Exploring the Learning Potential of Acoustic Design in 3D Virtual Environments

Liz Falconer, Julian Green
University of the West of England
United Kingdom

Abstract

This paper discusses the current research and understanding of the effects of audio environments on learning, with a particular focus on the affordances of the effects in computer-generated, 3D virtual environments and virtual worlds. Both sound and music have different effects on different people, and this has been one of the confounding issues in the search for a generalizable theory of the effects of audio environments on learning. This in turn has meant that the affordances of both soundscapes and musicscapes in education remain under-researched, particularly in multi-media environments such as virtual worlds. Using evidence from student experiences when studying an online Master’s programme that runs entirely in a virtual world, synthesized with research in the subject areas of virtual learning, music and sound, we argue for a change in approach; that, in the case of 3D online virtual worlds and similar environments, the personal control that users have over the environment in which they learn presents opportunities for choice of their own, personal soundscapes and musicscapes to enhance individual learning experiences. This also presents a fertile ground for research into immersion, engagement, simulation and virtual environments as facilitators of learning when utilising both ambient and contextual sound with music, or alone. We therefore propose research in this field.

1. Acoustic design and the nature of 3D virtual environments

“The home territory of soundscape studies will be the middle ground between science, society and the arts. From acoustics and psychoacoustics we will learn about the physical properties of sound and the way sound is interpreted by the human brain. From society we will learn how man behaves with sounds and how sounds affect and change his behaviour. From the arts, particularly music, we will learn how man creates ideal soundscapes for that other life, the life of the imagination and psychic reflection. From these studies we will begin to lay the foundations of a new interdiscipline – acoustic design.” [1]

In the past 20 years acoustic design principles have been applied in a range of subject areas but the effect of acoustic design generally, and soundscapes more specifically, on learning is a scant research area. For example, there are research papers that report on the acoustic design of classrooms and how that design might impact upon learning [2] and studies of how acoustic design might be incorporated into the curricula of subjects such as liberal arts [3], but there has been little research upon how a designed soundscape, including the use of music, might affect learning as a human activity.

One of the underlying problems with the current understanding of the effect of an audio environment on learning is the personal differences we all display in relation to our responses to music and sounds. Soundscapes are generally a function of the physical environment in which the learning is taking place and cannot easily be changed in the physical world, and the practical and technological difficulties in providing a choice of music and/or soundscapes for personal devices in the physical world makes it effectively impossible for students to choose their own acoustic environment. But, the affordances of 3D virtual world technologies do make those kinds of choices possible for all participants.

Three-dimensional virtual worlds, where users can create and adapt their own environments, have been available for more than 10 years; the most widely used in education are the virtual world technologies Second Life® and Open Sim® [4]. In these virtual worlds users represent themselves as avatars and can move around the environment, and interact with objects and other users, whilst controlling many aspects of the environment that is displayed on their screens, including sound. Ward & Sonneborn [5] introduce the construct of “individualized collaboration” in these environments, where “…Unlike face to face group activities in real world settings, in which the ambient conditions are largely the same for all participants, virtual worlds have properties that make it possible, in principle,
for individuals to personalize their experience even while interacting with others in collaborative groups.”

Sound and music are generally less well-discussed than visual affordances in relation to virtual environments when they are being used as environments for learning. Virtual worlds are undoubtedly visually rich spaces and lend themselves to simulation and role play for educational purposes, but the visual can dominate design considerations when these environments are being used for these kinds learning. For example, in their 2012 paper, Duncan et al [4] developed a taxonomy of virtual world usage in education from scrutiny of over 100 published academic papers and classified the finding into 6 categories, the first 5 of which largely map onto who, what, where, how and why questions. The 6th category maps onto research case studies and specific learning research. Whilst the visual and communication affordances of these environments featured prominently in the literature in all categories, there was no mention of ambient soundscapes, nor indeed the wider issue of acoustic design.

It is possible for creators of 3D virtual environments to design soundscapes that blend ambient sound, specific sounds and music, and for the user to choose which elements of the soundscape to enable or disable for their own personal experience. Through the use of specific channels, environment creators can provide choices of music that form part of a soundscape, a musicscape or a combination of both. In this paper we borrow the term ‘musicscape’ from Oakes et al [6], who use it to describe the musical environment in retail stores that is intended to elicit emotional and behavioural responses in customers. There are many similar considerations in using music to enhance learning as will be discussed in this paper.

The next section summarises some of the findings of research into the effects of sound and music on recall, memory and sense of presence when used in environments intended to encourage learning. These findings are then synthesized with findings relating to the nature of the effects of sound and music on emotional responses in different environments which leads us to propose a number of research questions.

2. Sounds, music and learning

We in the Education Innovation Centre at the University of the West of England have been running a Master of Arts programme in Education in Virtual Worlds since 2012, and the course team has accumulated hundreds of hours of teaching in virtual world environments. The programme of study runs entirely in a virtual world, with students attending in avatar form from countries around the world, including the UK, USA, Canada, Argentina, Greece, Germany, Australia and New Zealand. Our students are predominantly school and college/university educators wishing to extend their practice, and our experiences show us that sound and music can have significant effects on their learning experiences, affecting a range of responses such as emotion, immersion and recall [7]. Music forms an important part of the socialization aspect of their study, and attendance at virtual dances and parties is a popular way for students to consolidate their relationships with other students, and to explore the affordances of virtual environments in a more relaxed manner than in scheduled classes. It is important to important to state here that classes in the programme do not consist of students only sitting in virtual classrooms being lectured to by tutors. Most of the class time is spent in active learning, including students building their own learning environments, visiting a wide range of places and simulations in virtual worlds and working on group projects.

As an example of the creativity and imagination that virtual environments unlock in our students, the following is an excerpt from a student reflective journal, capturing his thoughts on how he will approach one of the assessments on the programme; to design and build a learning environment aimed at helping his students to achieve a particular set of learning outcomes. In this case, the subject of the assessment is an English literature class studying “Night” by Elie Wiesel. It is interesting that soundscapes are specifically mentioned in this reflective note.

“In many virtual spaces I have experienced a sense of immersion purely because of what surrounded me; the artefacts, the atmosphere, the soundscapes, the lighting; all combining to provide a sense of being there that was beyond that of a spectator; I was emotionally there. It is my view that this level of immersion would facilitate students connecting with a literary work such that discussions would provide far more meaningful and emergent learning than those conducted solely from a classroom environment. As Thomas states (2010, p.503) “engaged learning is an emergent property of learning spaces that are designed to provide affordances that actively encourage such engagement”. I want to investigate how a build can actively engage students in literature at an emotional level and move their discussions beyond the abstract.”

Our students’ experiences on the MA are not an isolated case. Studies over the past 15 years have demonstrated that audio cues in virtual environments can increase a user’s sense of presence and recall of objects and their locations in the spatial environment. For example, Dinh et al [8] found that increasing the modalities of sensory input, particularly auditory cues, can increase both the user’s sense of presence, and memory for objects, in a virtual environment. It
is particularly noteworthy that in their study, increasing the level of visual detail did not result in the same increases in sense of presence or recall of objects, demonstrating that virtual environments are not all about the visual. More recent studies have begun to focus upon the effects of different types of soundscape on aspects of recall and memory relating to awareness of detail in virtual environments. For example, Schmidt et al [9] found that spatial audio cues in a 3D virtual environment can be constructed to guide a user’s attention.

There is significantly more literature relating to the effects of music on learning than on the learning effects of soundscapes or acoustic design, and so the major themes are discussed in the following subsection.

2.1 Music and learning

The appreciation of music seems to be a uniquely human trait [10] and there is evidence that it has been a fundamental characteristic that has developed as humans have evolved. For example, Brown et al [11] have demonstrated that the limbic and paralimbic systems in the brains of humans are activated when listening to music. The limbic system supports functions that include behaviour, emotion and long-term memory and is considered to be one of the earliest parts of the human brain to have developed during our evolution; its activation when listening to music is suggestive that music and similar sounds have a long history in human development, linked to emotional rewards from a sense of community and communication and in the formation of memories. Whilst we would not argue that learning is only memory, the two are clearly associated.

The connection between music and learning has been widely researched, both from the point of view of learning to play music [12] and exposure to music, both actively and passively [13]. It can be difficult to find a consensus in the literature about the effects of music on learning, both in relation to listening to music as a precursor to learning, and listening whilst learning. In some studies, music is referred to as a “distraction” [14]. But this definition of distraction is in relation to traditional classroom learning skills that focus upon reading and writing, or concentrating upon the actions of teachers. Matloubi et al [15] suggest that this may be due to competition between music and verbal processing in the brain. Conversely, Ilie and Thompson [16] demonstrated that music listening may have a positive effect on creativity and speed of processing during learning activities, and that these effects are likely to be linked to positive mood and increased arousal.

In studies that focus upon the use of music in online environments that engage learners in more active techniques than reading and writing, positive correlations seem to be demonstrated between both listening to and hearing music (the former suggesting a deliberate act whilst the latter refers more to background music) and improved engagement and learning [17]. There is currently scant research on the effects of music on learning in 3D virtual environments specifically, but in an example of an investigation into the effects of background music and immersive display systems on memory of facts learned, Fassbender et al [18] found that participants in their study had a significantly higher recall of facts when learning in an immersive environment that included instrumental music, in comparison to the same environment with no music. But, different types of music have different characteristics; instrumental music has qualities that differ from vocal music, for example. It is therefore impossible to make a definitive statement about the effect of music on learning as if it were a single phenomenon.

We have discussed how positive effects of music on learning appear to be connected to positive mood and arousal, but personal taste plays an important part in the effects of music on any one individual. There are also differences in the effects of music on individuals with different personal characteristics. For example, research suggests that learners with disparate personality types react differently to music when used as an adjunct to learning. Dobbs et al [14] found that the performance of participants in their study who displayed introvert characteristics tended to be more negatively affected by music when carrying out complex cognitive tasks than those who displayed extravert tendencies. A postulated reason for this is that it supports the Eysenkian hypothesis of the difference in optimum cortical arousal in introverts and extraverts.

So, considering the problems outlined above, we argue that choice, both of whether or not to engage with a musicscape, and which musicscapes a learner might choose to engage with, offers the promise of an adaptable and appropriate approach to engaging with music for learning.

The picture is not completely clear from the findings of research to date. But, there seems to be significant evidence that using music in a 3D learning environment that concentrates upon visual cues and synthetic activities (e.g. exploration through the agency of an avatar), rather than reading and writing text, may enhance learning. If we are to capitalize upon the choice affordances of virtual environments we argue that it is important to understand the mechanisms behind musical perception and its ontological status, and that designing effective musicscapes for 3D virtual learning is dependent upon such understanding. The next section therefore discusses the nature of music.
3. The nature of music

To properly assess music in the context of musicscapes in 3D virtual environments, we will restrict our discussions here to instrumental music, i.e. music without words, rather than risk conflating this complex subject with language. Music has long been regarded as a purely abstract art form due to its apparent lack of representation - how do changes in air pressure manifest representational meaning or refer to the extra-musical? Sound and hence music is self-evidently produced, carried and transmitted through the medium of air and is executed and experienced (at least by the tympanic membrane) as changes in air pressure within a particular range of audible frequencies. But what makes these moving air waves into something that we call “music”? And further still, what is it that transforms this into something that often moves us to the core of our emotional being?

Roger Scruton [19] has written that sounds are secondary objects in a similar way that colours are secondary properties, whereby sounds are not the same as the things that emit them nor the regions of space from which they emanate. In this way, we describe music and its movement metaphorically; we employ spatial terms such as ‘high’, ‘low’, ‘dense’, ‘moving towards’, ‘away from’ and so on. Scruton writes that at the “…most basic apprehension of music there lies a complex system of metaphor, which is the true description of no material fact. [...] Take this metaphor away and you cease to describe the experience of music.”

We hear sounds differently that have been conceptualised in this metaphorical manner, otherwise all we would hear is a series of discrete, unconnected sounds that do not possess the gravitational attraction that we typically ascribe to them. Scruton continues: “to talk of tones is already to talk at the sophisticated level of musical phenomenology.” That music belongs uniquely to the intentional sphere, and not to the material realm means, for Scruton, that we cannot eliminate metaphor from our descriptions of music. And this is what makes the distinction between sound and music; our musical understanding is governed by concepts in response to the unfolding musical object. Music in this view, therefore, is something that is virtual.

To take two examples, the air pressure changes created by a sounding car alarm will typically engender annoyance or sudden fear, but not music. Conversely, listening to the opening bars of ‘Country Honk’ from the Rolling Stones album Let it Bleed (1969), and hearing the staccato car horn in amongst the rhythmic steel guitars, makes it the sort of aesthetic experience that is not created by hearing the same horn on our daily commute to work. The sound produced by the horn does not suddenly become music, but we may intend it as music because of its context. This notion of intentionality is fundamental to our current interest here in the enhancement of immersion and hence engagement within virtual environments, for it aligns itself with the suspension of disbelief necessary for the effective engagement within fictional worlds and synthetic environments.

Crucially, music may appear to emanate from within virtual environments, but in fact it is the one element that crosses freely between the virtual and the real. That is, the actual changes in air pressure arise and resonate in the physical spaces in which we ourselves occupy; it is physically upon us, around us and even within us. Accordingly music, itself a virtual element, paradoxically transcends the virtual; it crosses the illusory border for it was never there in the first place, merely implied. However, it is this ability that very likely affords music much of its emotional potency and certainly its immediacy in contrast to other forms of media.

In film music, computer games and other viable forms of virtuality, music is used to indicate the following: identity (leitmotifs); territory (virtually encloses its own musically-structured world); emotion and mood (use of faster tempi to excite and physically raise heart rate, minor scales for melancholy); style and period (use of musical genres and conventions to elicit a particular historical time or place); anamnesis (music that reminds us of an earlier personal time and is hence associative at an individual level). All of these aspects have the potential to create meaning for the individual, but it is music’s ability to cross into real-world territory that is of particular import. Indubitably, this explains some of music’s power, in certain multimedia contexts, to arouse such strong emotions. Consider the terrifying shower scene in Hitchcock’s Psycho (1960) or the stirring double bass theme in Spielberg’s Jaws (1975); without its context would the music in either case be so affecting? These examples serve to illustrate the employment of music to complement and amplify the mise-en-scene and montage depicting respectively the horrific slaying of the character Marion Crane in Psycho, and the suspenseful shark attacks in Jaws. In this way, the viewer experiences a third meaning; that is, the blend of music and image conflate to produce a new and hitherto unavailable effect. As such, the engaging power of the filmic medium, verisimilitude, immersion and our conviction are all intensified.

In discussing musical meaning, Edward T. Cone [20] commented “no context, no content”; we need, therefore, to know what the signifiers (the music) belong to. A particularly useful example of this is given by Nicholas Cook [21], who cites a television commercial depicting a Citroen sports car wending its way through a beautiful French landscape set to
the opening bars of the Mozart’s Overture to *The Marriage of Figaro*. He uses this example because it offers a very clear platform from which to analyse the ‘communicative context within which this meaning is realized’ - that is, how are the advertisers marketing the car and communicating its selling points. In short, what he demonstrates is the how the music’s attributes are imparted to the car through association: in particular, the music’s ‘liveliness’ and ‘precision’ now become about the engine and the road handling, respectively. Further still, the apparent paradox between the associated technology - the highly-engineered car - and the enduring symbol of high art - the music of Mozart - is cleverly solved through presenting this as an ‘ideal synthesis of art and technology’. Hence, the communicative structure of the commercial is realised through this unusual juxtaposition. We have, then, a further example of a third or ‘emergent’ meaning, the conflation of the visuals, subject and music not only affects the overall message but actually creates it.

This context may, however, be supplied in many forms. The amorphous nature of music attests to its uncanny knack of ‘marrying’ itself to multiple contexts. However, there may well be a limit here; consider for example the shower scene music in * Psycho* being supplanted with Henry Mancini’s Pink Panther theme. Horror immediately becomes farce.

### 3.1 Music and Territory

To continue our explication of territory, two related aspects are time and space. In creating territory (by being music the listener is already complicit in intending this), musical time may be said to be qualitatively different to clock time. Barbara Barry [22] comments that it would seem, when listening to music in an involved and concentrated way, our ‘rule of thumb’ feel for clock time is suspended or blocked out by the network of musical rhythms. If musical time can be described as the interaction between two distinct components, one being the style, harmonic characteristics and internal subdivisions of the piece of music, then the other is the responsive part played by the individual in the perception of time. Music in this view forges its own world; it becomes its own universe in the Deleuzian sense and it has the power to share this territorial universe with the other world. Similarly, Thomas Clifton [23] comments that musical space has significance because listeners find themselves there, as a place to take up what he calls ‘temporary habitation’. If we accept this phenomenological perspective, we can see that music has the capacity to delimit a space or territory for the individual who experiences it as such. And an effect such as this has obvious benefits when combined with 3D synthetic environments in mapping such an area.

### 3.2 Sound and Soundscapes

In a similar way to flamboyantly dressing one’s avatar in lavish costumes or designing exotic and preternatural 3D spaces in which to be present, it is also possible to customize various sounds, alter their intensities and affect the soundscape through user selection in virtual 3D environments. Clearly there is much about these environments in which to ‘lose’ and immerse oneself, but there are also some barriers. In gaming, for example, it is not yet possible to emulate the stimulation of all five senses. Instead, game designers have to suffice with the manipulation of the visual, sonic and haptic stimuli (to-date the latter is somewhat basic and limited in most domestic applications). Where the game designer wishes to invest the gaming experience with a sense of realism, she or he often resorts to an amplification and manipulation of the sonic information. Garner and Grimshaw [24] write that in order to achieve greater realism in gaming, the sonic dimension is required to simulate a sense of olfactory, tactile, and gustatory input via representational inference. Accordingly, they give a rather vivid example: “A visually rendered corpse alone may not trigger an olfactory sensation of decay, but the sound of buzzing flies and wriggling maggots around it has a much greater potential to do so”.

Naturally, film directors and their sound designers also use this technique to enhance realism or to create a particular aesthetic. A good example occurs in Martin Scorsese’s *Taxi Driver* (1975) when the protagonist Travis Bickle drops an effervescent tablet into his glass of water. The camera slowly zooms in towards the fizzing water with a concomitant increase in sound intensity and saturation, the shot switches back to a slow zoom on Travis’s face and gaze and then back to the source of the sound, using the same visual pace and rhythm segueing the two subjects and leaving no doubt about the character’s fixation - and hence ours - to the effervescing liquid. Now this example can be read in multiple ways, the turbulent bubbling (sonically and visually) as the camera reaches its maximum focus, symbolically points to the turmoil and disturbances within Travis’s world. However, what is of particular import for our discussions here is the magnification of sound and visual and the cross-modal excitation that results. The close magnification elicits a deeper immersion with the mise-en-scene, we can almost feel, haptically, the spray from the effervescing glass.

In the design of 3D virtual environments for learning simulations, an awareness and careful manipulation of such sonic augmentation can be very powerful and lead to a much greater immersion for the student or visitor. This assertion can be evidenced by our experiences where the lack of an effective soundscape was felt by students to detract
from the learning affordances of a simulated environment. At the University of the West of England we developed a simulated accident investigation for U.K. environmental health students to practice a full accident investigation. This activity forms an important aspect of the professional practice of U.K. environmental health practitioners in local authorities, and as such is a valuable learning opportunity for students training in this field. An accident scenario was developed in a virtual world and students role-played either witnesses to the accident, or investigators under the Health and Safety at Work Act 1974. The scenario is fully described, along with its evaluation, in Falconer’s paper [26]. Environmental sounds, such as those of fork-lift vehicles, reversing alarms, conveyor-machine noises were not included in this scenario, and during the evaluation phase students commented on how they could only use their sense of sight in undertaking the activity, and that using other senses such as hearing would have increased their sense of presence and immersion.

Participants using virtual environments have the opportunity to make environmental choices and can choose their own soundscapes and musicscapes, as we have previously explained; following our arguments above there is a viable potential for immersion and engagement from doing so. The individual and personal nature of musical experience demands that levels of engagement and immersion will be a consequence of one’s own history and experience, i.e. anamnesis, but we have shown a mechanism. In a related way, psychologists have shown how the emotional value of content improves transfer from short to long term memories [25]. Accordingly, it follows that if we are moved emotionally by something (perhaps the background music that we choose) then any associated emotion will be a consequence of one’s own history and experience, i.e. anamnesis, but we have shown a mechanism. In a related way, psychologists have shown how the emotional value of content improves transfer from short to long term memories [25]. Accordingly, it follows that if we are moved emotionally by something (perhaps the background music that we choose) then any associated experience should become more retrievable at a later stage. The theory here will hopefully help to build a foundation for such an empirical study.

4. Conclusions

The literature and experiences discussed in this paper raise a wide range of issues worthy of research. Here, by way of conclusion, we make 3 suggestions for research topics that might further illuminate the field and lead to subsequent, more detailed investigations.

- The attempt to identify all forms of sound as simply ‘sounds’, and then ask if these have an effect on learning, appears to us to lack the necessary granularity of approach to a complex subject. For example, in anecdotal discussions with students so far, we have found a significant difference between the distinctive and reinforcing effects of vocal and instrumental music. We suggest that the search for a generalizable effect of both soundscapes and musicscapes on learning, where learning is defined solely or largely as reading and writing, be focussed into a more subtle and sophisticated exploration of different kinds of soundscapes and musicscapes, and a wider range of learning activities, including situated and experiential forms in virtual environments.

- It is important to involve learners using virtual environments in early scoping stages of design, and initial research could usefully experiment with different uses of soundscapes and musicscapes in different virtual environments, to identify trends in choices made by learners and the effects of a range of sound and musicscapes on various forms of learning.

- From the point of view of technology, the introduction of 3D virtual headsets such as the Oculus Rift® is beginning to add a deeper sense of immersion to interactions in virtual environments, and future research could explore the effects of a range of soundscapes on learners using these types of device.

5. References


