interacting - e.g. plucking guitar strings vs. playing with plectrum vs. using a bow. In the same way, the user was given controllers in order to manipulate the object, and the resulting change in the object's properties would adjust the sound. There are 12 final object properties: the x, y, z components of, and resulting component vectors, translation, rotation and velocity.

### 3.2 Genome

A metaphor was needed for the representation of an instrument, and since the ultimate aim would be to evolve these instruments, representing them as a genome made the most sense. An example genome can be seen below. The range 0-127 was chosen as it corresponds well with the MIDI CC numbers. For boolean values, a value of 0 or 1 was sent.

$$\underbrace{1\ 0\ 1}_{\text{control params}} \quad \underbrace{64\ 64\ 64}_{\text{scale, metro, freq}} \quad \underbrace{64\ 64\ 64\ 64}_{\text{env params}} \quad \underbrace{64\ 64\ 64\ 64}_{\text{filter params}}$$

This interwoven system of audio inputs, object parameters, and specific genome structure is the core of REVOLVE.

### 3.3 Evolution

We concluded that the best way to combine GAs and musical interactions would be with a subjective fitness function, whereby the user specifies their preference for a specific instrument, and this is used to indicate whether an instrument will pass onto the next generation. An implementation of an evolutionary algorithm, consisting of basic crossover and mutation, spawns a new generation of instruments, which are then loaded by the user for use. This process is then repeated. In this case the mutation involved randomly changing a random parameter every few generations to within the range.

## 4. EXPERIMENTAL DESIGN

With a basic prototype completed, the next stage was to perform experiments in which users are asked to interact with REVOLVE, and their responses are recorded, transcribed, and analysed for recurring themes. The PRET A Rapporter framework (PRETAR) described by Blandford [1] gives a basic structure for designing, undertaking and reporting Semi-Structured Qualitative Studies (SSQSs).

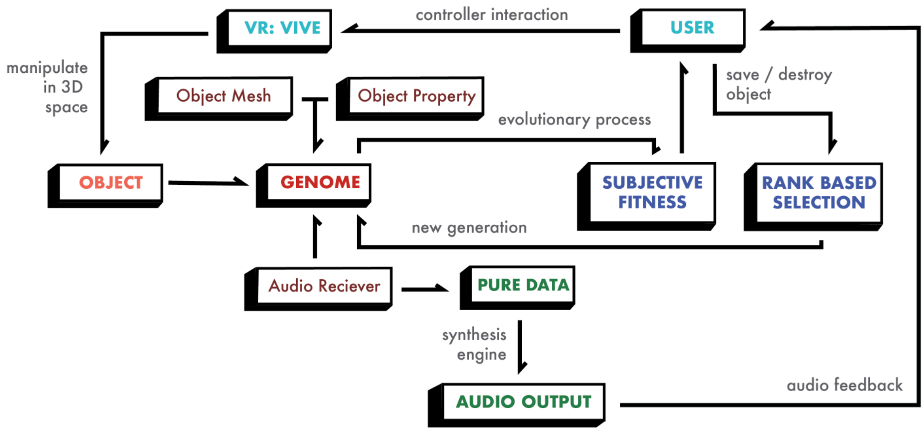


**Figure 2: A flow diagram of the system logic showing how the user's interaction effects the evolutionary process.**

## 4.1 Participants

As well as performing the interviews, we collected some minimal background data on the participants, asking their age, gender, musical and VR experience, in order to try and minimize any external factors that may influence the outcome of the experiment. This simple questionnaire contained questions with a simple scale, where the meaning of each of the responses was clearly indicated above. The experiments were run across one week. In total, there were 6 participants, with resulting audio spanning between 15-20 minutes for each. Therefore, approximately 90 minutes of audio data was collected and transcribed during the study. The numbers correspond to a scale given from 1, no experience to 5, expert-level experience.

| Subject ID | Age | Gender | Musical Experience | VR Experience |
|---|---|---|---|---|
| *VR01* | 22 | F | 3 | 3 |
| *VR02* | 22 | M | 4 | 4 |
| *VR03* | 22 | M | 2 | 5 |
| *VR04* | 21 | M | 2 | 4 |
| *VR05* | 21 | F | 3 | 3 |
| *VR06* | 26 | F | 3 | 4 |

**Table 1: A table of the participants of the experiments, and their collected demographic data.**

## 4.2 Data Collection

Users were given instructions to wear the headset, and information on how to use the controllers, and interact with the system safely. They were then asked to partake in the experience for around 15 minutes. As part of the information sheet, the users were instructed that they would need to 'think-aloud'. The 'think-aloud' vocalisation would happen whilst users were inside the experience, giving them the chance to relay any thoughts they may not remember or picked up on once they had stopped engaging with the experience. After this, the users took part in a Semi-Structured Interview, where there would be a rough topic guide.

# 5. DATA ANALYSIS

Thematic analysis is described by Guest [3] as a technique of analysing qualitative data through the targeting of recurring themes in data. This was performed upon, via the coding process, the transcribed data corpus from the study.

# 6. EVALUATING TELEPRESENCE

## 6.1 Results

In the following section we present the themes that arose whilst conducting the thematic analysis. 5 key themes arose: Control, Comparison to Real World, Immersion, General Usability, and Limitations.

### *6.1.1 Control*

Throughout the experiments, users referred to their feelings surrounding levels of control. In the experience, there were 3 main levels in which control could be exerted: The low-level direct control through the hardware controllers, the mid-level control through interactions that influence the sound, and the high-level, more abstract control in the overall evolutionary process. Many users felt like they didn't have the mid-level control, but users were left feeling in control from both a low-level nature, and a high level.

Overall, users became frustrated when they didn't have an optimal amount of control over a system. Looking at Steuer's [6] model of factors influencing telepresence, control falls under the 'interactivity' umbrella, under 'mapping'. From this model, a decreased level of control leads to less telepresence.

### *6.1.2 Analogies with Real World*

Many users identified parallels between the VR world and how they interacted with the experience. This flavoured their expectations of how the virtual experience would operate, particularly in relation to the sound, and how it operates in the real world. For example, many users expected the sound to emit from the cube itself, rather than the world.

### *6.1.3 Immersion*

Many users discussed various factors that led to them feeling immersed in the experience. In general, immersion seemed to be one of the strong positives of the experiment - users felt engaged with the experience. They pointed to Virtual Reality, and its implications, as a key factor in increasing the immersion. Virtual Reality offers this, especially when other senses are incorporated, thus increasing the 'breadth'

| VRMI | *DP1: Feedback and mapping* | *DP2: Latency* | *DP3: Cyber Sickness* | *DP4: Do not copy tech.* | *DP5: Interaction Natural / Magic* | *DP6: Ergonomy* | *DP7: Sense of presence* | *DP8: Body represent* | *DP9: Social Experience* |
|---|---|---|---|---|---|---|---|---|---|
| **REVOLVE** | Audio primary, visual secondary | Mentioned, not evaluated (None noted) | Didn't occur during experiments | New techniques | More natural, some magical | Nothing raised about discomfort with Vive | Not explored directly, but through immersion | Virtual representations of controllers | N/A |

**Table 2: Evaluation table of designed system using VRMI design principles by Serafin et al. [5].**

of 'vividness' in Steuer's [6] model. Some users had used other VR systems and had a less immersive experience than with our system.

## 6.2 Discussion

One of the aims of our project was to prototype and construct a novel system that incorporated Virtual Reality, Evolutionary Algorithms and Musical Interactions — ultimately to keep users engaged, endlessly fascinated, and progressively shape the instruments to their personal tastes. By this criteria, the project was a success.

Several key findings from the study were found. Firstly, users were left feeling frustrated when they had a lack of control, which overall decreased their level of telepresence. Ways to increase control include: increasing visual feedback, making interactions more explicit, and adding more obvious feedback from the music. Secondly, users' expectations of VR, taken from analogies with the real world, influenced their immersion with the system. Situations which didn't build upon, or felt unnatural in comparison to, real world interactions led to less immersion. Examples of these are: lack of audio spatialisation, a confined space of instrument sounds, as well as environmental and physical factors, such as gravity. Users also explored other areas by which their telepresence was increased, such as increased immersion directly from the VR system used, and the personalisation produced by the Evolutionary Algorithm. Users pointed towards this creating a state of 'flow'.

In contrast, not capitalising on the full breadth of the users' senses led to less immersion. The tactile system could have been used, with haptic feedback, and the visual system - a core component of VR - could have been used to reinforce the sound. Users were particularly excited by VR's potential to feed into the visual domain, with examples such as increased shape or colour-based feedback. Users also referred to expectations based on the way using the VR was framed: many described it as an experience, and noted that this increased their level of telepresence.

All of the above points are useful for creating VR musical interfaces in general, and will help contribute to future research.

## 6.3 Evaluation using VRMI Principles

As discussed earlier, Serafin et al. present 9 principles to evaluate Virtual Reality Musical Instruments. The evaluation of our system using these principles can be seen in Table 2.

From this evaluation model, our project has some strengths and weaknesses. Some key weaknesses that could be improved upon are: the presence, body representations, and the social experience factors of REVOLVE. However, using the Vive seemed to serve well in many areas, contributing strongly to the ergonomy, low latency, and lack of cybersickness. In our system we attempted to be innovative with the technologies and interactions, and this is seen through the natural and magical interactions, and the 'do not copy technology' principles. Overall, REVOLVE fared well against this model of evaluation, but points towards areas that could be improved upon.

# 7. FUTURE WORK

Procedural meshes, where an object's shape changes in real time, provide an interesting exploration point, and a study around which meshes work well with which instruments could aid in taking the experience to a new level. Also, increased visual feedback through gestures, UI, and visual metaphors could all be explored. We'd also like to incorporate haptic feedback, audio spatialisation and physical-based modelling. Machine Learning methods, such as neural networks, also offer an alternative method for increasing personalisation. The Evolutionary Algorithm used in the experiment was very one-dimensional - users could only affect the sound, and not the interaction. In future work we'd like to explore other axes of personalisation, such as interaction and mapping. A personalisation technique that works both explicitly and implicitly could also be trialled, by creating an implicit 'engagement factor' that could quantify how much a user is enjoying a certain instrument.

# 8. CONCLUSION

Overall, this paper outlined the process undertaken in building and constructing an interactive, evolutionary, virtual, musical experience - REVOLVE. It is shown that using this trio of fields - VR, Evolutionary Algorithms and Musical Interaction did increase telepresence, through the interactions and system immersion, thus proving our initial hypothesis correct. In addition, several practical improvements were found through the exploratory study, which will be incorporated into future versions of REVOLVE. Although the main focus of this paper was around musical interaction, the study also revealed key insights around factors that affect immersion levels in VR. It was paramount for users to have a direct feeling of control, and analogies to the real world allowed a user to interact more intuitively with the process, increasing their immersion, and therefore their telepresence.

# 9. REFERENCES


[1] A. Blandford, A. Adams, S. Attfield, G. Buchanan, J. Gow, S. Makri, J. Rimmer, and C. Warwick. Pret a reporter: evaluating digital libraries alone and in context. *Information Processing and Management, 2008.*

[2] S. Das, T. Franguiadakis, M. Papka, T. A. DeFanti, and D. J. Sandin. A genetic programming application in virtual reality. In *Fuzzy Systems, 1994*. IEEE, 1994.

[3] G. Guest, K. M. MacQueen, and E. E. Namey. Themes and codes. *Applied thematic analysis*, 2012.

[4] M. Keijzer, S. Lucas, E. Costa, and T. Soule. LNCS 3003 - Genetic Programming, 2004.

[5] S. Serafin, C. Erkut, J. Kojs, R. Nordahl, and N. C. Nilsson. Virtual reality musical instruments. *Proceedings of the Audio Mostly 2016 on - AM '16*, pages 266–271, 2016.

[6] J. Steuer. Defining Virtual Reality: Dimensions Determining Telepresence. *Journal of Communication*, 42(4):73–93, Dec 1992.