Multimodal Systems for Public Speaking: A Case in Support of a Positive Computing Approach

Fiona Dermody and Alistair Sutherland
School of Computing, Dublin City University, Dublin, Ireland
Fiona.Dermody3@mail.dcu.ie, Alistair.Sutherland@dcu.ie

Keywords: Multimodal Interfaces, Positive Computing, HCI, Public Speaking, Real-time Feedback.

Abstract: Positive Computing involves the utilisation of digital technology to foster psychological wellbeing and human potential. We will present an overview of Positive Computing and what it implies for multimodal systems for public speaking. The position of this paper is that a Positive Computing approach can make such systems more effective and improve user experience. We will focus on three of the tenets of Positive Computing viz. awareness, autonomy and stress-reduction. We will discuss different existing multimodal systems for public speaking within the context of Positive Computing.

1 INTRODUCTION

People who experience fear of public speaking tend to practise avoidance when it comes to communicating to or speaking in front of a group of people (Harris et al., 2002). A number of digital systems have been designed to help people with this fear and to enable them to get feedback on their public speaking before going in front of a live human audience. The position of this paper is that Positive Computing is an appropriate paradigm for multi-modal public speaking systems and we will make some recommendations based on it. In this paper, the following three tenets of Positive Computing will be presented, namely, self-awareness, autonomy and stress reduction. These tenets will be illustrated with reference to a multimodal positive computing system for public speaking (Dermody and Sutherland, 2016) and to other systems that have been developed in the field of computer-mediated communications (Batrinca et al., 2013; Schneider et al., 2015; Bubel et al., 2016).

2 MULTIMODAL SYSTEMS FOR PUBLIC SPEAKING

Many researchers have developed multimodal systems for public speaking. The term 'multimodal' refers to the fact that these systems can detect multiple speaking modes in the speaker such as their gestures, voice and facial expressions. These systems typically use a 3D body sensor such as the Microsoft Kinect to detect human body poses and motion. Most systems give feedback to the user on their performance. Feedback can be provided in different ways such as visual icons, text, haptic devices or through the reactions of a virtual audience. Feedback can be in real-time or retrospective, interruptive or continuous. The rest of this paper will discuss examples of these multimodal systems for public speaking to demonstrate the appropriateness of using Positive Computing as a paradigm for their development.

3 EXISTING MULTIMODAL SYSTEMS FOR PUBLIC SPEAKING

A number of multimodal public speaking systems have focused on awareness in the context of public speaking.

3.1 Haptic Feedback

AwareMe utilises a wristband that provides speakers with haptic and visual feedback as they are speaking on voice pitch, speaking rate and filler words (Bubel et al., 2016). Feedback is provided to the speaker through a vibrating wristband and through a coloured display as per Figure 1.

3.2 Virtual Audience

Cicero:Virtual Audience Framework utilises a virtual audience comprising of avatars to convey non-verbal feedback to speakers (Batrinca et al., 2013; Chollet et al., 2015a; Chollet et al., 2015b). As can be seen in Figure 2, feedback is relayed to the speaker by engaged or disengaged body poses of the virtual audience and through a coloured bar at the top of the screen.

3.3 Video

Presentation Trainer, see Figure 3 represents the user using video and provides them with real-time feedback on one nonverbal speaking modality at a time with a gap of at least six seconds between feedback displays (Schneider et al., 2015). The feedback provided by these systems can make users aware of their speaking behaviour and this awareness can aid in the development of communication skills.

Figure 1: AwareMe. Haptic feedback is provided to the speaker using a wristband (Bubel et al., 2016).

Figure 2: Cicero - Virtual Audience Framework. The audience responds to the speaker's performance using either engaged or disengaged body poses (Chollet et al., 2015a).

Figure 3: Presentation Trainer. Interruptive, textual feedback is provided to the user and the user is represented using live video (Schneider et al., 2015).

4 POSITIVE COMPUTING

The primary objective of Positive Computing is to foster psychological wellbeing and human potential (Calvo and Peters, 2015). Positive Computing can be described as the positive application of computing to real-world problems using a fusion of knowledge (or in conjunction with) theoretical frameworks from the humanities and the social sciences (Calvo et al., 2014; Calvo and Peters, 2015). There are three spheres of experience in which technology can impact on wellbeing-viz. external activity, technology environment and personal development, see Figure 4.

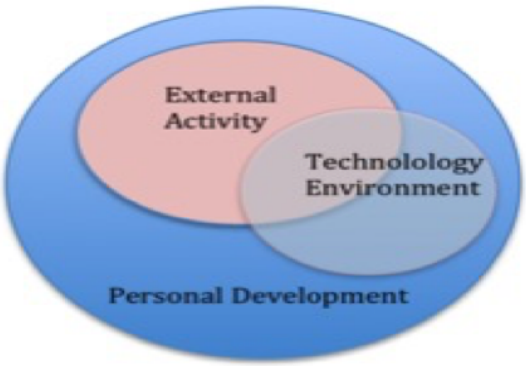

Figure 4: The Spheres of Positive Computing (Calvo and Peters, 2014).

In the context of this research, the external activity is the user's speaking ability, the technology environment is the system and personal development is the increased sense of competence at speaking and reduction in stress experienced when speaking. The objective of the system is to reduce the anxiety which a user experiences when speaking in public. From this objective, it follows that the system itself must not add

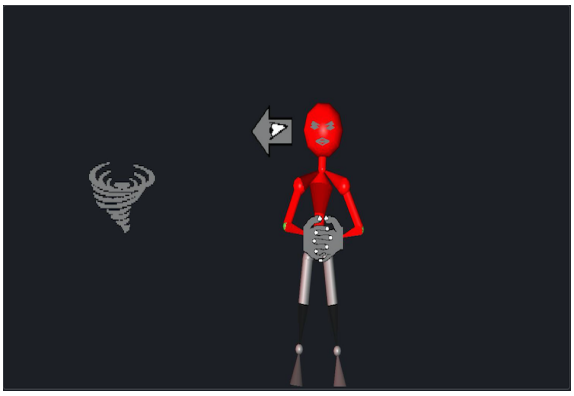

Figure 5: System Avatar with indicative visual feedback on gaze direction, agitation and hands touching.The avatar represents the user (Dermody and Sutherland, 2016).

Figure 6: System Video Stream with visual feedback on gaze direction and hands-touching. The user is represented in the video (Dermody and Sutherland, 2016).

to any anxiety already experienced by a user. Using the system should be an enjoyable experience. Furthermore, users should feel that using the system is beneficial for them in addressing any speaking anxiety. Feedback displayed by the system should be assimilated with minimum cognitive load on the users. If users are stressed trying to assimilate feedback from the system, they are unlikely to either react to it or find the system pleasant to use.

5 MOTIVATION FOR MULTIMODAL SYSTEMS FOR PUBLIC SPEAKING

Public speaking involves the interaction between the different speaking modalities voice, gestures and eye contact (Toastmasters International, 2008; Toastmasters International, 2011). The interaction between these modalities determines the level of engagement that a speaker forms with an audience. Multimodal feedback systems for communication skills development such as (Dermody and Sutherland, 2016; Chollet et al., 2015a; Schneider et al., 2015), provide users with feedback that gives them a choice to potentially adapt their communication behaviour as they speak. The assimilation of this feedback by a speaker could alert the user to any ineffective speaking traits that they may be exhibiting.

Prior to the development of digital systems for public speaking, the only way for a speaker to gain insight on their speaking was either to practise in front of a human mentor or in front of a mirror. Both of these can cause stress or anxiety for the user. The objective of multimodal systems for public speaking is to allow the user to gain this awareness in private without being exposed to stress or anxiety.

6 TENETS OF POSITIVE COMPUTING RELEVANT TO PUBLIC SPEAKING

We will now present the tenets of Positive Computing that are particularly relevant for Public Speaking.

6.1 Positive Computing and Stress-Reduction

The central premise of Positive Computing is designing for wellbeing which, includes freedom from stress and pressure. The objective of a Positive Computing system is to make it enjoyable for users to use. Users should not feel pressurized or dictated to when interacting with the system. Feedback presented by a system can impact on the levels of stress experienced by users.

6.1.1 Feedback Intensity

There have been investigations on how public speaking performance has been impacted by a speaker's sensitivity to feedback (Smith and King, 2004). They found that speakers, who were sensitive to feedback, displayed more positive speaking behaviours when feedback messages were of low intensity. By 'low intensity', (Smith and King, 2004) meant feedback that is 'worded in a manner that was less likely to be taken as a direct personal criticism'. They claim that feedback, which is focused on the task, results in improved performance. Conversely, harsh feedback that could be interpreted as severe or as a threat to self impedes performance. They differentiate between the two levels of feedback intensity using these examples.

'Your eye contact needs improvement. You don't appear to be looking out toward the audience as frequently as you should or maintaining the eye contact when you do look up. In your next speech, try to increase your eye contact'. This is an example of low-intensity feedback. The high-intensity feedback was worded thus: *'Your eye contact was bad in the speech. You rarely look up and when you do glance toward the audience, it's only for a moment. Your next speech requires significant improvement in eye contact'*.

They found that speakers who had a high sensitivity to feedback modified their speaking more when the feedback was of lower intensity. The authors posited that highly negative feedback causes the feedback sensitive learner to make negative attributions, to focus on meta-task issues, such as seeing feedback as punishment and to fail to modify behaviour (Smith and King, 2004). During a review of the multimodal system for public speaking, Presentation Trainer, experts found that the system should 'shift focus and become a tool to develop awareness of nonverbal communication, instead of correcting it' (Schneider et al., 2017).

6.2 Positive Computing and Autonomy

One of the core issues within the framework of Positive Computing, is developing and supporting autonomy (Calvo and Peters, 2016; Calvo and Peters, 2014). In the context of multimodal applications, autonomy is an important consideration in relation to the display of real-time feedback. A key design principle in Positive Computing 'has been to provide reflective rather than directive feedback 'consider this rather than do this' (Calvo and Peters, 2014, p. 71). For instance in (Dermody and Sutherland, 2016) users are given real-time feedback on their performance while they are rehearsing a speech. Real-time feedback is useful as it makes the user immediately aware of their speaking behaviour. It also links to the Positive Computing principle of autonomy which centres on the idea of a user having control over which feedback they choose to react to. As can be seen in (Dermody and Sutherland, 2018), real-time visual feedback is displayed to users in proximity to the area of the user's body it relates to. We can see some of these visual feedback icons in Figures 5 and 6. They make the user aware of their gaze direction and if they have been clasping their hands for a long period of time. However, the user has the autonomy or choice to adapt their speaking behaviour in response to the feedback displayed.

6.3 Public Speaking and Awareness

Making individuals aware of their behaviours, specifically habit behaviours, has been effective in reducing the behaviour. Studies examining the effectiveness of awareness training on tics and nervous habits suggest that a directive approach may be an unnecessary component to reverse the habit (Wiskow and Klatt, 2013). The effectiveness of awareness training for reducing vocal dysfluencies in public speaking with university students has been evaluated (Mancuso and Miltenberger, 2015). The study consisted of awareness training and competing response training. All the participants showed an immediate decrease in their use of vocal dysfluencies. The results indicate that nervous habits in public speaking can be effectively addressed reduced and reversed using awareness training. Moreover, the authors reported that the participants greatly decreased the nervous vocal habits during awareness training even before the competing-response training had commenced. They suggest that awareness training alone may be sufficient for decreasing distracting nervous behaviors in public speaking.

6.4 Positive Computing - Awareness and User Representation

In the Positive Computing framework, self-awareness is explained in the context of reflection and getting to know oneself. In regard to public speaking, this implies an awareness of how we appear to an audience while speaking. For instance, some speakers may not be aware of the importance of using gestures to engage an audience (Toastmasters International 2011a). In the context of Positive Computing, a key design principle has been to provide reflective rather than directive feedback (Calvo and Peters, 2014). This principle is pertinent to the way in which the user's body pose and movements are represented in multimodal systems.

Different approaches have been taken to represent users in multimodal systems for public speaking. Some use live video and others use either 2D or 3D avatars.

6.4.1 Avatars and Awareness

An avatar presents an abstract representation of the user and this form of abstract representation allows the user to see their body pose, gestures and facial expressions in 3D. (Kistler and André, 2015) surveyed users reactions to 2D avatars as shown in Figure 7. They found that the 2D avatars did not give a sense of depth and the users were not able to gauge their full

3D poses. For this reason, we would argue that a full 3D avatar is necessary.

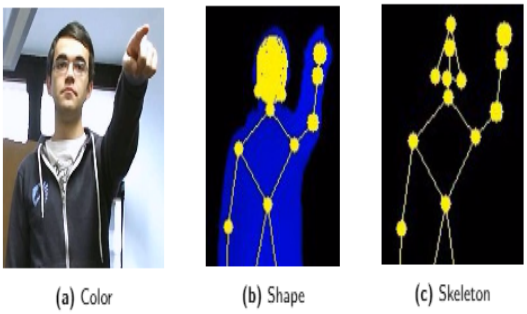

Figure 7: Different levels of representing the user's body (Kistler and André, 2015).

Avatars have been utilised in a mentor role for social skills training, for instance in (Gebhard et al., 2014). In their system, it was found that users responded more favorably to an avatar that was more understanding than one that made demands on them. Studies of fear of public speaking have shown that people do respond favourably to virtual agents 'even in the absence of two-way verbal interaction, and despite knowing rationally that they are not real' (Garau, 2006; Pertaub et al., 2002). Virtual agents have been used effectively in multimodal systems for public speaking, most notably, (Chollet et al., 2015b). In the aforementioned system, virtual agents were used to represent an audience that responded to the user's speaking performance. In (Dermody and Sutherland, 2016; Dermody and Sutherland, 2018), the avatar represents the user themselves .

6.4.2 Video and Awareness

Video can also be used to represent a user. However, not everyone reacts well to seeing themselves on video. 'The cognitive dissonance that can be generated from the discrepancies between the way persons think they come across and the way they see themselves come across can be quite emotionally arousing and, occasionally, quite aversive' (Dowrick and Biggs, 1983). Also they may become distracted by the physicality of their appearance. Their perceived level of physical attractiveness or lack thereof may become their focus, as they observe themselves on video, as opposed to their behaviour. If the person has a negative self-perception of themselves on video, then their reaction to the video will not be positive (Dowrick, 1999). An avatar's abstract representation could be advantageous because the user is less likely to be distracted by details of their physical appearance (Dermody and Sutherland, 2018). Dermody and Sutherland did a comparison of live video and 3D avatar. Nine out of ten users preferred the avatar because they found it less distracting and less stressful. However, one user said she preferred live video because the avatar made her feel "disembodied" (Dermody and Sutherland, 2018).

7 CONCLUSION

Our conclusions based on the argument above, is that an ideal system for public speaking would have the following characteristics.

It would have a full 3D avatar because this allows the user to gauge their full 3D body pose but does not distract them with details of their personal appearance.

It would give the user the option to use live video, if they choose, because some users have expressed a preference for this.

It would use reflective rather than directive feedback because this increases the user's self-awareness but gives the user the autonomy to choose whether to respond to the feedback or not.

It would use visual feedback displayed around the avatar because this is easier to assimilate than textual feedback and reduces the cognitive load and stress on the user.

These are consistent with the tenets of Positive Computing: self-awareness, autonomy and stress-reduction. Users could also potentially gain autonomy in a wider sense from developing a skill such as public speaking using a Positive Computing system. For instance, they may increase their success in education or enterprise by overcoming any speaking-related anxiety (Dwyer and Davidson, 2012; McCroskey et al., 1989; Harris et al., 2002).

REFERENCES

Batrinca, L., Stratou, G., and Shapiro, A. (2013). Cicero - Towards a Multimodal Virtual Audience Platform for Public Speaking Training. In Aylett, R., Krenn, B., Pelachaud, C., and Shimodaira, H., editors, *Intelligent Virtual Agents*, volume 8108 of *Lecture Notes in Computer Science*, chapter Cicero - T, pages 116–128. Springer Berlin Heidelberg, Berlin, Heidelberg.

Bubel, M., Jiang, R., Lee, C. H., Shi, W., and Tse, A. (2016). AwareMe: Addressing Fear of Public Speech through Awareness. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pages 68–73. ACM.

Calvo, R. A. and Peters, D. (2014). *Positive Computing: Technology for wellbeing and human potential*. MIT Press.

Calvo, R. A. and Peters, D. (2015). Introduction to Positive Computing: Technology That Fosters Wellbeing. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '15, pages 2499–2500, New York, NY, USA. ACM.

Calvo, R. A. and Peters, D. (2016). Designing Technology to Foster Psychological Wellbeing. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '16, pages 988–991, New York, NY, USA. ACM.

Calvo, R. a., Peters, D., Johnson, D., and Rogers, Y. (2014). Autonomy in technology design. *Proceedings of the extended abstracts of the 32nd annual ACM conference on Human factors in computing systems - CHI EA '14*, pages 37–40.

Chollet, M., Morency, L.-p., Shapiro, A., Scherer, S., and Angeles, L. (2015a). Exploring Feedback Strategies to Improve Public Speaking: An Interactive Virtual Audience Framework. *UbiComp '15: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 1143–1154.

Chollet, M., Stefanov, K., Prendinger, H., and Scherer, S. (2015b). Public Speaking Training with a Multimodal Interactive Virtual Audience Framework - Demonstration. *ICMI 2015 - Proceedings of the 2015 ACM International Conference on Multimodal Interaction*, pages 367–368.

Dermody, F. and Sutherland, A. (2016). Multimodal system for public speaking with real time feedback: a positive computing perspective. In *Proceedings of the 18th ACM International Conference on Multimodal Interaction*, pages 408–409. ACM.

Dermody, F. and Sutherland, A. (2018). Evaluating User Responses to Avatar and Video Speaker Representations A Multimodal Positive Computing System for Public Speaking. In *Proceedings of the 13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2018)*, volume HUCAPP, pages 38–43, Madeira. INSTICC.

Dowrick, P. W. (1999). A review of self modeling and related interventions. *Applied and Preventive Psychology*, 8(1):23–39.

Dowrick, P. W. and Biggs, S. J. (1983). *Using video: Psychological and social applications*. John Wiley & Sons Inc.

Dwyer, K. K. and Davidson, M. M. (2012). Is Public Speaking Really More Feared Than Death? *Communication Research Reports*, 29(2):99–107.

Garau, M. (2006). Selective fidelity: Investigating priorities for the creation of expressive avatars. In *Avatars at Work and Play: Collaboration and Interaction in Shared Virtual Environments*, pages 17–38. Springer.

Gebhard, P., Baur, T., Damian, I., Mehlmann, G., Wagner, J., and André, E. (2014). Exploring interaction strategies for virtual characters to induce stress in simulated job interviews. In *Proceedings of the 2014 International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '14, pages 661–668, Richland, SC. International Foundation for Autonomous Agents and Multiagent Systems.

Harris, S. R., Kemmerling, R. L., and North, M. M. (2002). Brief virtual reality therapy for public speaking anxiety. *Cyberpsychology & behavior : the impact of the Internet, multimedia and virtual reality on behavior and society*, 5(6):543–550.

Kistler, F. and André, E. (2015). How can i interact?: Comparing full body gesture visualizations. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*, CHI PLAY '15, pages 583–588, New York, NY, USA. ACM.

Mancuso, C. and Miltenberger, R. G. (2015). Using habit reversal to decrease filled pauses in public speaking. *Journal of applied behavior analysis*.

McCroskey, J. C., Booth Butterfield, S., and Payne, S. K. (1989). The impact of communication apprehension on college student retention and success. *Communication Quarterly*, 37(2):100–107.

Pertaub, D.-P., Slater, M., and 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.

Schneider, J., Börner, D., Rosmalen, P., and Specht, M. (2017). Presentation Trainer: what experts and computers can tell about your nonverbal communication. *Journal of Computer Assisted Learning*, 33(2):164–177.

Schneider, J., Börner, D., Van Rosmalen, P., and Specht, M. (2015). Presentation Trainer, your Public Speaking Multimodal Coach. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, pages 539–546. acm.

Smith, C. D. and King, P. E. (2004). Student feedback sensitivity and the efficacy of feedback interventions in public speaking performance improvement. *Communication Education*, 53(3):203–216.

Toastmasters International (2008). *Competent Communication A Practical Guide to Becoming a Better Speaker*.

Toastmasters International (2011). Gestures: Your Body Speaks. Available from: http://www.toastmasters.org.

Wiskow, K. M. and Klatt, K. P. (2013). The effects of awareness training on tics in a young boy with Tourette syndrome, Asperger syndrome, and attention deficit hyperactivity disorder. *Journal of applied behavior analysis*, 46(3):695–698.