When Room Size Matters: Acoustic Influences on Emotional Responses to Sounds

Ana Tajadura-Jiménez and Pontus Larsson

Division of Applied Acoustics, Chalmers University of Technology, Gothenburg, Sweden

Aleksander Väljamäe

Division of Applied Acoustics, Chalmers University of Technology, Gothenburg, Sweden;

Laboratory for Synthetic Perceptive, Emotive and Cognitive Systems Institute of Audiovisual Studies, Universitat Pompeu Fabra, Barcelona, Spain

Daniel Västfjäll

Division of Applied Acoustics, Chalmers University of Technology, Gothenburg, Sweden;

Decision Reserch, Eugene, OR, USA

Mendel Kleiner

Division of Applied Acoustics, Chalmers University of Technology, Gothenburg, Sweden

Author note

Ana Tajadura-Jiménez is now at the Psychology Department, Royal Holloway, University of London.

This work was carried out as part of the PRESENCCIA project, a European Union–funded Integrated Project under the IST Program (Project Number 27731). The second author of this paper would like to thank for the support from the Swedish Foundation for Strategic Research. We appreciate the helpful comments of two anonymous reviewers on a previous version of this article.

Corresponding concerning this article should be addressed to Ana Tajadura-Jiménez, Psychology Department, Royal Holloway, University of London. Egham, Surrey, TW20 0EX, UK. Tel. +44(0)1784 276551; Fax. +44(0)1784 434347. E-mail: ana.tajadura@rhul.ac.uk

Abstract

When people hear a sound (a “sound object” or a “sound event”) the perceived auditory space around them might modulate their emotional responses to it. Spaces can affect both the acoustic properties of the sound event itself and may also impose boundaries to the actions one can take with respect to this event. Virtual acoustic rooms of different sizes were used in a subjective and psychophysiological experiment that evaluated the influence of the auditory space perception on emotional responses to various sound sources. Participants (N = 20) were exposed to acoustic spaces with sound source positions and room acoustic properties varying across the experimental conditions. The results suggest that, overall, small rooms were considered more pleasant, calmer and safer that big rooms, although this effect of size seems to disappear when listening to threatening sound sources. Sounds heard behind the listeners tended to be more arousing and elicited larger physiological changes than sources in front of the listeners. These effects were more pronounced for natural, compared to artificial, sound sources, as confirmed by subjective and physiological measures.

Keywords: emotion, affect, emoacoustics, room acoustics, auditory virtual rooms

When Room Size Matters: Acoustic Influences on Emotional Response to Sounds

Although people are not often aware of it, the surrounding spatial context modulates their perception, cognition and emotion in their everyday life (Bitner, 1992). There are many types of rooms and spaces, which may be categorized in terms of size, shape, building material, but also in terms of the experiences they offer. Even an outdoor space has a spatial layout and may be considered as a "room". It has been hypothesized that processing of spatial features has developed over human evolution to afford long-time survival (Ulrich, 1983). In this regard, the perception of the surrounding space and emotional responses should be linked, since emotions serve to establish our position *vis-à-vis* our environment and keep a constant margin of safety surrounding our body (Levenson, 1994).

Research in the visual domain suggests that environmental features (including spatial aspects) are processed rapidly and automatically (Korpela, Klemettilä, & Hietanen, 2002). It has been proposed that humans may be adapted for life in open landscapes, that are rich in possibilities and that afford exploration, although for safety and maintaining resources, protected smaller spatial layouts may be preferred (Appelton, 1975). In addition, not only the spatial features (e.g. small vs. big space) but also the possible interaction with an object is important for the ensuing emotional reaction (Russell and Snoddgrass, 1987). Threatening objects such as other humans or animals invading a space will create a more arousing negative reaction in a smaller space compared to a larger space. However, the "invasion" of space might be less negative or even positive if the invader is seen positively. In fact, people want to keep close proximity (and hence prefer smaller spatial settings) to individuals with which they have strong positive emotional bond (Wolhwill, 1976).

Spaces can be also heard since humans can perceive reflected sound. For example, blindfolded subjects can distinguish the size of a room by using speech and other reflected sounds (McGrath, Waldmann, & Fernström, 1996). Other research shows that people can effectively match photographs of rooms with binaurally reproduced soundfields representing those rooms (Sandvad, 1999). Nevertheless, multisensory research suggests that seeing a room effectively influences how people perceive the room through their ears, and that vision dominates when the two modalities are unmatched (e.g., seeing a small room and hearing a big room; Larsson, Västfjäll, Olsson, & Kleiner, 2007).

Ample research has shown the effects of room dimensions and geometry on speech and music intelligibility, clarity, spaciousness and similar perceptual aspects (e.g., Cremer & Müller, 1982). However, few studies have focused on the influence of the acoustical space on emotional responses to sound events, although the fact that sound events can elicit a full range of emotional responses has been proved for a variety of sounds (e.g., Bradley & Lang, 1999). Despite this lack or research, room acoustic designers are aware of the fact that "good acoustics" require more than clarity of the sound. For example, a reverberant dining room might be nerve wracking and a highly damped room might be perceived as "dead" or "oppressive" (Knudsen, 1932). This links directly to the effects of room dimensions on mood and other emotional responses. While some research in this area has begun to look at the affect-based judgments of music and speech reproduced with different spatial acoustic properties (Kjellberg, 2004; Västfjäll, Larsson, & Kleiner, 2002), it is still unknown to what extent auditory space perception influences the emotional responses elicited by different sound events.

This paper directly addresses the influence of spatial determinants on auditory-induced emotions. We assume that our emotional response to sound events depends on the meaning that people attribute to events, which stems from an interaction between the sound source, the listener and the situational context (Tajadura-Jiménez, 2008). The auditory space is a constituent of this context because it affects the acoustic properties of the sound event itself and may also impose boundaries to the actions one can take with respect to the events. Therefore, we might hypothesize that a threatening sound, such as the growl of an aggressive dog, might elicit a substantially different emotional response if heard in a space that is small and enclosed as compared to an outdoor environment, where there might be more chances to escape. The emotional response might be also different if the sound comes from a region of the space in our visual field as compared to the situation when the same sound comes from a visually occluded location.

The present study investigates the emotional impact of different sound sources depending on the size of the space in which the listener is located and the perceived location of the source. Participants were exposed to different virtual sound sources, while room acoustic properties and sound source positions were varied across experimental conditions. To assess the effect of the meaning attributed to sound sources, both natural and artificial sounds were used. Artificial sounds in both continuous and discontinuous versions were used. The purpose of including discontinuous versions of the artificial sounds was to make these sounds more physically similar to natural sounds, which are mostly discontinuous. Emotions were assessed by self-report and by measuring changes in electrodermal activity (EDA) and facial electromyography (EMG, from the *Corrugator Supercilii* [CS] and *Zygomatic Major* [ZM] muscles). EDA is a sensitive and valid indicator for the lower arousal range (Boucsein, 1992), while the activity of the CS and ZM muscles is linked to unpleasant and pleasant emotions respectively (Andreassi, 1995). Emotions were characterized in terms of two continuous dimensions: *Valence* or *pleasantness* and *arousal* or *activation* (Russell, 1980). We expected:

> *Hypothesis 1: an interaction between the room size and the type of sound source on the emotional responses induced in listeners, with unpleasant events being perceived as more negative when the space surrounding the listener is reduced*
>
> *Hypothesis 2: more intense emotional responses for sound sources located behind the listeners, compared to sound sources located in front of the listener, especially, for threatening sources.*

Regarding Hypothesis 2 it should be mentioned that, overall, when no visual cues are available, people tend to locate sound sources at their back rather than at their front (Begault, 1994; Tajadura-Jiménez, Väljamäe, Kitagawa, & Ho, 2007). This might reflect a specific attention bias on auditory perception towards the space outside one's visual field, which might also exist at emotional level.

Methods

Participants

All participants (N=20, 8 women. M_{age}=30, range=21-66) had normal hearing and were naïve as to the purposes of the study.

Materials

Participants sat in a dark sound-attenuated room where the walls and ceiling were covered with black cloth drape, which made difficult for participants to estimate the size of the physical room.

Eight sounds were used: Four natural sounds (‘dog growling’, ‘duck quacking’, ‘woman screaming’, ‘man laughing’) and four artificial sounds (a single sinusoidal tone having a frequency of 262.6Hz (the note C); a string sound played in a major 7th dyad (the notes C and B, fundamentals 262.6Hz and 494.8Hz); and their discontinuous, 0.25s-intermittent, versions, where the 0.25s-silence periods contained only the reverberation of the sound). The sounds were chosen to be of different types (‘natural animal’, ‘natural human’, ‘artificial continuous’, ‘artificial discontinuous’) and emotional categories (for each type, one emotional ‘negative’ and one ‘neutral’, non-negative, sound source). Sounds were categorized into ‘negative’ or ‘neutral’ according to the emotional ratings given for similar sounds in previous research (Bradley & Lang, 1999). All sounds had an approximate duration of 5.5-6.5s.

Three rooms were acoustically modeled using the auralization software CATT (www.catt.se; models available upon request). Presenting participants with simulations of auditory environments allows for instant switching of the “virtual” size of the room from condition to condition, gaining experimental time[1]. The simulated rooms were a 400-seat, 4800m^3 concert hall with a predicted T-30 reverberation time of 1.88s (hereafter referred to as the ‘big’ room), a semi-open inner courtyard (i.e. which has no ceiling) about 6600m^3 (‘outdoor’ room; T-30 not applicable here) and a small 101m^3 studio/listening room (‘small’ room; T-30=0.36s at 1kHz). Source and listener positions were in all rooms separated by 2.5m and located approximately in the middle of the ‘small’ and ‘outdoor’ rooms, and close to the stage in the ‘big’ room. Two listener orientations were also rendered in each room: One with the listener facing the source and one where the listener is looking in the opposite direction, away from the source (‘front’ and ‘back’ source positions).

The B-format and Ambisonics techniques were used in the preparation and delivery of the virtual acoustics space via six identical loudspeakers (GENELEC 8030A–Active monitors) symmetrically located around the participant in the semi-anechoic room. The diameter of the installation was approximately 3m. The SpeakerDecoder-application in a Lake CP4 audio workstation was used for decoding the B-format to the speaker array. The loudspeakers were covered with a thin voile cloth drape, with very low sound insulation, to make them less visible without significantly affecting the sound. The sound level was approximately 75dBA, as measured at the participant's ear position.

No visual input was used, except from a small LCD-screen used for the self-reports, together with a gamepad. Presentation® software was used to control stimuli delivery.

In addition, a BIOPAC-*MP150* System was used to record the physiological signals of participants while listening to the sounds. The signals were sampled (at a rate of 3125Hz for EMG and 390.6Hz for EDA) and amplified. AcqKnowledge software controlled the digital data collection.

Procedure

Participants first completed a short practice block and then two experimental blocks, both containing all sound conditions presented randomly. After each sound, participants answered some questions that differed for each block. In the first block, participants rated their feelings towards the sounds using the Self-Assessment Manikin (SAM; Bradley & Lang, 1999), and a horizontal visual analog (VAS) scale, ranging from *"very safe"* to *"very unsafe"* (safeness). In the second block, participants were required to estimate the room size and the distance to the sound source using a horizontal VAS scale ranging from *"very big"* to *"very small"*, and from *"very far"* to *"very close"*, respectively. During the experiment, self-report and physiological data for each trial were collected.

Design and data analyses

A within-participants experimental procedure was used, with every participant experiencing all sound conditions. There were 48 possible sound conditions with a 4 x 2 x 3 x 2 factorial design (Sound Type ['natural animal', 'natural human', 'artificial continuous'. 'artificial discontinuous'] x Sound Emotional Valence ['negative', 'neutral'] x Room Size ['big', 'outdoor'; 'small'] x Sound Source Position ['front', 'back']).

Self-reported valence and arousal ratings were used as dependent variables for two multivariate ANOVAs (MANOVAs) where Wilks' Lambda was used as the multivariate criterion. The remaining data from the various different conditions ratings were submitted to repeated-measures analyses of variance (ANOVA). The physiological data were averaged for each condition across the two experimental blocks. Alpha level was fixed at .05 for all statistical tests. Greenhouse-Geisser correction was used to correct for unequal variances.

Physiological recordings were individually inspected for possible artifacts. EMG signals were band-pass filtered (10-400Hz; Andreassi, 1995). Change *z*-scores were calculated separately for each signal by subtracting the average response for each 1-s interval for the 7s following sound onset from the mean activity during the 1s preceding sound onset (baseline), yielding seven time intervals per sound (Dimberg, 1990). For the factorial analysis, the average of the EMG change scores from 0 to 5s following sound onset and the average of the EDA change scores from 2 to 7s were used.

Results

Manipulation check

This check was performed to explore whether participants' emotional experience and room size perception was successfully manipulated with our experimental stimuli. We first looked at the effect of the different sound events on emotional responses. Results revealed that sounds categorized as 'negative' were rated as more unpleasant and arousing, led to a greater feeling of unsafe situation and elicited a stronger physiological change (reflected in CS muscle activity and EDA), than their 'neutral' counterparts (see Table 1 for means and statistical values). The factor 'sound type' also had a significant effect and interacted with the factor 'sound emotional valence': The 'human negative' sound source was perceived as the most unpleasant, arousing and unsafe, and elicited the largest physiological response (CS muscle activity and EDA), followed by the 'artificial discontinuous negative', the 'animal negative' and the 'artificial continuous negative' sources. No significant effects were found for the ZM muscle activity.

Second, by looking at the estimations of room size made by participants, it was clear that participants identified the acoustically simulated 'big' room as the biggest room, followed by the 'outdoor' and the 'small' room (see Table 1)[2]. The estimation of distance to the sound source was related to the simulated room size, as participants perceived the sound source as most distant in the 'big' room.

[Table 1]

Hypothesis 1: Effects of Room Size

This hypothesis predicted an interaction between the room size and the type of sound source on the induced emotional responses. To explore this hypothesis we first looked at the overall emotional impact of 'room size' on participants. Sounds occurring in 'big' rooms were considered more arousing and unpleasant, and led to a greater feeling of unsafe situation, than those occurring in 'small' rooms (see Table 1).

Next, to directly test Hypothesis 1, we looked at the interaction of the factors 'room size' with 'sound type' and 'sound emotional valence'. Results showed that all factors interacted significantly. In particular, results (see Figure 1) showed that 'room size' interacted with 'sound emotional valence'. The effect of 'room size' described above was pronounced in the SAM ratings only for the 'neutral' sound sources, but not for those categorized as 'negative' ($F(4,74)=4$; $p=.005$, $\Lambda=.67$).

Results also showed that the elicited EDA significantly depended on the interaction of the factors 'room size' and 'sound type' ($F(4,75)=2.8$; $p=.032$): Natural sounds in the 'small' room elicited lower EDA than in the other rooms (see Figure 1). Paired *t*-test comparisons showed a significant higher EDA elicited by the 'animal' sound in the 'big' versus 'small' room ($t(19)=3$; $p=.007$), and in the 'outdoor' versus the 'small' room ($t(19)=2.3$; $p=.031$). The EDA elicited by the 'human' sound was also significantly higher in the 'outdoor' versus the 'small' room ($t(19)=2.7$; $p=.015$). For the artificial sound sources none of the paired *t*-test comparisons reached significance. No significant effects were found for the CS and ZM muscle activities.

Hypothesis 2: Effects of Sound Source Position

This hypothesis predicted that sound sources located behind the listeners would elicit more intense emotional responses, compared to sound sources located in front, especially for threatening sources. Results did not reveal an overall significant difference on participants' emotional responses to 'back' versus 'front' sound sources. However, results did show an interaction between the factors 'sound type' and 'sound source position' (see Figure 1). In particular, sources located behind the listeners were more arousing than sources in front of the listeners ($p<.05$; one-tailed), except for the 'artificial discontinuous' sound ($t(19)=2.1$; $p<.05$), and elicited a higher EDA ($F(2.6,50)=4.8$; $p=.007$) especially for the 'human' sounds. Paired *t*-test comparisons showed a significant difference between the EDA elicited by the 'human' sound located at the 'front' versus 'back' ($t(19)=-2.9$; $p=.009$). No paired *t*-test comparisons reached significance for the other sound sources. There was no significant effect on the CS and ZM muscle activities.

[Figure 1]

Discussion

The results of the current study highlight the influences of auditory space perception on emotional responses to sound. Smaller auditory-rooms were considered as more pleasant, calmer and safer than big rooms, although these differences seemed to disappear when unpleasant sound sources were present. In addition, sources perceived behind the listener were more arousing than sources that listeners faced to. Importantly, these effects were mainly pronounced for natural sound events.

Our Hypothesis 1 predicted an interaction between room size and the emotional valence of the sound source. In analogy to the visual case (e.g., Russell and Snodgrass, 1987), we expected unpleasant events to be perceived as more negative when the space surrounding the listener is reduced. Instead, our results showed that, overall, the big spaces (‘big’ and ‘outdoor’ rooms) were considered less safe and evoked more unpleasant and arousing emotional responses on listeners than small spaces, as registered by subjective and physiological measures. Moreover, this effect of room size only reached significance for the sound sources categorized as ‘neutral’, having little, non-significant, effect on the responses to ‘negative’ sources.

A few studies provide support for our unexpected findings. One shows that the difference between two settings (a nature trail or a busy urban street), which are perceived as having different restorative potential (i.e. potential to reduce stress or attention fatigue) in a low-danger condition, is eliminated when a source of danger is present (Herzog & Rector, 2009). Another study shows that neuroticism, which is typically accompanied by negative affect, is associated to a preference for smaller settings (Küller, 1971). Yet other studies have shown that a small reduction of reverberation time of unpleasant sounds presented during sleep reduces the number of arousal responses (Berg, 2001), while long reverberation times when listening to music and speech are perceived as most unpleasant (Västfjäll, *et al.,* 2002).

We propose a tentative interpretation for the disparity between our results and those obtained in the visual domain. Following a motivational approach to emotion (Cacioppo & Gardner, 1999), it has been proposed that affect is processed by two distinct systems: one for negative, threat-related information, and another for positive information. These two systems might respond to auditory spatial cues in different manners. If a threatening event takes place, the listener might only focus on the specific event that signals the alarm and calls for action.

Then the visual system might take lead, since it can provide the most of the information about the spatial surrounding. However, if no threats are detected, the auditory exploration of the surrounding environment might continue, and then, other auditory cues, such as the spatial ones, might gain importance. This explanation goes in line with the notion of the auditory system acting as an alarm system, whose main mission is to inform of potential threats as soon as they are detected (Juslin & Västfjäll, 2008). It is easy to keep track of all auditory changes in a small and intimate setting, and people might associate such room with a safer shelter. Clearly, more research combining visual and auditory room information is needed in order to understand these spatial influences on induced emotions.

We also predicted more intense emotions for sound sources located behind rather than in front of the listeners (Hypothesis 2). Results confirmed that sounds from ‘back’ locations, especially natural ones, led to an increase in subjective and physiological arousal. These results might suggest the existence of an auditory attention and emotion bias towards the space outside one’s visual field. This is in line with previous results showing a greater tendency to locate sound sources at our back when no visual cues are available (Begault, 1994; Tajadura-Jiménez, *et al.,* 2007). These findings also partially support our interpretation of the results deviating from Hypothesis 1. We suggest that auditory and visual systems complement each other: The former is in charge of detecting possible threats and alarming (Juslin & Västfjäll, 2008), in order to shift visual attention focus to obtain more detailed spatial information. This interaction of perceptual and affective systems may help to sustain a constant margin of safety around our body (see embodied emotion theories; e.g., Niedenthal, 2007; Tajadura-Jiménez, 2008).

Our results showed some different effects for natural and artificial sources. One possible explanation for this difference might be that human sensory systems are tuned to detect and identify natural sources, rather than artificial, synthesized sounds. It should be also noted that for the continuous artificial sounds it was more difficult for listeners to extract the reverberation room cues than in the other types of sounds, which might account for a lack of influence of the room size manipulation on the emotional response to these sounds.

We verified that our selection of ‘negative’ and ‘neutral’ sounds had the expected distinct emotional impact on listeners, as confirmed by subjective and physiological measures, and that participants were sensitive to the “virtual” change of auditory room size. However, a number of limitations should be pointed out when considering the present results. First, the limited selection of sound sources, room auralizations and source positions makes it difficult to generalize our results. Future work should therefore extend this work to explore other conditions. Second, inter-cultural studies might reveal different effects for people accustomed to spend more time outdoors. Third, it may be useful to perform a similar study where listeners are in different “physical” (and not only “virtual”) rooms, since awareness of being in the same room during the experiment might have influenced the results. Furthermore, it might be worth to explore the mentioned effects by creating situations where people can interact with the environment (e.g., gaming), in order to understand how humans might behave in real threat situations. Most importantly, future research needs to address the disparity between our results obtained via purely auditory stimulation and conflicting results from visual domain. Multimodal simulations using either virtual or real environments could offer possibilities to investigate multisensory interaction effects in room perception and emotional processing.

References

Andreassi, J. L. (1995). *Psychophysiology: Human behavior and physiological response* (3rd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Appelton, J. (1975). *The experience of landscape.* New York: Wiley.

Begault, D. R. (Ed.). (1994). *3-D Sound for virtual reality and multimedia*. San Diego, CA: Academic Press Professional, Inc.

Berg, S. (2001). Impact of reduced reverberation time on sound-induced arousals during sleep. *Sleep, 24*(3), 289-292.

Bitner, M. J. (1992). Servicescapes: The impact of physical surroundings on customers and employees. *Journal of Marketing, 56*(2), 57-71.

Boucsein, W. (1992). *Electrodermal activity*. New York: Plenum press.

Bradley, M. M., & Lang, P. J. (1999). *International affective digitized sounds (IADS): Stimuli, instruction manual and affective ratings* (Tech. Rep. No. B-2). Gainesville, FL: The Center for Research in Psychophysiology, University of Florida.

Cabrera, D. (2007). Control of perceived room size using simple binaural technology. In *Proceedings of the 3th International Conference on Auditory Display*, Montreal, Canada.

Cacioppo, J. T., & Gardner, W. L. (1999). Emotion. *Annu Rev Psychol, 50*, 191-214.

Cremer, L., & Müller, H. A. (1982). *Principles and applications of room acoustics*. New York: Elsevier Science.

Dalenbäck, B.-I. (1996) Room acoustic prediction based on a unified treatment of diffuse and specular reflection. The *journal of the Acoustical Society of America, 100*, 899-909.

Dimberg, U. (1990). Facial electromyography and emotional reactions. *Psychophysiology, 27*(5), 481-494.

Herzog, T. R., & Rector, A. E. (2009). Perceived danger and judged likelihood of restoration. *Environment & Behavior***.** 41(3), 387-401.

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: the need to consider underlying mechanisms. *Behavioral & Brain Sciences, 31*(5), 559-575; discussion 575-621.

Kjellberg, A. (2004). Effects of reverberation time on the cognitive load in speech communication: theoretical considerations. *Noise Health, 7*(25), 11-21.

Knudsen, V. O. (1932). *Architectural acoustics*. New York: J. Wiley & sons; London: Chapman & Hall, limited.

Korpela. K., Klemettilä, T., & Hietanen, J. (2002). Evidence for rapid affective evaluation of environmental scenes. *Environment & Behavior*, *34*, 478-494.

Küller, R. (1971). *Rumsperception* [Room perception]. Report. Lund Institute of Technology, Department of theoretical and applied Aesthetics, Lund, Sweden.

Larsson, P., Västfjäll, D., Olsson, P., & Kleiner, M. (2007). When what you see is what you hear: Auditory-visual integration and presence in virtual environments. In *Proceedings of the 10th Annual International Workshop on Presence*, Barcelona, Spain

Levenson, R. W. (1994). Human emotions: A functional view. In P. Ekman, Davidson, R.J. (Ed.), *The nature of emotion: Fundamental questions* (pp. 123-126). New York: Oxford University Press.

McGrath, R., Waldmann, T., & Fernström, M. (1996). Listening to rooms and objects. In *Proceedings of the 16th Audio Engineering Society International Conference*, Rovaniemi, Finland.

Niedenthal, P. M. (2007). Embodying emotion. *Science, 316*(5827), 1002-1005.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(6), 1161-1178.

Russell, J. A. & Snodgrass, J. (1987). Emotion and the environment. In *Handbook of Environmental Psychology*, Vol. 1 (pp. 245-280). New York: John Wiley and Sons.

Sandvad, J. (1999). Auditory perception of reverberant surroundings *The Journal of the Acoustical Society of America, 105*(2), 1193.

Tajadura-Jiménez, A. (2008). *Embodied psychoacoustics: Spatial and multisensory determinants of auditory-induced emotion* PhD thesis, Chalmers University of Technology, Gothenburg, Sweden.

Tajadura-Jiménez, A., Väljamäe, A., Kitagawa, N., & Ho, H. (2007). Whole-body vibration influences sound localization in the median plane. In *Proceedings of the 10th Annual International Workshop on Presence*, Barcelona, Spain.

Ulrich, R. S. (1983). Aesthetic and affective response to natural environment. In I. Altman & J. F. Wohlwill (Eds.), *Human behavior and environment: Advances in theory and research* (Vol. 6, pp. 85-125). New York: Plenum.

Västfjäll, D., Larsson, P., & Kleiner, M. (2002). Emotion and auditory virtual environments: Affect-based judgments of music reproduced with virtual reverberation times. *Cyberpsychology & Behavior, 5*(1), 19-32.

Wohlwill, J. F. (1976). Environmental aesthetics: The environment as a source of affect. In I. Altman & J. Wohlwill (Eds.), *Human behavior and environment: Advances in theories and research,* Vol. 1 (pp. 37-85). New York: Plenum Press.

Footnotes

[1] From all acoustical indicators, reverberation (or late sound reflection) has the strongest effect on room size perception, as has been confirmed for both real and modeled rooms, but also the temporal distribution and level of early reflections influence auditorily perceived room size (see Cabrera, 2007). The auralization software uses well-established and verified physical modeling techniques to simulate both early and late reflections (Dalenbäck, 1996); hence one can assume that the correct room size perception is achieved by the auralizations.

[2] It may seem unintuitive that the 'outdoor' room was rated smaller than the 'big' room, as the volume of the 'big' room is lower than the volume of the 'outdoor' courtyard room. However, there are two possible explanations to this result. First, the 'big' room has long, very audible reverberation and it is one of the strongest cues to room size perception. On the opposite, the 'outdoor' room has minimal reverberation. Second, there is a clear difference in the early horizontal plane reflections that also provide cues to room size. Due to the differences in geometry, these reflections arrive later in the 'big' room compared to the 'outdoor' room falsely suggesting to the listeners that the 'outdoor' room is smaller.

Figure 1. The influence of auditory perceived room size ('big', 'outdoors' or 'small') and sound source position ('back' vs. 'front') on emotional responses to different sound objects. The upper panels show the mean valence and arousal ratings (in a 9-point scale) ± *SEM*, while the lower panels show the mean electrodermal activity (EDA) ± *SEM* (*z*-score) for the different sound conditions used in the experiment. Sounds occurring in 'big' rooms were considered more arousing and unpleasant than those occurring in 'small' rooms, although this effect was pronounced only for 'neutral' sound sources (see upper left panel). Accordingly, natural sound sources in 'big' rooms also elicited higher EDA in listeners than those in 'small' rooms (see lower left panel). For most of the sound sources, location behind the listeners influenced arousal more than front location (see upper right panel; 'Cont.' stands for 'Continuous' and 'Disc.' for 'Discontinuous'.). Natural sound sources located behind the listeners also elicited higher EDA than when located in front (see lower right panel). EDA change *z*-scores were calculated by subtracting the mean activity between 2 and 7 s following stimulus onset from the mean activity during the 1 s preceding sound onset (baseline). See text for the details on statistical analysis and results.

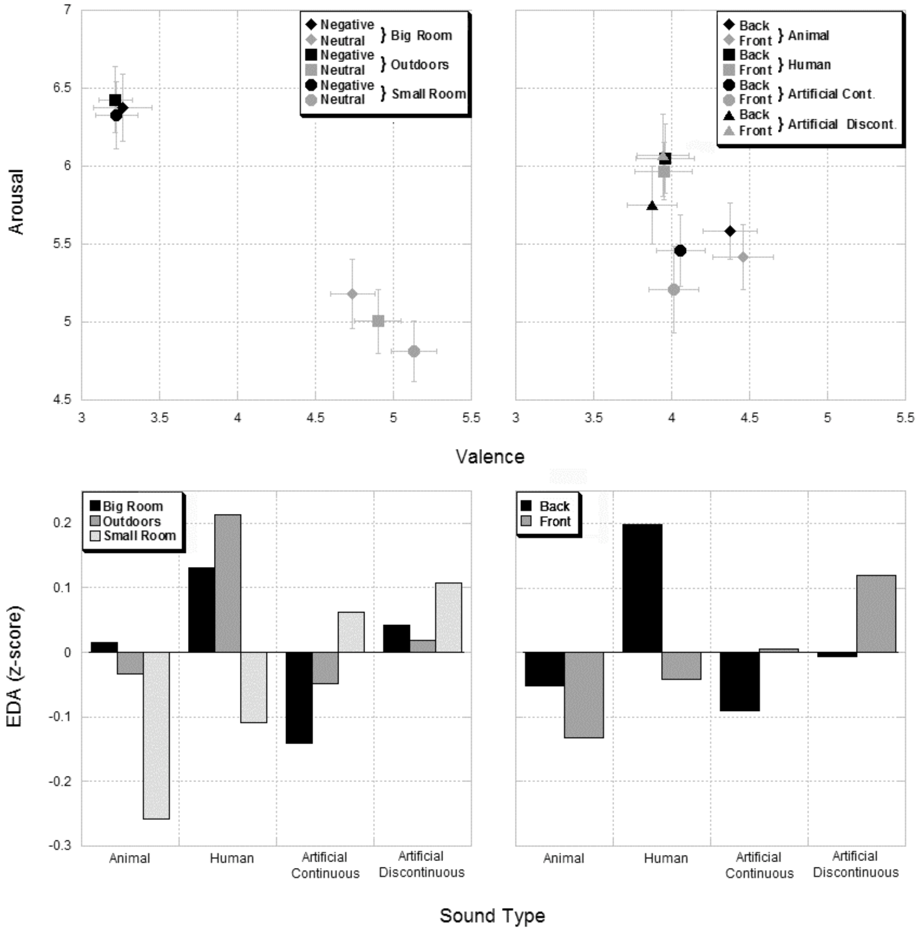

Table 1. Emotional effects induced by the different sound conditions.

Emotional effects of sound events

	Animal		Human		Artificial Continuous		Artificial Discontinuous		ANOVAs		
	Neg.	Neu.	Neg.	Neu.	Neg.	Neu.	Neg.	Neu.	Type (ST)	Emotional Valence (SEV)	ST x SEV
Measurements											
Valence	3.7 (.2)	5.1 (.3)	2.2 (.2)	5.7 (.4)	3.8 (.2)	4.3 (.2)	3.2 (.1)	4.6 (.2)	$F_{(6,112)}$=2.8 p=.014 Λ =.75	$F_{(2,18)}$=2.2 p=.001 Λ=.22	$F_{(6,112)}$ =8.5 p=.001 Λ=.47
Arousal	5.8 (.2)	5.2 (.2)	7.2 (.3)	4.8 (.3)	5.8 (.2)	4.8 (.4)	6.7 (.3)	5.2 (.3)			
Perceived Safeness	-123 (24)	33 (24)	-233 (28)	92 (32)	-94 (17)	21 (25)	-140 (18)	4 (23)	*n.s.*	$F_{(1,19)}$=54.3 p=.001	$F_{(2.3,43)}$=14.1 p=.001
EDA	-.08 (.06)	-.1 (.07)	.21 (.09)	-.07 (.07)	-.01 (.06)	-.07 (.06)	.12 (.06)	-.02 (.07)	$F_{(2.4,46)}$=2.8 p=.064	$F_{(1,19)}$=5.6 p=.029	$F_{(2.5,47)}$=5.8 p=.003
CS	-.02 (.07)	-.05 (.06)	.13 (.07)	-.04 (.12)	.11 (.07)	.13 (.07)	.07 (.08)	.03 (.08)	$F_{(2.2,42)}$=2.8 p=.067	$F_{(1,19)}$=4.3 p=.052	*n.s.*

Emotional effects of room size

	Big	Outdoors	Small	ANOVAs
Measurements				
Valence	4 (.1)	4.1 (.1)	4.2 (.1)	$F_{(2,26)}$=3.7, p<.05^{a}
Arousal	5.8 (.2)	5.7 (.2)	5.6 (.2)	$F_{(2,33)}$=3.9, p=.034^{a}
Perceived Safeness	-64 (12)	-55 (13)	-46 (13)	$F_{(1.9,36)}$=3.8, p=.034
EDA	184 (14)	30 (12)	-121 (24)	$F_{(2,26)}$=73.5, p=.001
CS	-41 (15)	-114 (12)	-115 (22)	$F_{(2,28)}$=8.5, p=.003

Note. The upper panel shows the effects of the different sound objects in mean valence and arousal ratings (in a 9-point scale); mean perceived safeness (in a VAS scale with corresponding coordinates ranging from –450 to 450); mean corrugator supercilii muscle activity (CS; *z*-score); and mean electrodermal activity (EDA; *z*-score). The lower panel shows the effects of the different room auralizations in mean valence and arousal ratings; mean perceived safeness ; mean perceived room size (in a VAS scale with corresponding coordinates ranging from –450 to 450); and mean perceived distance to the sound source (in a VAS scale with corresponding coordinates ranging from –450 to 450) . Parentheses give the standard errors for the mean. 'Neg.' stands for 'Negative', 'Neu.' for 'Neutral', 'ST' stands for 'Sound Type', 'SEV' for 'Sound Emotional Valence', *n.s.* for 'non-significant'.

[a] The multivariate test for the factor ‘room size’ revealed an effect close to significance (p = .084), which was significant when looking independently at the univariate tests of valence and arousal ratings.