Measuring the Subjective Unit of Discomfort using Speech Technology

Dwi HARTANTO[a,1], Ni KANG[a], Willem-Paul BRINKMAN[a],
Isabel L. KAMPMANN[b], Nexhmedin MORINA[b], Paul G.M. EMMELKAMP[b],
Mark A. NEERINCX[a,c]

[a] *Delft University of Technology, Netherlands*
[b] *University of Amsterdam, Netherlands*
[c] *TNO Human Factors, Netherlands*

Abstract. Current practice in Virtual Reality Exposure Therapy (VRET) is that therapists ask patients about their anxiety level by means of the Subjective Unit of Discomfort (SUD) scale. With an aim of developing a home-based VRET system, this measurement ideally should be done using speech technology. In a VRET system for social phobia with scripted avatar-patient dialogues, the timing of asking patients to give their SUD score becomes relevant. This study examined three timing mechanisms: (1) dialogue dependent (i.e. naturally in the flow of the dialogue); (2) speech dependent (i.e. when both patient and avatar are silent); and (3) context independent (i.e. randomly when patient is talking). Results of an experiment with non-patients (n=24) showed a significant effect for the timing mechanisms on the dialogue flow, user preference, and reported presence.

Keywords. virtual reality exposure therapy, social phobia, anxiety level measurement, SUD score, speech recognition, speech detector.

1. Introduction

Social phobia is an anxiety disorder characterized by a strong fear of being judged by other people and of being embarrassed. The Delft Remote Virtual Reality Exposure Therapy (DRVRET) [1] is a platform designed to treat the patients with these disorder by exposing them to semi scripted human-avatar dialogues in a virtual environment. During a session, therapists normally ask patients about their anxiety by giving a rating on the Subjective Unit of Discomfort (SUD) scale from zero ("no anxiety at all") to 10 ("the highest level of anxiety that you can imagine"). With the aim of developing a home-based VRET system, where no therapist is being present, this measurement ideally should be done automatically using speech technology. Therefore, a key question becomes the timing of asking for a SUD score as unexpected interruption might negatively affect patients' experience in a given situation. In this study, three timing mechanisms were examined: (1) dialogue dependent (i.e. naturally in the flow of the dialogue, e.g. just before the start of a new avatar question), (2) speech dependent (i.e. when both patient and avatar are silent), (3) context independent (i.e. randomly when patient is talking, testing a worst case scenario interruption).

[1] Corresponding Author. Email: d.hartanto@tudelft.nl

2. Method

The main part of the experiment consisted of three sessions with a virtual audience, talking about three out of four different topics (Democracy, France, Dogs and Penguins [2]). The presentation phase lasted between 2 and 3 minutes, after which avatars started the question and answer phase. All participants (n = 24) were exposed to three different timing mechanisms. After each session participants were asked to complete the Igroup Presence Questionnaire (IPQ), the Dialogue Experience Questionnaire (DEQ) [2] and a SUD Score Experience Questionnaire (SEQ), specially designed for this study.

3. Results

To study the effects of the timing mechanisms a series ANOVAs with repeated measures were conducted. A significant effect was found in the total SEQ score ($F(2,46) = 1065.24$; $p < 0.001$) and total DEQ score ($F(2,46) = 628.96$; $p < 0.001$). The total SEQ score suggested that participants rated the dialogue dependent timing mechanism ($M = 5.1$, $SD = 0.6$) as less interruptive than the speech dependent timing mechanism ($M = 18.9$, $SD = 2.2$), and the latter was again rated as less interruptive as the context independent timing mechanism ($M = 26.3$, $SD = 2.0$). The total DEQ score showed a similar pattern with regard to the dialogue experiences (dialogue dependent: $M = 172.\ 6$, $SD = 3.3$; speech dependent: $M = 163.0$, $SD = 4.5$; context independent: $M = 141.4$, $SD = 4.0$). Yet, an opposite pattern was found in the total IPQ score ($F(2,46) = 4.05$; $p = 0.024$). Participants rated presence highest for the context independent timing mechanism ($M = 42.8$, $SD = 3.5$), while again the speech dependent in the middle ($M = 42.6$, $SD = 3.5$) and lowest for dialogue dependent timing mechanism ($M = 42.2$, $SD = 3.5$). A side effect of the phenomenon called breaks in presence [3] might explain this.

4. Conclusion

Although potentially more development intensive, in cooperating the moment of asking for a SUD score into the flow of the dialogue seems to outperform other timing mechanism such as speech dependent and context independent timing mechanism.

Acknowledgment

This research is supported by the Netherlands Organization for Scientific Research (NWO), grant number 655.010.207.

References

[1] W.P. Brinkman, D. Hartanto, N. Kang, D. de Vliegher, I.L. Kampmann, N. Morina, P.M.G. Emmelkamp, M.A. Neerincx, A virtual reality dialogue system for the treatment of social phobia, *CHI'12: CHI'12 extended abstracts on human factors in computing systems* (2012).

[2] N. ter Heijden,, and W.P. Brinkman. Design and evaluation of a virtual reality exposure therapy system with automatic free speech interaction, *Journal of CyberTherapy and Rehabilitation*, **4**(2011), 41-55.

[3] M. Slater & A,J. Steed, A virtual presence counter, *Presence-Teleoperators and Virtual Environments*, **9**(2000), 413–434.