Using Expressive and Talkative Virtual Characters in Social Anxiety Disorder Treatment

Ana Paula Cláudio[1], Maria Beatriz Carmo[1], Augusta Gaspar[2,3] and Renato Teixeira[1]
[1]*LabMAg, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisboa, Portugal*
[2]*CIS- Centro de Investigação e Intervenção Social, ISCTE Instituto Universitário de Lisboa, 1649-026 Lisboa, Portugal*
[3]*Grupo de Psicologia, Faculdade de Ciências Humanas, Universidade Católica Portuguesa, 1649-023 Lisboa, Portugal*
{apc, bc}@di.fc.ul.pt, augusta.gaspar@iscte.pt, fc37057@alunos.fc.ul.pt

Keywords: Virtual Reality, Virtual Reality in Exposure Therapy (VRET), Social Anxiety, Fearing of Public Speaking, User Study, Nonverbal Behaviour Research.

Abstract: Social Anxiety affects a significant number of people, limiting their personal and social life. We describe an interactive Virtual Reality approach to the exposure therapy for social anxiety, resorting to virtual characters that exhibit combinations of facial and body expressions controllable in real-time by the therapist. The application described in this paper updates and significantly improves a former version: ameliorating the graphical quality of the virtual characters and providing them with the ability of articulating a set of sentences. The application executes in ordinary computers and it is easily used in counselling and research contexts. Although we have only resorted to free or very low cost 3D models of virtual humans, we adopted strategies to obtain an adequate final quality that we were able to validate with a significant number of observers. Moreover, a set of therapists tested the application and gave positive feedback about its potential effectiveness in Virtual Reality Exposure Therapy.

1 INTRODUCTION

Social phobia or Social Anxiety Disorder (SAD) is a human condition characterized by intense anxiety when the individual faces or anticipates public performance (APA, 2000). This condition can be very crippling in the personal, social and professional domains, as those bearing it withdraw from social contact; it also has a high comorbidity with depression (Stein, 2000). People with SAD fear negative social judgments and are hypervigilant for signals in other's behavior, thereby identifying faster and more efficiently than other people facial clues to threatening or negative content (Douilliez, 2012).

Therapy approaches to SAD include medication relaxation methods, and psychotherapy, mainly Cognitive-Behavioral Therapy (CBT). CBT produces the most efficient and persistent improvements, especially when it is applied as Exposure Therapy (ET) (Beidel, 2007) which consists in exposing the patient to the feared situation.

Virtual Reality (VR) has been used in ET since the early 90's, being called Virtual Reality applied to Exposure Therapy (VRET).

Several studies have concluded that VRET produces results that are similar to traditional exposure therapy (Klinger, 2004; Herbelin, 2005). VRET allows a precise control over the habituation (and extinction) to the fear of the phobic object and offers thus, several additional advantages over classic ET (which is based on images and later contact with *in vivo* situations).

Comparing with traditional ET, VRET presents some important advantages: i) it allows scenario configuration and interactions in order to fulfill each patient's needs and progress levels along the therapy; ii) it provides better preparation of the patient before facing a real life scenario, avoiding the risk of a premature exposure to a real situation; iii) it reduces the risk of taking steps backwards because of overreactions, allowing a more stable and progressive environment towards predictable and solid results; iv) it assures patient privacy.

The downside of VRET is the cost of the immersive virtual equipment (e.g., Head-Mounted Displays, CAVE) and sometimes the secondary effects reported by a few users (cybersickness) (Laviola, 2000).

This paper describes a VR approach to the treatment of SAD, specifically to the fear of public speaking - before a jury in an evaluation context, in a job interview or in some other similar scenario. The main idea is that in a session therapy, while the patient faces a jury of one-to-three virtual characters that show facial and body expressions, the therapist controls these characters according to the level of stress he/she wants to induce in the patient. This control is accomplished through an interface (visible only to the therapist) which, among other options, triggers specific facial and body movements that can be combined to convey neutral, positive or negative emotional content, and simulate various degrees of attention or lack of interest.

Before a therapy session, the therapist configures the scenario by choosing the characters and their look (hairstyle, clothes and glasses), their position at the table and also the appearance of the room (classical or modern furniture, different wall colors). The simulation scenario must be projected on a canvas or a wall, in such way that characters are displayed in real size, enhancing the immersive effect.

Our application is also a useful tool to the investigation of non-verbal behavior: the communication effect related to facial or body actions is not entirely known, and it is currently object of intense academic debate (Gaspar 2014; Russell, 1997). The application allows the individual visualization of a virtual human's face (henceforth VH) and supports many combinations of body and facial behavioral units. It also makes possible a fine tuning control over a single individual facial action, whilst others are kept constant. Evaluating the impact of these combinations in observers may clarify the communicative role of single and composed actions, giving clues about the treatment of social anxiety. It may also assist various other research lines in the field of non-verbal behavior.

The work described in this paper follows previous work from the same team. A first version of the application recreates an auditorium filled with virtual characters with controllable behaviors but with few facial expressions (Cláudio, 2013); a second version considers a simulation of a jury composed by a set of virtual characters with body and facial expressions controllable by the therapist in real time (Cláudio, 2014). The development of each version has been closely followed by a psychologist who has played a crucial role in the identification of facial and body expressions that potentially convey to the observer a positive, a neutral or a negative feedback.

This paper addresses the implements included in the present version, thus focusing the quality of the virtual models; we report results from a validation study concerning the VH's facial expressions, and the psychologist's assumptions underpinning the creation of the facial actions' menu. Usability tests were performed with experts (therapists) in order to study the suitability of the application and its potential effectiveness to VRET.

The document is organized as follows: next section presents some of the most relevant work in the area; section 3 describes our approach and the implemented application; section 4 reports the VH's facial expression validation process and the evaluation of the application by therapists; finally, section 5 draws conclusions and lines for future work.

2 RELATED WORK

The first VR application to treat public fear before an audience was presented in (North, 1998). This application included a scenario with up to 100 characters. During a therapy session, the therapist was able to vary the number of characters and their attitudes, using pre-recorded video sequences. Patient used an HMD, listening to the echo of his own voice.

Slater et al. created a virtual room with 8 characters with random autonomous behaviors, such as swinging the head and eye blinking (Slater, 1999). The initial study gathered 10 students, with different levels of difficulty about public speaking, and was extended later to include phobic and non-phobic individuals (Pertaub, 2001; Pertaub, 2002; Slater, 2006).

James et al. proposed a double scenario: one subway wagon populated by characters that express neutral behaviors, which is considered a non-demanding situation from a social interaction point of view; and a more demanding situation that took place in a bar with characters that look uninterested (James, 2003). The characters' behavior included eye gazing and pre-recorded sentences.

Klinger et al. conducted a study with 36 participants to evaluate changes in fear before public speaking during 12 sessions (Klinger, 2004). To recreate virtual characters, they used real people's photos in typical situations. Participants were divided in 2 groups: one treated with CBT and the other with VRET. A higher reduction in social anxiety was reported in the VRET group.

Herbelin published a 200 patient validation test, demonstrating that his platform fulfilled the requirements of therapeutic exposure in social phobia (Herbelin, 2005) and that clinic evaluation can be improved with integrated monitoring tools, such as eye-tracking.

All the referred approaches resorted to HMD equipment; in a study described by Pertaub et al. half of patients tried one of the virtual environments through a HMD, while the rest of the group used a desktop (Pertaub, 2002). Herbelin and Grillon, in addition to an HMD and a computer screen also used a big projection surface (Herbelin, 2005; Grillon, 2009).

Haworth et al. implemented virtual scenarios to be visualized simultaneously by patient and therapist in computer screens, possibly in different physical locations and over the internet (Haworth, 2012). Scenarios are oriented to patients suffering from acrophobia (fear of heights) or arachnophobia (fear of spiders). A Kinect is used to control patent body movements (url-Kinect). The few results of this study suggest that this type of low-cost solution is effective in these phobias.

3 VIRTUAL SPECTATORS

Virtual Spectators is our VRET approach to the fear of public speaking. The virtual scenario is, during a therapy session, the stage to a simulation controlled by the therapist and observed by the patient.

The application has two types of users: the therapist as an active user, and the patient as a passive user. The patient, while giving a speech in front of a set of virtual humans, receives stimulus from these characters; the therapist, who is observing patient's behavior and reaction to these stimulus, interacts with the application in order to modify the simulation accordingly, whether it is by varying the characters' behavior, or by triggering multiple events in the simulation scenario. The main initial scenario of the application is configured by the therapist as well.

The equipment required to use the application is: a computer, two sound columns, a projector and a canvas or wall used as a projection surface. The application generates two separate windows: the simulation window, that must be projected, and the interface window which is displayed in the therapist's computer. The sound columns must be close to the projection to increase the degree of realism of the simulation.

The equipment is inexpensive and it is easy to install; the projected image should contain the models of the VHs in real size, making it easy for the patient to get an immersion feeling. Moreover, it is possible to gather several people observing simultaneously the simulation, which can be valuable in research or, for instance, in the training of students. Additionally, all the unpleasant secondary effects reported by some users when using a RV equipment are eliminated.

The version here presented follows two others: the first version simulates an audience with virtual humans (Cláudio, 2013) and the second one a jury composed by one to three virtual characters (Cláudio, 2014). This paper addresses the improvements performed in this last version.

3.1 The Present Version

The main goal of this development stage in Virtual Spectators was to provide characters with high realism. Such goal had to be met without compromising application's performance, which had to respond in real time to the therapist's control.

Finding a balance between the characters' final appearance and the most critical aspects of the application (number of polygons of the meshes, the textures' resolution and the complexity of the illumination algorithms) was crucial to have the real time response. Early on we realized that it was not possible to improve the characters of the previous versions. So we adopted completely distinct models. After a series of tests, we found that the models generated using free tools, such as MakeHuman (url-MakeHumans) did not satisfied the aimed quality level. Thus we shifted our approach to combining different models (or parts of them) from the software Poser (url-Poser) and from online repositories. In such way, we produced two males and one female characters appearing different ages, and named John, Carl and Jessi (Figure 1, 2 and 3).

Figure 1: John model exhibiting two different facial expressions.

Next, we proceeded to the animation of these characters considering the characteristics and functionalities of the software tools involved in the implementation of the application: Blender (url-Blender) and a free version of Unity3D (url-Unity). In Unity, animations are obtained through rigging animation, i.e., by animating a skeleton, a hierarchical structure of interconnected bones.

The animation also comprises the process of skinning: association of a bone to a certain set of neighboring vertices. When a bone moves, it drags

Figure 2: Jessi model exhibiting two different hair styles and clothes.

Figure 3: Three virtual models exhibiting distinct facial expressions (from left to right; John, Carl and Jessi).

the corresponding vertices with it; the vertices closest to the bone suffer a bigger displacement than those furthest away. In Blender, the association vertices-bone and the weights that define the amplitude of the translation are tuned resorting to a functionality called weight painting.

In order to recycle as much as possible the animations used in the previous version, we proceeded to the integration of the skeleton of the old characters into the new ones. However, the base position of these skeletons does not match the position of the mesh in the new characters, which required a proper adjustment. Such adjustment led to the modification of the base position of the skeleton resulting in modifications in every animation that relies on the bones used in the animation. Thus, each animation had to be adjusted as well. Regarding skinning, this process is created from scratch, both for the face as for the body. Given that the three characters are anatomically different, this exhaustive process was repeated for every VH character.

3.2 Interface Functionalities

The current interface presents significant improvements over the previous one in terms of available functions, namely:

- Possibility of choosing between different scenarios. The application is ready to new scenarios that might be relevant to consider in the future.
- A drag-and-drop mechanism that allows the user to choose the characters: the user selects the picture that represents the character and drags it to the corresponding position in the table contained in the virtual scenario; the photo of the character turns grey as a sign that it is no longer available to be chosen.
- Choosing glasses for a character previously selected.
- Choosing a clothing style (formal or informal).
- Choosing a type of hairstyle between formal and informal (except for Carl that is bald). Choosing the hairstyle is independent from choosing clothes allowing multiple combination of options.
- Character preview: therapist has the possibility to visualize the selected and modified character with the previous options. This preview allows a close-up of its face.

3.3 Facial Expressions Interface

The interface that allows the user to control facial expressions is illustrated in Figure 4.

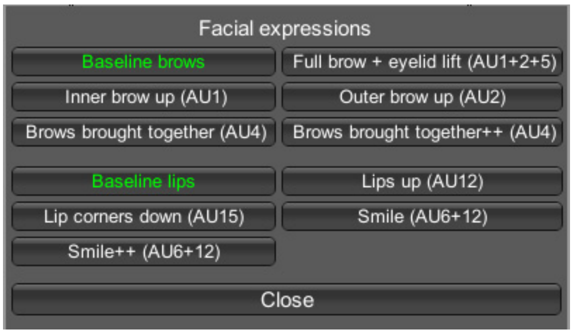

Figure 4: The interface to manipulate facial expressions.

It contains buttons that correspond to Action Units (AU) from Facial Action Coding System (FACS) (Ekman, 2002); in two AUs (AU4, AU12) there is an option to increase intensity. Our choice of facial element to include and the universe of expressions one can compose with them is based on current validated knowledge on the content of human facial behavior (for an updated review see (Gaspar, 2014)). Although there are applications, games and films today with a wide range of expressive behavior, the way expressive behavior is decoded by people and what exactly do real people convey with their facial and body behavior is still largely debated in emotion Psychology (see (Russell, 1997)) so we opted for a range of AUs and possible combinations, that has been most consensually derived from behavioral studies of spontaneous facial behavior (Gaspar, 2012; Gaspar, 2014).

Furthermore, these facial compositions can also be associated with several other body postures. Our selection of postures is based on cross-cultural studies of human non-verbal communication (Eibl-Eibesfeldt, 1989).

Figure 5 shows the same model with different AU combinations. The neutral face (upper left corner), with no AU activated and a negative-emotion expression, AU4 (upper right corner). In the lower right, we see a combination of AU4 and AU15, both related to negative affect. In the lower left corner we see a combination of AU4 and AU6+12 ("smile" button). The AU6+12 combination is the known as "Duchenne smile", and is consensually associated to positive affect more than any other smile form.

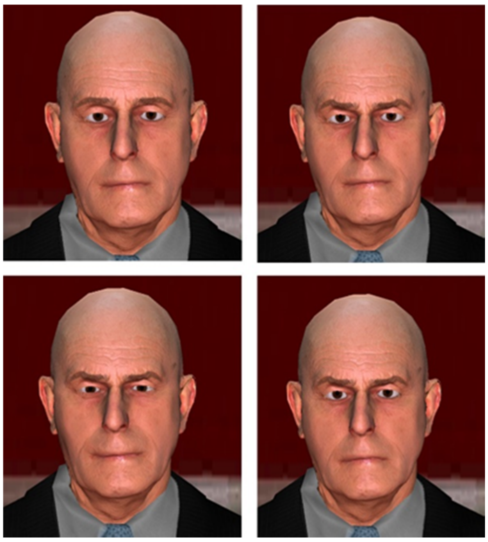

Figure 5: Clockwise from top: a VH with a baseline face (no AUs), displaying a frown (AU4), displaying a combination of AU4 and AU15, displaying a combination of AU4 and AU6+12 (the smile button).

3.4 Speaking Characters

Since our main goal is to attain realistic human-like characters, it was considered a very important task to include in the characters the ability of speaking. Verbal communication plays an important role in human interactions, reason why this requirement is important. However, to provide a VH with the ability to speak is a complex matter and difficult to implement, requiring complex algorithms as well as complex animation process. So, we conceived a simple but effective solution.

The idea was to define a set of speeches that each character could reproduce (with sound synchronized with lip movements) and that were controllable and triggered by the therapist during the simulation. On a first trial, we implemented this solution for a single character, John.

The main steps involved in this process are:

1. Record the intended set of speeches that will figure in the application. It can easily be accomplished using the microphone of a laptop.
2. Define and shape all mouth positions involved in the process of animation of the characters considering the sounds that need to be reproduced. On a first instance we defined 5 shapes/states for the Portuguese language: Base (baseline), A, E, O and U. This means that every sound can be obtained using only these 5 states.
3. Each shape corresponds to an animation. Therefore, this includes a process of animating the characters in order to be able to visually reproduce these different states. The only bones affected belong to the mouth area, namely, jaw and lips.
4. Decompose the sentence given as an input. This is the sentence that the character should be able to reproduce. Thus, any given sentence should correspond to a sequence of animations that the character has to process and verbalize. For example, the sentence "Hello world" should be translated to a language recognizable to the character and it would correspond to the sequence of animations (E,O,U,O,Base). This process involves the following steps:
   a) Decomposing the sentence into smaller segments. The division of the sentence is somewhat similar to division by syllabus. However, each segment must contain a single vowel. Given the example, it would be "He-llo-w-o-rld" (in this approach, letter w is considered as a vowel);
   b) Evaluate each segment according to the sounds each segment demands;
   c) Add the corresponding animation to the buffer of all animations that need to be played. For example, a segment "He" will correspond to state *E*, "llo" to *O*, and so on.
5. Reproduce sequentially all the animations. It is defined a time interval between each animation so that all animations can be synchronized with the sound that starts playing as the therapist triggers the event.

This approach provides a simple but realistic and attractive solution. It is a good starting point towards a more robust and complete approach. New states can be considered in order to include a wider set of sounds. As of now this solution only takes into account Portuguese language so the proper modifications are required for other languages.

4 EVALUATION AND VALIDATION

This sections describes two evaluation tasks that have been performed. A validation study was made to confirm the assumptions of the psychologist in the team during the creation of facial actions' menu. Additionally, we performed usability tests with a set of experts (therapists) to study the suitability of the application and its potential effectiveness in VRET.

4.1 Validation Phase

Beyond the theoretical framework that presided our choice of expressive elements to include in the VHs, derived from human behavior, it is necessary to validate the content in the expressive behavior displayed by the VH's. This step precedes writing a manual with guidelines for therapists who wish to design a comprehensive intervention plan with the different levels of positive/negative affect or intimidation that the VH´s may convey. For such purpose we have conducted a preliminary study of the facial behavior content with a normative sample of 38 voluntary participants (31F; 7M), recruited in two universities (ages 18- 25 yrs. old).

Participants were tested in groups of 10 and given instructions on their appraisal tasks. These were presented and controlled by an *E-Prime 2.1* programmed experiment. Experiments took place in a dim light room and consisted on participants watching and rating each of 28 animated clips (3" duration) with a close up view of a VH showing either a neutral or an expressive face) projected on a canvas 2-3 meters in front of them. We tested expressions comprised of the following AUs/AU combinations:

- Baseline eyebrows + Baseline lips
- Eyebrows brought together AU4
- Eyebrows brought together (more intense)++ (AU4)
- Full brow up and eyelid lift (AU1+2+5)
- Inner brow up (AU1)
- Outer brow up (AU2)
- Lips up (AU12)
- Smile (AU6+12)
- Intense smile ++ (AU6+12)
- Lip corners down (AU15)
- Eyebrows brought together++ (AU4) + Lips up (AU12)
- Eyebrows brought together++ (AU4) + Lip corners down (AU15)
- Inner brow up (AU1) + Lips up (AU12)
- Inner brow up (AU1) + Lip corners down (AU15)

Clips were presented randomly to control for order effects. The first task consisted of rating the clip on each of the Self-Assessment Manikin (SAM) (Bradley, 1994) scales, a non-verbal pictorial assessment technique that measures the arousal, pleasure, and dominance associated with a person's affective reaction to an observed stimulus. The second task consisted in choosing from a list of emotion labels (happy, angry, surprised, fear, other positive, other negative or neutral) the one that best suited the face just watched (see Figure 6 with screenshots used in the validation experiments).

Results were encouraging as they showed great convergence in both content attribution and emotional impact to many compositions, thus validating our content assumptions when choosing which expressive elements to include.

In regard to content, an independence X^2 test performed on the cross tabulation of composition and content, showed a highly significant association between images and label (X^2=2088.07; p<=.001; N=978), with 12 out of 28 clips presenting convergence above 75% (6 on the label angry, 4 on happy, 1 on surprised and 1 o neutral). Convergence was congruent with the expected content in all convergent pictures.

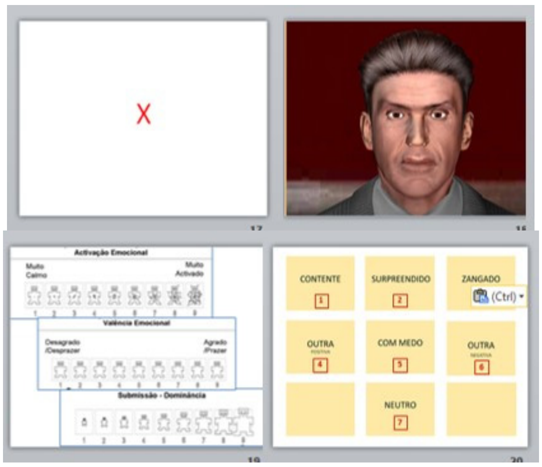

Figure 6: Types of screens in the user experience.

As to emotional impact, images generally did not elicit high **Arousal** (only 3 clips scored higher than the Median). **Valence** was highly convergent, fitting a narrow bell curve, with 100% of data below SD=1.9 and Mean=4.53, approaching the Median=5. Regarding **Dominance/control,** 19 clips scored higher than the Median, in an also centered distribution where Mean=5.16 approaching the

Median=5 and 100% of data fell below SD=1.9; the highest scores on feeling dominant occurred in response to target faces that had been convergent on positive content in the labelling task and resulted from exposure to the female avatar, for exact same compositions seen in male faces.

In conclusion, we were able to tabulate values for relevant affective impact parameters - a crucial step towards the setting up of a therapist's manual with validated content. It will be challenging to create compositions that elicit higher levels of arousal and Dominance/control. However, these may reveal much different when the application is tested with a clinical population, which is a future step in our research.

4.2 Evaluation Phase

Another study was conducted with the goal of evaluating the application on two main features: usability and VH realism

We recruited 6 therapists (5F; 1M) ages 34- 59 years old. Within this group, only 2 elements had not tried the previous version of the application. Tests were conducted individually with each therapist, in a dim light room (to improve visualization of images and feeling of immersiveness). The apparatus for tests was identical to that of typical application use: a laptop computer connected to an LCD projector displaying the image onto a projection canvas. Both client and server were executed in localhost. Whilst the therapist interface is shown in the laptop screen, the client interface is projected onto the canvas approaching real-life size. Two sound columns were connected close to the projection canvas.

The evaluation was divided into 4 distinct phases: character evaluation regarding realism (compared to the previous version), usability of the implemented functions, realism of the application as a whole, and an open answer questionnaire section. Each part was evaluated as the user performed each respective task.

As a result, in every section therapists considered the new characters to be more realistic than those of the previous version. The modifications made to the configuration interface were welcomed as improvements. Suggestions made to this interface were:

- increase the number of available scenarios and characters for higher versatility and the options to edit each character;
- add new body animations (such as, looking at the watch or touching the hair).

Each therapist was further asked to trigger a specific speech in the "John" character and none had difficulty performing this task. The favorite feature of therapists in the new version is the VH's ability to speak.

Finally, every therapist mentioned that if available, they would likely use the application in their therapy sessions.

5 CONCLUSION AND FUTURE WORK

One of the main questions related to the use of VRET by the therapists is the cost of the immersive equipment and some discomfort associated with its use. Having this in mind, we sought to implement a low cost solution that effectively assists in the treatment of SAD resorting to VRET in the specific situation of fear of public speaking, a common problem.

The approach we propose involves ordinary equipment: computer, sound system (two sound columns are enough), projector and canvas (or a wall) to project the simulations. The software for this project (with no budget whatsoever) was developed using freeware and free or very low cost 3D models. The main disadvantage of this approach is the difficulty in obtaining photo-realistic models.

Nonetheless, results from our content validation experiments, facial expressions in our present models are consistently interpreted by normative observers and are decoded according to expectations derived from emotion and expression science, which is indirect evidence of realism. These gives us the confidence to use the application to generate animation clips to be used in the research domain of non-verbal communication.

Tests performed with therapists further confirm that the current models are more realistic than those of the previous version. Speech articulation was seen as a major improvement, and in general the idea of using this tool in a clinical environment was welcomed with enthusiasm.

The main follow up steps will be: i) to develop and integrate in the application an artificial intelligence module toward the simulation of emotions; ii) afterwards, to validate the usefulness of the application with a clinical population in a therapeutic context.

ACKNOWLEDGMENTS

We acknowledge all the therapists and volunteers that participated in our user study with no reward. We thank our national funding agency for science, technology and innovation (*Fundação para a Ciência e a Tecnologia*- FCT) and the Research Unit LabMAg for the financial support (through PEst-OE/EEI/UI0434/2014) given to this work.

REFERENCES

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

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

Bradley, M. M. and LANG, P. J., 1994. Measuring Emotion: the Self-Assessment Manikin and the Semantic Differential. J. Behav. Therap.. and Exp. Psvchyatry, 25, 1, 49-59.

Cláudio, A.P., Carmo, M.B., Pinheiro, T., Esteves, F., 2013. A Virtual Reality Solution to Handle Social Anxiety. *Int. J. of Creative Interfaces and Computer Graphics*, 4(2):57-72.

Cláudio, A.P., Gaspar, A., Lopes, E., Carmo, M. B., 2014. Virtual Characters with Affective Facial Behavior. *In Proc. GRAPP 2014*, pp 348-355, SciTePress.

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 and 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.

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., 2014. 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 M. Pina and N. Gontier (Eds), The Evol. of Social Comm. in poserPrimates: A Multidisciplinary Approach, Interdisciplinary Evolution Research. New York: Springer (pp.101-145).

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., 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 and behavior, 6(3), 237-243.

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

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 tech. 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 and 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., Pertaub, D. P., Steed, A., 1999. Public speaking in virtual reality: Facing an audience of avatars. IEEE CGandA, 19(2), 6-9.

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

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

(last access to these sites on Dec 2014).

url-Blender http://www.blender.org/

url-Kinect http://www.microsoft.com/en-us/kinectforwindows/

url-MakeHuman http://www.makehuman.org.

url-Poser http://my.smithmicro.com/poser-3d-animation-software.html.

url-Unity http://Unity.com/