CyberPsychology & Behavior
Volume 11, Number 1, 2008

DOI: 10.1089/cpb.2007.0002

Self-Representation in Mediated Environments: The Experience of Emotions Modulated by Auditory-Vibrotactile Heartbeat

ANA TAJADURA-JIMÉNEZ, Lic.Tech.,[1] ALEKSANDER VÄLJAMÄE, Ph.D.,[1]
and DANIEL VÄSTFJÄLL, Ph.D.[1,2]

ABSTRACT

In 1890, William James hypothesized that emotions are our perception of physiological changes. Many different theories of emotion have emerged since then, but it has been demonstrated that a specifically induced physiological state can influence an individual's emotional responses to stimuli. In the present study, auditory and/or vibrotactile heartbeat stimuli were presented to participants ($N = 24$), and the stimuli's effect on participants' physiological state and subsequent emotional attitude to affective pictures was measured. In particular, we aimed to investigate the effect of the perceived distance to stimuli on emotional experience. Distant versus close sound reproduction conditions (loudspeakers vs. headphones) were used to identify whether an "embodied" experience can occur in which participants would associate the external heartbeat sound with their own. Vibrotactile stimulation of an experimental chair and footrest was added to magnify the experience. Participants' peripheral heartbeat signals, self-reported valence (pleasantness) and arousal (activation) ratings for the pictures, and memory performance scores were collected. Heartbeat sounds significantly affected participants' heartbeat, the emotional judgments of pictures, and their recall. The effect of distance to stimuli was observed in the significant interaction between the spatial location of the heartbeat sound and the vibrotactile stimulation, which was mainly caused by the auditory-vibrotactile interaction in the loudspeakers condition. This interaction might suggest that vibrations transform the far sound condition (sound via loudspeakers) in a close-stimulation condition and support the hypothesis that close sounds are more affective than distant ones. These findings have implications for the design and evaluation of mediated environments.

INTRODUCTION

One of the pioneers in emotional research, William James,[1] hypothesized that emotions are our perception of physiological changes in our body. Following James's line of thought, triggering bodily changes should lead to an emotional experience. In the classical study by Schachter and Singer,[2] injection of adrenaline made naïve participants more susceptible to a provoking situation. Another study[3] showed that external modification of participants' facial expression to frowning or smiling (e.g., holding a pen in their mouths) significantly modulated their evaluation of cartoons. More recent studies have attempted to evoke bodily changes by using external stimulation, mainly visual, with auditory domain mostly reduced to music studies (e.g., music assisted relaxation techniques).[4] How-

[1]Division of Applied Acoustics, Chalmers University of Technology, Gothenburg, Sweden.
[2]Department of Psychology, Göteborg University; Gothenburg, Sweden.

ever, other acoustic stimulation can also effectively produce bodily changes. For instance, pure tones or noise may alter facial electromyography activity;[5–6] Bradley and Lang[7] showed how listening to ecological sounds may influence heart rate and electrodermal and facial electromyography activity.

The present study focuses on the effect of sound on listeners' physiological arousal and how this alteration might influence a subsequent emotional experience. In particular, we wanted to investigate the impact of the perceived distance to sound in this paradigm. Our body serves as a reference frame when establishing our position in relation to surrounding events.[8] Hence, physical distance to objects modulates the *intensity* of emotional responses: close is arousing, intimate, engaging.[9] Surprisingly, few studies have explored the effect of the spatial dimension of sound on emotional reactions. Recently, it has been shown that multichannel audio rendering can enhance emotional responses to music.[10] Similarly, reverberation times in auditory virtual environments may influence self-reported emotional reactions.[11] In a study[12] comparing loudspeaker- and headphone-reproduction of news, no significant difference was found in elicited arousal, although the headphone listening was preferred over loudspeakers. For the present study, we chose heartbeat sounds because we expected them to magnify the disparity between sounds rendered close to or distant from the body. Heartbeat sounds belong to the *self-representation sounds* category, a particular case of ecological, everyday sounds that can be associated with a person's own body (e.g., breathing) and its embodied activity (e.g., footsteps).[13] These sounds increase body awareness in listeners, and we hypothesize that they may have a stronger potential for inducing an emotional experience, especially if perceived close to oneself. In previous research, attempts were made to influence listeners' actual heart rate by means of false auditory heartbeat.[14–16] The studies that succeeded in this task reported small variations (1 to 2 beats per minute) in participants' heart rate after listening to fast or slow heartbeats. In line with James's paradigm, these studies tried to relate physiological changes to emotional experience; most of these authors used monothematic slides with semi-naked human bodies to be rated on an attractiveness scale. Unfortunately, results from these studies have not been very conclusive, and they failed to show a significant correlation between changes in heart rate and participants' attributions of arousal. In this study, we tested three hypotheses:

Hypothesis 1: Fast rate heartbeat sounds will influence listeners' heart rate, which may impact their emotional experience. In this case, affective stimuli will be judged as more arousing, in line with the generally accepted phenomenon that links stimuli rate with increased arousal.[17]

Hypothesis 2: Sounds perceived close to oneself (headphones) will have a larger impact on physiology and picture judgment than distant ones (loudspeakers).

Hypothesis 3: Coupling auditory stimuli with synchronous vibrotactile stimulation will bias this multimodal stimuli toward oneself, thus increasing the impact both on physiology and affective experience.

MATERIALS AND METHODS

Participants

Twenty-four naïve participants, 18 male (mean age 24.4; $SD = 4.6$), took part in the experiment. All participants had normal hearing. They were informed that during the experiment they would be exposed to heartbeat sounds. The current study was conducted under approval of the local ethics committee.

Apparatus

Visual stimuli (768×576 pixels resolution and $33° \times 26°$ field of view) were presented on a flat projection screen placed at 1.7 meters distance. The heartbeat stimuli was delivered using a pair of headphones (SONY, ear-buds), a pair of loudspeakers (GENELEC 1029A, active motor), stereo bone-conduction transducers for bilateral stimulation (OIIDO Equipment), and/or a mechanical shaker mounted under the experimental chair. The loudspeakers were placed on the sides of the screen and hidden from participants' view. A Pasport PS-2105 Heart Rate Sensor was attached to an ear clip to measure participants' heart rate.

Stimuli and design

A series of 34 pictures (two used for instructions) were selected from the International Affective Picture System (IAPS)[18]—a set of normative pictures rated, in a two-dimensional affective scale, for *valence* (or pleasantness) and *arousal* (or activation). Stimuli were chosen according to their arousal (5 on a 9-point scale) and valence values (3 and 7 out of 9 for negative and positive pictures respectively). Two mono-files containing heartbeat sounds at two different rates, medium and high (60 and 110 beats per minute [bpm]), were synthesized. The sound files were played using headphones, loudspeakers, bone-conduction headset, or mechanical shaker. The heartbeat sound level was set to match the level of low speech (approximately 60 dBA). To assure simulation transparency, vibrations did not exceed 5 cm/s.[2] Due to the resonance frequency of the shaker, vibrational signal was band-limited to 40 to 100 Hz. Presentation software (www.neurobs.com) was used to deliver the stimuli. Auditory/vibrotactile stimuli lasted 50 seconds, and images were presented during the last 6 seconds. The factorial design was $4 \times 2 \times 2 \times 2$ (sound rendering type $\times$ vibrotactile stimulation on/off $\times$ rate: medium [60 bpm], fast [110 bpm] $\times$ picture type: positive/negative). Sound rendering types included no sound or silence (S), bone-conduction headset (BC), loudspeakers (LS), and headphones (HP).

Measures

During the experiment, participants' physiological changes were measured by means of the heart rate sensor (50 Hz sampling rate). Valence and arousal ratings for pictures (self-reported emotional experience) were collected by using the Self-Assessment Manikin (SAM), a 9-point pictorial scale developed by Lang.[19] Finally, after the experiment, a free-recall task was performed to assess the memory for the pictures presented. High-arousing events tend to be more memorable;[20] thus, this test served as a second measure of emotional experience.

Procedure

After a short training session (two trials), the actual experiment started. During each trial, heart rate was collected; after each stimulus, SAM scale was displayed on the screen and participants rated picture valence and arousal by using a keyboard. Stimuli were presented in randomized order, and participants could choose to pause after each stimuli presentation. After completing the experiment, a 3-minute free-recall task of the experimental photographs was performed. Finally, the experiment leader performed verbal probing concerning participants' sensations. Then, subjects were debriefed, thanked, and paid for their participation.

RESULTS

Physiological data, SAM ratings (self-reported valence and arousal), and memory performance results were subjected to $4 \times 2 \times 2$ (rendering type $\times$ vibration $\times$ rate) repeated measures ANOVA (Greenhouse-Geisser correction). Results for bone-conducted rendering served as evaluation of this novel technology and are not reported here.

Effects on physiology

Physiological data from the heart rate sensor were individually inspected for possible artifacts due to participant's movement. In each trial, only the 40 seconds before the picture presentation were used for data analysis. Hence, as physiological measures did not include the picture exposure time, data from positive and negative picture conditions were averaged. Due to artifacts, two out of 768 recordings (24 participants, 32 conditions) were omitted from the appropriate condition average. The sound rendering type condition had a significant effect on participants' heart activity ($F[2.6, 58.8] = 2.88$, $p = 0.05$) with means ($SE = 1.9$) increasing from 69.06 (S) to 69.7 (LS) and highest at 69.87 (HP). Least significant difference (LSD) adjusted pairwise comparison revealed a significant difference between silence and the sound conditions ($p < 0.05$). No significant effects of vibrotactile stimulation or rate were observed.

When inspecting the interaction between sound rendering type and vibration, significance was reached at $F(2.5, 57.9) = 5.24$, $p < 0.005$. Additional analysis showed that this effect was mainly caused by the loudspeakers condition ($p < 0.005$, $t[23] = 3.34$; see Figure 1).

Effects on self-reported emotional experience: SAM ratings

Appetitive and defensive motivational systems encompass different mechanisms.[21] For this reason, when assessing self-reported emotional experience, an extra variable, picture type (positive vs. negative), was included in the analysis, and effects for each picture type are reported whenever a differential effect was found. Valence and arousal were used as dependent variables for a multivariate ANOVA (MANOVA); Wilks's lambda was used as the multivariate criterion.

(a) Rate effect. In line with hypothesis 1, there was a significant effect of rate on self-reported emotional experience (Wilks's $\Lambda = 0.74$, $F[2, 22] = 3.82$, $p < 0.05$). Results showed that the increase of rate significantly increased arousal ratings ($F[1, 23] = 4.92$, $p < 0.05$) with means 4.48 ($SE = 0.22$), for medium rate, versus 4.85 ($SE = 0.24$), for fast rate. No significant effect was found for valence ratings.

(b) Rendering effect. Results revealed a significant effect of sound rendering on SAM ratings (Wilks's $\Lambda = 0.78$, $F[6, 136] = 3$, $p < 0.01$). Because a significant interaction was found with the variable picture type (Wilks's $\Lambda = 0.82$, $F[6, 136] = 2.4$, $p < 0.05$), effects are reported separately for positive and negative pictures. While for negative pictures the main effect of rendering was nonsignificant, for positive pictures MANOVA did show significance (Wilks's $\Lambda = 0.73$, $F[6, 136] = 3.9$, $p < 0.001$). Specifically, sound render-

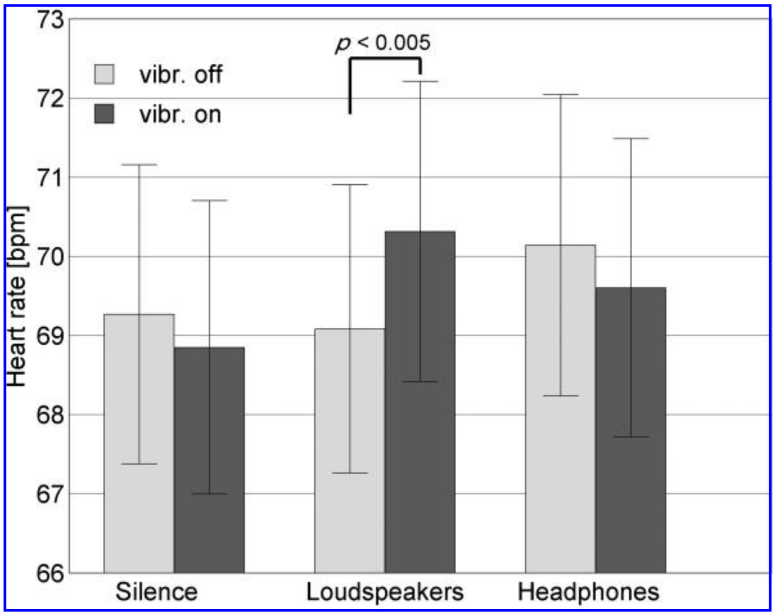

FIG. 1. Rendering and vibrotactile (vibration on versus off) effect on subjects' heart rate (beats per minute). The whiskers show the standard errors of the means.

ing significantly influenced both valence (at $F[2.9, 65.8] = 4.2$, $p < 0.01$) and arousal ($F[2.6, 59.9] = 4.9$, $p < 0.01$) ratings. The addition of sound had a clear effect on the self-reported emotional responses: pictures were rated as "more positive" and "more arousing" after auditory stimulation. Means were 6.35 (S; $SE = 0.16$), 6.64 (HP; $SE = 0.15$), and 6.73 (LS; $SE = 0.17$) for valence and 3.94 (S; $SE = 0.3$), 4.61 (HP; $SE = 0.23$), and 4.64 (LS; $SE = 0.24$) for arousal.

(c) Interaction between vibrotactile stimulation and distance to sound. The effect of vibration alone did not reach significance. Due to the significant interaction of sound rendering and vibration observed at the physiological level, we further searched for the same trends on the picture judgments. A closer look on the two-dimensional space of SAM ratings for the positive pictures showed a trend similar to the one in physiological data (Figure 2). Vibrations increased the difference between medium and fast heart rate stimulation for the headphones condition ($p = 0.067$, $t[23] = 1.9$ for sound only and $p < 0.011$, $t[23] = 2.8$ for auditory-vibrotactile stimulation). More importantly, vibrations interacted with loudspeaker rendering, bringing the difference between medium and fast heart rate to the headphone level ($p = 1$, $t[23] = 0$ for sound only and $p = 0.06$, $t[23] = 2$ for auditory-vibrotactile stimulation).

Effects on memory

On average, 13 ($SD = 3.24$) of 32 pictures were recalled. Results from the memory test were submitted to $4 \times 2 \times 2 \times 2$ (sound reproduction $\times$ vibration $\times$ rate $\times$ picture type) ANOVA. Analysis showed a significant effect of rate on memory ($F[1, 23] = 6.95$, $p = 0.015$). Pictures were more memorable when rate was higher (means for picture recall rate were 43.5%, $SE = 2.7\%$, versus 37.5%, $SE = 2\%$).* In addition, picture type also showed an insignificant trend ($F[1, 23] = 3.11$, $p = 0.09$), with positive pictures being more memorable than negative ones (means were 44% versus 37%; $SE = 2.9\%$). The effects of the different sound rendering types and vibrotactile stimulation on memory performance did not reach significance.

DISCUSSION

In the present study, we investigated the influence of the perceived distance to auditory and auditory-vibrotactile heartbeat stimuli on participants' affective state. This influence was assessed by measuring both the changes on participants' heart rate and on their subsequent judgments of an emotional experience (seeing affective pictures). Performance on a memory task served as a further indicator of the effect of the manipulation of distance to stimuli.

Presenting heartbeat sounds resulted in a small but significant increase in participants' heart rate with respect to the silence condition. In addition, heartbeat sound also significantly influenced emotional re-

*100% picture recall rate would mean that all participants remembered the picture in that sound/vibrotactile condition.

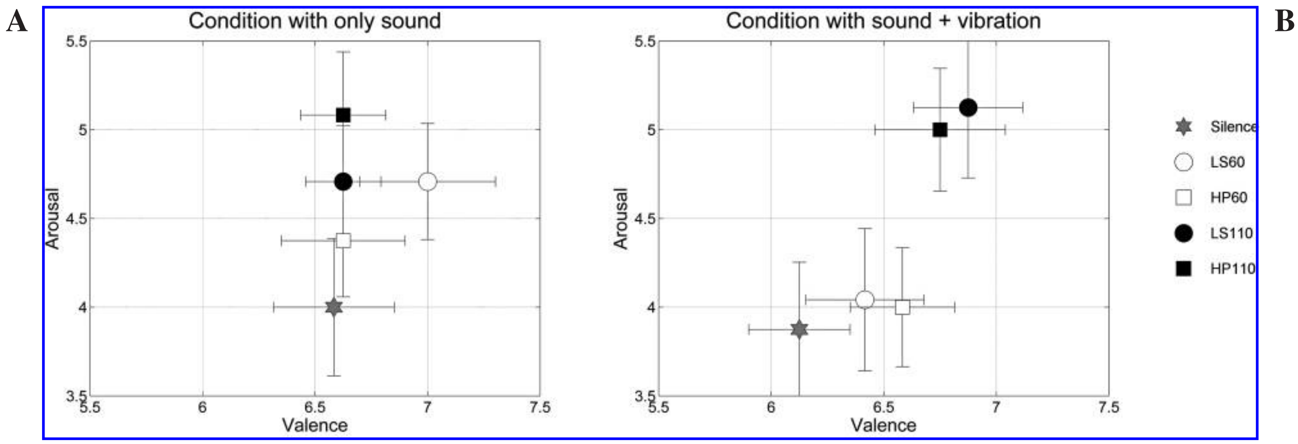

FIG. 2. Rate effect on valence and arousal judgments of positive pictures for the conditions with only sound (**A**) and combined auditory and vibrotactile stimulation (**B**). Valence and activation are rated in a 9-point scale (the whiskers show the standard errors of the means). 'LS' refers to loudspeakers and 'HP' to headphones; '60' refers to the condition with medium rate heartbeat stimuli (60 beats per minute) and '110' to the fast rate heartbeat stimuli (110 beats per minute).

sponses to affective visual stimuli, with positive pictures rated as "more positive" and "more arousing" after auditory stimulation. In agreement with our first hypothesis, fast-rate heartbeat sound resulted in higher arousal ratings and enhanced pictures' memorability in the postexperimental free-recall task. These results corroborate the data from Bolls et al.[20] showing that the arousal dimension plays a stronger role than valence in making stimuli more memorable.

We hypothesized that stimuli perceived as close to oneself (headphones) are more likely to alter physiology and emotional experience than far stimuli (loudspeakers). We did not find direct support for these hypotheses; however, the concurrent presentation of vibrotactile stimulation with heartbeat sound significantly influenced participants' physiology, mainly for loudspeaker rendering conditions. We suggest that this physiological effect was translated into the affective response to emotional pictures. The observed trends (see Figure 2) on the self-reported valence and arousal seem to confirm this hypothesis, especially when examining responses to positive pictures. In auditory-only conditions, the effect of the heartbeat sound on participants' arousal was stronger when sound was perceived closer to oneself (headphone reproduction). When vibrations were simultaneously presented, both loudspeaker and headphone conditions resulted in similar arousal ratings. In other words, vibrations seemed to equalize the affective power of heartbeat sounds between the loudspeaker- and headphone-based renderings. Therefore, our hypothesis that the close spatial location of arousing stimuli would be more affective than the distant one received partial support, especially if assuming that vibration did transform the far sound condition (sound via loudspeakers) in a close-stimulation condition (cf. tactile 'capture' of audition in the study by Caclin et al.[22–23]).

In summary, the presented results suggest that the emotion-eliciting power of auditory modality is influenced by the perceived distance to sound. Future research should clarify whether the distance cues alone may be responsible for affect modulation or, in addition, it may be also caused by stimuli recognition as a representation of one's own body. We believe that self-representation sounds (e.g., heartbeat, breath, footsteps) may play an important role in the design of virtual environments, since they are an inseparable part of users' multimodal virtual body, a self-avatar.[13,24] Previous research showed that visual cues constituting a self-avatar (e.g., body parts or shadow) significantly increased presence experience in virtual environments. Therefore, self-representation sounds can be seen as an auditory part of such body-centered interaction.

Arousing or intense emotional experiences tend to be described as more engaging,[25] thus increasing the feeling of presence in virtual environments.[26] E-learning environments might benefit from arousing and positive experiences because they facilitate memory for events and encourage users to go on with the tasks.[27] In the area of health and telemedicine, understanding users' ongoing emotional state is essential, for instance, in applications for fighting against stress[28] or fear (e.g., of public speaking).[29] Understanding and including affective components in virtual and mediated environments helps to establish an efficient human–computer communication[30] and influences the way users interact in these environments.

ACKNOWLEDGMENTS

This work was carried out as part of the PRESENCCIA project, a European Union–funded Integrated Project under the IST Programme (Project Number 27731).

REFERENCES

1. James W. (1890). *The principles of psychology*. New York: Holt.
2. Schachter S, Singer JE. Cognitive, social and physiological determinants of emotional state. Psychological Review 1962; 69:379–99.
3. Strack F, Martin LL, Stepper, S. Inhibiting and facilitating conditions of the human smile. Journal of Personality and Social Psychology 1988; 54:768–77.
4. Pelletier CL. The effect of music on decreasing arousal due to stress: A meta-analysis. Journal of Music Therapy 2004; 41:192–214.
5. Kjellberg A, Skoldstrom B, Tesarz M, Dallner M. Facial EMG responses to noise. Perceptual and Motor Skills 1994; 79:1203–16.
6. Jäncke L, Vogt J, Musial F, Lutz K, Kalveram KT. Facial EMG responses to auditory stimuli. International Journal of Psychophysiology 1996; 22:85–96.
7. Bradley MM, Lang PJ. Affective reactions to acoustic stimuli. Psychophysiology 2000; 37:204–15.
8. Damasio A. Fundamental feelings. Nature 2001; 25:781.
9. Reeves B, Nass C. (1996). *The media equation: how people treat computers, television, and new media like real people and places*. New York: Cambridge University Press.
10. Västfjäll D. The subjective sense of presence, emotion recognition, and experienced emotions in auditory virtual environments. CyberPsychology & Behavior 2003; 6:181–8.
11. Västfjäll D, Kleiner M, Larsson P. Emotion and auditory virtual environments: affect-based judgments of music reproduced with virtual reverberation times. CyberPsychology & Behavior 2002; 5:19–32.
12. Kallinen K, Ravaja N. Comparing speakers versus headphones in listening to news from a computer—individual differences and psychophysiological responses. Computers in Human Behavior 2007; 23:303–17.
13. Väljamäe A, Larsson P, Västfjäll D, Kleiner M. Sound representing self-motion in virtual environments enhances linear vection. Presence: Teleoperators and Virtual Environments 2008; 17 (in press).
14. Valins S. Cognitive effects of false heart-rate feedback. Journal of Personality & Social Psychology 1966; 4:400–8.
15. Woll SB, McFall ME. The effects of false feedback on attributed arousal and rated attractiveness in female subjects. Journal of Personality 1979; 47:214–29.
16. Crucian GP, Hughes JD, Barrett AM, Williamson DJ, Bauer RM, Bowers D, et al. Emotional and physiological responses to false feedback. Cortex 2000; 36:623–47.
17. Scheirer J, Picard R. Affective objects. MIT Media Laboratory Perceptual Computing Section, Technical Report No 524, 1999.
18. Lang PJ, Bradley MM, Cuthbert BN. International affective picture system (IAPS): affective ratings of pictures and instruction manual. Tech Report A-6 1997. University of Florida.
19. Lang PJ. (1980). Behavioral treatment and bio-behavioral assessment: computer applications. In: Sidowski JB, Johnson JH, Williams TA, eds., *Technology in mental health care delivery systems*. Norwood, NJ: Ablex Publishing, pp. 119–37.
20. Bolls PD, Lang A, Potter RF. The effects of message valence and listener arousal on attention, memory, and facial muscular responses to radio advertisements. Communication Research 2001; 28:627–51.
21. Lang PJ, Bradley MM, Cuthbert BN. Emotion, motivation, and anxiety: brain mechanisms and psychophysiology. Biological Psychiatry 1998; 44:1248–63.
22. Caclin A, Soto-Faraco S, Kingstone A, Spence C. Tactile "capture" of audition. Perception & Psychophysics 2002; 64:616–30.
23. Tajadura-Jiménez A, Väljamäe A, Kitagawa N, Ho H. (2007) Whole-body vibration influences sound localization in the median plane. Proceedings of the 10th Annual International Workshop on Presence, Barcelona, Spain, October 25–27, 2007.
24. Slater M, Usoh M. (1994). Body centred interaction in immersive virtual environments. In Magnenat Thalmann N, Thalmann D, eds., *Artificial life and virtual reality*. London: Wiley, pp. 125–148.
25. Watson D, Tellegen A. Toward a consensual structure of mood. Psychological Bulletin 1985; 98:219–35.
26. Freeman J, Avons SE, Pearson DE, Ijsselsteijn WA. Effects of sensory information and prior experience on direct subjective ratings of presence. Presence: Teleoperators & Virtual Environments 1999; 8:1–13.
27. Graf C, Niebuhr S, Kohler K. (2006) Enhancing business software through fun-of-use: a pattern-based approach. Workshop on the role of emotion in human–computer interaction. HCI 2006.
28. Millard N, Hole L. (2008). In the moodie: using "affective widgets" to help contact centre advisors fight stress. In: Peter C., Beale R., eds., *Affects and emotion in human-computer interaction*. LNCS, vol. 4868. Springer, Heidelberg.
29. Pertaub DP, Slater M, Barker C. An experiment on public speaking anxiety in response to three different types of virtual audience. Presence: Teleoperators & Virtual Environments 2002; 11:68–78.
30. Picard RW, Daily SB. (2005). Evaluating affective interactions: alternatives to asking what users feel. Workshop on evaluating affective interfaces: innovative approaches. CHI 2005.

Address reprint requests to:
Ana Tajadura-Jiménez
Div. of Applied Acoustics
Chalmers University of Technology
Chalmers Room Acoustics Group
Sven Hultins gata 8a 41296, Gothenburg, Sweden

E-mail: ana.tajadura@ta.chalmers.se