JUSTIN CHRISTENSEN

Virtual Distance and Soundstage, and their Impacts on Experienced Emotional Valence

Introduction

For this paper, I borrow theories from animal ethology research, which suggest that varying levels of valence are experienced in response to being varying distances from a predator. Adopting ideas from the theory of refinement, I examine whether these survival instincts have relevance for a music listening experience, focusing on how a listener's virtual distance from the signal source might alter their experienced emotional valence of music. As well, I investigate whether changes to distance cues affect all of the presented basic emotion categories (happy, sad, fearful, calm) similarly, or if changes to distance cues affect each of the basic emotional categories differently. I also add a third distance condition, including trials that combine audio and haptic stimulation in the nearfield. This follows from the theory that multimodal experiences are often deemed to be more trustworthy signals.

Links between Aesthetic Emotions and Survival Emotions

Prior to 2000, the bulk of music research focused on music as a cultural invention. As a result, 'biomusicology' with its investigations into the biological functions of music is still a relatively new and growing field.[1] In 2007, Frijda and Sundararajan presented a theory of refinement[2] as an elaboration on James' theory of emotions,[3] where aesthetic emotional responses are seen as refined forms of more coarse survival responses. 'Refined' for this theory means that the emotions do not obviously manifest themselves in overt behaviors such as fight or flight responses. This theory of refinement suggests that emotional responses induced by aesthetic response are the same as emotions aroused by survival instincts, differing mainly in regards to the magnitude of the emotional response, and their resultant behavioral and physiological

---

[1] Nils Lennart Wallin and Björn Merker, *The Origins of Music* (MIT Press, 2001.

[2] Nico H. Frijda and Louise Sundararajan, "Emotion Refinement: A Theory Inspired by Chinese Poetics," *Perspectives on Psychological Science* 2, no. 3 (September 2007), 227–41, doi:10.1111/j.1745-6916.2007.00042.x.

[3] William James, "II.—What Is an Emotion?," *Mind*, no. 34 (1884), 188–205.

responses. This fits well with Patel's "transformative technology of the mind" theory. Patel states:

> "neuroscientific research suggests that music is an invention that builds on diverse pre-existing brain functions, rather than a trait that originated via processes of natural selection (…) growing evidence from neuroscience also suggests that music is biologically powerful, meaning that it can have lasting effects on non-musical abilities…"[4]

As further evidence for music's connection to everyday goal-directed emotions, Peretz has hypothesized that music's universal appeal is likely due to music being "…particularly well suited (or designed) to invade emotion circuits that have evolved for emotional vocalizations."[5]

While musical emotions are often considered personal and enigmatic, emotional responses to music within a 'basic emotion' framework (happy, sad, fearful) are remarkably invariant across listeners,[6] and are already easy to distinguish by children already by the age of six.[7] Schmidt and Trainor also found that emotional valence (attractiveness vs. averseness) in music is also significantly correlated to specific brain activation patterns as measured by EEG.[8] Several other EEG studies by Altenmüller et al.,[9] Flores-Gutiérrez et al.,[10] and Tsang et al.[11] also support these results for hemispheric lateralization of function in music listening, where "Positive emotions were related to a preponderance of left frontal activation whereas negative emotions resulted in a more bilateral fronto-temporal activation with preponderance of the right hemisphere."[12] These results show that there are numerous benefits for using music to study the neural correlates of emotion. Stefan Koelsch has listed a number of them that are applicable for the current study.

---

[4] Aniruddh D. Patel, "Music, Biological Evolution, and the Brain," in *Emerging Disciplines*, ed. C. Levander & C. Henry, (Texas, Rice University Press: 2010), 91–144.

[5] Isabelle Peretz, William Aubé, and Jorge L. Armony, "Towards a Neurobiology of Musical Emotions," *The Evolution of Emotional Communication: From Sounds in Nonhuman Mammals to Speech and Music in Man*, (Oxford University Press, 2013), 277.

[6] Ibid.

[7] Joseph G. Cunningham and Rebecca S. Sterling, "Developmental Change in the Understanding of Affective Meaning in Music," *Motivation and Emotion* 12, no. 4 (December 1988), 399–413, doi:10.1007/BF00992362.

[8] Louis A. Schmidt and Laurel J. Trainor, "Frontal Brain Electrical Activity (EEG) Distinguishes Valence and Intensity of Musical Emotions," *Cognition & Emotion* 15, no. 4 (2001), 487–500, doi:10.1080/0269993004200187.

[9] Eckart Altenmüller et al., "Hits to the Left, Flops to the Right: Different Emotions during Listening to Music Are Reflected in Cortical Lateralisation Patterns," *Neuropsychologia* 40, no. 13 (2002): 2242–56, doi:10.1016/S0028-3932(02)00107-0.

[10] Enrique O. Flores-Gutiérrez et al., "Metabolic and Electric Brain Patterns during Pleasant and Unpleasant Emotions Induced by Music Masterpieces," *International Journal of Psychophysiology* 65, no. 1 (July 2007), 69–84, doi:10.1016/j.ijpsycho.2007.03.004.

[11] C. D. Tsang et al., "Frontal EEG Responses as a Function of Affective Musical Features," *Annals of the New York Academy of Sciences* 930, no. 1 (June 2001), 439–42, doi:10.1111/j.1749-6632.2001.tb05764.x.

[12] Altenmüller et al., "Hits to the Left, Flops to the Right," 2242.

> "…(1)Music is capable of evoking *strong emotions* (…) (5) Both listening to music and making music can evoke emotions, enabling investigators to study interactions between emotion and *action*. (6) Music can be used to study the *time course* of emotional processes, with regard to both short-term emotional phenomena (in the range of seconds) and longer-term emotional phenomena (…) (7) It appears that, with regard to human evolution, music is originally a social activity. Therefore, music is well suited to study interactions between emotion and *social factors*. [emphasis in the original]"[13]

Emotional Valence and Distance

The present study explores how acoustic processing of musical examples (from differing basic emotional categories: happy, sad, fearful, calm) to simulate acoustic spaces, where the virtual sound sources appear at varied virtual distances from the listener, might evoke distinct emotional approach/withdrawal motivations in subjects, seen through anterior brain asymmetry of induced EEG event-related oscillations (EROs). Emotional valence has often been tied to action tendencies towards rewards (appetitive stimuli) and away from danger or punishment (aversive stimuli). Valence falls under appraisal theories of emotion, where "A common tenor of these theories is that (external or internal) phenomena, circumstances, actions, individuals, or objects are evaluated as "good", that is, contributing to achieving a goal, or "bad," that is, obstructing the achievement of a goal".[14] The use of virtual distance for this study is inspired by work in animal ethology,[15] and by the work of David Huron.[16]

Animal ethology research has shown that there are changes to the levels of both physiological arousal and action tendency motivations (valence) that result from varying an animal's distance from a predator. Bracha[17] presents the theory that a freeze response of an animal is activated when flight or fight are no longer effective options, when a danger is too close and too oppressive to effectively escape from. Relatedly, according to the "distance-dependent defense hierarchy"[18], a close predator induces a fight response, while a moderately close predator induces a flight response. Combined, these two theories present three different distance associations for three of the *F*'s of a *fight-or-flight* response. In addition to these categories of fear responses with their distinct distance associations, anticipatory anxiety responses can frequently be induced in response to distant perceived threats.

---

[13] Stefan Koelsch, *Brain and Music* (John Wiley & Sons, 2012), 203.

[14] Stefan Koelsch, "Music-Evoked Emotions: Principles, Brain Correlates, and Implications for Therapy," *Annals of the New York Academy of Sciences* 1337, no. 1 (March 2015), 193, doi:10.1111/nyas.12684.

[15] John Maynard Smith and David Harper, *Animal Signals* (Oxford: Oxford University Press, 2003).

[16] David Huron, "Affect Induction through Musical Sounds: An Ethological Perspective," *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 370, no. 1664 (March 2015): 20140098, doi:10.1098/rstb.2014.0098; David Huron, "Understanding Music-Related Emotion: Lessons from Ethology," 2012, 23–28.

[17] H. Stefan Bracha, "Freeze, Flight, Fight, Fright, Faint: Adaptationist Perspectives on the Acute Stress Response Spectrum.," *CNS Spectrums* 9, no. 09 (September 1, 2004), 679–85, doi:10.1017/S1092852900001954.

[18] Stanley C. Ratner, "Immobility of Invertebrates: What Can We Learn?," *The Psychological Record*, (1977).

Anxiety responses are often considered ameliorated fear responses, neither as immediately aversive nor inducing as much arousal in the sympathetic nervous system as fear.[19] In comparing fear responses, the "distance-dependent defense hierarchy" found that the emotional quality of a fight response was both more negatively motivating and caused higher levels of arousal in the prey than the emotional quality of a flight response.[20] In a slight variation from these other anxiety and fear responses, which show a direct negative correlation between the averseness experienced and the arousal levels exhibited, a fright response displays an acutely negative valence with a reduced arousal level. This combination is likely the result of the prey's intention to reduce movement as much as possible as part of the freeze response.[21]

Related to ethological action tendencies, Isobel Peretz considers action tendencies and movement to be key components of emotional reactions to music. "Much music is composed with very specific action tendencies in mind – getting people to sing, to dance, march and fight, or work and play together."[22] While these action tendencies for music are unrelated to animal fear response action tendencies, following the theory of refinement mentioned earlier, music's action tendencies can be seen to be more refined versions of the coarse animal survival instincts. This is further supported by music having similar reward motivated brain activations towards food[23] and drugs,[24] and away from threats.[25] Juslin and Västfjäll suggest that music uses mimesis (or emotional contagion) to express emotional valence, as music can mimic the 'natural emotional attributes' that are found in emotional speech signals (as has been investigated in Lundqvist et al.[26], Peretz[27], and Lima & Castro[28]).[29] As a result, I chose the musical selections for this investigation to follow

---

[19] Bruce Duncan Perry and Maia Szalavitz, *The Boy Who Was Raised as a Dog: And Other Stories from a Child Psychiatrist's Notebook: What Traumatized Children Can Teach Us about Loss, Love, and Healing* (New York: Basic Books, 2006).

[20] Ratner, "Immobility of Invertebrates."

[21] Stefan Bracha, "Freeze, Flight, Fight, Fright, Faint: Adaptationist Perspectives on the Acute Stress Response Spectrum."

[22] Isabelle Peretz, "The Nature of Music from a Biological Perspective," *Cognition*, The Nature of Music, 100, no. 1 (May 2006), 22–23, doi:10.1016/j.cognition.2005.11.004.

[23] Dana M. Small et al., "Changes in Brain Activity Related to Eating Chocolate," *Brain* 124, no. 9 (September 2001), 1720–33, doi:10.1093/brain/124.9.1720.

[24] Anne J. Blood and Robert J. Zatorre, "Intensely Pleasurable Responses to Music Correlate with Activity in Brain Regions Implicated in Reward and Emotion," *Proceedings of the National Academy of Sciences* 98, no. 20 (September 2001), 11818–23, doi:10.1073/pnas.191355898.

[25] Nathalie Gosselin et al., "Amygdala Damage Impairs Emotion Recognition from Music," *Neuropsychologia* 45, no. 2 (2007), 236–44, doi:10.1016/j.neuropsychologia.2006.07.012; Patrik N. Juslin and Daniel Västfjäll, "Emotional Responses to Music: The Need to Consider Underlying Mechanisms," *Behavioral and Brain Sciences* 31, no. 05 (October 2008), 559–75, doi:10.1017/S0140525X08005293.

[26] Lars-Olov Lundqvist et al., "Emotional Responses to Music: Experience, Expression, and Physiology," *Psychology of Music*, (October 2008), doi:10.1177/0305735607086048.

[27] Peretz, Aubé, and Armony, "Towards a Neurobiology of Musical Emotions."

[28] César F. Lima and São Luís Castro, "Speaking to the Trained Ear: Musical Expertise Enhances the Recognition of Emotions in Speech Prosody," *Emotion* 11, no. 5 (2011), 1021–31, doi:10.1037/a0024521.

[29] Patrik N. Juslin, László Harmat, and Tuomas Eerola, "What Makes Music Emotionally Significant? Exploring the Underlying Mechanisms," *Psychology of Music*, (August 2013), 1-25, doi:10.1177/0305735613484548.

the natural attributes of emotional speech patterns. This will be discussed more in the methodology section below.

For evaluating the valence aspect (approach withdrawal motivations) of emotions evoked in the subjects of this study, I measured the subjects with EEG. In past EEG studies, both approach/withdrawal motivations[30] and positive or negative valence[31] have been found to be correlated with EEG hemispheric asymmetry. Activity in the left prefrontal cortex has been associated with approach motivations and positive valence, while right prefrontal cortex activity has been associated with withdrawal motivations and negative valence. Connected with this, Rutherford and Lindel have associated positive emotions, such as happiness, with a motivation to move towards this cause of happiness. They similarly have associated negative emotions, such as fear or sadness, [32] with a motivation to withdraw from the situation that causes this negative emotion.[33]

The hemispheric asymmetry in the brain linked with emotional valence is observed by measuring *alpha lateralization* (the inverse power in the Alpha band[34] of neural oscillations, now defined as 8-12 Hz.), as lower Alpha power has been connected with greater neural activity in the corresponding brain region.[35] These hemispheric asymmetries measured by alpha lateralization have already found to be a good fit with self-reported affective responses both for positive and negative film clips[36] and for music examples,[37] so this gives promise for this methodology. Recently, Kragel and LaBar[38] also found distinct dimensional mapping between differing emotional states through FMRI testing, which is quite promising for a more precise neural mapping of emotional states.

While these results all support significant correlations between localizations of brain activation patterns and emotional valence, the emotions perceived often differ from the emotions evoked in the listeners. Zentner and Scherer state:

---

[30] Eddie Harmon-Jones and J.J.B. Allen, "Anger and Frontal Brain Activity: EEG Asymmetry Consistent with Approach Motivation despite Negative Affective Valence," *Journal of Personality and Social Psychology* 74, no. 5 (1998): 1310–16, doi:10.1037/0022-3514.74.5.1310.

[31] Robert E. Wheeler, Richard J. Davidson, and Andrew J. Tomarken, "Frontal Brain Asymmetry and Emotional Reactivity: A Biological Substrate of Affective Style," *Psychophysiology* 30, no. 1 (January 1993), 82–89, doi:10.1111/j.1469-8986.1993.tb03207.x.

[32] With the exception of anger, which is a negative, but approach motivated emotion.

[33] Helena J. V. Rutherford and Annukka K. Lindell, "Thriving and Surviving: Approach and Avoidance Motivation and Lateralization," *Emotion Review* 3, no. 3 (July 2011), 333–43, doi:10.1177/1754073911402392.

[34] Other frequency bands of neural oscillations are named Delta (0-4 Hz.), Theta (4-8 Hz.), Beta (12-30 Hz.), Gamma (30-80 Hz.)

[35] DB Lindsley and JD Wicke, "The Electroencephalogram: Autonomous Electrical Activity in Man and Animals," *Bioelectric Recording Techniques*, 1974.

[36] Andrew J. Tomarken, Richard J. Davidson, and Jeffrey B. Henriques, "Resting Frontal Brain Asymmetry Predicts Affective Responses to Films," *Journal of Personality and Social Psychology* 59, no. 4 (1990): 791–801, doi:10.1037/0022-3514.59.4.791; Wheeler, Davidson, and Tomarken, "Frontal Brain Asymmetry and Emotional Reactivity."

[37] Tsang et al., "Frontal EEG Responses as a Function of Affective Musical Features"; Altenmüller et al., "Hits to the Left, Flops to the Right."

[38] Philip A. Kragel and Kevin S. LaBar, "Multivariate Neural Biomarkers of Emotional States Are Categorically Distinct," *Social Cognitive and Affective Neuroscience*, (March 2015), nsv032, doi:10.1093/scan/nsv032.

> "The aforementioned research tradition made valuable contributions to an understanding of music-specific affects—for example, by pointing to the possibility that canonical emotion labels may not do justice to the emotions evoked by music".[39]

They also found that while positive emotions are aroused similarly to how they are perceived, negative emotions are more often perceived than felt. The studies by Gabrielsson,[40] and Kallinen and Ravaja[41] both support this view, with findings that sad and fearful music are still quite likely to evoke positive emotional valence in listeners.

The trustworthiness of a signal: multimodal vs. unimodal

As an added component to my explorations with this study, I added a multimodal nearfield condition to the stereo audio nearfield and stereo audio virtually distant conditions. I did this as I was interested in studying whether the addition of a multimodal component could increase the trustworthiness or closeness of the signal. I only added the haptic stimulation to the nearfield condition, since tactile stimulation is limited by the length of a subject's limbs, and thus this biological limitation might further increase the trustworthiness of a close signal, and add to a greater sense of nearness.

As one means for deciding which stimuli animals should motivate themselves towards or away from, animals judge the credibility of the signal given off by stimuli in their environment. Scott-Phillips has described a signal as:

> "Any act or structure that (i) affects the behaviour of other organisms; (ii) evolved because of those effects; and (iii) which is effective because the effect (the response) has evolved to be affected by the act or structure."[42]

As a result, trustworthy signals given off by animals are highly social encounters, benefitting both the sender and receiver of the signal. These signals are often multimodal, as trustworthiness can be heightened through having trustworthy performances registered across multiple perceptual modalities.[43]

Among these trustworthy signals, we can include many human emotional expressions. Juslin usefully summarized Scherer's definition of emotion[44] as:

---

[39] Marcel Zentner, Didier Grandjean, and Klaus R. Scherer, "Emotions Evoked by the Sound of Music: Characterization, Classification, and Measurement," *Emotion* 8, no. 4 (2008): 496, doi:10.1037/1528-3542.8.4.494.

[40] Alf Gabrielsson, "Emotion Perceived and Emotion Felt: Same or Different?," *Musicae Scientiae* 5, no. 1 suppl (September 2002), 123–47, doi:10.1177/10298649020050S105.

[41] Kari Kallinen and Niklas Ravaja, "Emotion Perceived and Emotion Felt: Same and Different," *Musicae Scientiae* 10, no. 2 (September 2006), 191–213, doi:10.1177/102986490601000203.

[42] T. C. Scott-Phillips, "Defining Biological Communication," *Journal of Evolutionary Biology* 21, no. 2 (March 2008), 388, doi:10.1111/j.1420-9101.2007.01497.x.

[43] Ibid.

[44] Klaus R. Scherer, "Psychological Models of Emotion.," in *The Neuropsychology of Emotion.*, ed. Joan Borod (New York: Oxford University Press, 2000), 137–62.

> "…emotions are relatively brief and intense reactions to goal-relevant changes in the environment that consist of many subcomponents: cognitive appraisal, subjective feeling, physiological arousal, expression, action tendency, and regulation…"[45]

Emotions have evolved and are an influential part of social practice as they reliably convey information of interest to both the sender and the receivers of these emotional expressions through changes in speech prosody, speech rate, facial expression, changes in gestures, general body language and action tendencies.[46]

For sound localization, humans use micro-time deviations between the ears, loudness cues, spectral information, as well as correlations from early and late reflections.[47] Compiled together, these features help give cues to localize a sound source. The more these features have a combined agreement with past experiences of sound localizations, the more easily the listener adapts to them and accepts them as being trustworthy.[48] The soundstage of a recording is the recreation of a virtual musical experience in a 3-dimensional space. From the recording, the listener is often able to localize the individual sound sources in this virtual space, and derive cues for their distance from the source.[49] This current study explores how these distance cues can alter the evoked emotional response for musical examples with diverse approach/withdrawal motivations.

My hypothesis is that musical stimuli should cause stronger valenced responses in the nearfield than at a distance regardless of whether they are negative or positive. The motivation to approach appetitive stimuli in the nearfield should be exhibited more strongly in the nearfield than from a distance. Likewise, motivations to withdraw from aversive stimuli should be exhibited more strongly in the nearfield than from a distance. Thus, music experienced as being negatively valenced at a distance should be even more negatively valenced in nearfield, and music that is experienced as having a positive valence at a distance should be even more positively valenced in nearfield.

Method

Subjects

This study included 12 participants with a mean age of 24 years, including 7 males and 5 females, with 8 of them having had over 5 years of musical training. The data from two subjects were removed. One subject self-reported as being left-handed, and since handedness has in the past shown an effect on alpha lateralization, their results were removed prior to data analysis. The EEG results of a second subject included

---

[45] Patrik N. Juslin and Petri Laukka, "Communication of Emotions in Vocal Expression and Music Performance: Different Channels, Same Code?," *Psychological Bulletin* 129, no. 5 (2003), 770–814, doi:10.1037/0033-2909.129.5.770.

[46] Klaus R. Scherer and Paul Ekman, *Approaches To Emotion* (Psychology Press, 2014).

[47] Jens Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localization* (MIT Press, 1997).

[48] P. Zahorik, E. Brandewie, and V. P. Sivonen, "Auditory Perception in Reverberant Sound Fields and Effects of Prior Listening Exposure," *Principles and Applications of Spatial Hearing*, (2011), 24–34.

[49] See also Mads Walther-Hansen, *Sound Events, Spatiality and Diegesis – The Creation of Sonic Narratives in Music Productions*. p. 29-46 in this issue.

frequent and severe artifacts in the AF4 electrode, so their results were also removed prior to analysis. Correspondingly, the final collected results were drawn from 6 male and 4 female participants.

Apparatus

This experiment was conducted using a stereo setup with KH120A speakers and SVS sub. For creating a haptic response, a Clark synthesis TST239 tactile transducer was used, and it was connected to the stereo through a stereo to mono bridge (with applied resistors) and then passed through its own 200 watt amplifier. The EEG device used was a 14 channel emotiv epoc headset with software development kit, consistent with the international 10-20 EEG electrode placement,[50] referenced with a left-hand CMS (Common Mode Sense active electrode[51]).

The music used for the experiment was: 1. a slightly sped-up and transposed-up excerpt of *Hoppipolla* from Sigur Ros (happy), 2. an excerpt from Arvo Pärt's, *Cantus in memoriam Benjamin Britten* (sad), 3. A slowed-down excerpt of Imogen Heap's *hide and seek*, stretched to 8 times the original duration by using PaulStretch[52] (calm), 4. *Streets of Pripyat* from the *S.T.A.L.K.E.R.: Shadow of Chernobyl* soundtrack (fear). These musical examples were collected from a number of online sources[53] that rated them as being exemplars of their respective emotional categories. I initially chose a few exemplars from each of these 4 emotional categories, and further selected from these based on their psychoacoustic features, based on appraisal theories of emotion.[54] Furthermore, none of the musical selections contained either sung or spoken text. Both the *Hoppipolla* and *hide and seek* excerpts were taken from karaoke versions, with the solo voice stripped from the track.

Distant versions of the musical excerpts were prepared with distance-related head-related transfer functions in Ircam Spat[55] to convey a sense of distance. I processed the distant sources through Spat at a radiated distance of 20.52 meters from the virtual listener, with a reduced room presence (seen in figure 1 below) and a reduced room live-

---

[50] Richard W Homan, John Herman, and Phillip Purdy, "Cerebral Location of International 10–20 System Electrode Placement," *Electroencephalography and Clinical Neurophysiology* 66, no. 4 (April 1987): 376–82, doi:10.1016/0013-4694(87)90206-9.

[51] The CMS is a reference electrode which all other electrodes are measured against to help reduce movement artifacts.

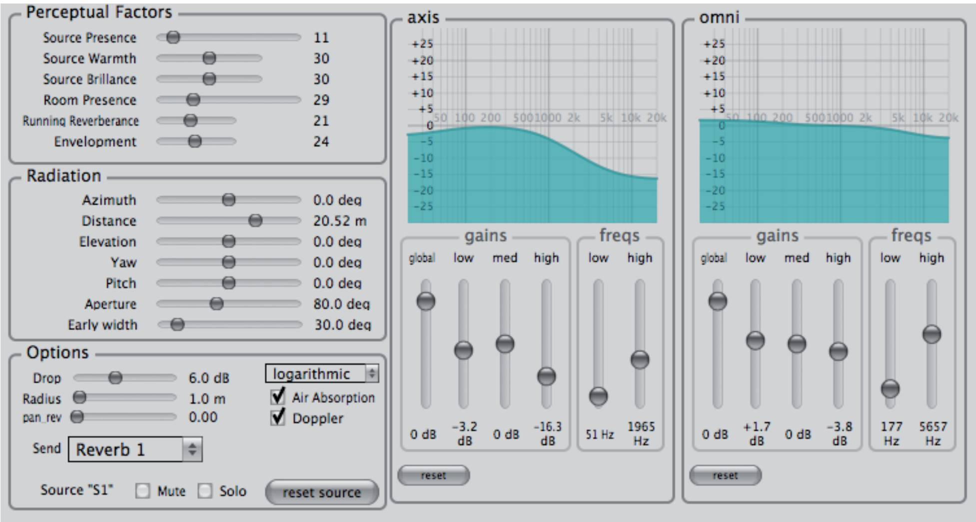
[52] Computer application for extreme stretching of digital audio files.

[53] Among the online lists were:
http://www.nme.com/photos/50-most-uplifting-songs-ever/279564 ;
http://www.npr.org/sections/deceptivecadence/2010/09/27/130157375/the-saddest-music-in-the-world-7-tunes-to-make-you-tear-up ;
http://www.thefourohfive.com/music/news/article/top-10-horror-video-game-music-themes ;
https://michelsabbagh.wordpress.com/2015/08/17/7-video-game-soundtracks-that-give-me-goose-bumps/ ; Paulstretch calm ambience discussed in B Sturm, "Generation of ambient background textures from a given waveform," (master thesis, Aalborg University Copenhagen, 2011).

[54] David I. Leitman et al., "Getting the Cue: Sensory Contributions to Auditory Emotion Recognition Impairments in Schizophrenia," *Schizophrenia Bulletin* 36, no. 3 (May 2010), 545–56, doi:10.1093/schbul/sbn115.

[55] Jean-Marc Jot and Olivier Warusfel, "Spat: A Spatial Processor for Musicians and Sound Engineers," In *CIARM: International Conference on Acoustics and Musical Research* (1995).

ness. These values were lowered, as I felt this gave the distant sounds a slightly more realistic feel. All other parameters were left at their standard preset values. Figure 1 below shows the configuration for perceptual factors in Spat, used for the creation of the virtual distant listening situation. All of these parameters remained unchanged for the near-field setting, except for the distance setting (under Radiation), which was changed from 20 meters to 1 meter. Spat automatically alters other parameters (such as source presence) to match the change of distance. After the excerpts were prepared with Spat, they were normalized to -18dB LUFS (Loudness Units relative to Full Scale[56]). Perceived loudness has shown to have a strong effect on evoked emotional responses to music,[57] so the aspect of loudness as a contributing factor to distance was removed to minimize the perception of loudness itself as a possible extraneous experience variable. This changing of the virtual distance of the recording is largely considered a change to the 'soundstage' of the recording, altering the depth and richness of the original recording to give a new soundstage. For this experiment, the original soundstage was not removed from the recording prior to adding new distance cues with Spat.


Perceptual Factors
Source Presence 11
Source Warmth 30
Source Brillance 30
Room Presence 29
Running Reverberance 21
Envelopment 24

Radiation
Azimuth 0.0 deg
Distance 20.52 m
Elevation 0.0 deg
Yaw 0.0 deg
Pitch 0.0 deg
Aperture 80.0 deg
Early width 30.0 deg

Options
Drop 6.0 dB
Radius 1.0 m
pan_rev 0.00
logarithmic
☑ Air Absorption
☑ Doppler
Send Reverb 1
Source "S1" ☐ Mute ☐ Solo

axis
gains: global 0 dB, low -3.2 dB, med 0 dB, high -16.3 dB
freqs: low 51 Hz, high 1965 Hz

omni
gains: global 0 dB, low +1.7 dB, med 0 dB, high -3.8 dB
freqs: low 177 Hz, high 5657 Hz

Figure 1: Perceptual configuration of Spat for virtually distant listening situation.

Design

The experiment contained 3 distance conditions (far, near, near with haptic stimulation), 4 emotion conditions (happy, calm, sad, fear), and each musical example was presented in each condition to each listener as a repeated measures test. The order of the musical excerpts was randomized according to a Latin square to minimize order effect.

[56] Esben Skovenborg, "Loudness Range (LRA) – Design and Evaluation," *Audio Engineering Society Convention* 132 (April 2012): 1–12.

[57] Laura-Lee Balkwill and William Forde Thompson, "A Cross-Cultural Investigation of the Perception of Emotion in Music: Psychophysical and Cultural Cues," *Music Perception: An Interdisciplinary Journal* 17, no. 1 (October 1999), 43–64, doi:10.2307/40285811.

Each listener's alpha lateralization was taken against his or her own baseline, which was calculated as an average from a combination of the EEG recording period prior to testing, as well as an average of the alpha lateralization during the testing period. This way, I could measure the emotional valence as motivations negatively away from or forwards from their baseline.

Procedure

EEG data was recorded in OpenViBE[58] at a rate of 128 Hz. These EEG recordings were supported with subject feedback responses. The EEG recordings were then analysed through the use of EEGlab[59] and ERPlab[60] in Matlab. Data was passed through a 2-way least squares FIR bandpass filter, using a high-pass filter of 180 poles, attenuated to -6 dB at 2.5 Hz. The EEG signals were then segmented into six 6 second epochs (768 samples at 128 Hz) per musical excerpt. These epochs were chosen from the last 45 seconds of each musical excerpt (which are roughly 60 seconds in duration for each), to allow a training period for the listeners to better reach the desired evoked emotional state. Much of the work of preparing an EEG signal for analysis are methods of noise reduction or artifact removal, or methods for increasing the signal-to-noise ratio. Here, the high-pass FIR reduces noise from eye-movement and muscle activity, while the epoch segmentation increases the signal-to-noise ratio.

Artifacts were then corrected with a combination of AMICA and wavelet noise reduction. AMICA makes an independent component analysis (ICA[61]) decomposition of input data, running multiple ICA models for best fit. Once AMICA separated the EEG signals into individual components, I ran a stationary wavelet transform (SWT[62]) wavelet noise reduction on the EEG components to remove artifacts. This SWT noise reduction used a HAAR wavelet wave shape, 5 levels, and was thresholded with a Rigorous SURE algorithm (an adaptive thresholding based on Stein's Unbiased Risk Estimator). The HAAR wavelet shape gives a good compact representation of certain types of artifacts (e.g. eye blinks) and the adaptive thresholding relies on the Gaussian distribution of EEG signals. As a result, the SWT keeps the artifact and removes the EEG signal. A clean EEG residual signal is recovered by subtracting the artifact signal from the original. I implemented my wavelet noise reduction mostly following the method-

---

[58] Yann Renard et al., "OpenViBE: An Open-Source Software Platform to Design, Test, and Use Brain–Computer Interfaces in Real and Virtual Environments," *Presence: Teleoperators and Virtual Environments* 19, no. 1 (February 2010), 35–53, doi:10.1162/pres.19.1.35.

[59] Arnaud Delorme and Scott Makeig, "EEGLAB: An Open Source Toolbox for Analysis of Single-Trial EEG Dynamics Including Independent Component Analysis," Journal of Neuroscience Methods 134, no. 1 (March 2004), 9–21, doi:10.1016/j.jneumeth.2003.10.009.

[60] Javier Lopez-Calderon and Steven J. Luck, "ERPLAB: An Open-Source Toolbox for the Analysis of Event-Related Potentials," *Frontiers in Human Neuroscience* 8 (April 2014), doi:10.3389/fnhum.2014.00213.

[61] ICA attempts to separate independent signals from the mixed signals in each electrode through blind source separation. As seen in: Pierre Comon, "Independent Component Analysis, A New Concept?," *Signal Processing*, Higher Order Statistics, 36, no. 3 (April 1994), 287–314, doi:10.1016/0165-1684(94)90029-9.

[62] G.P. Nason and B.W. Silverman, "The Stationary Wavelet Transform and Some Statistical Applications," in *Wavelets and Statistics*, ed. Anestis Antoniadis and Georges Oppenheim (Springer New York, 1995), 281–99, doi:10.1007/978-1-4612-2544-7_17.

ology of Akhtar et al.[63] Finally, a liberal threshold of (-200 and 200 µVolt) was used to discard extreme values. This combination of SWT and AMICA is a very powerful artifact removal tool, especially for the removal of eye blinks and movement artifacts.

The Hemispheric asymmetry scores were then computed by subtracting the PSD (power spectrum density) of channels in the left hemisphere from their homologs in the right hemisphere.[64] For this, I used an implementation of Welch's method to compute the PSD.[65] I based my measurements from the combination of log-transformed results from the F3/4, F7/8, and AF3/AF4 EEG electrodes. These results were then saved into bins and summed for all of the participants. Following this, the results were indexed by their categories of virtual distance and evoked basic emotion to allow for statistical comparisons of variance against the results from other categories of distance and emotion.

Results

A two-way repeated measures ANOVA was run to determine the effect of different virtual distances for music evoking different emotions (happy, sad, calm, fear). All of the music listening situations were normally distributed ($p > .05$), as assessed by Shapiro-Wilk's test of normality on the studentized residuals. The data also contained no outliers, as assessed by no studentized residuals greater than ± 3 standard deviations. Three data points fall outside of ± 2 standard deviations, and they can be seen in the boxplot below in Figure 2. The Shapiro-Wilk's test is run to ensure that the results come from a normally distributed population, and that noise is not a major contributing factor that skews the final results.

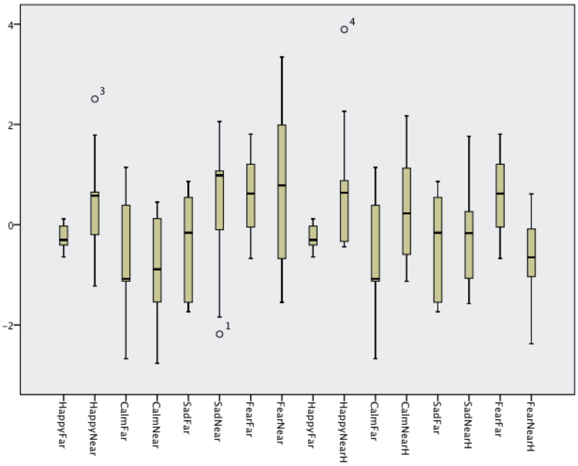


Figure 2: Boxplot of mean alpha lateralization with their standard deviations across all 12 conditions. Distant conditions are repeated on right to provide symmetry

---

[63] Akhtar, Muhammad Tahir, Wataru Mitsuhashi, and Christopher J. James. "Employing spatially constrained ICA and wavelet denoising, for automatic removal of artifacts from multichannel EEG data." *Signal Processing* 92, no. 2 (2012): 401-416.

[64] Cheryl L. Dickter and Paul D. Kieffaber, *EEG Methods for the Psychological Sciences* (SAGE, 2013).

[65] "…Welch's method improves the reliability of the PSD by dividing a discrete time signal into a subset of shorter segments, applying the DFT to each of the segments in the subset and then averaging the results ." in Ibid, p. 106.

When not using haptic stimulation in the nearfield listening situations, there was no statistically significant ($p < .05$) two-way interaction between distance and emotional quality of the music, $F= 0.742$, $p = .537$. However, there was a significant variance between the emotional conditions F=3.673, p=.026. Mauchly's test of sphericity indicated that the assumption of sphericity was met for the interactions of the emotional conditions, $\chi^2(2) = 0.442$, $p = .363$. This suggests that changes to distance cues do not affect the presented basic emotion categories (happy, sad, fearful, calm) differently from one another.

In figure 3 below, one can see that the individual emotional conditions move nearly in parallel with one another (with the exception of the calm condition). When this occurs, even when there is a reasonable distance between the conditions, the interaction effect is not likely to be statistically significant. Significant interactions between conditions are much more likely to be seen when the conditions take different courses.

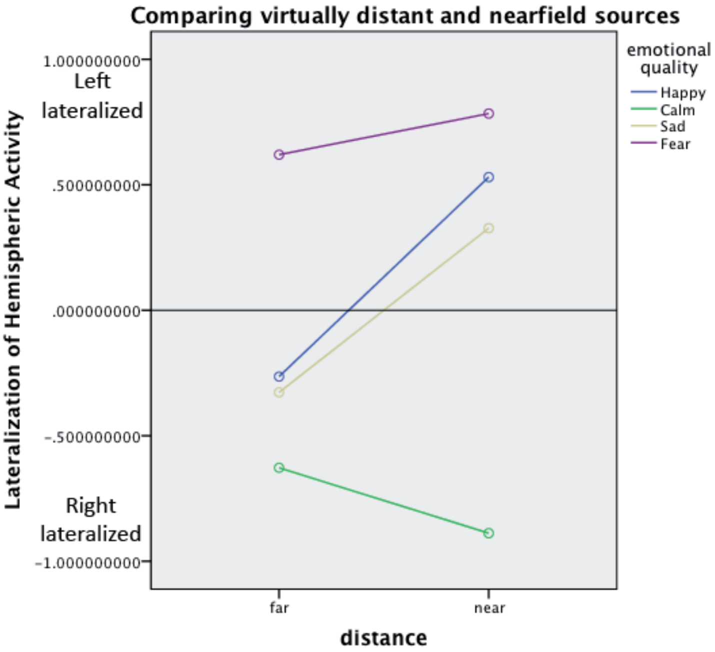


Figure 3: Comparison of alpha lateralization for musical examples of different emotional qualities in virtually distant and nearfield listening situations.

In contrast to the results without haptic stimulation, the inclusion of haptic stimulation in the nearfield condition is consistent with the hypothesis that changes to distance cues do affect the presented basic emotion categories (happy, sad, fearful, calm) differently. As seen in figure 4 below, when using haptic stimulation in the nearfield listening situation, there was a statistically significant ($p < .05$) two-way interaction between changes of distance and emotional quality, $F = 3.994$, $p = .019$. Mauchly's test of sphericity indicated that the assumption of sphericity was met for the two-way interaction, $\chi^2(2) = 0.626$, $p =.681$.

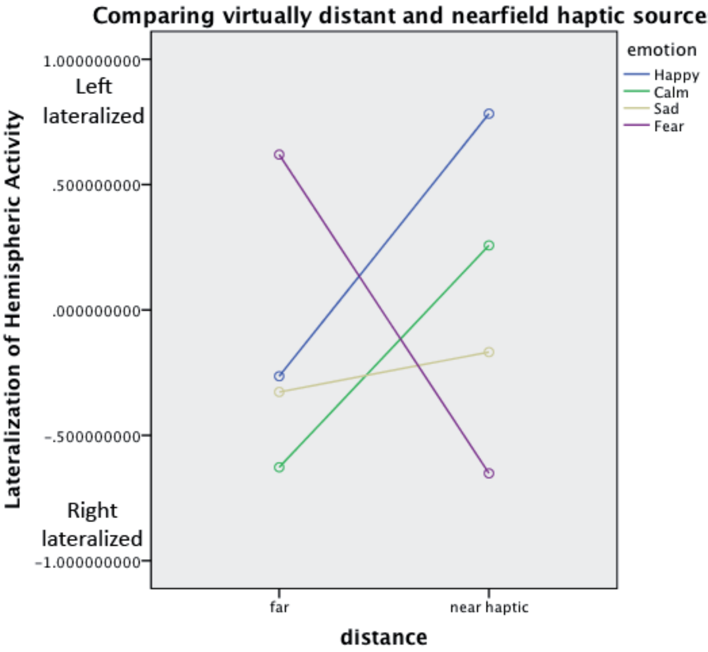


Figure 4: Comparison of alpha lateralization for musical examples of different emotional qualities in virtually distant and nearfield haptic listening situations.

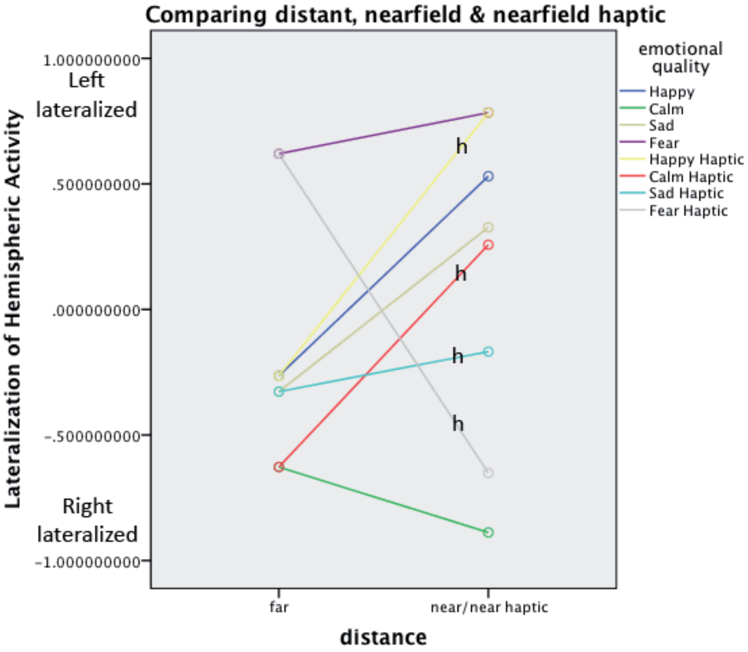


Figure 5: Comparison of all listening conditions together

With the inclusion of haptic stimulation in the nearfield condition, alpha lateralization did not significantly ($p < .05$) change over time for the happy condition between distant and near-field conditions, $F= 0.182$, $p = .052$. Neither did alpha lateralization significantly change over time for the calm condition between distant and near-field conditions, $F= 2.406$, $p = .159$ or for the sad trial, $F= 0.075$, $p = .791$. However, alpha lateralization did significantly change over time in the fear trial, $F= 8.447$, $p = .02$. The differences between the emotions at a virtually distant listening position were not significant, F=2.768, p=.064, nor at a virtually near-field listening position, F=2.225, p=.111. These results suggest that with a small sample size it is difficult to determine the impact of changes of distance due to individual variance, especially for differences of small statistical power. The results also imply that the statistical power (F) between distant and near-field conditions can vary greatly depending upon the basic emotional category presented. The statistical power for fear was strongest at F=8.447, followed by calm (F=2.406), happiness (F=0.182), and finally sadness (F=0.075). The strong statistical power for the fear listening situation was a large contributing factor in achieving a significant result.

There exists "…many methodological problems for measuring physiological indices (e.g., individual differences, time-dependent nature of measures, difficulties in providing effective stimuli)."[66] This does not even take into account non-brain related artifacts, such as electromyogram (EMG) and electrooculogram (EOG) that very often exist in an EEG signal and can make analysis difficult. Given the error variance that these methodologies produce, it is arguably more problematic for the no-specificity hypothesis.[67] With the significance exhibited in the larger condition interactions, and the significance of the fear trial, in addition to the various very nearly significant conditions (p<.05), this suggests further work in this area with larger sample sizes and better controlled conditions are warranted.

Discussion

If emotional responses to music at a nearfield are to be considered to be more strongly valenced than virtually-distant music sources, then the zero-point baseline does not seem to be a tipping point that separates positively-valenced experiences from negatively-valenced experiences. This is especially the case when comparing the distant condition with the nearfield including haptic stimulation condition. If one only examines the standard deviations away from the mean, then both individual results and combined results show that both of the near conditions diverge further away from their means than the distant condition. The average standard deviation for the individual evoked responses at a distance is .2865, while nearfield is .45175, and nearfield haptic is .38225. The average standard deviation from the mean for combined results at a distance is .537, while nearfield is .742, and nearfield haptic is .611. This is

---

[66] Patrik N. Juslin and Petri Laukka, "Communication of Emotions in Vocal Expression and Music Performance: Different Channels, Same Code?," *Psychological Bulletin* 129, no. 5 (2003), 770–814, doi:10.1037/0033-2909.129.5.770, p. 772.

[67] Ibid.

supported when looking at the graphs (see figures 3 and 4), which show an opening up between the distant (left) and near (right) conditions. Making post-hoc speculation, there is the possibility that the zero axis baseline for valence is closer to -.5. This would present results consistent with the hypothesis in that emotional valence for virtual sources in the nearfield are more intense than emotional valence for sources from a distance.[68] This conjecture suggests possibilities for more detailed future work.

One clear finding is that the results are not consistent with their status as purported exemplars for evoking certain discrete emotional categories as described on online lists. These results, with their discrepancies from the expected evoked valences, are also supported by many of the subjective responses of the post-exam debriefing. A number of the subjects described their experiences of the music in terms of how it would be placed in a movie. The fear music was considered by many to 'anticipate' something big or bad to happen. The happy music was often heard as 'triumphant', or 'trying to be triumphant'. The sad excerpt was described as 'cinematic grief', while the calm excerpt evoked feelings of 'floating' or 'religion'. These mismatched results are also supported by earlier EEG music research, where the perceived emotions did not fully match evoked emotional responses.[69]

In contrast with these discrepancies, the results of the nearfield haptic condition match closely with the expected evoked emotional valence for their respective emotional categories. Does this support the idea that multimodal signals are more trustworthy? These results, while very speculative, suggest future work in this area is warranted. However, before these results can be considered more than speculation, they need to be tested against more clearly quantifiable measures of perceived emotion. In this study, I used qualitative research methods, discussing descriptions of perceived emotions in a post-exam debriefing. This encourages highly individualized answers, which does not easily allow for ranking or quantifiable comparisons. For this reason, future work should include either a 23-item self-report scale immediately following each excerpt, such as in Stephens et al.,[70] or as continuous ratings on a computer application with a 2-dimensional emotional matrix, such as in Egerman et al.[71] One of the reasons for my hesitation in advocating for this theory is that there is another possible interpretation for these results. The spectral envelope of the fear excerpt produced the strongest response from the tactile transducer, so the aberrant result in the nearfield haptic condition for the fear excerpt could likely result from perceived overstimulation from a novel source. None of the subjects had previously experienced lis-

---

[68] In support of this post hoc theory, Altenmüller et al exhibits a somewhat similar skew in their results especially from their male participants. Seen in: Altenmüller et al., "Hits to the Left, Flops to the Right."

[69] Schmidt and Trainor, "Frontal Brain Electrical Activity (EEG) Distinguishes Valence and Intensity of Musical Emotions."

[70] Chad L. Stephens, Israel C. Christie, and Bruce H. Friedman, "Autonomic Specificity of Basic Emotions: Evidence from Pattern Classification and Cluster Analysis," *Biological Psychology*, The biopsychology of emotion: Current theoretical and empirical perspectives, 84, no. 3 (July 2010): 463–73, doi:10.1016/j.biopsycho.2010.03.014.

[71] Hauke Egermann et al., "Probabilistic Models of Expectation Violation Predict Psychophysiological Emotional Responses to Live Concert Music," *Cognitive, Affective, & Behavioral Neuroscience* 13, no. 3 (2013): 533–53, doi:10.3758/s13415-013-0161-y.

tening to music with a tactile transducer prior to testing, so it is possible that above a certain threshold the haptic stimulation would be felt as weird or overpowering and something to move away from.

Since the original soundstages were not stripped from the recordings, another attractive interpretation of the results without haptic stimulation is also possible. The calm excerpt was greatly stretched prior to processing with Spat so its original soundstage would have also been distorted for the nearfield listening condition. All of the other excerpts largely kept their original soundstages, so these results could instead indicate more positively valenced results for the original soundstage when compared to a more distant version of itself.[72] In future, the soundstage of the original should be stripped and replaced with a virtual nearfield (instead of layering the new soundstage over the original) to rule out this possibility.

Conclusions

The evoked responses for the nearfield conditions diverge further away from the mean than in the distant condition, which suggests a correlation between the virtual distance and the strength of valence experienced. Another interpretation results from comparing the original 'intended' soundstage with a more distant version, where the intended soundstage is experienced as having a more positive valence than its more distant version for most conditions. Furthermore, the multimodal audio haptic version most closely matches the intended evoked emotion. This could suggest an increased signal trustworthiness when using a multimodal signal. Further work will be needed with larger sample sizes and more controlled soundstages to adequately resolve the possible implications of the results. Other aspects that would be good to explore in future work are: an inclusion of multiple virtual distances, the addition of haptic stimulation to a distant condition to attempt to destabilize the 'trustworthiness' of the distance perceived, independent variations to the early and late delay features of the reverb, and a possible inclusion of a more external environment that might better conform with animal ethological approaches.

---

[72] The vast majority of the subjects did not recognize most of the excerpts. From 10 subjects, only 2 subjects recognized the happy excerpt, 2 recognized the sad excerpt, 1 recognized the sad excerpt, and none recognized the fear excerpt. As a result, I would be very reluctant to suggest that these responses resulted from connections to memories of past listening experiences.