Check for updates

OPEN ACCESS

EDITED BY Daniel Thalmann, Swiss Federal Institute of Technology Lausanne, Switzerland

REVIEWED BY Judith Molka-Danielsen, Molde University College, Norway Sabrina Gado, Julius-Maximilians-University of Würzburg, Germany

*CORRESPONDENCE Teresa Schmidt-Peter, teresa.schmidt@psychologie.uniregensburg.de

RECEIVED 07 January 2025 ACCEPTED 19 May 2025 PUBLISHED 18 June 2025

CITATION

Schmidt-Peter T, Wechsler TF, Kroczek LOH and Mühlberger A (2025) The effects of different variants of eye-tracking-based feedback of attentional processes during virtual social interactions. *Front. Virtual Real.* 6:1556898. doi: 10.3389/frvir.2025.1556898

COPYRIGHT

© 2025 Schmidt-Peter, Wechsler, Kroczek and Mühlberger. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

The effects of different variants of eye-tracking-based feedback of attentional processes during virtual social interactions

Teresa Schmidt-Peter*, Theresa F. Wechsler, Leon O. H. Kroczek and Andreas Mühlberger

Department of Psychology, Clinical Psychology and Psychotherapy, University of Regensburg, Regensburg, Germany

Introduction: Enhanced self-focused attention plays an important role in the maintenance of Social Anxiety Disorder (SAD). Therefore, changing attentional processes is a major target in cognitive behavioral therapy (CBT) and recent approaches apply Virtual Reality (VR) behavioral exercises to change these processes. A promising approach to enhance such VR exposure-based exercises is implementing eye-tracking-based feedback.

Methods: This experimental study investigates which characteristics of gazerelated feedback lead to a positive valence and an increase in focused attention on social stimuli. Additionally, we examine differential effects in low (LSA) vs highly socially anxious (HSA) individuals. Overall, 50 participants, who were grouped into LSA and HSA according to the median split of the SPIN, were instructed to hold eye contact with virtual agents until they received feedback either in the form of a smile, a positive tone, or a praise. Furthermore, the required duration of maintaining eye contact with virtual agents to receive feedback was manipulated. The feedback variants were evaluated during and after the experiment *via* ratings, and participants' gaze was measured *via* eye tracking.

Results: Results revealed that the smile feedback was perceived as more pleasant and elicited more eye contact in a subsequent test phase than the praise, which was associated with higher valence than the tone. In addition, LSA participants rated the social feedback variants (smile, praise) as significantly more pleasant than HSA participants, who showed reduced sensitivity to positive social feedback.

Discussion: These findings suggest that socially rewarding feedback is more effective in LSA individuals and may not generalize to those with high social anxiety. Future research should therefore explore further feedback variants within individuals with SAD to further refine and optimize VR-based attentional interventions for enhanced therapeutic outcomes.

KEYWORDS

virtual reality, social anxiety, self-focused attention, social interaction, eye tracking, feedback, exposure therapy

1 Introduction

The main characteristic of SAD is pronounced and persistent fear, anxiety, or avoidance of social situations with the possibility of being scrutinized by others or of embarrassment (DSM-5; Falkai, 2018). With a 12-month prevalence of about 2% (Wittchen and Hoyer, 2011) and a median lifetime prevalence of about 6% (Fehm et al., 2005) for Europe and a lifetime prevalence of 12%–13% (Kessler et al., 2005; 2012) in the USA, SAD is one of the most common mental illnesses. People with SAD suffer from a reduced quality of life and functional impairments in social domains as well as severe impairments in school and professional areas of life (Fehm et al., 2005).

Maintenance of social anxiety has been linked to avoidance behavior, self-focused attention, safety behaviors, or lack of social skills (e.g., Rapee and Heimberg, 1997). Furthermore, cognitive behavioral models on SAD (Clark and Wells, 1995b; Heimberg et al., 2010; Hofmann, 2007; Rapee and Heimberg, 1997) highlight the role of distorted attentional focus for the maintenance of the disorder. If people with social anxiety are in a feared situation, they show increased self-attention by observing and checking their appearance and especially their physiological symptoms in detail (Clark and Wells, 1995). The self-focused attention enhances perception of one's own fear reaction and prevents a realistic processing of the situation and the behavior of others because of the lack of attention toward the surrounding. Furthermore, affected individuals draw biased conclusions about how they appear to others based on internal information such as physical symptoms of anxiety or distorted thoughts about their observable self (Clark and Wells, 1995).

The hypervigilance-avoidance hypothesis (Mogg et al., 1987) proposes external attentional biases to contribute to the maintenance of social anxiety. This hypothesis postulates an initial hypervigilance and subsequent avoidance of social threat cues and some empirical studies support this hypothesis (Bögels and Mansell, 2004). Recent studies on biased attentional focus using eye tracking have shown that people with SAD avoid making eye contact (Hessels et al., 2018; Horley et al., 2003), and looking at emotional faces (Chen et al., 2015) or faces in general (Moukheiber et al., 2010). According to these studies making eye contact can be a major challenge for people with social anxiety (Radke et al., 2013; Wieser et al., 2009), resulting in impairments in establishing contact with potential interaction partners.

In addition to attentional biases, people with social anxiety show deficits in pro-social behavior by not responding appropriately to their interaction partner. Consequently, they could be evaluated as less sympathetic, confirming their central fears of being disliked (Vrijsen et al., 2010). On this basis, the experienced fear is reflected in nonverbal characteristics such as nervous and restless behavior, frowning, and a lower emotional mimicry of the interaction partner's smile (Heerey and Kring, 2007). All these behaviors might contribute to the maintenance of social anxiety.

CBT is recommended as the method of choice for the treatment of SAD according to, e.g., the German S3 guideline (Bandelow et al., 2021) as well as the British NICE guideline (National Institute for Health and Care Excellence, 2017), and is widely supported by current research (e.g., Arch and Craske, 2009; Mayo-Wilson et al., 2014). A central component of CBT is exposure. During exposure patients are being confronted with a feared object or situation until distress has decreased significantly and/or until a violation of dysfunctional expectancies about the feared object or situation is reached (Craske et al., 2008). Meanwhile, exposure in virtual reality (VRE) is increasingly being used, and has become an important therapeutic instrument for the treatment of SAD (Kampmann et al., 2016). In their meta-analysis comparing VRE and in vivo exposure, Wechsler et al. (2019) found a medium effect for the superiority of in vivo over VR exposure for SAD. In this context, the authors concluded that a combination of VR exposure with cognitive elements might be beneficial for the treatment of SAD (Caponnetto et al., 2021; Emmelkamp et al., 2020; Wechsler et al., 2019). Accordingly, based on the results of previous metaanalyses (Chesham et al., 2018; Horigome et al., 2020; Wechsler et al., 2019) VRE therapy as a stand-alone treatment for social anxiety still requires optimization.

Based on the current research on the treatment of attentional bias in social anxiety, treatment approaches based on the cognitive model by Clark and Wells (1995) and Heimberg et al. (2010) conduct exposure for social anxiety in the form of experimental tasks, which are based on modulating the patients' attentional focus and/ or correcting their dysfunctional beliefs. During such experimental tasks, the adverse effects of self-focused attention are demonstrated and subsequently changed to externally focused attention (Clark and Wells, 1995a). Therefore, certain attentional trainings have already been established to increase the effects of CBT (e.g., Bögels and Mansell, 2004; Feiler and Powell, 2016; Wechsler et al., 2021).

In this context, Wechsler et al. (2021) published the first study on attentional training in combination with VRE, in which socially anxious individuals trained to change their attentional focus away from the self to external neutral or social objects during a speech. Results showed that the virtual attentional training achieved a shortterm reduction in self-reported anxiety measures, body sensations, and further secondary outcome measures Wechsler et al. (2021).

A further promising approach to optimize attentional trainings in VR is utilizing eye-tracking-based feedback of overt attention. This approach has already been investigated in VRE of Specific Phobias (Wechsler et al., 2023). According to current models the promotion of focused attention on the phobic stimuli *via* feedback, the induction of positive affect, and positive reinforcement during exposure could enhance fear extinction and inhibitory learning (Pearce and Mackintosh, 2010; Zbozinek et al., 2015; Zbozinek and Craske, 2017). Therefore, in social anxiety, promoting a gaze focus on social cues may target dysfunctional beliefs and reduce selffocused attention (Feiler and Powell, 2016; Wells and Papageorgiou, 1998).

To implement gaze-based feedback for treatment of social anxiety, there is a need to find suitable feedback variants. Several studies already explored the effectiveness of different types of feedback such as verbal (e.g., praise, acknowledgment) and nonverbal feedback (e.g., smiling, nodding) in educational (Hattie and Timperley, 2007; Wisniewski et al., 2020) and therapeutic contexts (Del Giacco et al., 2020; Skipper and Douglas, 2012). Smiling is considered a primary reinforcer due to its evolutionary basis as social signal that inherently conveys positive emotion and fosters bonding (Kroczek and Mühlberger, 2023; Ekman et al., 1990; Kraut and Johnston, 1979). Its effectiveness depends on timing (e.g., Kroczek and Mühlberger, 2023) and context (e.g., Kohls et al., 2009), therefore

verbal feedback such as praise is used as a further social reinforcer (Matyjek et al., 2020). Feedback in form of praise also plays a crucial role in therapeutic context and is used as reinforcer to promote desired behavior and therapeutic alliance (Rubel et al., 2015; Wisniewski et al., 2020). However, with regard to our target group, it should be noted that both smile and praise are social reinforcers that could be misinterpreted or perceived as threatening by socially anxious individuals, which could be associated with a reduction of the positive reinforcing effect (Gilboa-Schechtman et al., 2014; Gilboa-Schechtman and Shachar-Lavie, 2013). On this basis, it seems important to consider non-social feedback such as reinforcing tones (e.g., Altmeyer et al., 2022; Naal-Ruiz et al., 2022; Yang et al., 2018). In addition to its use in behavioral experiments (e.g., operant conditioning), the clear and precise character of tones is also used in technological applications (e.g., video games, learning programs) (Altmeyer et al., 2022; Linderkamp, 2009; Skinner et al., 1974).

Delay has been considered as a primary determinant of the effectiveness of a feedback or reinforcer (Lattal, 2010). Several studies have already investigated the timing of feedback and its effects on learning and behavior (e.g., Opitz et al., 2011; Smith and Kimball, 2010). While previous studies typically investigate delayed feedback after the desired behavior has been performed, our study involves continuous behavior until feedback is provided. Consequently, the duration of the required maintenance of eye contact with virtual agents determinates the delay of the feedback related to the initiation of eye contact. Considering the attentional biases of socially anxious individuals, it may be assumed that both very short and very long delays would lead to less effective engagement with the feedback. Short delays may not allow participants enough time to process or fully engage with feedback, potentially reducing its impact, whereas long delays might decrease the salience of the feedback, leading to increased difficulty and anxiety (e.g., Clark and Wells, 1995; Hessels et al., 2018; Moukheiber et al., 2010). Therefore, a medium delay duration would establish an optimal balance, providing enough time for participants to engage with and process the feedback while maintaining its salience and relevance. Learning theories would support this view, emphasizing the importance of timely and salient corrective feedback for optimal expectancy violation and effective attentional retraining (e.g., Craske et al., 2014; Pittig et al., 2016).

Based on the current state of research on the treatment of social anxiety, the aim of our broader project is to optimize attentional training in VR for patients with SAD by implementing eye-trackingbased feedback. As one important step, the current study attempts to investigate suitable feedback variants as one central component. Here, participants' attentional focus on the instructed social stimuli (virtual agents) is feedbacked using visual-social (smile), auditorysocial (praise) or auditory non-social (tone) stimuli. Thus, we are particularly interested in which characteristics of gaze-based feedback realized via eye tracking lead to a positive experience and to an increase in focused attention on social stimuli (primary research questions). Based on the current research (Kohls et al., 2009; Liu et al., 2021; Matyjek et al., 2020; Rubel et al., 2015), as a primary hypothesis we assume that the feedback variants differ significantly from each other in terms of their perceived valence. Regarding the differences we assume that the *smile* is perceived as more pleasant than the tone and the praise. Furthermore, we assume that the praise is perceived as more pleasant than the tone. In addition, we expect that the feedback variants differ significantly from each other in terms of increasing the focus of attention on social stimuli. The smile leads to a stronger increase in focused attention on social stimuli than the other feedback variants. Here, we also assume that the praise leads to a stronger increase in focused attention on social stimuli than the tone. Furthermore, in regard of the duration of maintaining eye contact with virtual agents, we assume that the different durations of maintenance differ significantly from each other regarding their perceived valence. We expect that maintaining eye contact for a medium duration (5.0 s) is perceived more pleasant than for a short (3.0 s) and a long duration (7.0 s). As the eye-tracking-based feedback of attentional processes is developed to be used in psychotherapy for SAD, it is of particular interest whether the effects differ between LSA and HSA individuals (secondary research question).

2 Materials and methods

2.1 Participants

We included participants between 18 and 65 years with normal or corrected-to-normal vision (including participants wearing glasses or contact lenses). Exclusion criteria were self-reported neurological (e.g., epilepsy, migraine) or mental illnesses (except SAD) and severe physical impairments. The sample size required to identify a medium effect size of $\eta_p^2 = 0.09$ with regard to the main research question is 52 participants with a power of 1- $\beta = 0.80$ to be achieved (determined for main effect of feedback variant and alpha = 0.05). As no relevant effect sizes have been reported in the literature so far, we used previous research on feedback variants as a reference and assumed a medium effect as relevant outcome.

Participants were recruited through advertisements at the university of Regensburg. Overall, 53 adults participated in the study. All gave written informed consent and psychology students received credit points for participation. The experimental procedure was approved by the ethics committee of the University of Regensburg (Reference number: 23-3376-101).

A total of three participants had to be excluded due to too many missing items in the questionnaires (n = 1) or due to technical problems during the experiment (n = 2). The remaining 50 participants were included in the analyses (mean age = 25.53 years, SD = 7.01, range 19–64, 36 female). For the gaze data analysis two additional data sets had to be excluded because of missing data (n = 48).

Regarding our secondary research question whether the effects differ between LSA and HSA individuals a screening instrument for social anxiety (Social Phobia Inventory; SPIN; Stangier and Steffens, 2002) was used to divide the sample into two groups by median split (median = 17). This resulted in a group of LSA (SPIN mean = 9.67, SD = 3.12, n = 24) and a group of HSA individuals (SPIN mean = 25.74, SD = 7.78, n = 26). Although the preregistration had suggested a cutoff score of 19 for distinguishing individuals with elevated social anxiety, we opted for a median split to ensure equal group sizes within our non-clinical sample. This approach also allowed us to contextualize the suggested threshold score, as the observed median was close to the recommended cutoff.

2.2 Study design

This study was conducted in the lab of the Department of Psychology, Clinical Psychology and Psychotherapy, at the University of Regensburg and implemented as within-subject design with the experimental factors feedback variant (smile vs tone vs praise), and maintenance (short vs medium vs long), and the between-subject factor social anxiety (LSA vs HSA). Participants had to engage in a social interaction with a virtual agent by maintaining eye contact for 3.0, 5.0, or 7.0 s (manipulated as factor maintenance). After holding eye contact for the full amount of time, participants received feedback: either a smile, a tone, or a praise. All combinations of feedback variants and maintenance durations occurred a total of twelve times each. Ratings of valence and arousal elicited by the feedback variants were measured as dependent variables. All measurements, statistical models, hypotheses, and the procedure of this study were preregistered prior to creation of data (https://doi.org/10.17605/ OSF.IO/PRB3Z).

2.3 Apparatus and materials

The virtual environment was generated using the Unreal Game Engine (v 4.27, Epic Games, Raleigh, United States) and Blender (v 2.79). A virtual room based on a seminar room at the University of Regensburg was created. Two male and two female virtual agents were created using the MetaHuman Plugin (Epic Games, Raleigh, United States). The gender of the virtual agents was systematically varied across blocks: all participants encountered male agents in block one and three, and female agents in block two. The virtual agents were animated sitting on a chair at a round table with a neutral facial expression. They were scripted to look around and to direct their gaze toward the participant as soon as the participant makes eye contact. Participants were immersed into VR via a headmounted display (HMD, Vive Pro Eye, Taoyuan, Taiwan). The HTC Vive motion controller was used for the ratings. The distance between the participants and the right and the left virtual agent was the same in each case. A black cross straight ahead on the wall in front of the participants served as fixation cross. Participants' gaze behavior was measured continuously via eye tracking. To trigger the feedback mechanism, a pre-defined target area had to be focused by the participants for a pre-specified amount of time. This target area was centered around the eyes of the agents concerning the eye region up to the eyebrows and nose (size: $21.5 \text{ cm} \times 26 \text{ cm} \times 10 \text{ cm}$). To initiate eye contact participants had to gaze on the agents' eye region for a minimum of 0.3 s. From that time on the total time of eye contact was measured. After a dwell time of either 3.0, 5.0 or 7.0 s (maintenance) feedback was given. However, if the participants shift their gaze away (e.g., look at the body of the agent or environment) for more than 0.5 s during the maintenance period, the agent looked away and eye contact had to be made and maintained again for the full amount of time.

Based on previous research, we chose *visual social* feedback in the form of a smile realized by the focused virtual agent. For *auditory social* feedback, a praise was chosen. The praise consisted of the recording of a male voice ("very good"), which was played through the headphones from the background (off), rather than being spoken by the virtual agent. As *non-social* feedback we decided to use a positive tone. On the basis that evaluation of sounds can depend on individual preferences (e.g., Liu et al., 2021; Schäfer et al., 2013), sounds with a positive connotation according to the literature (Altmeyer et al., 2022; Naal-Ruiz et al., 2022; Yang et al., 2018) were rated by eight researchers in regard of their valence. Finally, we decided to use the sound effect "positive bell" by Altmeyer et al. (2022), which was also sounded from the off.

2.4 Measures

Before the start of the experiment, demographic data (e.g., sex, age, deployment, or study program), and severe physical impairments were collected *via* self-report. Social anxiety was measured using the Social Phobia Inventory (SPIN, Connor et al., 2000, German version; Stangier and Steffens, 2002) and the Social Interaction Anxiety Scale (SIAS; Mattick and Clarke, 1998; German version: Heinrichs et al., 2002).

Besides these self-reported measures, participants' attentional focus was measured in its frequency and duration (dwell-time) during the experiment *via* eye tracking. Therefore, an eye tracker integrated into the HMD (HTC Vive Pro Eye) was used. Gaze data was recorded with a custom-built Unreal Engine plugin developed at our department using the SRanipal SDK (HTC Vive) and interpreted with regard to predefined regions of interest (ROIs) in the virtual environment. ROIs included the virtual agents' eye areas, the heads, and the bodies below the heads of the virtual agents, and the rest of the virtual environment (nonsocial stimuli, e.g., fixation cross, table, mobile phone, wall). Furthermore, electrodermal activity was recorded but is not in the focus of this study.

To measure the perceived valence and arousal of the different feedback variants, the Self-Assessment Manikin (SAM, Bradley and Lang, 1994) was used both during the experiment (online affect ratings) and after the experiment (post training affect ratings). While the online affect ratings refer to the evaluation of the combination of feedback variant and duration of maintaining eye contact (e.g., "How pleasant did you find the last feedback?"/"Wie angenehm empfanden Sie das letzte Feedback?"), the individual feedback variants and durations of maintaining eye contact were evaluated separately in the post training affect ratings (e.g., "How pleasant did you find the tone after making eye contact?"/"Wie? angenehm haben sie den Ton im Anschluss an ihren Blickkontakt empfunden?"). Ratings were given on a 9-point Likert-scale ranging from 1 (low arousal, highly unpleasant) to 9 (high arousal, highly pleasant). The online affect ratings occurring directly after feedback, measured perception of the feedback variants directly in the current moment and enabled an assessment and analysis of the interaction between feedback variant and duration of maintenance of eye contact. Therefore this measure is primarily used to answer the research questions. The post training affect ratings allow a more distanced (and prompted) evaluation in retrospect and reflect an overall assessment that can be taken into account when interpreting the results.

After the experiment, physical and social presence were assessed by the Multimodal Presence Scale (MPS; Makransky et al., 2017; German version; Volkmann et al., 2018) and affective state by adaptive versions of the Brief Measure of Positive and Negative

| Sample characteristics | HSA (n = 26) | LSA (n = 24) | Fª/X ² /FE | p |
|--|---------------|--------------|-----------------------|---------|
| Sociodemographic variables | | | | |
| Female, sex, n (%) | 20 (77) | 16 (67) | 1.65 | 0.420 |
| Age, M (SD) years | 25.69 (8.87) | 25.61 (4.63) | 0.00 | 0.968 |
| General qualification for university entrance, n (%) | 25 (96) | 22 (92) | 3.12 | 0.211 |
| University degree | 8 (31) | 12 (50) | 2.11 | 0.551 |
| Other | 18 (69) | 12 (50) | | |
| Study course, n (%) | | | | |
| Psychology | 13 (50) | 15 (63) | 2.10 | 0.350 |
| Other | 13 (50) | 9 (37) | | |
| Visual aid, n (%) | 15 (58) | 16 (67) | FE | 0.570 |
| Is carried, n (%) | 8 (31) | 6 (14) | FE | 0.757 |
| Mental health variables | | | | |
| SPIN, M (SD) | 25.74 (7.78) | 9.67 (3.12) | 74.98 | < 0.001 |
| SIAS, M (SD) | 28.12 (12.16) | 15.50 (6.26) | 20.59 | < 0.001 |
| VR-related variables | | | | |
| MPS: Physical Presence, M (SD) | 2.87 (0.66) | 2.60 (0.76) | 1.77 | 0.190 |
| MPS: Social Presence, M (SD) | 2.64 (0.74) | 2.37 (0.87) | 1.42 | 0.239 |

TABLE 1 Sociodemographic, health, and VR-related characteristics for the analysis sample.

Note. The table displays mean values (*M*) and standard deviations (*SD*) for metric variables and absolute (*n*) and relative (%) frequencies for categorial variables. F-values and p-values from ANOVA conducted to examine group differences in metric variables, and X²-values and p-values from chi-square-test, or p-values from Fisher's exact tests (FE), both conducted to examine group differences in categorial variables.

^aDegrees of freedom (1, 49).

SPIN, social phobia inventory; SIAS, social interaction anxiety scale; MPS, multimodal presence scale.

Affect (PANAS; Watson et al., 1988; German version; Breyer and Bluemke, 2016) referred to how the participants feel about the different feedback variants in retrospect. While asking participants to consider the appropriateness of the feedback variants within a therapeutic intervention (attentional training in virtual social situations), a 9-point Likert scale ranging from 1 (not at all suitable) to 9 (very suitable) was used. At the end, participants had the opportunity to note comments relating to the experiment. Table 1 displays sociodemographic, health, and presence in VR characteristics for the analysis sample of 50 participants.

2.5 Procedure

Before the start of the experiment, informed written consent was given and questionnaires on demographic data and social anxiety were filled in. To increase familiarity with the rating procedure, the SAM and its included images were shown and explained to the participants. The experiment consisted of a total of three blocks of 36 trials each. There were breaks of self-determined length between blocks.

At the beginning of the VR-scenario participants had 2 minutes to visually explore the virtual environment and familiarize themselves with the VR motion controller, which was exclusively used for providing ratings. This familiarization phase did not involve physical movement but allowed participants to look around the virtual room while seated on a chair. After familiarization with the VR environment, participants' eyes were individually calibrated for the eye tracking system to ensure optimal conditions for the gazebased feedback. Following the calibration phase, participants received further instructions *via* the display or the headphones. Within the VR-scenario participants found themselves in a virtual room, sitting in a triangle with two virtual agents around a round table (Figure 1a).

Using an introductory instruction on the HMD, for which the participants could take as much time as necessary, the task of the following experiment was explained to the participants again. The instruction was followed by an acoustic request *via* headphones to direct the gaze to the fixation cross. After 2 seconds of focusing on the cross, the actual experiment began. Gaze instructions (*"eyecontact with person on the right/left"*), feedback variants (smile, tone, praise), and duration of maintenance (short, medium, long) followed a balanced randomized trial procedure. Before the beginning of the first trial, an idle time of 6 seconds served as a baseline survey of spontaneous gaze behavior of the participants, after which the first gaze instruction began. The participants were then instructed to make and maintain eye contact with the right or left person until feedback is given. One of the three feedback variants



(smile, tone, or praise) occurred after a maintenance duration of 3.0 (short), 5.0 (medium) or 7.0 s (long). After the feedback presentation it always took 6.0 s until the next trial started. During this period the spontaneous gaze behavior of the participants was recorded again *via* eye tracking, while, gaze behavior of the virtual agents was independent of the gaze behavior of the participants, i.e., agents randomly looked around. Every trial started with an instruction to look at the fixation cross followed by the procedure described above. Per block, a total of nine ratings (online affect ratings) of valence and arousal elicited by the previous feedback trial were assessed (Figure 1b, see demo video 10.6084/m9.figshare.28914335).

After the three blocks were completed, participants were asked to fill out a questionnaire on social and physical presence. To additionally assess the affect in retrospect, for each individual feedback variant and maintenance duration valence and arousal were assessed (post training affect ratings) and the participants' positive and negative affect in regard of the feedback variants was assessed in retrospect using the PANAS. After the appropriateness of the feedback variants for further therapeutic interventions has been assessed, some general questions and comments related to the experiment were answered. The total duration of the experiment was about 65 min.

3 Statistical analyses

All statistical analyses were performed using SPSS 29 (IBM). The α -level was set to 0.05. All measures were analyzed using repeated measurement ANOVAs with *feedback variant* (three levels: smile, tone, praise) and *maintenance* (three levels: short, medium, long) as within-subject factors and *social anxiety* (two levels: LSA, HSA) as between-subject factor. In case of significant main effects of feedback variant or maintenance, pairwise comparisons were calculated. In



are indicated in the figure.

case of significant interaction effects of feedback variant x maintenance, *post hoc* t-tests (paired) were conducted, and in case of effects involving the factor *social anxiety* additional independent sample t-tests were conducted to check for group differences. Furthermore, ANOVAs were conducted separately for LSA and HSA individuals to check for significant differences between the feedback variants, and maintenance durations separately in both groups (in addition to our preregistration).

Violations of sphericity were corrected using Greenhouse-Geisser correction. If the Levine test indicated inhomogeneity of variances, corrected values for independent-sample t-test following the Welch correction were reported. For the t-tests, Cohen's *d* was calculated as effect size (interpretation: $d \ge 0.20$ as small, $d \ge 0.50$ as medium, and $d \ge 0.80$ as large). For repeated measures ANOVAs, partial eta squared served as effect size (interpretation: $\eta_p^2 > = 0.01$ as small, $\eta_p^2 > = 0.06$ as medium, and $\eta_p^2 > = 0.14$ as large).

In terms of measuring participants' attentional processes, the 6 seconds of free viewing time at the beginning of the experiment served as baseline measurement. Mean fixation duration was calculated for social and non-social environment and served as reference value (see Supplementary Text S1). For each measurement of fixation duration after receiving feedback the difference to the baseline was calculated.

4 Results

4.1 Affect ratings

4.1.1 Online valence ratings

The overall valence of the feedback variants was moderate (M = 5.87, SD = 1.32; total range: 1–9) (Supplementary Table S1). Descriptively, online valence ratings were higher for both social feedback variants (smile and praise) than for the non-social variant (tone) and were higher for LSA than for HSA individuals, while valence ratings did not show any descriptive differences across durations of maintenance (short, medium, and long) (Figures 2a,b).

A mixed ANOVA with the factors *feedback variant*, *maintenance*, and *social anxiety* resulted in a significant main effect of *social anxiety*, F(1,48) = 4.78, p = 0.034, $\eta_p^2 = 0.09$, and significant interaction effects of *feedback variant x social anxiety*, F(1.65, 79.35) = 4.86, p = 0.015, $\eta_p^2 = 0.09$, and *feedback variant x maintenance*, F(3.45, 165.78) = 2.65, p = 0.043, $\eta_p^2 = 0.05$. There were no significant main effects of feedback variant (p = 0.081) and maintenance (p = 0.803), and no significant interaction effect between *maintenance x social anxiety* (p = 0.989) (see Supplementary Table S2).



Concerning the significant main effect of *social anxiety* and to follow-up on the interaction of *feedback variant x social anxiety*, *post hoc* t-tests (independent) between HSA and LSA for each feedback variant were conducted. Results revealed significantly higher valence values in LSA than in HSA individuals for the smile, t (48) = 2.95, p = 0.005, d = 0.82, and praise, t (48) = 2.02, p = 0.049, d = 0.57, but not for the tone, t (48) = 0.47, p = 0.642, d = 0.13. Additionally, *post hoc* t-tests within LSA showed that the tone was perceived as less pleasant than the smile, t (23) = 4.34, p < 0.001, d = 0.89, and the praise, t (23) = 2.63, p = 0.015, d = 0.54, by LSA individuals. Within HSA individuals no significant differences between the feedback variants have been found (p > 0.620).

Post-hoc t-tests to follow-up on the interaction of *feedback variant x maintenance* revealed that the tone was rated more pleasant after a long than after a medium maintenance duration, t (49) = 2.55, p = 0.014, d = 0.23. Furthermore, for medium maintenance duration of eye contact the praise was rated as more pleasant than the tone, t (49) = 3.86, p < 0.001, d = 0.57. For all other variables, no significant differences were found (Figure 3) (see Supplementary Tables S3 and S4).

In summary, LSA perceived the social feedback variants (smile and praise), but not the non-social variant (tone), as significantly more pleasant than HSA individuals. Significant differences between the smile and the tone as well as the praise and the tone were only observed within the LSA group.

4.1.2 Online arousal ratings

Overall, the online arousal values regarding the feedback during the experiment were quite low (M = 3.35, SD = 1.54; total range: 1–9) (see Supplementary Table S1). Descriptively, the feedback variants seem to differ from each other as well as LSA show lower arousal

values than HSA individuals, but not depending on the feedback variant (see Supplementary Figures S1a,b). No differences are obvious for the maintenance duration.

A mixed ANOVA with the factors *feedback variant*, *maintenance*, and *social anxiety* was conducted with respect to the arousal ratings during the experiment. There was a significant main effect of *feedback variant*, F(1.32, 96) = 20.33, p < 0.001, $\eta_p^2 = 0.30$, but no main effect of *maintenance* (p = 0.158) and *social anxiety* (p = 0.355), and no interaction effect between *feedback variant x maintenance*, *feedback variant x social anxiety*, *maintenance x social anxiety*, and *feedback variant x maintenance x social anxiety* (all ps > 0.390) (Supplementary Table S2).

In regard of the main effect of *feedback variant*, *post hoc* pairwise comparisons showed significantly greater arousal elicited by the smile compared to the praise, t (48) = 13.57, p < 0.001, d = 0.68, and the tone, t (48) = 4.30, p < 0.001, d = 0.61. Between the praise and the tone, there was no significant difference (p = 0.458).

In summary, arousal was rated lowest for the tone, higher for the praise and highest for the smile. The maintenance duration of eyecontact as well as group membership (LSA vs HSA) did not influence arousal ratings.

4.2 Post training affect ratings

The following results refer to the SAM-Ratings assessed after the intervention (post training affect ratings). In contrast to the online affect ratings, the individual feedback variants and maintenance durations were prompted and assessed separately. Therefore, these factors will be analyzed separately (no interactions of the two factors can be analyzed).

4.2.1 Post training valence ratings

The valence ratings at post training were moderate (M = 6.19, SD = 1.54, total range: 1–9) (see Supplementary Table S5). Figures 2c,d display that valence ratings descriptively differ both between the feedback variants and the maintenance durations. In regard of the social feedback variants (smile, and praise) the valence values of LSA tend to be descriptively higher than those of HSA individuals.

A mixed ANOVA with the factors *feedback variant* and *social anxiety* was conducted. Results revealed a significant effect of *feedback variant*, F(2,98) = 5.90, p = 0.004, $\eta_p^2 = 0.12$, but no significant main effect of *social anxiety* (p = 0.162). The interaction between *feedback variant x social anxiety* missed significance (p = 0.063) (Supplementary Table S2). Pairwise comparisons were calculated to follow up on the main effect of *feedback variant*. The smile and the praise were both rated as more pleasant than the tone, t(49) = 2.67, p = 0.005, d = 0.37; t(49) = 2.29, p = 0.10, d = 0.33, respectively. Between the smile and the praise there was no significant difference (p = 0.181).

Moreover, an ANOVA with the factor *maintenance* and *social anxiety* was conducted. The analysis revealed a significant main effect of *maintenance*, F(2,98) = 28.03, p < 0.001, $\eta_p^2 = 0.45$, but no main effect of *social anxiety* (p = 0.124), and no significant interaction between *maintenance* x *social anxiety* (p = 0.246) (Supplementary Table S2). To follow up the main effect of *maintenance*, pairwise comparisons showed that the short and the medium maintenance durations were perceived as more pleasant than the long maintenance, t (49) = 6.11, p < 0.001, d = 0.86; t (49) = 2.67, p = 0.005, d = 0.37. Furthermore, the medium maintenance duration, t (49) = 5.45, p < 0.001, d = 0.73. Further exploratory analyses are reported in Supplementary Texts S2 and S3.

In summary, we found significant differences between the feedback variants and maintenance durations independent of social anxiety. The smile was perceived as most pleasant, followed by the praise, and the tone. Feedback, which appeared after a short duration of maintenance of eye contact was perceived as most pleasant in retrospect, followed by the medium, and the long maintenance duration. In contrast to the online valence ratings no group differences were found.

4.2.2 Post training arousal ratings

The post training rated arousal elicited by the feedback variants and the maintenance durations was as well at a low level (M = 4.15, SD = 1.99; range 1–9). Independent of social anxiety, differences between the feedback variants can descriptively be recognized (see Supplementary Figures S1; Supplementary Table S5).

To check for significant differences between the feedback variants, arousal ratings were analyzed using a mixed ANOVA with the factors *feedback variant* and *social anxiety*. The ANOVA revealed a significant main effect of *feedback variant*, *F* (2, 98) = 12.86, p < 0.001, $\eta_p^2 = 0.21$, but no significant main effect of *social anxiety* (p = 0.324) or interaction between *feedback variant* x *social anxiety* (p > 0.671) (see Supplementary Table S2). Pairwise comparisons were calculated to follow up on the main effect of *feedback variant*. Results revealed higher arousal values for the smile compared to the tone, *t* (49) = 4.47, *p* < 0.001, *d* = 0.63, and the praise, *t* (49) = 3.41, *p* = 0.001, *d* = 0.48. There were no significant differences between the tone and praise (p = 0.10).

In addition, a mixed ANOVA with the factors *maintenance* and *social anxiety* was conducted. There was a significant main effect of *maintenance*, F(2,98) = 3.71, p = 0.028, $\eta_p^2 = 0.07$, but no significant main effect of *social anxiety* (p = 0.231) or *maintenance x social anxiety* interaction (p = 0.999) (see Supplementary Table S2). According to the results of the pairwise comparisons, the short maintenance duration elicited higher arousal values than the medium maintenance duration, t (49) = 3.20, p = 0.001, d = 0.45. Between the short and long maintenance duration, there were no significant differences (all ps > 0.103).

In summary, the highest arousal was elicited after receiving the smile compared to the tone and the praise. According to the participants' post training affect ratings, they indicated more emotionally arousal after a short than after a medium duration of maintaining eye contact with virtual agents.

4.3 Appropriateness for attention trainings

After the training, participants were asked to consider the appropriateness of the feedback variants within a therapeutic intervention. A mixed ANOVA of the ratings with the factors *feedback variant* and *social anxiety* was conducted and revealed a significant main effect of *feedback variant*, *F* (2, 98) = 20.84, *p* < 0.001, $\eta_p^2 = .30$, but no significant effects including the factor *social anxiety* (all *ps* > 0.590).

Paired t-tests showed higher appropriateness values for the smile in comparison to the tone, t (49) = 5.51, p < 0.001, d = 0.81, and the praise, t (49) = 2.60, p = 0.012, d = 0.36. Also, the praise has been considered as more appropriate than the tone, t (49) = 4.16, p < 0.001, d = 0.58 (see Supplementary Table S6).

In summary, the participants (independent of social anxiety) would prefer the social feedback variants for therapeutic interventions, with smiling being rated most appropriately.

4.4 Positive and negative affect

We additionally examined differences in participants' positive and negative affective responses in regard to the different feedback variants after the experiment. To check for significant differences between the feedback variants and between HSA and LSA, a mixed ANOVA with the factors *feedback variant* and *social anxiety* was conducted for positive affect (PA) and negative affect (NA).

The ANOVA for PA revealed a significant main effect of *feedback variant*, *F* (2, 98) = 3.30, p = 0.041, $\eta_p^2 = 0.06$. No significant effects including the factor *social anxiety* were found (all *ps* > 0.641). To follow up on the main effect of *feedback variant*, *post hoc* t-tests were conducted. For PA paired t-tests showed higher values for the smile than for the tone, *t* (49) = 2.11, *p* = 0.020, *d* = 0.30. Between the tone and the praise, there were no significant differences (*p* > 0.430).

The ANOVA for NA revealed a significant main effect of *feedback variant*, *F* (2, 98) = 4.76, *p* = 0.011, η_p^2 = 0.09, and a significant interaction between *feedback variant x social anxiety*, *F* (2, 98) = 3.35, *p* = 0.039, η_p^2 = 0.06, but no significant main effect of *social anxiety* (*p* = 0.311).



To follow up on the interaction effect, independent sample t-tests were conducted. Results revealed a significant difference between LSA and HSA individuals regarding the smile feedback, t (48) = 1.89, p = 0.030, d = 0.53, with higher negative affect values for the smile within HSA. Groups did not differ significantly from each other in regard of the other feedback variants (all ps > 0.340) (see Supplementary Table S7).

In summary, PA in relation to the feedback variants was rated highest for the smile followed by the praise, and the tone. Interestingly, for the smile feedback HSA reported higher NA than LSA individuals.

4.5 Attentional processes

To investigate whether there was an increase in focused attention on virtual agents' eye-areas after receiving feedback, an ANOVA for fixation duration with the factors *feedback variant* and *social anxiety* revealed a significant effect of *feedback variant*, *F* (2,92) = 5.58, p = 0.007, $\eta_p^2 = 0.11$, but no significant effect of *social anxiety* and no significant interaction between *feedback variant x social anxiety* (all ps > 0.650). To follow up on the main effect of *feedback variant*, *post hoc* t-tests were conducted, which showed that receiving the smile after maintaining eye contact elicited more eye contact with virtual agents directly after the feedback than the tone and praise, t (47) = 2.84, p = 0.007, d = 0.41; t (47) = 2.83, p = 0.007, d = 0.41. There were no significant differences between the tone and the praise (p = 0.826) (Figure 4).

5 Discussion

Within this study, we investigated which characteristics of eyetracking-based feedback of attentional processes lead to a positive valence and to an increase in focused attention on social stimuli. Additionally, it was of interest whether the effects differ between individuals with low and higher social anxiety. The sample, which was divided into two groups regarding social anxiety, had to engage in a social interaction with virtual agents by making eye contact and maintaining that eye contact for a pre-specified amount of time. By constantly directing the focus of attention to the eye region of the virtual agent, feedback was automatically triggered. The results of this study inform about attentional processes and serve as indication to choose an eligible feedback variant for further implementing an attentional training with eye-tracking-based feedback of attentional processes in VR for people with SAD.

5.1 Summary and interpretation of main results

The results of the main outcome, the online valence ratings, revealed a significant interaction between feedback variant and social anxiety. Therefore, LSA individuals perceived the social feedback variants (smile, and praise), but not the non-social variant (tone) as more pleasant than HSA. Thus, the valence values for the smile, and the praise as feedback were higher for LSA than HSA individuals. The lower valence values of HSA individuals regarding the social feedback variants are in line with the affective-reactivity-hypothesis, which claims that socially anxious individuals, compared to individuals with low social anxiety, respond more negatively to positive stimuli (e.g., smiles) and positive events (e.g., positive feedback) (Alden et al., 2008; Gilboa-Schechtman et al., 2014; Gilboa-Schechtman and Shachar-Lavie, 2013). These findings can also be supported by the results of our exploratory analyses of the post training valence ratings (see Supplementary Text S2) as well as the results of the PANAS questionnaire on NA. Here, the smile feedback was perceived as

more pleasant by LSA compared to HSA individuals, whose NA scores were higher regarding the smile feedback. Reduced pleasantness of social feedback variants in socially anxious individuals is consistent with previous research on social anxiety (e.g., Clark and Wells, 1995; Gilboa-Schechtman et al., 2014; Hayward et al., 2018), however our results were obtained in a different setting, feedback was triggered by one's own behavior and was intended to serve as a positive reinforcer. Finally, these results are also in line with the assumption of fear of positive evaluation in socially anxious individuals (Kashdan and Collins, 2010; Weeks and Howell, 2014) in addition to the typical characteristic of fear of negative evaluation in social anxiety (see Clark and Wells, 1995). Assuming that both the smile and the praise could be interpreted more as positive social evaluation than a tone, then the social feedback variants could be associated with more distress for HSA individuals. Thus, fear of positive evaluation should also be considered when selecting an eligible feedback variant for attentional training for individuals with SAD.

In addition to the observed interaction effect between feedback variant and social anxiety, follow-up analyses revealed that significant differences between feedback variants (smile and tone, praise and tone) were only present within the LSA group, while no such differentiation emerged within the HSA group. Given that the LSA group can be considered a normative baseline in terms of emotional processing of feedback, the lack of variation in the HSA group is particularly informative. These findings suggest that individuals with high social anxiety may perceive feedback as generally less pleasant in the context of socially demanding situations and/or may feel generally observed and scrutinized by the feedback itself, which leads to increased discomfort (e.g., Falkai, 2018; Gilboa-Schechtman et al., 2014). This uniformity in valence ratings within HSA individuals highlights the importance of carefully selecting feedback variants for attentional training in this population. Therefore, future research is needed, on the one hand, to systematically examine a broader range of feedback variants (e.g., positive feedback-symbol combined with tone and praise) within HSA individuals, and on the other hand, to examine whether the same undifferentiated rating pattern can be confirmed in this group.

While the online valence ratings revealed no main effect of feedback variant, such an effect could be found in the post training valence ratings. The ratings conducted after the experiment may reflect the overall impression of the participants with regard to the feedback variants and could be affected by request, social desirability, or memory biases. Nevertheless, it should be mentioned that in line with our expectations, the smile feedback was perceived as most pleasant, followed by the praise, and finally the tone. While smile and praise are clear positive social reinforcers (Kircher, 2019; Wittchen and Hoyer, 2011), the perception and evaluation of tones can differ between individuals (e.g., Ansorge and Leder, 2011). Looking at the range (2-8) of the valence ratings of the tone, interpersonal differences can be assumed. While some participants described the tone as annoying and monotonous, others described it as clearly pleasant. Accordingly, selecting a tone that also addresses both genders equally is a major challenge. Based on the assumption that a positive tone could be used very well for attentional trainings due to its conciseness (Kopp and Mandl, 2014) and that it would represent a valid alternative for social feedback variants, future research should take a closer look at testing further tones. The preference for smile feedback may also be explained by differences in modality and source localization. While the smile was directly displayed on the virtual agent and thus perceived as an immediate social response, the praise and the tone were presented *via* headphones from an external source, with no link to the agent. This lack of spatial anchoring may have reduced their perceived valence, especially in a VR setting focused on visual attention.

As we are interested in the characteristics of eye-tracking-based feedback, we should also consider the results concerning the duration of maintenance of eye contact, which determines the timing of the feedback. According to the online affect ratings, contrary to our expectations participants did not notice any differences between the short, medium, and long duration of maintaining eye contact. In contrast, the exploratory retrospective evaluation of the participants showed that maintaining eye contact with virtual agents for a short duration, respectively receiving feedback after 3.0 s, was perceived as more pleasant than the medium and long duration of maintenance. The latter was associated with the lowest valence values. The fact that receiving feedback after a short duration of maintaining eye contact was perceived as most pleasant is in line with previous research, which claims that immediate feedback leads to greater certainty and sense of agency compared to long-delayed feedback (Wen, 2019). Especially for socially anxious individuals, waiting for feedback could lead to uncertainty and doubts as to whether the task has been carried out correctly. Due to the pronounced avoidance behavior (e.g., Clark and Wells, 1995) and the difficulty in making and maintaining eye contact in socially anxious individuals (e.g., Hessels, 2018), it is plausible that the long duration of maintaining eye contact was associated with increased difficulty and consequently discomfort (Kopp and Mandl, 2014). Of course, the question arises at this point why no differences between the maintenance durations were found in the online affect ratings. While the post training valence ratings explicitly assessed the individual durations of maintaining eye contact with virtual agents, the online ratings were a more implicit evaluation of the durations, as the participants were not explicitly asked about. It can therefore be assumed that the participants did not notice any major differences in the duration during maintaining eye contact with virtual agents. Since there are also no differences between HSA and LSA individuals, the duration of maintaining eye contact could be chosen flexibly with regard to the exposure rationale.

Furthermore, we investigated which feedback variants lead to an increase in focused attention on virtual agents. Even though there were no differences between LSA and HSA, we found significant differences between the feedback variants. Consequently, and in line with our hypotheses, participants showed significantly more eye contact with virtual agents directly after receiving the smile compared to the praise and the tone. As the smile feedback is visual and came directly from the virtual agent, this feature could have led to an increase in focused attention compared to the auditory variants, which were sounded from the off. Because attention towards social cues is a main goal of the treatment of social anxiety, implementing feedback associated with the target might be helpful. The attention towards the eye regions of virtual agents was very similar between HSA and LSA. Here, we assume that the lack of differences between HSA and LSA individuals could be explained by the learning rate and change of expectancy during the experiment. The total of 108 trials always followed the same pattern, so that the participants could relatively quickly learn that nothing serious happens either during the eye contact or the free viewing period of 6 seconds.

On the basis that we are looking for an eligible feedback variant for subsequent attentional training in VR for people with SAD, our participants were asked to evaluate the appropriateness of the feedback variants for such training in the context of a performance situation (giving a speech). In accordance with the (post training) valence ratings, most participants preferred the smile feedback. Nevertheless, this result should be viewed with reservation and only as an initial estimation. Our participants became familiar with the feedback variants in the context of a social interaction with making eye contact, and only had to imagine an attentional training in the context of a speech to evaluate the suitability of the feedback variants. Even if we are looking for a feedback variant that is suitable for virtual attentional training in various situations, it should be noted that feedback in our paradigm could have had a different effect than, for example, during a speech.

5.2 Limitations

As one important limitation, technical problems with the eye tracker in a total of six participants should be mentioned. Although these participants made eye contact with the instructed agent and did not avoid it, the fixation measured via eye tracking was sometimes not immediately recognized (M = 3.5 times per block), so that the process of making eye contact and receiving feedback consequently took longer. Based on the results of the online valence ratings on the interaction between feedback variant and maintenance, it can be assumed that the evaluation of the feedback variants was largely unaffected by this problem (see Figure 3). In contrast, the perceived duration of maintaining eye contact of these participants could have been influenced by this limitation, especially the long duration of 7.0 s could have been perceived as even longer. In summary, we assume that the predefined area (eye region of the virtual agent up to the eyebrows and nose) was defined a little too strictly, so it could make sense, to extend the hitbox surrounding the eye regions of the virtual agents to the entire head in subsequent experiments, so that the gaze of the participants is recognized as such without error.

Secondly, when participants averted their gaze during the fixation period, they were instructed (*via* headphones) to not look away and to fixate on the agent again (see Supplementary Text S6). Such information can lead to insecurity, especially for HSA individuals, or even be perceived as punishment. Since this study is investigating approach feedback, this type of avoidance feedback could also be omitted in future experiments.

Furthermore, it should be mentioned that we investigated the characteristics (visual/auditory, social/non-social, verbal/non-verbal) of feedback as a primary research question, but the division of feedback features did not include all variants. While there was an auditory social (praise) and auditory non-social (tone) feedback variant, we used a visual social (smile) variant, while a

visual non-social feedback variant was missing. Consequently, one should consider whether a visual non-social feedback variant (e.g., a thumbs-up symbol) or even combinations of visual and auditory feedback (e.g., smile with tone) should be tested again.

Additionally, our secondary research question aimed to determine whether there are differences in the effects between LSA and HSA individuals. Since we examined a non-clinical sample, we divided them into two groups by median split (median = 17) and interpreted the results accordingly. With a total score of 0–68, the authors of the SPIN consider a score of 19 to be most suitable for distinguishing between persons with and without SAD (Stangier and Steffens, 2002). While we were able to investigate a sample of variable and also higher social anxiety, we cannot generalize the results to patients having SAD. Therefore, the next step should be to apply the feedback variants to a sample diagnosed with SAD.

Finally, the duration and effort of the task should be mentioned. Looking at the overall arousal ratings, these were at a rather low level. Participants had to make eye contact with virtual agents a total of 108 times, so monotony and boredom as influencing factors could not be excluded. However, if we consider our exploratory analyses, which show no differences in the valence ratings between the three blocks (see Supplementary Text S5), we can assume that the valence ratings (primary outcome) remained stable across the blocks and consequently, the feedback did not lose its reinforcing character. Furthermore, for the further use of the feedback in attentional training, e.g., as part of a speech scenario, boredom should be less relevant.

5.3 Conclusion

This study compared effects of feedback variants on affect ratings and attention. We could confirm that the social feedback variants (smile, and praise), but not the non-social variant (tone) were perceived as more pleasant by LSA compared to HSA individuals. However, no significant differences between the feedback variants were observed within the HSA group regarding perceived valence, which was rated moderately, ranging from neutral to positive. Therefore, we identified eligible feedback variants for the further development of attentional training in VR for the treatment of individuals with speech anxiety and SAD, comprising of gaze interaction elements and measuring of overt attention *via* eye tracking. Using gaze-related feedback, exposure towards social stimuli will be promoted and external attention focus in social anxiety will be enhanced.

Data availability statement

The raw data supporting the conclusions of this article are accessible in a public repository (https://doi.org/10.17605/OSF. IO/PRB3Z).

Ethics statement

The studies involving humans were approved by Ethics committee of the University of Regensburg (Reference number:

23-3376-101). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

TS-P: Writing – original draft, Writing – review and editing. TW: Writing – original draft, Writing – review and editing. LK: Writing – original draft, Writing – review and editing. AM: Writing – original draft, Writing – review and editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. TS's PhD project is financially supported by the Friedrich Naumann Foundation for Freedom. Otherwise, the author(s) received no financial support for the research, authorship, and/or publication of this article.

Acknowledgments

We thank Lena Altmann for her help in data acquisition as well as Alexander May and Andreas Ruider for their help in developing the VR environment, the VR scripts, and the integration of eye tracking.

References

Alden, L. E., Taylor, C. T., Mellings, T. M. J. B., and Laposa, J. M. (2008). Social anxiety and the interpretation of positive social events. *J. Anxiety Disord.* 22 (4), 577–590. doi:10.1016/j.janxdis.2007.05.007

Altmeyer, M., Hnatovskiy, V., Rogers, K., Lessel, P., and Nacke, L. E. (2022). Here comes No boom! The lack of sound feedback effects on performance and user experience in a gamified image classification task. *Proceedings of the CHI Conference on Human Factors in Computing Systems* 1–14. doi:10.1145/3491102.3517581

Ansorge, U., and Leder, H. (2011). "Auditive wahrnehmung," in *Wahrnehmung und Aufmerksamkeit*. Editors U. Ansorge and H. Leder (VS Verlag für Sozialwissenschaften), 119–134. doi:10.1007/978-3-531-92094-8_9

Arch, J. J., and Craske, M. G. (2009). First-line treatment: a critical appraisal of cognitive behavioral therapy developments and alternatives. *Psychiatric Clin. N. Am.* 32 (3), 525–547. doi:10.1016/j.psc.2009.05.001

Bandelow, B., Aden, I., Alpers, G. W., Benecke, A., Benecke, C., Beutel, M. E., et al. (2021). Deutsche S3-Leitlinie Behandlung von Angststörungen. Version 2. Available online at: www.awmf.org/leitlinien/detail/II/051-028.html (Accessed January 10, 2025).

Bögels, S. M., and Mansell, W. (2004). Attention processes in the maintenance and treatment of social phobia: hypervigilance, avoidance and self-focused attention. *Clin. Psychol. Rev.* 24 (7), 827–856. Article 7. doi:10.1016/j.cpr.2004.06.005

Bradley, M. M., and Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. J. Behav. Ther. Exp. Psychiatry 25 (1), 49–59. doi:10.1016/0005-7916(94)90063-9

Breyer, B., and Bluemke, M. (2016). *Deutsche version der positive and negative affect schedule PANAS (GESIS panel)*. Mannheim, Germany: Zusammenstellung sozialwissenschaftlicher Items Skalen (ZIS). doi:10.6102/ZIS242

Caponnetto, P., Triscari, S., Maglia, M., and Quattropani, M. C. (2021). The simulation game—virtual reality therapy for the treatment of social anxiety disorder: a systematic review. *Int. J. Environ. Res. Public Health* 18 (24), 13209. doi:10.3390/ ijerph182413209

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

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frvir.2025.1556898/ full#supplementary-material

Chesham, R. K., Malouff, J. M., and Schutte, N. S. (2018). Meta-analysis of the efficacy of virtual reality exposure therapy for social anxiety. *Behav. Change* 35 (3), 152–166. doi:10.1017/bec.2018.15

Clark, D. M., and Wells, A. (1995). A cognitive model of social phobia. In Social Phobia: diagnosis, assessment, and treatment. New York, NY: Guilford Press, 69–93.

Clark and Wells. (1995a). Social phobia: diagnosis, assessment and treatment (S. 69-93). New York: Guilford Press.

Clark and Wells. (1995b). A cognitive model of social Phobia. In Editors, Heimberg, R., Liebowitz, M., Hope, D. A., and Schneier, F. R. (Hrsg.). New York: Guilford Press.

Connor, K. M., Davidson, J. R. T., Churchill, L. E., Sherwood, A., Weisler, R. H., and Foa, E. (2000). Psychometric properties of the social phobia inventory (SPIN): new self-rating scale. *Br. J. Psychiatry* 176 (4), 379–386. doi:10.1192/bjp.176.4.379

Craske, M. G., Kircanski, K., Zelikowsky, M., Mystkowski, J., Chowdhury, N., and Baker, A. (2008). Optimizing inhibitory learning during exposure therapy. *Behav. Res. Ther.* 46 (1), 5–27. doi:10.1016/j.brat.2007.10.003

Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T., and Vervliet, B. (2014). Maximizing exposure therapy: an inhibitory learning approach. *Behav. Res. Ther.* 58, 10–23. doi:10.1016/j.brat.2014.04.006

Del Giacco, L., Anguera, M. T., and Salcuni, S. (2020). The action of verbal and nonverbal communication in the therapeutic alliance construction: a mixed methods approach to assess the initial interactions with depressed patients. *Front. Psychol.* 11, 234. doi:10.3389/fpsyg.2020.00234

Ekman, P., Davidson, R. J., and Friesen, W. V. (1990). The Duchenne smile: emotional expression and brain physiology: II. *J. Personality Soc. Psychol.* 58 (2), 342–353. doi:10.1037/0022-3514.58.2.342

Emmelkamp, P. M. G., Meyerbröker, K., and Morina, N. (2020). Virtual reality therapy in social anxiety disorder. *Curr. Psychiatry Rep.* 22 (7), 32. doi:10.1007/s11920-020-01156-1

Falkai, P. (2018). Diagnostisches und statistisches manual psychischer Störungen DSM-5[®]. Editor T. Kircher, H.-U. Wittchen, M. Döpfner, W. Gaebel, W. Maier, and W. Rief, (Hrsg.; 2. Aufl.). Hogrefe. doi:10.1026/02803-000 Fehm, L., Pelissolo, A., Furmark, T., and Wittchen, H.-U. (2005). Size and burden of social phobia in Europe. *Eur. Neuropsychopharmacol.* 15 (4), 453–462. doi:10.1016/j. euroneuro.2005.04.002

Feiler, A. R., and Powell, D. M. (2016). The role of self-focused attention and negative self-thought in interview anxiety: a test of two interventions. *Int. J. Sel. Assess.* 24 (2), 132–149. doi:10.1111/ijsa.12136

Gilboa-Schechtman, E., Shachar, I., and Helpman, L. (2014). "Evolutionary perspective on social anxiety," in *Social anxiety* (San Diego, CA, USA: Elsevier), 599–622. doi:10.1016/B978-0-12-394427-6.00021-2

Gilboa-Schechtman, E., and Shachar-Lavie, I. (2013). More than a face: a unified theoretical perspective on nonverbal social cue processing in social anxiety. *Front. Hum. Neurosci.* 7, 904. doi:10.3389/fnhum.2013.00904

Hattie, J., and Timperley, H. (2007). The power of feedback. *Rev. Educ. Res.* 77 (1), 81-112. doi:10.3102/003465430298487

Hayward, D. A., Pereira, E. J., Otto, A. R., and Ristic, J. (2018). Smile! Social reward drives attention. J. Exp. Psychol. Hum. Percept. Perform. 44 (2), 206–214. doi:10.1037/ xhp0000459

Heerey, E. A., and Kring, A. M. (2007). Interpersonal consequences of social anxiety. *Journal of Abnormal Psychology* 116 (1), 125–134. doi:10.1037/0021-843X.116.1.125

Heimberg, R. G., Brozovich, F. A., and Rapee, R. M. (2010). "A cognitive behavioral model of social anxiety disorder," in *Social anxiety* (San Diego, CA, USA: Elsevier), 395–422. doi:10.1016/B978-0-12-375096-9.00015-8

Heinrichs, N., Rapee, R. M., Alden, L. A., Bögels, S., Hofmann, S. G., Oh, K. J., et al. (2002). Kognitive aspekte sozialer angst: die social interaction anxxiety scale (SIAS) und die social phobia scale (SPS). *Diagnostica* 48 (4), 253–259. doi:10.1026//0012-1924.48. 4.253

Hessels, R. S. (2018). How social interactions influence gaze behavior: A review. Infant Behavior and Development 52, 62–80. doi:10.1016/j.infbeh.2018.05.005

Hessels, R. S., Holleman, G. A., Cornelissen, T. H. W., Hooge, I. T. C., and Kemner, C. (2018). Eye-contact takes two – autistic and social anxiety traits predict gaze behavior in dyadic interaction. *J. Exp. Psychopathol.* 9 (2), jep.062917. doi:10.5127/jep.062917

Hofmann, S. G. (2007). Cognitive factors that maintain social anxiety disorder: a comprehensive model and its treatment implications. *Cogn. Behav. Ther.* 36 (4), 193–209. doi:10.1080/16506070701421313

Horigome, T., Kurokawa, S., Sawada, K., Kudo, S., Shiga, K., Mimura, M., et al. (2020). Virtual reality exposure therapy for social anxiety disorder: a systematic review and meta-analysis. *Psychol. Med.* 50 (15), 2487–2497. doi:10.1017/S0033291720003785

Horley, K., Williams, L. M., Gonsalvez, C., and Gordon, E. (2003). Social phobics do not see eye to eye. J. Anxiety Disord. 17 (1), 33–44. doi:10.1016/S0887-6185(02)00180-9

Kampmann, I. L., Emmelkamp, P. M. G., and Morina, N. (2016). Meta-analysis of technology-assisted interventions for social anxiety disorder. *J. Anxiety Disord.* 42, 71–84. doi:10.1016/j.janxdis.2016.06.007

Kashdan, T. B., and Collins, R. L. (2010). Social anxiety and the experience of positive emotion and anger in everyday life: an ecological momentary assessment approach. *Anxiety, Stress and Coping* 23 (3), 259–272. doi:10.1080/10615800802641950

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., and Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the national comorbidity survey replication. *Archives General Psychiatry* 62 (6), 593. doi:10. 1001/archpsyc.62.6.593

Kessler, R. C., Petukhova, M., Sampson, N. A., Zaslavsky, A. M., and Wittchen, H.-U. (2012). Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States: anxiety and mood disorders in the United States. *Int. J. Methods Psychiatric Res.* 21 (3), 169–184. doi:10.1002/mpr.1359

Kircher, T. (2019). "Grundlagen," in *Hrsg., Kompendium der Psychotherapie.* Editor T. Kircher (Berlin Heidelberg, Germany: Springer), 5–16. doi:10.1007/978-3-662-57287-0_2

Kohls, G., Peltzer, J., Herpertz-Dahlmann, B., and Konrad, K. (2009). Differential effects of social and non-social reward on response inhibition in children and adolescents. *Dev. Sci.* 12 (4), 614–625. doi:10.1111/j.1467-7687.2009.00816.x

Kopp, B., and Mandl, H. (2014). "Aspekte der Feedbacknachricht," in *Feedback und rückmeldungen: theoretische Grundlagen, empirische Befunde, praktische Anwendungsfelder*. Editors H. Ditton and A. Müller (Münster, Germany: Waxmann Verlag), 151–158.

Kraut, R. E., and Johnston, R. E. (1979). Social and emotional messages of smiling: an ethological approach. *J. Personality Soc. Psychol.* 37 (9), 1539–1553. doi:10.1037/0022-3514.37.9.1539

Kroczek, L. O. H., and Mühlberger, A. (2023). Time to smile: how onset asynchronies between reciprocal facial expressions influence the experience of responsiveness of a virtual agent. *J. Nonverbal Behav.* 47 (3), 345–360. doi:10.1007/s10919-023-00430-z

Lattal, K. A. (2010). Delayed reinforcement of operant behavior. J. Exp. Analysis Behav. 93 (1), 129–139. doi:10.1901/jeab.2010.93-129

Linderkamp, F. (2009). Operante methoden. editors S. Schneider and J. Margraf (Hrsg.), *Lehrbuch der Verhaltenstherapie*, Berlin Heidelberg, Germany: Springer. 209–220. doi:10.1007/978-3-540-79545-2_13

Liu, J., Xu, J., Wu, Z., Cheng, Y., Gou, Y., and Ridolfo, J. (2021). Soundscape preference of urban residents in China in the post-pandemic era. *Front. Psychol.* 12, 750421. doi:10.3389/fpsyg.2021.750421

Makransky, G., Lilleholt, L., and Aaby, A. (2017). Development and validation of the Multimodal Presence Scale for virtual reality environments: a confirmatory factor analysis and item response theory approach. *Comput. Hum. Behav.* 72, 276–285. doi:10. 1016/j.chb.2017.02.066

Mattick, R. P., and Clarke, J. C. (1998). Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. *Behav. Res. Ther.* 36 (4), 455–470. doi:10.1016/S0005-7967(97)10031-6

Matyjek, M., Meliss, S., Dziobek, I., and Murayama, K. (2020). A multidimensional view on social and non-social rewards. *Front. Psychiatry* 11, 818. doi:10.3389/fpsyt.2020. 00818

Mayo-Wilson, E., Dias, S., Mavranezouli, I., Kew, K., Clark, D. M., Ades, A. E., et al. (2014). Psychological and pharmacological interventions for social anxiety disorder in adults: a systematic review and network meta-analysis. *Lancet Psychiatry* 1 (5), 368–376. doi:10.1016/S2215-0366(14)70329-3

Mogg, K., Mathews, A., and Weinman, J. (1987). Memory bias in clinical anxiety. J. Abnorm. Psychol. 96 (2), 94–98. doi:10.1037/0021-843X.96.2.94

Moukheiber, A., Rautureau, G., Perez-Diaz, F., Soussignan, R., Dubal, S., Jouvent, R., et al. (2010). Gaze avoidance in social phobia: objective measure and correlates. *Behav. Res. Ther.* 48 (2), 147–151. doi:10.1016/j.brat.2009.09.012

Naal-Ruiz, N. E., Alonso-Valerdi, L. M., Ibarra-Zarate, D. I., Serrano-Cena, A., and Navas-Reascos, G. (2022). Mexican validation of the International Affective Digitized Sounds second edition (IADS-2) and additional sounds. *Sci. Rep.* 12 (1), 21824. doi:10. 1038/s41598-022-26320-w

National Institute for Health and Care Excellence (2017). Social anxiety disorder: recognition, assessment and treatment (CG159). Available online at: https://www.nice.org.uk/guidance/cg159.

Opitz, B., Ferdinand, N. K., and Mecklinger, A. (2011). Timing matters: the impact of immediate and delayed feedback on artificial language learning. *Front. Hum. Neurosci.* 5, 8. doi:10.3389/fnhum.2011.00008

Pearce, J. M., and Mackintosh, N. J. (2010). Two-process theory: A framework for understanding attention in learning. *Attention and associative learning: From brain to behaviour*. Editor C. J. Mitchell and M. E. Le Pelley (Oxford University Press). 41–70.

Pittig, A., Van Den Berg, L., and Vervliet, B. (2016). The key role of extinction learning in anxiety disorders: behavioral strategies to enhance exposure-based treatments. *Curr. Opin. Psychiatry* 29 (1), 39–47. doi:10.1097/YCO.00000000000220

Radke, S., Roelofs, K., and De Bruijn, E. R. A. (2013). Acting on anger: social anxiety modulates approach-avoidance tendencies after oxytocin administration. *Psychol. Sci.* 24 (8), 1573–1578. doi:10.1177/0956797612472682

Rapee, R. M., and Heimberg, R. G. (1997). A cognitive-behavioral model of anxiety in social phobia. *Behav. Res. Ther.* 35 (8), 741–756. doi:10.1016/S0005-7967(97)00022-3

Rubel, J., Ehrlich, T., and Lutz, W. (2015). Feedback-Ansätze in der Psychotherapie. Psychother. Im. Dialog 16 (04), 30-34. doi:10.1055/s-0041-105247

Schäfer, T., Sedlmeier, P., Städtler, C., and Huron, D. (2013). The psychological functions of music listening. *Front. Psychol.* 4, 511. doi:10.3389/fpsyg.2013.00511

Skinner, B. F., Skinner, B. F., and Skinner, B. F. (1974). Die Funktion der Verstärkungin der Verhaltenswissenschaft: Contigencies of reinforcement (Studienausgabe). Seattle, Washington, USA: Kindler.

Skipper, Y., and Douglas, K. (2012). Is no praise good praise? Effects of positive feedback on children's and university students' responses to subsequent failures. *Br. J. Educ. Psychol.* 82 (2), 327–339. doi:10.1111/j.2044-8279.2011.02028.x

Smith, T. A., and Kimball, D. R. (2010). Learning from feedback: spacing and the delay-retention effect. J. Exp. Psychol. Learn. Mem. Cognition 36 (1), 80–95. doi:10. 1037/a0017407

Stangier, U., and Steffens, M. (2002). Social phobia inventory (SPIN) - deutsche fassung. Main: Psychologisches Institut der Universität Frankfurt am.

Volkmann, T., Wessel, D., Jochems, N., and Franke, T. (2018). German Translation of the Multimodal Presence Scale. *In Mensch und Computer 2018 - Workshopband*. Bonn: Gesellschaft für Informatik e.V. 1–4. doi:10.18420/MUC2018-MCI-0428

Vrijsen, J. N., Lange, W.-G., Becker, E. S., and Rinck, M. (2010). Socially anxious individuals lack unintentional mimicry. *Behav. Res. Ther.* 48 (6), 561–564. doi:10.1016/j. brat.2010.02.004

Watson, D., Clark, L. A., and Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *J. Personality Soc. Psychol.* 54 (6), 1063–1070. doi:10.1037/0022-3514.54.6.1063

Wechsler, T. F., Kocur, M., Schumacher, S., Rubenbauer, M., Ruider, A., Brockelmann, M., et al. (2023). Looking fear in the eye: gamified virtual reality exposure towards spiders for children using attention based feedback. *Clinical Child Psychology and Psychiatry* 29(3), 1121–1136. doi:10.1177/13591045231194103

Wechsler, T. F., Kümpers, F., and Mühlberger, A. (2019). Inferiority or even superiority of virtual reality exposure therapy in phobias? a systematic review and quantitative meta-analysis on randomized controlled trials specifically comparing the

efficacy of virtual reality exposure to gold standard *in vivo* exposure in agoraphobia, specific phobia, and social phobia. *Front. Psychol.* 10, 1758. doi:10.3389/fpsyg.2019. 01758

Wechsler, T. F., Pfaller, M., Eickels, R. E. v, Schulz, L. H., and Mühlberger, A. (2021). Look at the audience? A randomized controlled study of shifting attention from selffocus to nonsocial vs. social external stimuli during virtual reality exposure to public speaking in social anxiety. *Front. Psychiatry* 12, 751272. doi:10.3389/fpsyt.2021.751272

Weeks, J. W., and Howell, A. N. (2014). "Fear of positive evaluation: the neglected fear domain in social anxiety," in J. W. Weeks (Eds.), the wiley blackwell Handbook of social anxiety disorder (1. Aufl., S. 433-453) (Hoboken, NJ: Wiley). doi:10.1002/9781118653920.ch20

Wells, A., and Papageorgiou, C. (1998). Social phobia: effects of external attention on anxiety, negative beliefs, and perspective taking. *Behav. Ther.* 29 (3), 357–370. doi:10. 1016/S0005-7894(98)80037-3

Wen, W. (2019). Does delay in feedback diminish sense of agency? A review. Conscious. Cognition 73, 102759. doi:10.1016/j.concog.2019.05.007

Wieser, M. J., Pauli, P., Alpers, G. W., and Mühlberger, A. (2009). Is eye to eyecontact really threatening and avoided in social anxiety? an eye tracking and psychophysiology study. J. Anxiety Disord. 23 (1), 93–103. doi:10.1016/j.janxdis. 2008.04.004

Wisniewski, B., Zierer, K., and Hattie, J. (2020). The power of feedback revisited: a meta-analysis of educational feedback research. *Front. Psychol.* 10, 3087. doi:10.3389/fpsyg.2019.03087

Wittchen, H.-U., and Hoyer, J. (2011). Klinische Psychologie and Psychotherapie: Mit ... 126 Tabellen (2nd ed., revised and expanded). Berlin Heidelberg, Germany: Springer.

Yang, W., Makita, K., Nakao, T., Kanayama, N., Machizawa, M. G., Sasaoka, T., et al. (2018). Affective auditory stimulus database: an expanded version of the International Affective Digitized Sounds (IADS-E). *Behav. Res. Methods* 50 (4), 1415–1429. doi:10. 3758/s13428-018-1027-6

Zbozinek, T. D., and Craske, M. G. (2017). The role of positive affect in enhancing extinction learning and exposure therapy for anxiety disorders. *J. Exp. Psychopathol.* 8 (1), 13–39. doi:10.5127/jep.052615

Zbozinek, T. D., Holmes, E. A., and Craske, M. G. (2015). The effect of positive mood induction on reducing reinstatement fear: relevance for long term outcomes of exposure therapy. *Behav. Res. Ther.* 71, 65–75. doi:10.1016/j.brat.2015.05.016