Auditory-visual virtual environment for the treatment of fear of crowds

Marine Taffou[1,2,3,4,5,6,7,*], Jan Ondřej[8], Carol O'Sullivan[8], Olivier Warusfel[1,2,3], Stéphanie Dubal[4,5,6,7], Isabelle Viaud-Delmon[1,2,3]

[1]CNRS UMR 9912, STMS, F-75004, Paris, France
[2]IRCAM UMR 9912, STMS, F-75004, Paris, France
[3]Sorbonne Universités, UPMC Univ Paris 06, UMR 9912, STMS, F-75004, Paris, France
[4]Inserm, U 1127, F-75013, Paris, France
[5]CNRS, UMR 7225, F-75013, Paris, France
[6]Sorbonne Universités, UPMC Univ Paris 06, UMR S 1127, F-75013, Paris, France
[7]Institut du Cerveau et de la Moelle épinière, ICM, Social and Affective Neuroscience (SAN) Laboratory, F-75013, Paris, France
[8]Trinity College Dublin, Dublin, Ireland

ABSTRACT
Fear of crowds is a symptom found in agoraphobia (among others) and which has both visual and auditory relevant components. In order to examine the potential interest of auditory-visual virtual reality to treat fear of crowds, we investigated the efficiency of an auditory-visual virtual environment in evoking affective reaction. We conducted an evaluation test with healthy participants sensitive to the fear of crowds to explore the influence of auditory-visual virtual environment on affective reactions. Our application involves both high fidelity visual stimulation and interactive 3D sound. Our specific presentation of animated crowds of humans creates an environment that is highly arousing, suggesting that manipulating the sensory presentation together with the spatial location of stimulation might provide a way to modulate affective reactions.

General Terms
Measurement, Experimentation, Human Factors

Keywords
Fear of crowds, anxiety, multisensory integration, distance, spatial audition, emotion, therapy

1. INTRODUCTION
Virtual reality-based exposure therapy for phobias aims at reducing fear in response to the anxiogenic situation using a gradual confrontation to simulation of the real-life feared situation. Several phobias have been successfully treated with virtual reality-based therapy [e.g. 1, 2, 3]. Specifically, some studies showed a reduction of fear in patients suffering from panic disorder with agoraphobia after exposure sessions in virtual environments depicting a shopping mall, a bus or a subway train [4, 5]. In these studies, whereas an effort is made in rendering accurate visual stimulation, auditory rendering is often neglected. The auditory stimuli are often absent and if present are not rendered in 3D and are not interactive. Consequently, the auditory aspects of the phobia are often underexploited in the therapy. Thus, it is not yet clear how virtual reality involving multisensory stimulation impacts the treatment.

In a natural environment, emotional information is often conveyed via multiple senses and the interaction between sensory inputs from all modalities can influence perception and behavior in multiple ways. Moreover, auditory stimulation in virtual environment are known to improve immersion and feeling of presence [6, 7]. Thus, high fidelity auditory inputs may be of great importance to evoke, with virtual environments, affective reactions that are similar to those experienced in real-life.

The present study aimed at exploring the influence of multisensory stimulation, involving 3D interactive auditory stimuli, on affective reactions. We designed an auditory-visual virtual environment for the treatment of fear of crowds. Fear of crowds is a symptom of agoraphobia, which has a relevant acoustic aspect. Thus, it provides an ideal target to study how auditory and visual stimulation interact to produce affective reactions.

2. METHODS
Upon arrival, all participants provided informed consent to take part in the experiment, which was previously approved by the local ethical committee. Each participant first became acquainted with the equipment and the navigation mode (training) before navigating in a virtual environment containing crowds. Then, participants completed several questionnaires and were asked by the experimenter to comment on their experience (debriefing).

VRIC'15, April 8–10, 2015, Laval, France.

[*]Corresponding author. E-mail: marine.taffou@ircam.fr

2.1 Participants

Participants were recruited on the basis of a questionnaire exploring the fear of crowds (possible range for the score on this questionnaire: 0-45). Two hundred and twenty-eight individuals (121 women) completed this questionnaire. A distribution of scores was obtained and served as a basis to select participants to the current experiment. Twelve individuals (nine women; age = 23.3 ± 2.9), whose scores on the fear of crowd questionnaire were higher than 22 (80^{th} centile), were selected from this pool.

2.2 Virtual environment

The virtual environment reproduces the outdoor environment of the Trinity College Dublin campus composed of buildings, alleys and vegetation. Animated virtual individuals, referred to as humanoids, can be placed in this environment. The auditory component of the virtual environment consists of human speech and of an ambient audio environment composed of bird sounds and of urban activity.

In the virtual scene used for the exposure to virtual crowds, twelve crowds of 96 humanoids were distributed in the virtual environment. They all were static and could be presented through the visual channel only (unimodal visual: a silent crowd), through the auditory channel only (unimodal auditory: a crowd with talking humanoids that is hidden by a big flag) or through both the auditory and visual channels (bimodal auditory-visual: a crowd with talking humanoids).

2.3 Virtual reality setup

The virtual reality setup was installed in an acoustically damped and soundproof recording studio. The visual scenes were presented on a 300 x 225-cm^2 stereoscopic passive screen (see Figure 1). Users wore polarized stereoscopic viewing glasses. The auditory scenes were presented through Sennheiser HD650 headphones and the sound stimuli were processed through binaural rendering using non-individual Head Related Transfer Functions (HRTF) of the LISTEN HRTF database (http://recherche.ircam.fr/equipes/salles/listen/) previously selected as best-fitting HRTF for a majority of participants in different experiments involving binaural rendering [8, 9]. The ambient audio environment was rendered through virtual ambisonic sources and binaural audio rendering. Head movements were tracked using an ART optical system so that visual stereo and 3D sounds were appropriately rendered with respect to the users' position and orientation. The participants were equipped with a 3D mouse to navigate in the virtual environment. With this device, they could control both rotations and translation within the virtual scene.

2.4 Questionnaires and Interview measures

The state portion of the State Trait Anxiety Inventory (STAI) [10] was used before and after completion of the total experimental protocol. A 22-item cybersickness scale [11] and the presence questionnaire from the I-group [12] were presented at the end of the navigation in the virtual environment. Discomfort ratings were collected during all of the immersion in the virtual environment using Subjective Units of Distress (SUD) [13].

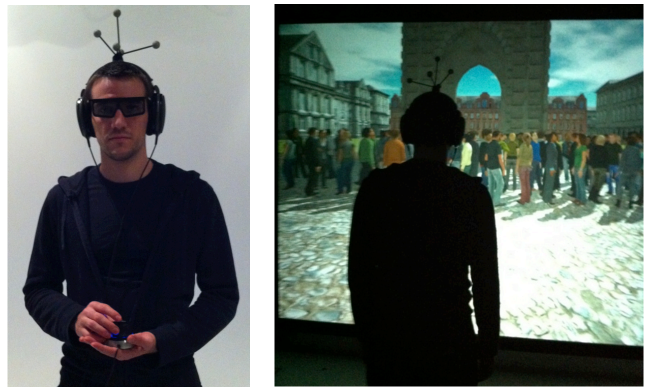

Figure 1. Virtual reality setup. (left) A participant equipped with polarized glasses, headphones and a tracking device. (right) A participant immersed in the virtual environment.

2.5 Procedure

First, participants completed the state portion of the STAI. Then, all participants took part in a training session completed in a virtual scene, which was humanoid-free. During training, the experimenter interacted with the participant in order to assist him/her in his/her first navigation.

After training, participants were immersed in the virtual scene with the aim to expose them to virtual crowds. Participants were instructed to follow numbered flags in order to explore the virtual scene. During the exploration, they encountered the different crowds, presented pseudo-randomly either in a unimodal or bimodal way. At each encounter with a crowd, participants had to rate their level of discomfort when they were at 8m and at 2m from the crowd.

After the navigation in the virtual environment, participants completed the cybersickness questionnaire and the presence questionnaire from the I-group. Finally, they completed the state portion of the STAI a second time.

3. RESULTS AND DISCUSSION

One participant did not complete the protocol because of high cybersickness. The measures collected on the eleven remaining participants are presented in Table 1.

There was no significant difference between state anxiety scores pre- and post-navigation in the virtual environment containing crowds (two-tailed non-parametric Wilcoxon T test for matched samples: $p = 0.919$). This could be due to the fact that the experiment was conducted with a non-phobic sample and/or to the presentation order of the different crowds that did not follow a progressive pattern of exposure.

We conducted two-tailed non-parametric Wilcoxon T tests for matched samples on the SUDs reported by participants during navigation in the virtual environment. At 2m from the crowd, participants reported higher discomfort levels in response to auditory-visual crowds compared to unimodal crowds ($p = 0.003$). This was not the case at 8m from the crowd ($p = 0.09$).

Table 1. Questionnaire scores and SUDs during navigation

Variable	Mean (SD)
Questionnaires scores	
State anxiety PRE	35.6 (8.0)
State anxiety POST	34.2 (9.4)
cybersickness	7.6 (6.6)
presence	50.3 (14.7)
SUDs during navigation	
At 8m	
SUD in response to unimodal crowds	24.4 (22.9)
SUD in response to bimodal crowds	27.4 (25.8)
At 2m	
SUD in response to unimodal crowds	42.3 (29.8)
SUD in response to bimodal crowds	50.5 (31.2)

Our results suggest that our auditory-visual virtual environment is highly arousing. Moreover, they suggest that manipulating the sensory presentation together with the spatial location of stimulation might provide a good way to modulate affective reactions. These findings underline the importance of rendering auditory as well as visual stimulation with high fidelity. Adding 3D, interactive audio in a virtual environment allows the manipulation of the spatial characteristics of auditory-visual stimulation, which provides more flexibility to control the emotional impact of the stimuli. Altogether, these results depict auditory-visual virtual reality as a promising tool for the treatment of fear of crowds.

4. ACKNOWLEDGMENTS

This research was supported by the EU FP7-ICT-2011-7 project VERVE (http://www.verveconsortium.eu/), grant n°288910. This work was performed within the Labex SMART (ANR-11-LABX-65) supported by French state funds managed by the ANR within the Investissements d'Avenir programme under reference ANR-11-IDEX-0004-02. The research leading to these results has also received funding from the program "Investissements d'avenir" ANR-10-IAIHU-06.

5. REFERENCES

[1] Botella, C., Bretón-López, J., Quero, S., Baños, R., & García-Palacios, A. (2010). Treating Cockroach Phobia With Augmented Reality. *Behavior Therapy*, *41*(3), 401–413.

[2] Garcia-Palacios, A., Hoffman, H., Carlin, A., Furness, T. A., & Botella, C. (2002). Virtual reality in the treatment of spider phobia: a controlled study. *Behaviour research and therapy*, *40*(9), 983–93.

[3] Rothbaum, B. O., Hodges, L. F., Kooper, R., Opdyke, D., Williford, J. S., & North, M. M. (1995). Effectiveness of computer-generated (virtual reality) graded exposure in the treatment of acrophobia. *American Journal of Psychiatry*, *152*(4), 626–628.

[4] Botella, C., Villa, H., Baños, R. M., Quero, S., Alcañiz, M., & Riva, G. (2007). Virtual Reality Exposure in the Treatment of Panic Disorder and Agoraphobia : A Controlled Study †. *Clinical Psychology & Psychotherapy*, *14*, 164–175. doi:10.1002/cpp

[5] Villa Martin, H., Botella, C., Garcia-Palacios, A., & Osma, J. (2007). Virtual Reality Exposure in the Treatment of Panic Disorder With Agoraphobia: A Case Study. *Cognitive And Behavioral Practice*, *14*(1), 58–69. doi:10.1016/j.cbpra.2006.01.008

[6] Hendrix, C., & Barfield, W. (1996). The sense of presence within auditory virtual environments. *Presence-Teleoperators and Virtual Environments*, *5*, 274–289.

[7] Viaud-Delmon, I., Warusfel, O., Seguelas, A., Rio, E., & Jouvent, R. (2006). High sensitivity to multisensory conflicts in agoraphobia exhibited by virtual reality. *European psychiatry : the journal of the Association of European Psychiatrists*, *21*(7), 501–8. doi:10.1016/j.eurpsy.2004.10.004

[8] Moeck, T., Bonneel, N., Tsingos, N., Drettakis, G., Viaud-Delmon, I., & Alloza, D. (2007). Progressive Perceptual Audio Rendering of Complex Scenes. In *Proceedings of the 2007 Symposium on Interactive 3D Graphics and Games, April 30-May 02, 2007*. Seattle, Washington.

[9] Sarlat, L., Warusfel, O., & Viaud-Delmon, I. (2006). Ventriloquism aftereffects occur in the rear hemisphere. *Neuroscience letters*, *404*(3), 324–9. doi:10.1016/j.neulet.2006.06.00

[10] Spielberger, C. D., Gorsuch, R. L., Lushene, P. R., Vagg, P. R., & Jacobs, A. G. (1983). *Manual for the State-Trait Anxiety Inventory (Form Y)*. Palo Alto, CA: Consulting Psychologists Press.

[11] Viaud-Delmon, I., Ivanenko, Y. P., Berthoz, A., & Jouvent, R. (2000). Adaptation as a sensorial profile in trait anxiety: a study with virtual reality. *Journal of Anxiety Disorders*, *14*(6), 583–601.

[12] Schubert, T., Friedmann, F., & Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. *Presence Teleoperators and Virtual Environments*, *10*, 266–281.

[13] Wolpe, J. (1973). *The practice of behavior therapy (2nd ed.)*. New York (NY): Pergamon.