The Feedback Layer from MAARS: An Audio Layer from the Multilayered Affect-Audio Research System for VR Learning Environments

Alvaro F Olsen  
NYU CREATE Lab  
370 Jay Street  
Brooklyn NY, USA  
al.olsen@nyu.edu

Agnieszka Roginska  
NYU MARL Lab  
35 W. 4th Street  
New York NY, USA  
roginska@nyu.edu

Jan L Plass  
NYU CREATE Lab  
196 Mercer Street  
New York NY, USA  
jan.plass@nyu.edu

ABSTRACT

How does audio impact emotions experienced by learners? This is a highly relevant question since emotions play an important role in learning. Learning materials, such as images and sounds, induce emotions during learning, and, inversely, our emotions influence what information we attend to [26]. The focus of the present paper is on the use of audio as emotional design in immersive virtual reality learning environments. We will introduce the Multilayered Affect-Audio Research System (MAARS) for learning environments in virtual reality [18], and then discuss the use of auditory displays in the Feedback layer from MAARS.

1. INTRODUCTION

Positive emotions have been found to provide better conditions to solve problems [12][20], enhance memory performance [6], improve reasoning [12], facilitate a mental state for relationships to be made between ideas [12], promote creativity, flexibility, divergent thinking [15], and enhance learning outcomes [29]. A new area of emotional design has recently emerged that is concerned with the question how the design of learning materials can be used to induce emotions that are conducive to learning [24]. Understanding how emotions mediate learning is pivotal to how we design learning environments. Most of this research has focused on designing multimedia learning environments that use emotional designs to induce positive emotions via visuals and sound [8][23][25].

Virtual reality technology provides an unprecedented level of opportunities for emotional design [9]. The two key cognitive affordances of virtual reality are presence and agency [16], which are supported by the sophisticated levels of fidelity of visual and auditory content. Yet, a meta-analysis has shown that little is known about designing learning experiences in immersive virtual reality [16]. Research in primary and secondary education [19], higher education [1][17] and professional training [13][30] suggests that virtual reality learning environments increase the level of engagement and participation from learners on learning tasks [2].

2. FRAMEWORKS FOR AUDIO DESIGN

Current frameworks for audio design in video games are not yet standardized [4][10], and moreover, they do not address games, simulations or virtual reality for learning. Current frameworks for audio design in video games are adopted in virtual reality with the added value of 3D and immersive audio. These frameworks cluster audio in various ways, the classic approach is to categorize sound as ambience, music and sound effects, while other approaches separate ambience and music, and some others include instruction and avatar sounds [10].

The IEZA framework proposed by Huiberts and Van Tol is a step towards a standard workflow for audio design and production in video games, as it clusters sound in two dimensions, Diegetic/Non-diegetic sounds in one dimension and Setting/User activated sounds in the other. Each combination produces the four main categories Interface, Effects, Zone and Affect sounds [10], however, the IEZA framework does not address various constructs of interest in learning science such as Instruction, Personalization and Accessibility.

This is the core reason why we propose the Multilayered Affect-Audio Research System (MAARS) as a framework for research of audio, its affective impact on the learner by its various properties such as composition, synthesis and processing, and its affordances and interactions on learning environments in virtual reality [18].

3. MAARS

The Multilayered Affect-Audio Research System for Virtual Reality Learning Environments (MAARS) organizes sound into five layers, namely, Feedback, Experience, Instruction, Social and Data. The layers cluster auditory displays based on their context and function within the learning environment. In essence, this system facilitates the operationality of the various auditory displays, sound sources and their affective value [18].

Figure 1: Multilayered Affect-Audio Research System (MAARS)

Each one of these layers has two affective dimensions based on Russell’s circumplex model of emotions that views emotions as a two-dimensional construct composed of valence in one
dimension and arousal in the other [27]. The valence
dimension indicates whether the emotion is negative or
positive, while the arousal dimension indicates the intensity or
level of activation for any given emotion.

Figure 2: Russell’s Circumplex model of emotions

Our goal is to present to the user sound designs that target
valence and arousal, that is, playing sounds that evoke positive,
negative or neutral emotions, and in addition, playing sounds
with various degrees of intensity, such as to elicit a state of
relaxation or excitement.

Figure 3: Operationalization of emotional designs based on
Russell’s circumplex model of emotions

Furthermore, the MAARS system provides additional
operationality support for audio synthesis and audio
processing, which allows us to compare between procedural
versus non-procedural audio and audio formats including
stereo, binaural, and multi-channel. See Figure 4 for the
schematic of the audio layer operationalization system.

Figure 4: The synthesis, processing and emotional design
operationalization audio workflow of MAARS

Finally, each layer provides additional accessibility support
and a level of personalization that can furthermore cater
individual learners and their various auditory thresholds. Due
to the nature of the system, a full fledged learning environment
would in addition feature ways for the learner to personalize
their auditory landscape by providing means to turn layers’
volume level down or up, and access various accessibility
features such as voice commands, subtitles or haptic feedback
for the deaf or hard of hearing community, commonly referred
in the literature as DHH [13].

4. FEEDBACK LAYER

The feedback layer from MAARS provides a level of control
and precision to the various auditory displays that constitute
this category. The feedback layer contains sounds associated
with user actions. These would include non-diegetic sounds
emitted by the interface and/or performance scores and
diegetic sounds emitted by interactive game objects presented
in the main scene.

Figure 5: The Feedback Layer from MAARS

User actions or user events in the Feedback layer can include
but are not limited to the controller’s trigger pull/release event,
raycast hover/leave events, drag and drop events, push and pull
events, rotating clockwise/counter-clockwise events,
throw/catch events, teleportation and locomotion, and user
interface events such as button presses, panel swipes, zoom ins
and zoom outs, edit/delete/create records, etc.

Figure 6: A sample of user events associated to the audio in
the feedback layer

5. SOUND DESIGN FOR EMOTION

Once the user events are identified the next step is to design
sounds using empirically validated sound features that impact
the listener emotionally in the context of learning. Various
studies in sound and music for learning have demonstrated the
effect of music in the classroom and have made progress
distinguishing what sound constituents are the right kind for a
number of learning contexts [30] [21][14][28].

Figure 7: Russell’s circumplex model of emotions linked with
music constituents found to evoke emotions.
Figure 7 provides a schematic representation of the circumplex model of emotions proposed by Russell (1980) along with the sound constituents linked to the corresponding emotion that they mediate based on and according to various studies that have been conducted on music in the classroom. [7][14][30][28][21].

6. EVALUATION AND NEXT STEPS

The next step for MAARS is to run empirical studies to gauge the possibility that emotion induction via sound design in virtual reality learning environments is attainable and viable in practice. The study would contain sound designs that target negative and positive emotions with various degrees of arousal potential. The feedback layer auditory displays are hypothesized to increase learner’s engagement and positive emotions as well as to provide the information necessary for learners to operate the learning environment. For this, we have designed a series of studies that will test these theories and that will be using surveys and non-invasive biometric sensors to measure psychophysiological responses to the auditory stimuli. We will be using the International Affective Digitized Sounds or IADS [31] to calibrate our sensing tools and will triangulate these measures with surveys, such as the Self-Assessment Manikin (SAM) and the Positive and Negative Affect Schedule (PANAS), along with behavioral data from game logs. We hope to find evidence that indicates that our sound designs evoke an emotional response and that such response is congruent with our IADS controls and self-evaluation measures.

7. REFERENCES


Quality of Expression, Color, Shape, and Dimensionality of Game Characters. Learning and Instruction, 7, 101194. Doi: 10.1016/j.learninstruc.2019.01.005, 2020.


