Virtual Characters with Affective Facial Behavior

Ana Paula Cláudio[1], Augusta Gaspar[2], Eder Lopes[1] and Maria Beatriz Carmo[1]
[1]LabMAg, Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal
CIS- Centro de Investigação e Intervenção Social, ISCTE Instituto Universitário de Lisboa, Lisboa, Portugal

Keywords: Virtual Humans, Social Anxiety Disorder, Facial Expressions, Body Postures, Exposure Therapy, Nonverbal Behaviour Research.

Abstract: This paper describes an application that generates a simulation of a jury with one to three virtual humans capable of exhibiting facial and body expressions controllable in real-time. This control is performed through an interface that combines facial action elements (AU-Action Units, as described in the Facial Action Coding System), with upper-body postures. The level of detail of this control offers a range of possible combinations to obtain emotional expressions. Besides offering the possibility of controlling the postures of the virtual characters, the application allows the user to choose, among a pre-defined set, the virtual characters that compose the jury, and is able to introduce in the simulation some extra, potentially distractive or annoying events. We envisaged two contexts for using this application: i) assisting psychotherapists in exposure therapy of patients suffering from anxiety of public speaking, particularly in front of a jury in an assessment situation; ii) supporting nonverbal behaviour research carried out by psychologists. The development of the application has been closely monitored by a psychologist that is part of our team. This application is a low-cost approach, which uses only free software and models and resorts to common equipment; it is easy to install and use by people without expertise in informatics. So far, we have performed an evaluation with therapists in the first application context, obtaining encouraging results.

1 INTRODUCTION

Social Anxiety Disorder (SAD), or Social Phobia, is a condition characterized by intense anxiety whenever the individual is faced with a situation of public performance or exposure or even the bare anticipation of such a situation (APA, 2000). When forced to face them they report the experience as a torture.

This condition cripples one's social, professional personal life, presenting itself with a high co-morbidity with depression (Stein and Kean, 2000). SAD patients fear negative judgment, and believe to be so judged. Their perception of others emotional signals is similar to that of other people when the signals are neutral or characteristic of positive affect. But regarding signals that may carry negative, threatening content, they are faster, more effective detectors (Douilliez et al., 2012) thereby evincing their hypervigilance for social threat.

Therapies used to treat SAD often include medication (mostly anti-depressants), cognitive-behavioral therapy, psychotherapy and relaxation techniques. By far, cognitive-behavioral therapy is the one that has shown the most efficient and persistent results, especially in the form of exposure therapy (see Beidel and Turner, 2007).

Exposure therapy comprises today, and since the early 90's, approaches based on Virtual Reality (VR). This is called Virtual Reality Exposure Therapy (VRET). These approaches have been reported to produce similar outcomes to those of traditional exposure counseling (Klinger et al., 2004; Herbelin, 2005). VRET enables an accurate control of the process of habituation (and extinction) of fear to a phobic object, and thus it entails many additional advantages over classic exposure therapy, which involved imagery and subsequent contact with real life situations: a) VRET enables the customization of scenarios and interactions to meet the needs and progress of different patients, and those of each patient throughout the treatment period; b) it allows an optimization of patient preparedness before facing reality, thereby avoiding the risk of a miscalculated premature exposure to reality; c) by reducing the risk of throwbacks and excessive reactions, it provides a more stable, progressive, securing environment for treatment, that may ensure a predictable and consolidated patient

Cláudio A., Gaspar A., Lopes E. and Carmo M..
Virtual Characters with Affective Facial Behavior.
DOI: 10.5220/0004677803480355
In *Proceedings of the 9th International Conference on Computer Graphics Theory and Applications* (GRAPP-2014), pages 348-355
ISBN: 978-989-758-002-4

outcome; d) finally it protects the patients privacy. The typical major drawbacks of VRET are the financial cost of purchasing virtual immersive equipment (Head-Mounted Displays, CAVEs) and the secondary effects manifested by some users (*cyber sickness*) (LaViola, 2000).

This paper describes an application that generates a jury of virtual humans (VHs) displaying facial and body actions controllable in real-time. Control is achieved through an interface that (among other functions) supports the combination of facial action elements with upper-body postures, enabling a wide range of compositions that may convey neutral, positive or negative emotional content, as well as attention or lack of interest. The application can be used to assist psychotherapists in the exposure therapy of patients with fear of public speaking or other type of public performance, mainly in front of a jury in an assessment situation.

The application is also a useful tool in nonverbal behavior research: There are currently several applications that depart from the assumption that the content of facial behavior and other communication elements is fully known, when it is not (Gaspar and Esteves, 2012; Gaspar et al. in press), so a customizable tool that enables composing different constellations of units of facial and body behavior whilst controlling the effect of context and other variables on perception and emotional reactions is a most useful, non-biased tool. The effect of changing a single action in a face whose actions are otherwise kept constant can be dramatic (see Figure 4); assessing its impact on observers, may clarify the communicative role of unitary or composite actions, and assist in establish objective parameters for exposure counseling as well as for the advancement of research in nonverbal behavior.

Our application simulates a jury that can be set to comprise 1-3 characters. The one character jury simulates an interview situation whereas the three character jury resembles more an MSc. dissertation defense or a contest jury. As to setting alone, there are multiple customization options in the application – from room features to the placement of different VH jury members, their clothing and other qualities of their looks before setting their behavior. This application is designed to be projected in a canvas or wall, at a distance whereby VHs approach real size and their background occupies an entire wall, creating a strong immersive effect. The team's psychologist played a decisive role in sorting behavior units to include in the application. Inclusion relied on two premises –validity of the behavior signal value and behavior appropriateness to SAD therapy.

The work herein described follows from previous work in which our team designed a set to expose patients to a large auditorium type virtual audience (Cláudio et al., 2013; Cláudio et al., 2012). The distinctive features of the current work are the level of detail that can be customized for the virtual humans' aspect and behavior (facial expressions, gaze orientation and posture) and the context, which implies a much closer interaction between user and VHs. The current work was also developed with a different tool, a freeware version of Unity 3 (url-Unity), a software tool for the production of video-games.

In both the current and the auditorium application we aimed to develop low cost, "portable" and ergonomically optimized VR tools that could be widely used in counseling and research.

This document is organized as follows: next section presents some of the most relevant related work; then we describe our approach and the evaluation performed with therapists, and finally we draw conclusions and present future work.

2 RELATED WORK

VRET has been applied since the 1990's to treat diverse phobic conditions. And targeting social anxiety or social phobia, namely the fear of public speaking, poses extra challenges: the inclusion of credible virtual characters into the VR scenario.

North et al. (1998) presented the first VR application to treat fear of public speaking - it included a scenario with an auditorium resembling a university auditorium and it could be filled with up to 100 characters. During the session the therapist could vary the number of characters and their attitudes using pre-recorded video sequences. The patient used a head-tracked HMD with an attached loudspeaker to hear the echo of his own voice reverberating in the auditorium.

Slater at al. (1999) created a public speaking simulation in a virtual seminar room with 8 characters sat in semi-circle exhibiting random autonomous behaviour such as head-nods, blinking and twitches. The initial study, with 10 students with different levels of difficulty in public speaking was later extended to include phobic and non-phobic individuals (Pertaub et al., 2001; 2002; Slater at. al, 2006). Pertaub et al. (2002) recreated 3 different types of audience behavior: an emotionally neutral audience that remained static throughout the talk, a positive audience showing friendly and appreciative behaviour, and a negative audience, displaying hostile and bored expressions throughout the talk. For the positive and negative scenarios 10 different audience behaviours were scripted. Based on the observation of the actual progress of the subject's speech, the therapist could decide the start time for each script; but the script order was predetermined.

Herbelin et al. (2002) used snapshots of eyes to recreate a virtual human audience. The snapshots were placed in concentric circles around the user and the system allowed pre-setting the number of circles, the number of snapshots and the distance to the user in order to change induced anxiety.

Anderson et al. (2003) created a virtual podium suitable for reading, placed in front of virtual curtains; when these open, one of two sets may appear: a seminar room with 5 people around a table or an auditorium with 25 people, made of video-clips of real humans. Along the treatment, the patient is exposed to the most feared situation.

James et al. (2003) proposed a two-fold scenario: an underground train with characters expressing neutral behaviors (a non-socially demanding environment) and a bar where characters seemed aloof and detached – a socially demanding situation. The behavior of characters included eye gaze and pre-recorded sentences. Authors reported that the bar situation was more likely to trigger anxiety in phobic subjects than the underground train.

Klinger et al. (2004) conducted a study that examined changes in the fear of public speaking in 36 participants over 12 sessions. For the virtual characters they used simple billboards on which pictures of real people involved in daily situations were projected. Participants were divided into 2 groups, one treated with traditional CBT and another with VRET. The VRET included 4 virtual environments simulating social situations involving performance (eg, public speaking), inter-personal interaction (eg, a dinner conversation), assertiveness (eg, defending an idea) and evaluation (eg, talk while being observed). Patients in the VRET group were reported to show a larger reduction in social anxiety than patients in the CBT group. Herbelin (2005) published a validation study with 200 patients, confirming that his VR platform met the requirements of VRET therapeutic exposure for social phobia. He also proved that it is possible to improve clinical assessment with monitoring tools integrated in the application, such as eye-tracking.

All the referred approaches have used head-tracked HMD; in the study described by Pertaub et al. (2002), half of the patients experienced the virtual seminar room through a HMD and the remainder on a desktop. Herbelin (2005) and Grillon (2009) besides using HMD and a simple computer screen have also used a large back-projection screen.

Haworth et al. (2012) implemented virtual scenarios to be visualized simultaneously by patient and therapist, possibly in different physical locations and communicating via a network. The scenarios target acrophobia and arachnophobia patients and are visualized in simple desktop screens; the patient sees only the hands and feet of his avatar in the virtual world whilst the therapist has no corresponding avatar in the virtual environment. A Kinect is used to track the movement of the patient's body. The few results of this study so far seem to demonstrate that this type of low-cost solution is effective for these particular phobic situations.

Vanni et al. (2013) report a survey on virtual environments devoted to treating the fear of public speaking and in their words "VEs represent an emerging, promising tool to carry out exposure treatment, better than imaginary exposure and, potentially, as good as in vivo exposure, in triggering anxious reactions". They point out that the visual fidelity of the virtual character is not of major importance, as behaviour is what really triggers emotional reactions. Thus, the VH´s facial expressions, gaze direction, and varied other actions are the crucial features of VRET.

The main distinctive characteristics of our application, besides the low-cost of the projection method, are: a) the high level of control that the user has of the simulation, made possible by a special purpose user friendly interface that allows real-time control of facial expressions and body postures of the 3D VHs in the jury; b) being a low cost solution – this is a feature that is meant to ease use and dissemination; c) it eliminates the often reported drawback of side effects of immersive VR equipment, such as motion sickness and nausea.

3 THE INTERACTIVE VIRTUAL JURY

3.1 Requirements' Gathering

The application herein described follows from a previous work designed to expose patients with fear of public speaking to a virtual auditorium inhabited by controllable virtual humans (Cláudio et al. 2013; Cláudio et al. 2012). The present work, required a scenario with considerably fewer characters: a virtual jury simulating an assessment situation. This was challenging due to the proximity between user (i.e. in therapy, the therapist or patient, in research, the observer) and the virtual characters, creating major concerns about their appearance and behaviour.

In both applications, the main goal was to create low cost, “portable” and ergonomically viable VR tools that could be widely used to assist therapists in the intermediary stages of exposure therapy treatment, between an *in imagino* exercise - imagined exposure situation - and *in vivo* situation - exposure to a real situation.

To fulfil this requirement, the apparatus we adopted for the application is comprised of a computer (with a mid-range graphics card), a projector and a canvas, and two sound columns. The application provides two separate windows: the simulation window, which is projected and displays the virtual jury, and the interface window, which is displayed in the computer screen and used to configure and to control the simulation. The sound is synchronized with events happening in the simulation.

This structure has the advantage of using inexpensive and easily installed equipment while, by projecting an image on a large screen, generating a feeling of immersion. Furthermore, it can be observed simultaneously by several people, which may be advantageous in education and training or in collective research sessions aimed at studying appraisal and reactions to the VHs behavior.

Comparing with our previous work, the distinctive requirements posed by the current work are the higher quality of the VH models and the level of detail that can be customized and controlled: gaze orientation, facial expressions and posture.

Requirements common to both applications, besides the low-cost, are:

- Simulations controllable in real-time by the therapist. The most important condition to perform a successful exposure therapy is triggering degrees of discomfort in the patient that are similar to those experienced in real situations, i.e., causing in the patient the feeling of presence (Herbelin, 2005).
- Two distinct windows. One containing the simulation, the other containing the interface to control the content of the previous one.
- Support of therapy sessions taking place with therapist and patient in the same room. This maintains proximity during the therapy, without risking the dehumanizing of treatment. A session is centered in a simulation that is controlled by the therapist and watched by the patient during a proposed task. The therapist, attentive to patient performance, controls stimuli intensity, triggering specific events in the simulation.
- Friendly and easy to learn therapist interface; easy to install and use for people without expertise in informatics.
- Configuration of the scenario prior to the simulation.
- Induce in the observer a strong sense of presence. The models of the characters in the jury and their behavior should be compelling and capable of looking at the patient.

To implement the whole application we used the free version of Unity, as previously mentioned. The VH models were free of cost using MakeHuman (url-MakeHuman); to adjust and animate the VH models and to model some assets of the scenario, we used Blender (url-Blender). As to hardware, we used a computer with processor: 2x Intel ® CORE ™ 2 Duo E8400@3.00GHz, 4GB Memory, Graphics Card Quadro FX 1700/PCI/SSE2 (a mid-range mobile graphics solution). We also resorted to a projector, a projection screen (or blank wall) and two speakers installed close to the projected image.

3.2 The Configuration and Simulation Interfaces

The application generates a jury simulation with 1-3 VHs capable of exhibiting facial and body expressions controllable in real-time. It displays two windows: one for the simulation and another for the user interface. In the snapshot of the simulation window in Figure 1 all characters are distracted; while one of them uses the laptop, another is whispering to the character sitting beside.

Figure 1: A snapshot of the simulation window.

The interface window displays different contents, depending on the time of use: i) the configuration interface to customize the scene and choose the virtual characters; ii) the simulation interface to control the characters and the events during the simulation; iii) the facial expression interface to control the facial actions of characters.

These interfaces are illustrated and described in the following subsections.

3.2.1 The Configuration Interface

In the configuration interface it is possible to choose the type of scenario, classic or modern, and the color of the walls and also, from a set of models, the characters of the jury, 1 to 3, and their corresponding position sitting in front of the observer. There are models of both genders, with different clothes and skin colors; any model can wear glasses. There are also 3 sliders that can be used to calibrate the camera, i.e. to adjust its position relatively to the observer's position.

Along the configuration phase, all choices made using this interface are immediately exhibited in the simulation window which is also visible.

3.2.2 The Simulation Interface

The simulation interface can be seen in Figure 2. Choices performed in this interface have effect in real time in the simulation window.

To manage interface complexity, the application supports user control of a character at a time, while the others exhibit a pre-defined behavior previously chosen. We call these control modes, interactive mode and automatic mode, respectively. Throughout the simulation the user is free to choose different characters to perform interactive control.

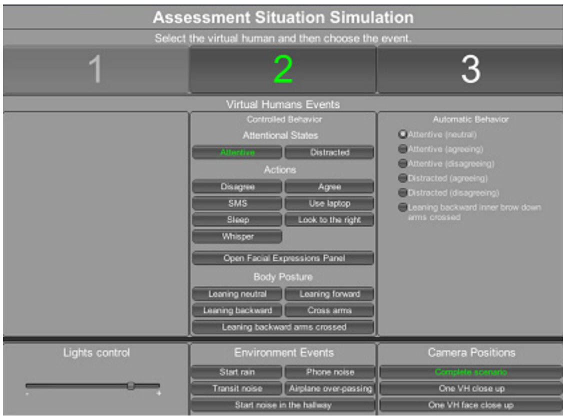

Figure 2: The simulation interface.

Figure 2 shows the simulation interface that corresponds to a previously configured jury of 2 elements, one on the middle (position number 2) and one on the right (position number 3). The first character is in interactive mode of control while the second is in automatic mode. Notice the differences in the available options; the interactive mode is significantly more complex.

There are six automatic modes that can be chosen alternatively: *Attentive* (*neutral*), *Attentive* (*agreeing*), *Attentive* (*disagreeing*), *Distracted* (*agreeing*), *Distracted* (*disagreeing*) and *Leaning backward, inner brow down*. They correspond to predefined animations executed by the VHs; the application is prepared to an easy addition of animations to the automatic mode of control. At the bottom there are three types of functionality (Figure 2, from left to right): i) a slider to control light intensity; ii) buttons to trigger sound events in the environment (plane flying over, phone ringing, traffic in the outside, rain, a conversation in the hallway); iii) buttons to control the position and zooming of the camera. This last set of buttons offers three alternatives: to visualize all characters (default), only the body and the face of the VH in interactive mode or only its face. This feature is particularly suitable to focus the observer’s attention on a particular character.

The interface area to control the HV in interactive mode (see position number 2 in Figure 2) has three areas of buttons: i) defining two possible states of attention (*Attentive, Distracted*); ii) those that define what we called by *Actions* (*Disagree*, *Agree, SMS* –answers to a text message that has just arrived– *Use laptop, Sleep, Look to the right, Whisper* –starts a conversation with the character sitting next, as shown in Figure 1) and, finally, a button to open the facial expressions menu, explained in the next subsection. At the bottom, there are buttons to control body postures: *Leaning neutral*, *Leaning forward*, *Leaning backward*, *Cross arms*, *Leaning backward arms crossed*. Some, like waving the head as a sign of agreement (*Agree* button) have a predetermined execution time, while others, like *Use laptop* are executed as long as the user desires. It is possible to combine some of the animations like, for instance, *Cross Arms* and *Agree*.

3.2.3 The Facial Expressions Interface

As seen in Figure 3, there is a facial expression control panel which pops out after clicking the button “Facial Expressions”. The menu comprises single elements of facial expressions and not full facial expressions – i.e. buttons provide the so called facial action units or AUs. AUs were taken from the *Facial Action Coding System* or FACS (Ekman et al., 2002) and in two AUs (AU4; AU12) there is an option for 2 intensity levels. At this stage of development it is extremely difficult to reproduce all 5 levels described in FACS for human spontaneous behavior. We have chosen only a few FACS AUs, as we based our work in only the most validated facial expression- context associations from studies of behavior and perception (Gaspar, Esteves and Arriaga, in press), as there has been great controversy surrounding the assumption that certain facial expressions convey given discrete emotions without support from studies of spontaneous facial expression (eg Russell and Fernandez-Dolls, 1997). These AUs can be combined and their simultaneous insertion into the VH's face is what creates its facial expression. These facial compositions can be combined with various options of body orientation and posture, thereby creating full face-body interactors. Postures and orientations were chosen according to relevant content findings in human nonverbal communication (Eibl-Eibbesfeldt, 1989).

The facial expression control panel includes actions though to convey positive affect - AU6+12, – with two intensity levels as options, activated by the buttons "Smile (AU6+12)" and "Smile++ (AU6+12)", and AU12 alone, activated by the "Lips up (AU12)" button. It also includes actions often processed as negative by observers, and associated in the sender with either attention (AU4, AU5, AU2, AU1+2) or negative affect (AU4, AU1+4 and AU15). The frown (AU4), a movement that brings eyebrows closer, forming wrinkles in between, and lowering the inner corners of eyebrows as well, which in humans is produced by the contraction of the *Corrugator* muscle, is particularly relevant and thus the two intensity levels– the buttons "Brows brought together (AU4)" and "Brows brought together (AU4)++". The menu also includes two buttons for baseline brows and baseline mouth, which allow the user to set the VH back to a virtually Emotion-inexpressive face.

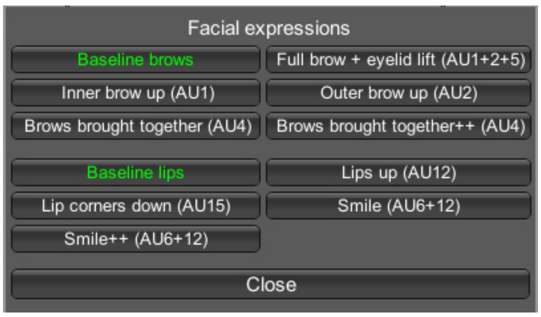

Figure 3: The facial action units'(AUs) panel, which allows for the composition of patterns (gestalten) of AUs or facial expressions. The buttons turn green when the respective AUs are active in the VH's face.

Figure 4 displays the same VH model with different combinations of AUs. The baseline face, with no AUs activated (top-left) and a frown, AU4 (top-right). At the bottom right, we see a combination of AU4 and AU15, both associated with negative affect. At the bottom left, we see a combination of AU4 and AU6+12 (the smile button). The combination AU6+12 is the defining feature of the "*duchenne smile*" which is supposed to convey positive affect, more so and more honestly than any other smile. The threatening effect frown is mitigated and possibly the frown is here perceived with other frequent context – attention.

Figure 4: Clockwise from top: a VH with a baseline face (no AUs), displaying a frown (AU4), displaying a combination of AU4 and AU15, displaying a combination of AU4 and AU6+12 (the smile button).

4 EVALUATION

To assess the potential utility of this application in the domain of VRET, in exposure therapy targeting patients suffering from anxiety of speaking in front of a jury, we performed an evaluation with six therapists (2M ages 41 and 59; 4F ages 29-45), familiarized with the use of exposure therapy and with no prior contact with the application.

The test had two successive phases. The first, with ca. 15 min, was a period of familiarization with the tool. In the second phase, which took 15-20 minutes, the therapists were asked to assume that they needed to use exposure therapy for a regular therapeutic cycle, to treat a patient suffering from a high level of anxiety regarding assessment in front of a jury. The therapist had to use the application to rehearse a first VR therapy session to be used in the early stage of treatment, and a second VR therapy session to be used in a later stage of treatment. After each session, the therapists answered some questions posed by the interviewer and gave scores to the experienced functionalities, using a scale from 1-bad to 5-excellent.

The overall classification to the ergonomics of the functionalities in the configuration interface was excellent and all users rated the number of VHs in the jury adequate. As to the ergonomics of the available functionalities in the simulation window, scores were all 5 except one: the ease in changing the facial expression (M=4.5; SD= 0.58).

Table 1: Scores obtained in aspects concerning the scenario, the events and the VH in the simulation.

	Average	SD
The realism of the virtual characters	3.75	0.50
The realism of the facial expressions	3.75	0.50
The realism of the body postures	4.25	0.50
The realism of the automatic behaviors	4.25	0.50
The realism of the scenario	4.5	0.58
The interest of the simulated events	4.25	0.96

Table 1 shows average rates and standard deviation (SD) in other aspects concerning the scenario, events and the VHs in the simulation. The realism of the scenario had the highest score (M= 4.5; SD=0.58), whilst the lowest score was obtained in the realism of the virtual characters and in the realism of the facial expressions (M= 3.75SD= 0.5).

Regarding the level of difficulty in using the prototype, psychotherapists unanimously classified it as simple, and said they would be willing to use this application to support a session of exposure therapy.

The best rated features in the simulation were: i) the real time control of jury behavior and the ease doing so; ii) the possibility of combining facial expressions and body postures, to convey emotion relevant information and the wide range of combinations; iii) being able to choose having characters nodding to show agreement or the reverse by shaking the head, to induce positive or negative reinforcement; iv) the credibility of the behavior – generating scenes similar to those of a real jury. We also received suggestions to make the VH's appearance more flexible (e.g. more outfit and age options) and- specifically to include an older men wearing a suit and a tie.

5 CONCLUSIONS AND FUTURE WORK

The user tests were aimed at evaluating the application from the perspective of a psychotherapist counsellor leading an exposure therapy session to treat anxiety in a patient who fears speaking in front of a jury.

The user-therapists were unanimous in stating that they would use this application to support an exposure therapy session. This indicates that our low-cost solution has a real potential for use in that context, even in the absence (and perhaps also because of it) of immersive, special purpose VR hardware. We acknowledge the need to validate the utility of the application with a clinical population and to establish standard values of stress reactions (eg. physiological measures of stress and self-reports) and emotional content interpretation in the normative population. This validation is in preparation. Despite this (temporary) shortcoming, we feel quite optimistic about the real therapeutic potential of the application, considering the feedback of the therapists that evaluated it. The validation will be decisive in confirming/refining the virtual humans most adequate and effective behaviour units toward the patients' progress.

The features with lower scores in the evaluation highlight the need to improve the VH models. This was actually expected, given the quality limitations of the freeware models. Notwithstanding, the realism can only be improved up to the limit of maintaining the rendering of the simulation in real-time, and most importantly, the virtual fidelity of the models has been shown to be less relevant than the behaviour displayed by the characters (Vanni, 2013). This conclusion is supported by tests with real patients suffering from fear of public speaking, so the aesthetical appraisals of therapists or researchers are secondary to the purpose. Our VH display signals with ecological validity, as these patients are indeed biased towards certain behaviour signals in their audiences, such as frowns (Esteves, 1999).

The application has also great potential for research, allowing to investigate the function of facial and body actions thought to express emotional content, by measuring responses and interpretation to different constellations of signals, and systematically test their impact, controlling for the effects of single components. We are currently improving the application at two levels: diversity of the avatars appearance (physiognomy and clothing), and the neutrality of avatars physiognomy. It is worth mentioning that physiognomy always affects appraisal, no matter how "neutral" facial features are. There is no face 100% neutral be it brow shape, gender or attractiveness (Adams et al., 2012). So, validation in the wider population should allow us to let us sort the effects of signals and physiognomy.

More scenarios will be developed to extend the application (eg with a group of people in an informal reunion in a bar). An artificial intelligence module concerning the simulation of emotions will be developed and fully integrated in the application.

We feel that the implemented scenarios are useful in the academic world, assisting for example, graduate and undergraduate students, who frequently seek the university psychologist in relation to their anxiety toward public presentations.

To keep updated on our progress please consult the project webpage (url-VirtualSpectators).

ACKNOWLEDGEMENTS

We thank the Portuguese Foundation for Science and Technology and the R&D unity LabMAg for the financial support given to this work under the strategic project Pest OE/EEI/UI0434/2011.

REFERENCES

Adams Jr., R. B, Nelson, A. J., Soto, J. A., Hess, U., Kleck, R. E., 2012. Emotion in the neutral face. *Cognition and Emotion, 26 (3)*, 431-441.

APA- American Psychiatric Association, 2000. DSM-IV-TR. *American Psychiatric Publishing, Inc.*.

Anderson, P., Rothbaum, B., Hodges, L. F., 2003. Virtual reality exposure in the treatment of social anxiety. *Cognitive and Behavioral Practice*, 10(3), 240-247.

Beidel, D. C., Turner, S. M., 2007. *Shy Children, Phobic Adults: Nature and Treatment of Social Anxiety Disorder*. Washington, DC: APA, 2nd edition.

Cláudio, A. P., Carmo, M. B., Esteves, F., Pinheiro, T., Lima, J., 2012, Virtual environments to cope with anxiety situations - Two case-studies. In IEEE *Information Systems and Technologies* (CISTI), 7th Iberian Conference, pp. 1-4.

Cláudio, A. P., Carmo, M. B., Pinheiro, T., Esteves, F., Lopes, E., 2013. Virtual Environment to Treat Social Anxiety. In A. *Marcus (Ed.): DUXU/HCII 2013, Part II, LNCS 8013*, *Springer, Heidelberg*, 442-451.

Douilliez, C. Yzerbyt, V., Gilboa-Schechtman, E., Philippot, P., 2012. Social anxiety biases the evaluation of facial displays: Evidence from single face and multi-facial stimuli. *Cog & Emo*, 26(6): 1107-1115.

Eibl-Eibesfeldt, I., 1989. *Human Ethology*. NY: Aldine de Gruyter.

Ekman, P, Friesen, WV., Hager, J C., 2002. *Facial action coding system*. Salt Lake City, UT: Research Nexus.

Esteves, F, 1999. Attentional bias to emotional facial expressions. *European Review of Applied Psychology / Revue Européenne de Psychologie Appliquée, Vol 49(2)*, 91-97.

Gaspar, A., Esteves, F., 2012. Preschoolers faces in spontaneous emotional contexts – how well do they match adult facial expression prototypes? *Int J Behav Dev. 36(5)*, 348–357.

Gaspar, A., Esteves, F., Arriaga, P. (in press). On prototypical facial expressions vs variation in facial behavior: lessons learned on the "visibility" of emotions from measuring facial actions in humans and apes. In *O. Pombo and N. Gontier (Eds), The Evolution of Social Communication in Primates: A Multidisciplinary Approach, Interdisciplinary Evolution Research. NY: Springer.*

Grillon, H., 2009. *Simulating interactions with virtual characters for the treatment of social phobia*. Doctoral dissertation, EPFL.

Haworth, M. B., Baljko, M., Faloutsos, P., 2012. PhoVR: a virtual reality system to treat phobias. In *Proceedings of the 11th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry*, 171-174.

Herbelin, B., Riquier, F., Vexo, F., Thalmann, D., 2002. Virtual reality in cognitive behavioral therapy: a preliminary study on social anxiety disorder. In *Proceedings of the 8th VSMM2002.*

Herbelin, B., 2005. *Virtual reality exposure therapy for social phobia*. Doctoral dissertation, EPFL.

James, L. K., et al., 2003. Social anxiety in virtual environments: Results of a pilot study. *Cyberpsychology & behavior, 6(3)*, 237-243.

Klinger, E., et al., 2004. Virtual reality exposure in the treatment of social phobia. *Studies in health technology and informatics, 99,* 91.

Klinger, E, et al., 2005. Virtual reality therapy versus cognitive behavior therapy for social phobia: A preliminary controlled study. *Cyberpsychology & behavior, 8(1)*, 76-88.

LaViola Jr, J. J., 2000. A discussion of cybersickness in virtual environments. *ACM SIGCHI Bulletin, 32(1)*, 47-56.

North, M. M., et al., 1998. Virtual reality therapy: an effective treatment for the fear of public speaking. *The International Journal of Virtual Reality 3,* 1-6.

Pertaub, D. P., Slater, M., Barker, C., 2001. An experiment on fear of public speaking in virtual reality. *Studies in health technology and informatics*, 372-378.

Pertaub, D. P., Slater, M., Barker, C., 2002. An experiment on public speaking anxiety in response to three different types of virtual audience. *Presence: Teleoperators & Virtual Environments, 11(1)*, 68-78.

Russell, J. A., Fernandez-Dols, J. M., 1997. What does a facial expression mean? In *The psychology of facial expression*, pp. 3–30. New York, NY: Cambridge University Press.

Slater, M., et al., 2006. A virtual reprise of the Stanley Milgram obedience experiments. *PloS one, 1(1), e39.*

Slater, M., Pertaub, D. P., Steed, A., 1999. Public speaking in virtual reality: Facing an audience of avatars. IEEE CG&A, *19(2)*, 6-9.

Stein, M.B. and Kean, Y.M., 2000. Disability and Quality of Life in Social Phobia: Epidemiologic Findings. *Am J Psychiatry 157*: 1606–1613.

Vanni, F., et al., 2013. A survey on virtual environment applications to fear of public speaking. *European Review for Medical and Pharmacological Sciences*, *17*(12), 1561-1568.

Unity http://Unity.com/

Blender http://www.blender.org/

MakeHuman http://www.makehuman.org

VirtualSpectators https://labmag.ul.pt/projects/virtual-environments-to-treat-social-anxiety