Computers in Human Behavior 76 (2017) 128–134

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Full length article

Potential of virtual reality as a diagnostic tool for social anxiety: A pilot study

Martin Dechant [b], Sabine Trimpl [a], Christian Wolff [b], Andreas Mühlberger [a], Youssef Shiban [a, *]

[a] Institute of Psychology, Department of Clinical Psychology and Psychotherapy, University of Regensburg, Germany
[b] Department of Media Informatics, University of Regensburg, Germany

ARTICLE INFO

Article history:
Received 16 October 2016
Received in revised form
21 June 2017
Accepted 2 July 2017
Available online 3 July 2017

Keywords:
Social anxiety disorder
Virtual reality
Eye tracking
Skin conductance response
Diagnostics

ABSTRACT

The potential of virtual reality for diagnosing anxiety disorders has been explored to a lesser extent than its use in psychotherapy so far. The current study applied virtual environments as an innovative diagnostic tool for social anxiety and social anxiety disorder and examined differences between low- and high-socially-anxious participants on the basis of specific psychological parameters and recordings of gaze behavior. Out of 119 subjects, 19 low- and 18 high-socially-anxious participants were selected by a social anxiety questionnaire. During the completion of eight tasks in a virtual train and waiting room scenario, the skin conductance response and gaze behavior was monitored. The findings indicate that analyzing fixation durations of faces in a virtual social situation is even more suitable for distinguishing low- and high-social-anxious participants than investigations of skin conductance responses.

1. Introduction

Virtual reality (VR) is a promising tool, which can simulate the complexity of experiences of the real world in a controllable and reproducible manner in the laboratory and activate fear networks in anxious subjects (Mühlberger, Wieser, & Pauli, 2008; Shiban, Brütting, Pauli, & Mühlberger, 2015; Shiban, Pauli, & Mühlberger, 2013). While the advantages of the application of VR in psychotherapy have already been illustrated (for a review see Freeman et al. (2017); Krijn, Emmelkamp, Olafsson, and Biemond (2004); Meyerbröker and Emmelkamp (2010); Opriş et al. (2012)), clinical-psychological diagnostics in virtual surroundings are still in the early stages of development. Virtual social scenarios have mainly been generated for therapeutic purposes thus far.

Typically, participants had to give speeches in front of a virtual audience (Harris, Kemmerling, & North, 2002; Slater, Pertaub, Barker, & Clark, 2006). However, additional social settings exist: e.g. a virtual clothing shop (Brinkman, Van der Mast, & de Vliegher, 2008, pp. 85–88), a virtual elevator (Mühlberger et al., 2008) or a virtual application interview (Grillon, Riquier, Herbelin, & Thalmann, 2006).

The current study assessed the application of virtual technology for psychological diagnostics in the case of social anxiety disorder. To simplify the analysis of eye movements, the VR was viewed over a desktop system as opposed to head-mounted displays (HMDs). Hence, problems such as motion sickness or distortion of data due to prevalent gesture-based interactions did not emerge (Kim, Kim, & Nam, 2016).

Train and waiting room scenarios were established in the framework of a behavior test to track the gaze behavior and skin conductance response (SCR) and to expand the results of commonly used virtual social paradigms. The essential feature of social phobia is an intense fear of social situations in which the individual is exposed to scrutiny by others such as social interactions, being observed and performing in front of others (American Psychiatric Association, 2013). In this experiment, we aimed at testing these main criteria. The train scenario includes triggering elements of social interaction and performance as the participants have to make eye contact with the porter and to complete a calculation task, and the waiting room scenario, which requested participants to make eye contact, to ask for a seat and resulted in being rejected, targeted fear of social interactions, observance and rejection. Furthermore, situations which include

* Corresponding author.
E-mail addresses: martin.dechant@ur.de (M. Dechant), Sabine.Trimpl@stud.uni-regensburg.de (S. Trimpl), christian.wolff@ur.de (C. Wolff), andreas.muehlberger@ur.de (A. Mühlberger), Youssef.Shiban@ur.de (Y. Shiban).

http://dx.doi.org/10.1016/j.chb.2017.07.005
0747-5632/

performance in public and meeting new people represent social situations that are feared most in social phobics (Crome & Baillie, 2014).

There is evidence that the measurement of the SCR can help distinguish between high- and low-socially-anxious people. A slower habituation and more spontaneous fluctuations of palmar skin conductance responses towards social stimuli were shown for socially-anxious individuals compared to controls (Lader, 1967). Dimberg, Fredrikson, and Lundquist (1986) revealed that the slower habituation and the amplitude of the SCR were positively correlated with the severity of social anxiety. Furthermore, a higher skin conductance level was reported for high-socially-anxious individuals during the imagination of socially-threatening stimuli, which were described by an auditory script (McTeague et al., 2009). Nonetheless, others reported results contradictory to the above (Edelmann & Baker, 2002; Mauss, Wilhelm, & Gross, 2003; Mauss, Wilhelm, & Gross, 2004).

Gaze behavior examinations revealed that, when observing faces, healthy individuals show a consistent pattern of eye movements. Characteristically, attention is paid to salient elements of faces, like eyes, nose and mouth; thus, the gaze movements represent a reversed triangle (Mertens, Siegmund, & Grüsser, 1993; Walker-Smith, Gale, & Findlay, 1977). The eyes are especially focused on as they can provide crucial information about the emotions of an interaction partner (Lundqvist, Esteves, & Ohman, 1999). Earlier studies found a deviating behavior in the case of social anxiety disorder. Affected individuals avoided the eye regions of faces (Hofmann, Gerlach, Wender, & Roth, 1997; Wieser, Pauli, Alpers, & Mühlberger, 2009). Chen, Thomas, Clarke, Hickie, and Guastella (2015) reported that socially-anxious individuals giving speeches fixated longer on non-social regions, defined as regions around the faces of people in a video-taped audience.

Consequently, we assumed that high-socially-anxious individuals have a higher SCR in the behavior test in the VR than low-socially anxious controls.

Furthermore, it was expected that high-socially-anxious individuals exhibit shorter fixation durations on the faces of virtual agents.

2. Method

2.1. Participants

A total of 119 test subjects were screened online using the Liebowitz Social Anxiety Scale (Liebowitz, 1987). Exclusion criteria were the engagement in current psychotherapy or a diagnosis in any module of the Mini International Neuropsychiatric Interview (MINI; German version: (Ackenheil, Stotz, Dietz-Bauer, & Vossen, 1999)), excluding Module G (social phobia).

A Kolmogorov-Smirnov test for normality approved that the sample distribution for anxiety scores in the online questionnaire did not differ significantly from a normal distribution ($p > 0.05$). The final selection of participants was based on the scores from the online questionnaire. The severity of social anxiety was judged by the reported cutoff-values for the self- and clinical-administered-format of the Liebowitz Scale (Mennin et al., 2002; Rytwinski et al., 2009). The authors gave a cutoff-value of 30 points for the differentiation of people with and without social anxiety disorder and 60 points for a generalized social anxiety disorder. The implementation of these cutoff-values proved to be difficult in our sample due to numerous canceled appointments by participants. Hence, individuals with a total score less than 33 were assigned to the low-anxious group and subjects with a total score greater than 55 to the high-anxious group. An overview of the demographic variables and scores in the online questionnaire is given in Table 1.

2.2. Apparatus and materials

The Liebowitz Social Anxiety Scale (Liebowitz, 1987; Stangier & Heidenreich, 1996) was administered as a screening tool. It is considered a reliable and valid measure of social anxiety disorder (Fresco et al., 2001).

Interviews: The German version (5.0.0) of the MINI (Sheehan et al., 1998) from Ackenheil et al. (1999) was applied to capture psychic axis-I-disorders in DSM-IV and ICD-10. The VR was generated with the Game Engine Unity3D (Version 5.2.2.f1). Physiological data was digitally amplified (V-Amp 16, Brain Products GmbH, Gilching, Germany) and continuously monitored (Brain Vision Recorder Software, Version 1.20, Brain Products GmbH, Germany).

Two different virtual environments were presented (see Fig. 1): In the first scenario of the behavior test, a porter stood in the front part of the train, handed a ticket to the participants as they walked towards him and established eye contact. The second scenario took place in the waiting room of a doctor's office where a receptionist sat in a corner behind two desks. Four patients waited on chairs next to each other. They slightly moved their heads. During the interactions with the participants they made no further movements. Two of the agents were female and there was a younger and an older avatar for each sex. At the end, a doctor dressed in doctor's overalls entered the scene.

All instructions from the virtual investigator and comments from the avatars were made possible by prerecordings. A speech recognition software enabled users to interact with the virtual agents. Eye movements were monitored binocularly with a TobiiEyeX-Eye Tracker (Tobii AB, Danderyd, Sweden) throughout the behavior test. Because of the low tracking frequency (60 Hz), a chin rest was not necessary. For the skin conductance response, two electrodes (Ag/AgCl, Ø = 8 mm) were attached with electrode cream to the thenar muscle of left hand (TD − 246, PAR

Table 1
Demographic variables and data from the Online-LSAS.

Variable	Total sample (N = 119)		Analysis sample (n = 37): low-anxious group (n = 19)		Analysis sample (n = 37): high-anxious group (n = 18)		*t*	*df*	*p*	*d*
	M	*SD*	*M*	*SD*	*M*	*SD*				
Age	21.00	2.98	23.11	2.94	20.44	2.62	2.90	35	0.006	0.97
LSAS	43.24	18.81	24.05	5.29	67.11	10.53				
	n	%	n	%	n	%			p^a	
Gender [female]	95	79.8	10	52.6	15	83.3			0.079	
Profession [student]	117	98.3	18	94.7	17	94.4			0.522	

Note. M = mean, *SD* = standard error, *df* = degrees of freedom, *d* = Cohen's *d*, n = quantity, p^a = Fisher's exact *t*-test, two-tailed.

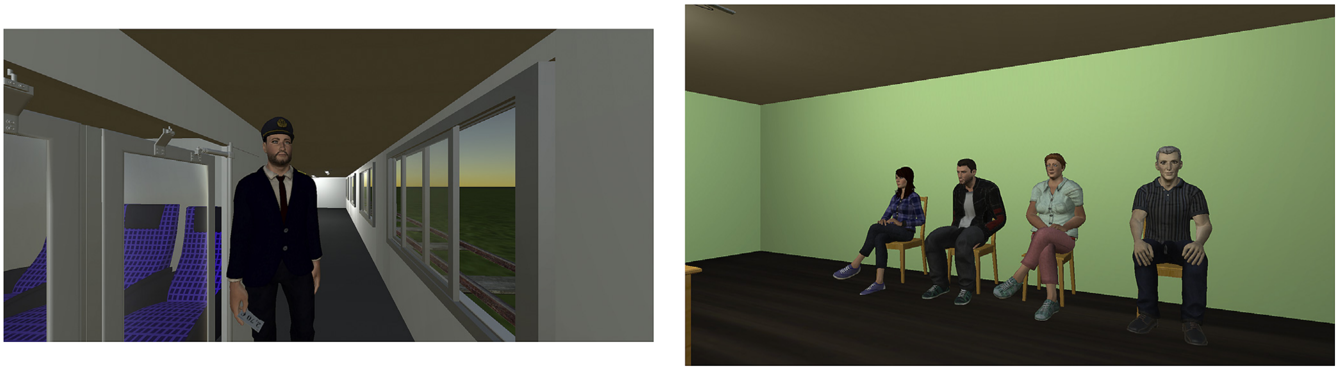

Fig. 1. Virtual train (left) and waiting room scenario (right).

Medizintechnik GmbH, Berlin, Germany) as the right hand was needed for navigation with the arrow keys and mouse. The participants' subjective anxiety was assessed by a rating scale from 0 to 100, which was presented on the desktop.

2.3. Experimental design

The between-subject factor (low- vs. high-anxious) was operationalized by the individuals' score in the online questionnaire. The SCR was assessed to capture the physiological responses to social interactions and was analyzed for three phases (calculation, answer, silence) of the train scenario and five phases of the waiting room scenario (eye contact with the virtual patients, approach, refusal, post-refusal, silence). Moreover, the gaze behavior was recorded in the time interval “answer” in the train and “approach” in the waiting room. Finally, participants estimated their expected and real fear on a scale from 0 to 100.

2.4. Procedure

At first, a real-person investigator conducted the MINI. The interview duration varied between 15 and 25 min. Subsequently, the sensors and the eye tracking device were calibrated and the experiment was started.

During the experiment in VR, the participants sat in a quiet room in front of the monitor. Great effort was taken to reduce real world distractions during the experiment. For example the examiner was sitting out of the participant's field of view. Furthermore, participants were instructed to focus on the monitor only. In the beginning, participants were instructed via headset to buy a ticket at the porter. After successfully making eye contact, each participant had to solve four multiplication or addition tasks. Eye contact could only be established when a trigger point was reached, which was signaled aurally. This way, the determined interaction distance equaled approximately 50 cm in the real world.

As soon as the porter named the amount of change, which he owed the participant following his own calculations, he asked 10 s later whether this amount was correct. Each test subject then had 10 s to react. The participants could answer *right* or *wrong* or state the correct amount. The expected fear before each calculation and real fear after each calculation had to be reported through mouse clicking on the overlaid scale. There was no time limit for the ratings.

Afterwards, the test subjects were supposed to induce eye contact with each virtual patient in the waiting room of the doctor's office and to ask each one whether he would give up his seat. In line with the train scene, trigger points were present in front of each agent leading to interaction distances similar to about 50 cm in the real world. The interaction order was arranged by the instructions given. After having managed to make eye contact, the participants were directed to ask for a seat. They had a time interval of 10 s to do so. All four agents refused to stand up and were annoyed (Avatar one to four: 19 s, 16 s, 16 s, 15 s). After each denial, the test subjects had again a time interval of 10 s to comment, e.g. apologize. Similar to the train scenario, the expected and real fear had to be rated before and after each interaction. In the end, a friendly doctor solved the seat problem telling the test subject that he was next, as his inoculation was a fast procedure.

2.5. Data analysis

For all analyses, the significance level was set at two-tailed $\alpha = 0.05$ and the four analog tasks in each of the scenes were averaged. Cohen's *d* or partial eta squared (η_p^2) are reported as effect sizes. In the train scene, the phases *calculation* (10 s after the calculation of the porter), *answer* (10 s after the question of whether the calculation was correct) and *silence* (phase after the participant reported his fear; time interindividual variable) were defined. The waiting room scene was split in the phases *eye contact with the virtual patients* (10 s after successfully making eye contact), *approach* (10 after the instruction to ask for a seat), *refusal* (unfriendly reaction of the agents ($1 = 19$ s, $2 = 16$ s, $3 = 16$ s, $4 = 15$ s)), *post-refusal* (10 s in which the test subject could react to the refusal of a seat) and *silence* (phase in which participants could report their fear; time interindividual variable).

SCR: SCR was analyzed using Brain Vision Analyzer. A 1 Hz cutoff filter was applied. Segments were created 2000 ms before and 7000 ms after each marker position. Minima were marked automatically in an interval from 0 to 1000 ms and maxima from 1000 to 7000 ms. If necessary, corrections were made manually. After subtracting maxima and minima, z- and T-values for the SCR magnitude were calculated.

Repeated-measures ANOVAS with the within-subject-factor time (train: calculation, answer, silence; waiting room: eye contact with the virtual patients, approach, refusal, post-refusal, silence) were computed for the SCR. The Greenhouse-Geisser correction was applied in case the sphericity assumption was violated.

For a significant factor of time, post-hoc tests with Bonferroni corrected levels of significance were done. Two participants of the low-anxious group and one high-anxious individual were excluded because their raw data were outliers in five or more measure points.

Eye Tracking: Areas of interests (AOIs) were defined as hair, forehead, eyes, nose, mouth and body. Fixation times for the AOIs and remaining areas (environment) were exported for 10 s for

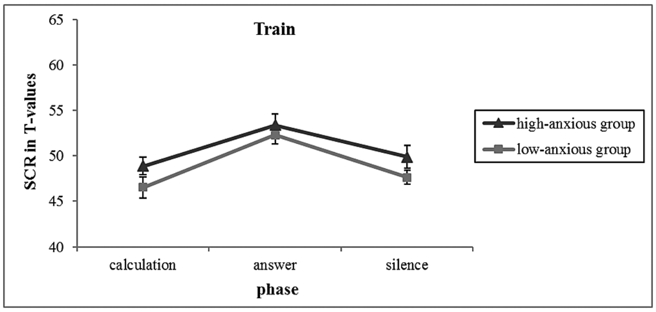

Fig. 2. Skin conductance response for low- and high-anxious individuals for the phases of the train scene. *Note.* Error bars represent standard error of the mean.

answer and *approach*. To increase the reliability of our measurements, we calculated cumulative values for the fixation time on the faces (hair, forehead, eyes, nose, mouth). Eye tracking data was evaluated using two-tailed t-tests for independent samples.

Fear rating: Repeated-measures ANOVAs with the within-subject-factor time (expected fear, real fear) were carried out for fear ratings.

3. Results

3.1. SCR

Train: As visible in Fig. 2, the SCR of the high-anxious group was higher than in the low-anxious group for all three phases. Firstly, it increased after the participants were asked for their opinion about the correctness of a calculation, and decreased afterwards. This was reflected in the main effects of the group, $F(1,32) = 5.04$, $p < 0.001$, $\eta_p^2 = 0.28$, and time, $F(2,64) = 12.31$, $p < 0.001$, $\eta_p^2 = 0.28$, in a repeated-measure ANOVA. There was no reliable interaction between group and time, $F(2,64) = 0.21$, $p = 0.811$, $\eta_p^2 = 0.007$. Post-hoc-tests revealed significant differences between the phases *calculation* and *answer*, $p = 0.001$, and *answer* and *silence*, $p < 0.001$.

Waiting room: In the waiting room scenario, it was observed that the SCR increased in the beginning when test subjects were supposed to verbally approach the agents. However, the denial of the participants' request was not followed by a further increase. In both groups, the SCR decreased during *refusal*. While the SCR in the high-anxious group increased slightly in the next phases thereafter, the SCR decreased further in the low-anxious group and only increased in *silence* (Fig. 3). “A repeated-measures ANOVA showed a statistically relevant effect of group, $F(1,32) = 5.04$, $p = 0.032$, $\eta p^2 = 0.14$, and time, $F(2.75, 87.93) = 70.53$, $p < 0.001$, $\eta p^2 = 0.69$. The SCR for the group of high-anxious individuals was lower than for the low-anxious ones. Only during approach the SCR for high-anxious individuals was higher ($M = 62.37$, $SD = 6.21$) than for low-anxious controls ($M = 61.98$, $SD = 6.86$).” A reliable interaction between time and group was not present, $F(2.75, 87.93) = 0.95$, $p = 0.412$, $\eta_p^2 = 0.029$. Post-hoc-comparisons revealed that the SCR during the phase *eye contact with the virtual patients* differed significantly from *approach* ($p < 0.001$), *refusal* ($p = 0.021$) and *post-refusal* ($p < 0.001$). Additionally, the SCR was significantly higher during *approach* than during *refusal*, *post-refusal* and *silence* (all *ps* < 0.001).

3.2. Eye tracking

Once the participants were asked to judge the correctness of the porter's calculations, low-anxious individuals focused their gaze for a longer amount of time on the porter's face than high-anxious test subjects, who concentrated their gaze more on the body and the environment (Fig. 4). There were significant differences for the percentage of fixation time on the face, $t(35) = 2.68$, $p = 0.011$, $d = 0.88$, and body, $t(35) = -3.29$, $p = 0.002$, $d = -1.08$. Regarding viewing the environment, the groups did not vary significantly, $t(35) = -0.45$, $p = 0.659$, $d = -0.15$. There was no discrepancy in the percentage of eye fixation between the groups, $t(35) = 0.82$, $p = 0.416$, $d = 0.28$.

As Fig. 5 shows, participants of the low-anxious group fixated on the agents' faces longer than high-anxious individuals when they had to ask for a seat. Furthermore, low-anxious test subjects looked for a shorter amount of time at the body and environment. There were significant differences regarding the fixation percentage for face, $t(35) = 3.00$, $p = 0.005$, $d = 0.99$, and body, $t(35) = -2.95$, $p = 0.006$, $d = -0.97$, but not for environment, $t(35) = -1.48$, $p = 0.146$, $d = -0.49$. The comparison of the fixation time on the

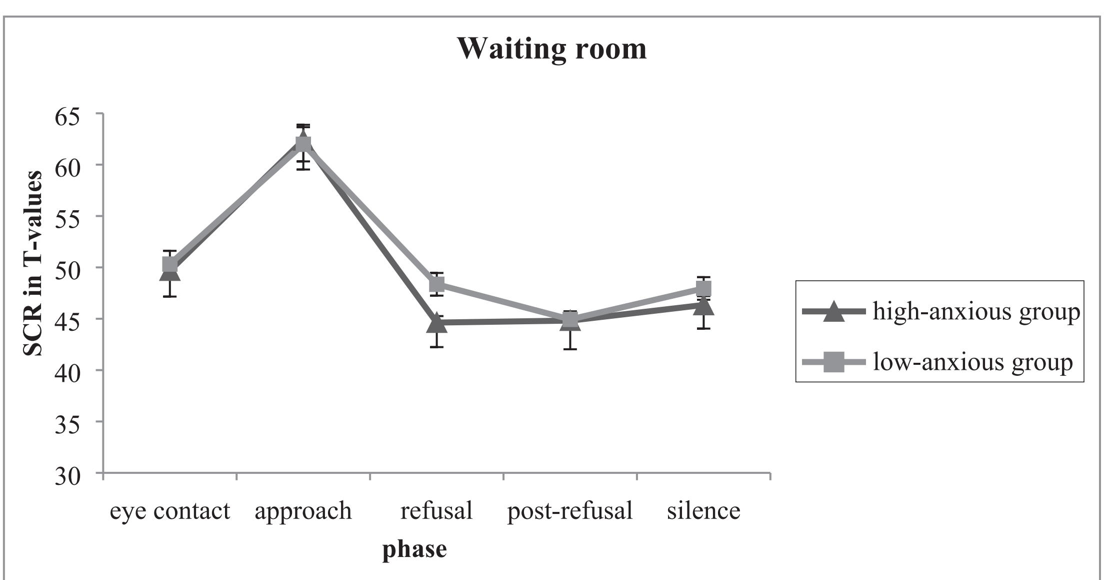

Fig. 3. Skin conductance response for low- and high-anxious individuals for the phases of the waiting room scene. *Note.* Error bars represent standard error of the mean.

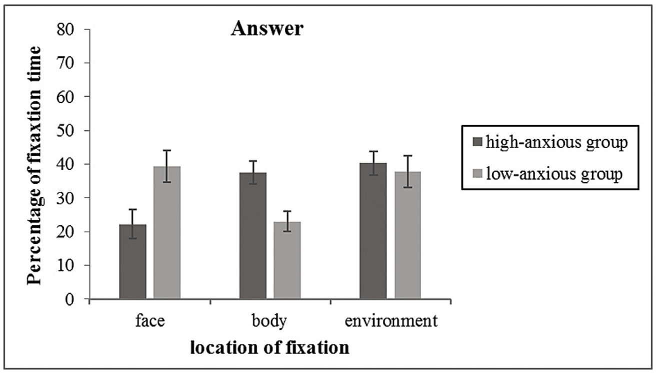

Fig. 4. Percentage of fixation time of face, body and environment for low- and high-anxious individuals for the phase *answer* in the train scenario. *Note.* Error bars represent the standard error of the mean.

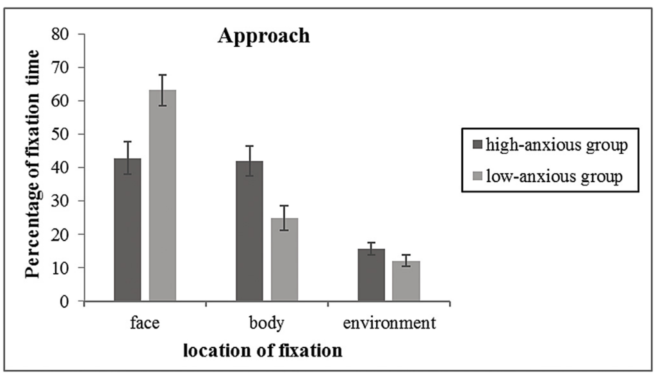

Fig. 5. Percentage of fixation time of face, body and environment for low- and high-anxious individuals for the phase *approach* in the waiting room scenario. *Note.* Error bars represent the standard error of the mean.

eyes between low- and high-anxious individuals was also significant, $t(35) = 4.06$, $p < 0.001$, $d = 1.34$, with low-anxious test subjects fixating on the eyes significantly longer.

3.3. Fear rating

Train: The means of the reported expected and real fears were higher for high-anxious participants. However, there were only minor disparities between expected and real fears within both groups (see Fig. 6). An ANOVA revealed a significant effect of group, $F(1,35) = 16.15$, $p < 0.001$, $\eta_p^2 = 0.32$, but not of time, $F(1,35) = 0.10$, $p = 0.755$, $\eta_p^2 = 0.003$, and the interaction between group and time, $F(1,35) = 0.02$, $p = 0.881$, $\eta_p^2 = 0.001$.

Waiting room: Consistently, means of the reported expected and real fears were significantly higher in the high-anxious group. In both groups, the expected fear was higher than real fear (Fig. 7). An ANOVA showed a significant effect of group, $F(1,35) = 20.30$, $p < 0.001$, $\eta_p^2 = 0.38$, and time, $F(1,35) = 14.24$, $p = 0.001$, $\eta_p^2 = 0.29$. The interaction between group and time was not significant, $F(1,35) = 0.16$, $p = 0.690$, $\eta_p^2 = 0.005$.

4. Discussion

To overcome the existing imbalance of the application of VR regarding social anxiety disorder for psychotherapeutic purposes, the current study investigated the suitability of virtual technology for diagnosing social anxiety disorder by implementing a behavior test in a virtual environment. The main goal of this research was to examine differences between low- and high-social-anxious based on SCR and gaze recordings.

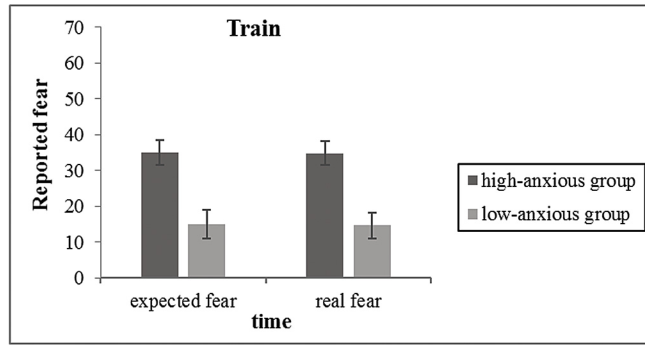

Fig. 6. Means of the reported expected and real fear of all four tasks of the train scenario for low- and high-anxious individuals. *Note.* Error bars represent the standard error of the mean.

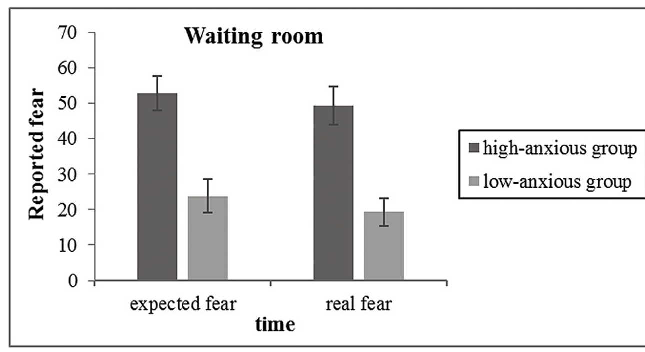

Fig. 7. Means of the reported expected and real fear of all four tasks of the waiting room scenario for low- and high-anxious individuals. *Note.* Error bars represent the standard error of the mean.

A higher SCR in high-anxious participants was only observed in the train scenario. In the waiting room, the SCR of the low-anxious group was higher. This discrepancy was unexpected. Nonetheless, SCR appears as a measure of arousal less sensitive for emotional valence (Boucsein, 2012). Therefore, not only anxiety but also arousal caused by a new scenario could have determined the physiological reactions. Furthermore, an already existing autonomic overarousal in patients suffering from anxiety disorders could have resulted in a lower reactivity of SCR (Lader, 1975). The SCR was consistently highest for *answer* and *approach* for both groups. Our results indicate that the participants' arousal rose up until these direct speaking situations and that the subjective stress was highest for these time intervals. However, this result can possibly be attributed to breathing influences (Greco & Baenninger, 1991).

Moreover, the high-anxious group concentrated their gaze for a significantly shorter amount of time on the faces of the avatars (Chen et al., 2015). The so-called hypervigilance-avoidance-hypothesis (Mogg, Bradley, de Bono, & Painter, 1997) postulates that anxious individuals pay attention to threatening stimuli first and then ultimately avoid them, which serves as a defensive reaction to reduce anxiety. A study conducted by Pflugshaupt et al. (2005) confirmed this pattern in spider phobia. Spider phobics showed a faster detection of spiders but subsequently fixated further from spiders than controls. Referring to this theoretical background, the avoidance of facial regions could have been a coping method for emerging social anxiety for high-anxious participants. An eye tracking study of Wieser et al. (2009) supports this assumption.

However, a shorter fixation time on the eye regions was only shown for the phase *approach*. Interestingly, we found longer fixation times on the body regions but not the environment whenever high-anxious participants fixated on the faces for a significantly shorter period of time. In sum, gaze behavior has proven to be a suitable predictor for social phobia, whereas SCR was a less reliable measurement of social anxiety. According to our results, social anxiety seems to affect avoidance behavior (gaze) to a greater extent than physiological responses (SCR). Gaze behavior in virtual reality can be regarded as a potential diagnostic tool for detecting social anxiety disorder. In future research, the applicability of multimodal measures such as cortisol, heart rate or neuronal correlates as diagnostic tools of social anxiety could be investigated.

As expected, the high-anxious group reported higher expected and real fears before and after all tasks. This is in line with results from Chen et al. (2015). The authors revealed higher fear ratings before and after giving a speech in front of a prerecorded audience for high-anxious participants in comparison to a control group. Yet, only in the waiting room scenario was the fear overestimated. It is possible that fear ratings were easier for test subjects in the train scenario. Solving arithmetic problems could have been more familiar to them whereas, due to reasons of politeness, asking for a seat might have been a totally new experience and therefore difficult to judge. It was likely more familiar to them to solve or check math arithmetic problems. Consistent with the waiting room scenario, Chen et al. (2015) also found higher fear ratings before a speech as opposed to after. Overall, there are various reasons why people have a social fear. Rejection, scrutiny or performance pressure are all cognitive parts of social anxiety; however, we do not exactly know what participants were afraid of as we did not ask. This should be considered for further research.

Test subjects belonging to the low-anxious group were significantly older than in the high-anxious group. Yet, we assume that the mean difference of 2.67 years between the groups did not influence our findings. It has to be taken into account that the paradigms chosen for the behavior test were not discussed in earlier research, which limits their comparability. Both paradigms were social situations wherein the train scene included performance attributes and the waiting room scene was rather oriented on social interaction with strangers. Because of the various content of the scenes, we presumed that the order of their presentation did not affect our results and therefore we presented all participants the train scenario first. Through the application of a desktop VR system, the test subjects could continuously perceive the laboratory around them.

Even though the degree of immersion seems to be apparently smaller for a desktop system than for HMDs, it is important to consider that not only is the technical equipment a crucial factor for experiencing presence, but also that the perceived emotions are decisive (Diemer, Alpers, Peperkorn, Shiban, & Mühlberger, 2015). Using a HMD system instead of a monitor setting is a considerable option. However, as our study was designed as a pilot study, we chose to use the monitor setting at this stage of our research due to simplicity and importantly, due to its quicker implementabiliy at clinics. However, at a later point, once this study is repeatedly validated and further research behind the working mechanisms is warranted a more immersive system such as a HMD based one could be implemented.

We accepted the possibility of reducing the feeling of presence by interrupting the scene with a visual scale. Nevertheless, the fear rating was an essential measure. To minimize the extent of the interruption, participants could constantly see the virtual environment in the background of the overlying scale and were still instructed per headphones. Our study showed that analyses of gaze behavior in a virtual environment are a promising first step for diagnosing social anxiety disorder. Yet, it has to be investigated further to what extent an avoidance of facial regions is specific for social anxiety disorder. Therefore, it is important to include clinical control groups in future research.

Author note

This study is part of the cooperation between the Department of Clinical Psychology and Psychotherapy and the Department of Media Informatics at the University of Regensburg (Play2Change). No potential competing financial interests exist.

Acknowledgements

The authors would like to thank Mario Seebauer, B.Sc., for technical support. In addition, we would like to thank the students who participated in this study.

References

Ackenheil, M., Stotz, G., Dietz-Bauer, R., & Vossen, A. (1999). *Deutsche Fassung des Mini-International Neuropsychiatric Interview*. München: Psychiatrische Universitätsklinik München.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Boucsein, W. (2012). *Electrodermal activity*. New York, NY: Springer.

Brinkman, W.-P., Van der Mast, C., & de Vliegher, D. (2008). *Virtual reality exposure therapy for social phobia: A pilot study in evoking fear in a virtual world*. Delft University of Technology. Retrieved from http://mmi.tudelft.nl/pub/wpbrinkman/hci4tel_brinkman_vandermast_devliegher.pdf.

Chen, N. T., Thomas, L. M., Clarke, P. J., Hickie, I. B., & Guastella, A. J. (2015). Hyperscanning and avoidance in social anxiety disorder: The visual scanpath during public speaking. *Psychiatry Research, 225*(3), 667–672. http://dx.doi.org/10.1016/j.psychres.2014.11.025.

Crome, E., & Baillie, A. (2014). Mild to severe social fears: Ranking types of feared social situations using item response theory. *Journal of Anxiety Disorders, 28*(5), 471–479.

Diemer, J., Alpers, G. W., Peperkorn, H. M., Shiban, Y., & Mühlberger, A. (2015). The impact of perception and presence on emotional reactions: A review of research in virtual reality. *Frontiers in Psychology, 6*, 26. http://dx.doi.org/10.3389/fpsyg.2015.00026.

Dimberg, U., Fredrikson, M., & Lundquist, O. (1986). Autonomic reactions to social and neutral stimuli in subjects high and low in public speaking fear. *Biological Psychology, 23*(3), 223–233. http://dx.doi.org/10.1016/0301-0511(86)90001-3.

Edelmann, R. J., & Baker, S. R. (2002). Self-reported and actual physiological responses in social phobia. *The British Journal of Clinical Psychology, 41*(Pt 1), 1–14. http://dx.doi.org/10.1348/014466502163732.

Freeman, D., Reeve, S., Robinson, A., Ehlers, A., Clark, D., Spanlang, B., et al. (2017). Virtual reality in the assessment, understanding, and treatment of mental health disorders. *Psychological Medicine*, 1–8.

Fresco, D. M., Coles, M. E., Heimberg, R. G., Liebowitz, M. R., Hami, S., Stein, M. B., et al. (2001). The Liebowitz social anxiety scale: A comparison of the psychometric properties of self-report and clinician-administered formats. *Psychological Medicine, 31*(6), 1025–1035. http://dx.doi.org/10.1017/S0033291701004056.

Greco, M., & Baenninger, R. (1991). Effects of yawning and related activities on skin conductance and heart rate. *Physiology & Behavior, 50*(5), 1067–1069. http://dx.doi.org/10.1016/0031-9384(91)90439-U.

Grillon, H., Riquier, F., Herbelin, B., & Thalmann, D. (2006). Virtual reality as a therapeutic tool in the confines of social anxiety disorder treatment. *International Journal on Disability and Human Development, 5*(3), 243–250. http://dx.doi.org/10.1515/IJDHD.2006.5.3.243.

Harris, S. R., Kemmerling, R. L., & North, M. M. (2002). Brief virtual reality therapy for public speaking anxiety. *Cyberpsychology & Behavior, 5*(6), 543–550. http://dx.doi.org/10.1089/109493102321018187.

Hofmann, S. G., Gerlach, A. L., Wender, A., & Roth, W. T. (1997). Speech disturbances and gaze behavior during public speaking in subtypes of social phobia. *Journal of Anxiety Disorders, 11*(6), 573–585. http://dx.doi.org/10.1016/S0887-6185(97)00040-6.

Kim, J.-W., Kim, H.-J., & Nam, T.-J.. (2016). M.Gesture: An acceleration-based gesture authoring system on multiple handheld and wearable devices. Paper presented at the proceedings of the 2016 CHI conference on human factors in computing systems, Santa clara, California, USA.

Krijn, M., Emmelkamp, P. M., Olafsson, R. P., & Biemond, R. (2004). Virtual reality exposure therapy of anxiety disorders: A review. *Clinical Psychological Review, 24*(3), 259–281. http://dx.doi.org/10.1016/j.cpr.2004.04.001.

Lader, M. H. (1967). Palmar skin conductance measures in anxiety and phobic states. *Journal of Psychosomatic Research, 11*(3), 271–281. http://dx.doi.org/10.1016/0022-3999(67)90021-9.

Lader, M. (1975). *The psychophysiology of mental illness*. Routledge & Kegan Paul.

Liebowitz, M. R. (1987). Social phobia. In D. F. Klein (Ed.), *Modern problems of pharmacopsychiatry* (Vol. 22, pp. 141–173). Basel: Karger. http://dx.doi.org/10.1159/000414022.

Lundqvist, D., Esteves, F., & Ohman, A. (1999). The face of wrath: Critical features for conveying facial threat. *Cognition & Emotion, 13*(6), 691–711. http://dx.doi.org/10.1080/026999399379041.

Mauss, I., Wilhelm, F., & Gross, J. (2003). Autonomic recovery and habituation in social anxiety. *Psychophysiology, 40*(4), 648–653. http://dx.doi.org/10.1111/1469-8986.00066.

Mauss, I., Wilhelm, F., & Gross, J. (2004). Is there less to social anxiety than meets the eye? Emotion experience, expression, and bodily responding. *Cognition and Emotion, 18*(5), 631–642. http://dx.doi.org/10.1080/02699930341000112.

McTeague, L. M., Lang, P. J., Laplante, M. C., Cuthbert, B. N., Strauss, C. C., & Bradley, M. M. (2009). Fearful imagery in social phobia: Generalization, comorbidity, and physiological reactivity. *Biological Psychiatry, 65*(5), 374–382. http://dx.doi.org/10.1016/j.biopsych.2008.09.023.

Mennin, D. S., Fresco, D. M., Heimberg, R. G., Schneier, F. R., Davies, S. O., & Liebowitz, M. R. (2002). Screening for social anxiety disorder in the clinical setting: Using the Liebowitz social anxiety scale. *Journal of Anxiety Disorders, 16*(6), 661–673. http://dx.doi.org/10.1016/S0887-6185(02)00134-2.

Mertens, I., Siegmund, H., & Grüsser, O. J. (1993). Gaze motor asymmetries in the perception of faces during a memory task. *Neuropsychologia, 31*(9), 989–998. http://dx.doi.org/10.1016/0028-3932(93)90154-R.

Meyerbröker, K., & Emmelkamp, P. M. G. (2010). Virtual reality exposure therapy in anxiety disorders: A systematic review of process-and-outcome studies. *Depression and Anxiety, 27*, 933–944. http://dx.doi.org/10.1002/da.20734.

Mogg, K., Bradley, B. P., de Bono, J., & Painter, M. (1997). Time course of attentional bias for threat information in non-clinical anxiety. *Behavior Research and Therapy, 35*(4), 297–303. http://dx.doi.org/10.1016/S0005-7967(96)00109-X.

Mühlberger, A., Wieser, M. J., & Pauli, P. (2008). Visual attention during virtual social situations depends on social anxiety. *Cyberpsychol Behav, 11*(4), 425–430. http://dx.doi.org/10.1089/cpb.2007.0084.

Opriş, D., Pintea, S., García-Palacios, A., Botella, C., Szamosközi, Ş., & David, D. (2012). Virtual reality exposure therapy in anxiety disorders: A quantitative meta-analysis. *Depress Anxiety, 29*, 85–93. http://dx.doi.org/10.1002/da.20910.

Pflugshaupt, T., Mosimann, U. P., von Wartburg, R., Schmitt, W., Nyffeler, T., & Müri, R. M. (2005). Hypervigilance–avoidance pattern in spider phobia. *Journal of Anxiety Disorders, 19*(1), 105–116.

Rytwinski, N. K., Fresco, D. M., Heimberg, R. G., Coles, M. E., Liebowitz, M. R., Cissell, S., … Hofmann, S. G. (2009). Screening for social anxiety disorder with the self-report version of the Liebowitz Social Anxiety Scale. *Depression and Anxiety, 26*(1), 34–38. http://dx.doi.org/10.1002/da.20503.

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., … Dunbar, G. C. (1998). The mini-international neuropsychiatric interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry, 59*(Suppl 20), 22–33. quiz 34–57.

Shiban, Y., Brütting, J., Pauli, P., & Mühlberger, A. (2015). Fear reactivation prior to exposure therapy: Does it facilitate the effects of VR exposure in a randomized clinical sample? *Journal of Behavior Therapy and Experimental Psychiatry, 46*, 133–140. http://dx.doi.org/10.1016/j.jbtep.2014.09.009.

Shiban, Y., Pauli, P., & Mühlberger, A. (2013). Effect of multiple context exposure on renewal in spider phobia. *Behaviour Research and Therapy, 51*(2), 68–74. http://dx.doi.org/10.1016/j.brat.2012.10.007.

Slater, M., Pertaub, D. P., Barker, C., & Clark, D. M. (2006). An experimental study on fear of public speaking using a virtual environment. *CyberPsychology & Behavior, 9*(5), 627–633. http://dx.doi.org/10.1089/cpb.2006.9.627.

Stangier, U., & Heidenreich, T.. (1996). Liebowitz-Soziophobie-Skala - German version. Unpublished manuscript. Johann Wolfgang Goethe-Universität, Frankfurt am Main.

Walker-Smith, G. J., Gale, A. G., & Findlay, J. M. (1977). Eye movement strategies involved in face perception. *Perception, 6*(3), 313–326. http://dx.doi.org/10.1068/p060313.

Wieser, M. J., Pauli, P., Alpers, G. W., & Mühlberger, A. (2009). Is eye to eye contact really threatening and avoided in social anxiety?—An eye-tracking and psychophysiology study. *Journal of Anxiety Disorders, 23*(1), 93–103. http://dx.doi.org/10.1016/j.janxdis.2008.04.004.