Seeking neutral: A VR-based person-identity-matching task for attentional bias modification – A randomised controlled experiment

Lichen Ma[a,*], Anne-Wil Kruijt[a], Anna-Karin Ek[b], Gustaf Åbyhammar[b], Tomas Furmark[b], Gerhard Andersson[c,d], Per Carlbring[a]

[a] Department of Psychology, Stockholm University, Frescativägen, 114 19 Stockholm, Sweden
[b] Department of Psychology, Uppsala University, P.O. Box 256, 751 05 Uppsala, Sweden
[c] Department of Clinical Neuroscience, Karolinska Institute, Solnavägen 1, 171 77 Solna, Stockholm, Sweden
[d] Department of Behavioral Sciences and Learning, Linköping University, 581 83 Linköping, Sweden

A B S T R A C T

Background: Attentional bias modification (ABM) aims to reduce anxiety by attenuating bias towards threatening information. The current study incorporated virtual reality (VR) technology and 3-dimensional stimuli with a person-identity-matching (PIM) task to evaluate the effects of a VR-based ABM training on attentional bias and anxiety symptoms.
Methods: One hundred participants with elevated social anxiety were randomised to four training groups. Attentional bias was assessed at pre- and post-training, and anxiety symptoms were assessed at pre-training, post-training, 1-week follow-up, and 3-month follow-up.
Results: Change in anxiety did not correlate with change in bias ($r = -0.08$). A repeated-measures ANOVA showed no significant difference in bias from pre- to post-ABM, or between groups. For anxiety symptoms, a linear mixed-effects model analysis revealed a significant effect of time. Participants showed reduction in anxiety score at each successive assessment ($p < .001$, Nagelkerke's pseudo $r^2 = 0.65$). However, no other significant main effect or interactions were found. A clinically significant change analysis revealed that 4% of participants were classified as 'recovered' at 3-month follow-up.
Conclusions: A single session of VR-based PIM task did not change attentional bias. The significant reduction in anxiety was not specific to active training, and the majority of participants remained clinically unchanged.

1. Introduction

Social anxiety disorder (SAD) is a common mental health problem that impairs social functioning and reduces quality of life (Barrera and Norton, 2009; Saris et al., 2017). Both psychological treatments such as cognitive behavioral therapy (Carpenter et al., 2018) and pharmacological treatments (Jakubovski et al., 2019) for SAD have shown high efficacy and can result in positive long-term outcomes (Mayo-Wilson et al., 2014). However, barriers such as inaccessibility to therapy (e.g. therapist shortage or geographical distance), high cost, long wait time, and stigmatisation can prevent SAD sufferers from seeking treatment (Hedman et al., 2016). In order to overcome these barriers, there has been an ongoing effort to develop treatment options that are accessible, effective, and acceptable for SAD patients (Heeren et al., 2015b; Lindner et al., 2017).

1.1. Attentional bias modification

Attentional bias modification (ABM) for SAD operates on the assumption that dysfunctional anxiety is caused by the preferential allocation of attention towards socially threatening information (Cisler and Koster, 2010; McNally, 2018). A number of studies have reported such an attentional bias in anxious individuals (e.g. Amir et al., 2008; Andersson et al., 2006; De Voogd et al., 2014). The rationale for ABM is that if the attentional bias underlying problematic anxiety can be attenuated via training, there will also be an associated reduction in anxiety symptoms (Bar-Haim, 2010; Koster et al., 2009; MacLeod and Mathews, 2012).

The effectiveness of ABM as a viable treatment option for anxiety disorders remains a contentious topic despite the large body of literature (see Cristea et al., 2015, 2017; Grafton et al., 2017; Kruijt and Carlbring, 2018; McNally, 2018; Mogg and Bradley, 2016). While early

---

studies in the field have reported significant reduction in attentional bias and anxiety symptoms following ABM training (e.g. Amir et al., 2008; Dandeneau et al., 2007; MacLeod et al., 2002), more recent studies have failed to replicate these results (e.g. Boettcher et al., 2013; Carlbring et al., 2012; Heeren et al., 2015a; Ma et al., 2019). Meta-analyses on ABM studies have produced conflicting findings, with some researchers concluding ABM to be an effective therapeutic tool for anxiety disorders (Linetzky et al., 2015), while others question the reliability and validity of the existing evidence supporting such a claim (Cristea et al., 2015).

1.2. Innovative ABM

Proponents of ABM maintain that anxiety symptom change cannot occur without a change in attentional bias. Therefore, the task used (the *procedure* of ABM) must successfully modify attentional bias (the *process* of ABM) in order for the training to be of any therapeutic value (Grafton and Macleod, 2016; Grafton et al., 2017). The most commonly used task in ABM research is the dot-probe task (MacLeod et al., 1986), which serves both as a measurement task and a bias modification task (by introducing a training contingency that encourages a shift in attention away from threatening information). However, in light of inconsistent results from studies using the dot-probe task as the ABM training procedure, many researchers have emphasised the need to develop and validate alternative ABM tasks (Bar-Haim, 2010; Van Bockstaele et al., 2014). One criticism against the dot-probe task is that it is very repetitive. The concern is that if the participant loses focus during the ABM training, they will be less likely to achieve bias modification and symptom reduction (Heeren et al., 2015b).

One way to improve ABM is to make the training more dynamic and engaging. Notebaert et al. (2015) developed the person-identity-matching (PIM) task based on the card game 'Snap', where participants were asked to make a judgement on whether two faces displaying the same expression (angry or happy) belong to the same individual. In the attend-happy task, participants were instructed to only focus on the identities of the happy faces to encourage attentional shift away from threat. In the attend-angry task, participants were asked to focus their attention on the angry faces instead. The PIM task also differed from traditional dot-probe tasks in that feedback of the correct response was provided to the participant. The authors reported that after ABM training using the PIM task, participants in the attend-happy condition showed reduced attentional bias towards threat compared to those in the attend-angry condition. Furthermore, participants in the attend-happy condition also showed less negative mood shift in a stressor task compared to those in the attend-angry condition.

Another potential strategy to increase task engagement is the incorporation of new technology. Urech et al. (2015) carried out a proof of concept study where ABM training was delivered inside a virtual environment. This virtual reality (VR) based ABM successfully induced a shift in attentional bias, along with a reduction in anxiety. VR-based treatment provides a great deal of control, since the experimenter can modify the therapy environment and how the stimuli are presented at will. The highly controlled environment also ensures consistent delivery of the treatment. The immersive nature of the VR environment can potentially increase task engagement and ecological validity. Furthermore, if a VR-based treatment can achieve equal clinical outcomes as face-to-face therapy, the non-reliance on clinician coupled with increasing accessibility of VR programs could mean wider distribution and lower costs compared to treatment at a clinic (Lindner et al., 2017).

1.3. Current study

The aim of the current study was to test the effectiveness of a single-session, VR-based PIM task in reducing attentional bias and social anxiety in participants with elevated trait anxiety recruited from the general population. Participants were randomly assigned to one of four experimental groups undergoing PIM training towards either neutral or disgust, with either 2D or 3D stimuli. Attentional bias was measured pre- and post-training using a dot-probe task. Self-reported anxiety symptoms were assessed at pre-training, immediately post-training, at 1-week follow-up, and at 3-month follow-up. We hypothesised that, at post-training and follow-up assessments, (i) participants in the neutral PIM groups would have lower attentional bias and anxiety scores compared to those in the disgust PIM groups; (ii) participants who received training with 3D stimuli would have lower attentional bias and anxiety scores compared to those who received training with 2D stimuli.

2. Method

2.1. Participants

One hundred participants were recruited from the general population between June and October 2017. The study was advertised on websites, newspapers, and national radio. Potential participants were directed to visit the study website iTerapi (Vlaescu et al., 2016), where they could learn more about the study and register an account to be screened for eligibility.

Inclusion criteria were: (i) score 30 or above on the Liebowitz Social Anxiety Scale, self-report (LSAS-SR), indicating probable SAD (Rytwinski et al., 2009); (ii) normal depth perception; (iii) fluent Swedish speaker; and (iv) at least 18 years of age. Exclusion criteria were: (i) any psychological treatment/counselling within the past 90 days; (ii) any change in psychopharmacological medication within the past 90 days (with the exception of as-needed medications such as beta-blockers); and (iii) Depression and suicidal ideation (as indicated by a total score of 14 or higher, and/or a score greater than 0 on the suicide item of the Patient Health Questionnaire (PHQ-9; Kroenke et al., 2001)). The study was approved by the Regional Ethical Review Board in Stockholm, Sweden.

### 2.2. Self-reported measures

The primary outcome measure was social anxiety assessed by the LSAS-SR (Fresco et al., 2001). The LSAS-SR is a 24-item questionnaire that taps into two dimensions of social anxiety: performance anxiety (13 items) and social situations (11 items). Participants first indicated how much fear is associated with the situation described by each item using a 4-point Likert scale. The same 24 items were rated again to indicate how much avoidance is associated with each situation. The LSAS-SR has demonstrated good test-retest reliability, structural validity, and internal consistency (Baker et al., 2002).

Secondary outcome measures included the Patient Health Questionnaire (PHQ-9; Kroenke et al., 2001) for depression, Generalised Anxiety Disorder 7-item scale (GAD-7; Spitzer et al., 2006), the Difficulties in Emotion Regulation Scale-16 (DERS-16; Bjureberg et al., 2016), and Brunnsviken Brief Quality of Life Inventory (BBQ; Lindner et al., 2016). All measures were in Swedish. The DERS-16 and BBQ were originally developed in Swedish. Translated versions of the LSAS-SR, PHQ-9, and GAD-7 have all been validated and used in previous studies on clinical populations (e.g. Hansson et al., 2009; Hedman et al., 2010; Johansson et al., 2013).

2.3. Attentional bias assessment and modification program

2.3.1. Apparatus

The VR-ABM program was developed by Mimerse (https://mimerse.com). The VR hardware used was the Oculus Rift consumer version headset, and response input was recorded using a wired Xbox 360 controller. The experiment ran on a Corsair Tortuga computer with 4Ghz Intel Core i7 processor and NVIDIA GeForce GTX 1080 graphics card.
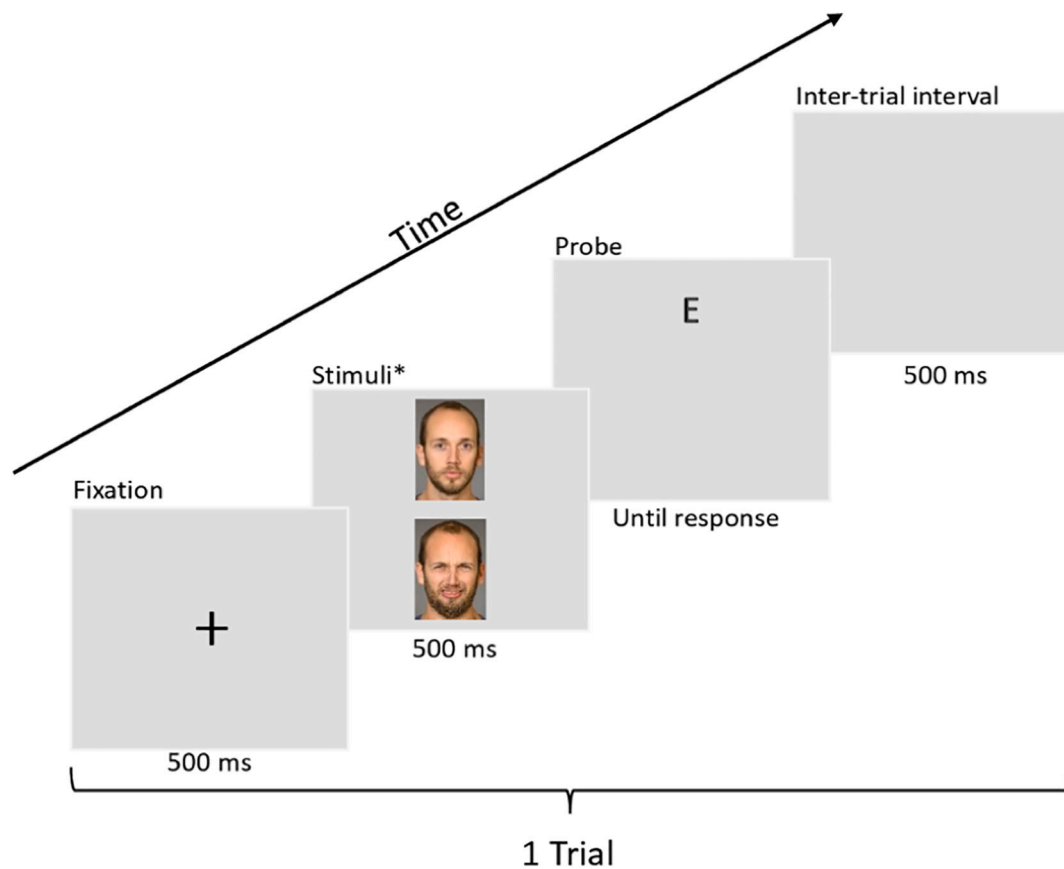
Fig. 1. Example trial of a dot-probe task. Each trial began with a fixation cross appearing on screen for 500 ms. After the fixation cross, two faces from the same individual showing a neutral expression and a disgusted expression appeared on the screen (arranged vertically) for 500 ms. The position of the neutral/disgust expressions was counterbalanced, so that each expression appeared with equal frequency on top or bottom. After the faces disappeared, a probe (letter 'E' or letter 'F', with equal frequency) would appear randomly in the location previously occupied by a neutral expression or a disgusted expression with equal frequency. Participants were instructed to identify the probe as quickly as possible by pushing the controller joystick left (for 'E') or right (for 'F'). A 500 ms inter-trial interval took place before a new trial began.
Note. The BP4D-Spontaneous database is proprietary, therefore the actual stimuli used are not permissible to print in publications. The faces shown in this example comes from the Umeå University Database of Facial expressions (Samuelsson et al., 2012). Examples of the VR environment (as seen on a computer monitor) can be found in Supplementary Materials.

2.3.2. Stimuli

The facial stimuli used in the current study were selected from the BP4D-Spontaneous Database (Zhang et al., 2014). A total of 32 individuals (50% female) each showing a neutral expression and a disgusted expression were included in the stimuli set, with a total of 64 expressions. Two sets of stimuli were created from these 64 images, one set being two dimensional (2D) and another set three dimensional (3D). The 2D images have a resolution of 1040 × 1392 pixels.

Meta-analytical studies have reported mixed findings regarding the moderating effect of stimulus modality on ABM (Jones and Sharpe, 2017). For instance, some studies found pictures to be more effective than words in changing bias (Beard et al., 2012), while others have found the opposite (Hakamata et al., 2010). Different facial expressions have been used in ABM studies, including anger, sadness, and disgust. The decision to use disgust as the socially threatening facial expression was because disgust underpins many complex emotions that are closely associated with social anxiety, such as shame, humiliation, and rejection (Amir et al., 2003; Phillips et al., 1998).

2.3.3. Attentional bias measurement (dot-probe task)

The dot-probe task was used to measure attentional bias (see Fig. 1 for task description).

Trials in which the probe appeared behind the disgusted expression were *congruent*. Trials in which the probe appeared behind the neutral expression were *incongruent*. A bias index was calculated by comparing a participant's average reaction time in incongruent trials versus congruent trials.

$$\text{Bias index} = \text{Mean}(\text{RT}_{\text{incongruent}}) - \text{Mean}(\text{RT}_{\text{congruent}})$$

A positive bias index indicated that the participant reacted faster to probes when they appeared behind disgusted faces, while a negative bias index indicated a faster reaction to probes behind neutral faces.

2.3.4. Attentional bias modification (person-identity-matching task)

Bias modification was carried out using a PIM task adapted from Notebaert et al. (2015); see Fig. 2 for task description).

In the neutral PIM, participants were instructed to identify whether the two neutral faces belonged to the same individual. Conversely, in the disgust PIM, participants were instructed to identify whether the two disgusted faces shown belonged to the same individual.

2.4. Procedure

Interested participants were registered and screened on the study website. Eligible participants were invited to book a VR session at Stockholm University, and provided written informed consent upon arrival. The participants were randomised to the four experimental groups (2D neutral, 2D disgust, 3D neutral, and 3D disgust). Before the VR session began, participants completed the pre-training assessment questionnaires (LSAS-SR, PHQ-9, GAD-7, DERS-16, and BBQ).
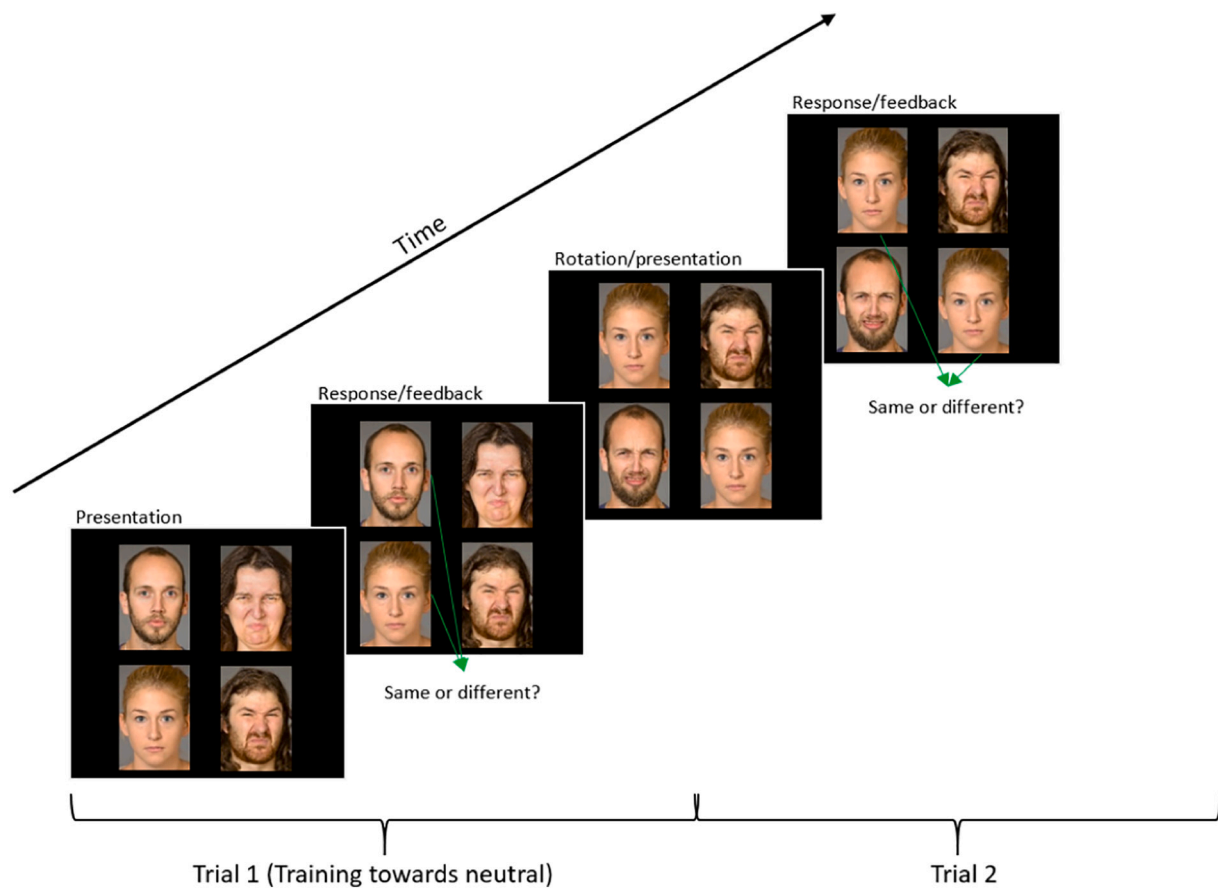
Fig. 2. Example trial of a person-identity-matching task. Two pairs of faces were presented simultaneously, forming a 2 × 2 grid of 4 faces. Each pair consisted of two individuals, one displaying a neutral expression and the other expressing disgust. In the neutral variant of this PIM task, participants were instructed to ignore the faces with a disgusted expression and only pay attention to the neutral faces in order to identify whether the two neutral faces belonged to the same individual (identity match). The participants used the controller joystick to respond by pushing it left for 'yes/same' and right for 'no/different'. If the participants responded correctly, the background of the VR environment would flash green, whereas if they responded incorrectly it would flash red. A new trial began with the top pair of faces rotating out of the grid, the bottom pair rotating to the top, and a new pair of faces appearing at the bottom.

To minimise the impact of participant cancellation on data collection, randomisation was done on VR training sessions rather than individual participants. Sessions were pseudorandomised in blocks of 4, 8, or 12 to the four experimental groups using R. Since the order of group affiliation was pre-designated, whenever a participant failed to come to an appointment, the group affiliation for the cancelled session (and subsequent sessions) would transfer to the next participant. The VR data were linked to each participant by their participant ID, which the experimenters manually input into the programme at the start of the session. Since the experimenters also needed to select the correct task for each condition before training begins, they were not blind to the experimental conditions.

The VR session began with a quick visual acuity check inside the VR environment to ensure that all participants could see the images clearly. The participants then familiarised themselves with the dot-probe task by completing a tutorial (five consecutive correct responses to probes). Baseline attentional bias was measured using 100 trials of dot-probe task. All bias measurements were carried out using 2D stimuli, regardless of what stimuli were used in the PIM training.

After bias measurement, the participants underwent another tutorial to learn the PIM task. Upon successfully finishing the tutorial (five consecutive correct responses), they completed two blocks of ABM training (190 trials each) with a self-paced break between blocks. Depending on their group affiliation, the participants received ABM with: (i) 380 trials of disgust PIM with 2D stimuli; (ii) 380 trials of neutral PIM with 2D stimuli; (iii) 380 trials of disgust PIM with 3D stimuli; or (iv) 380 trials of neutral PIM with 3D stimuli. After the

training phase, attentional bias was measured again using 100 trials of dot-probe. The participants finished the VR session by filling out the LSAS-SR again to assess their social anxiety post-training. Post-ABM questionnaires (LSAS-SR, PHQ-9, GAD-7, DERS-16, and BBQ) were sent to the participants for follow-up assessments at seven days and 90 days after the VR session.

2.5. Statistical analyses

All statistical analyses were performed in R (version 3.6.1; R Core Team, 2019).

3. Results

All 100 participants completed pre-ABM and post-ABM assessment of anxiety and bias. For follow-up measures, nine participants failed to complete the 1-week follow-up and seven participants failed to complete the 3-month follow-up. For the bias measurement data, trials were discarded if they (i) were error trials; (ii) had a response time < 200 ms or > 2000 ms; or (iii) had a response time that was beyond 2 standard deviations from the individual's mean response for each trial type (congruent/incongruent). Five participants were excluded from analyses as they had more than 20% of their trials discarded for at least one of the trial types in either the pre- or the post-training bias measurement task (see Fig. 3; for details of the data cleaning procedure, please refer to analysis script). None of the groups differ on any demographic characteristics or measures at baseline except for bias index at pre-
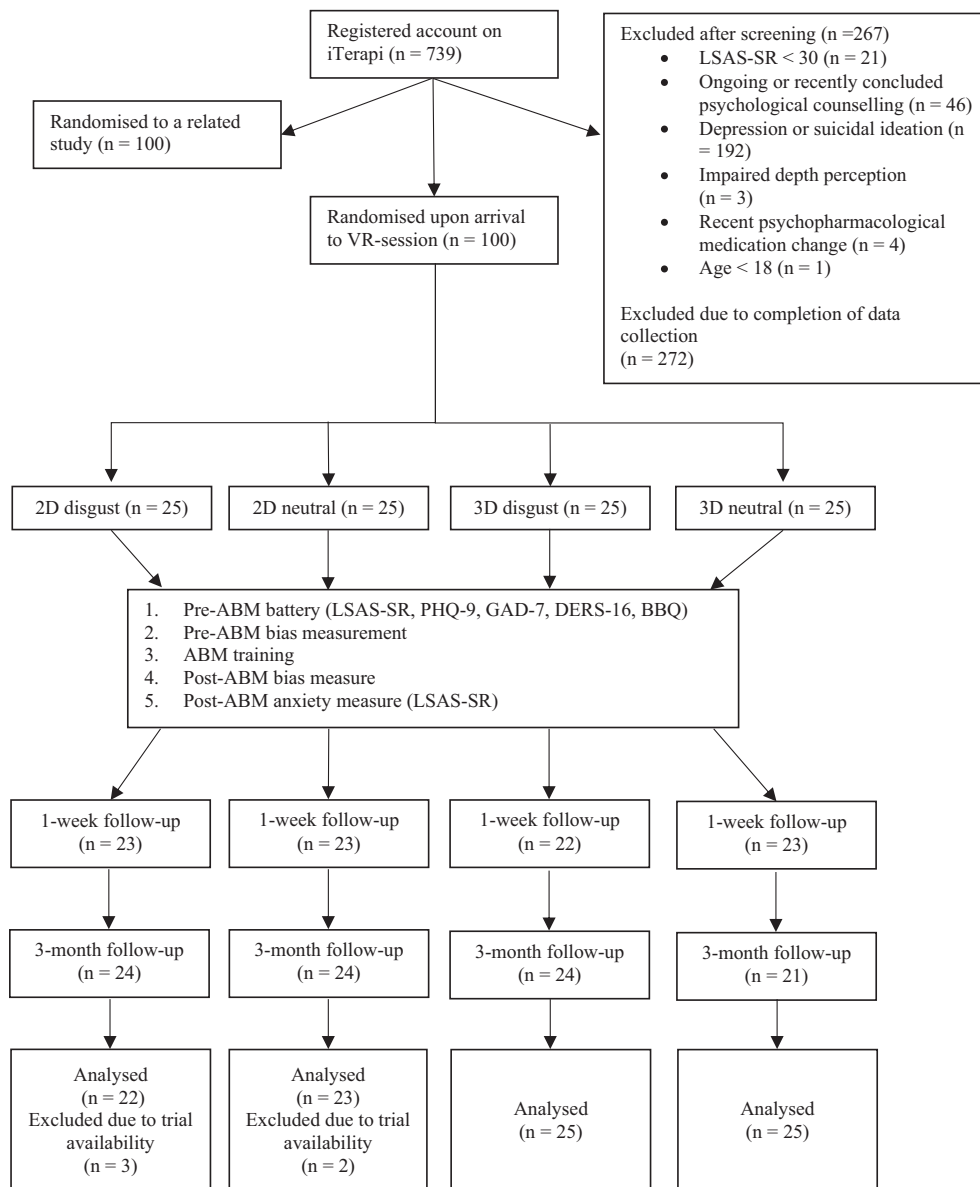
Fig. 3. Overview of participant flow throughout the study.

training (see Table 1).

3.1. Association between bias and anxiety symptoms

The relationship between attentional bias and anxiety symptoms was explored using simple correlations (Fig. 4). No significant correlations were found between bias index and LSAS-SR scores at pre-ABM ($r = -0.02$, $p = .88$) or post-ABM ($r = 0.03$, $p = .78$), or between bias change and anxiety score change ($r = -0.08$, $p = .44$).

3.2. Bias change

To evaluate whether the ABM training successfully induced a change in attentional bias, a 2×4 repeated measures ANOVA with the four groups as between-subjects factor and time (pre- vs. post-training) as within-subjects factor was performed. The results showed no significant difference in bias between pre- and post-training ($F_{(1,91)} = 0.37$, $p = .544$). No group differences were observed ($F_{(3,91)} = 1.40$, $p = .248$). No interaction between group and time was found ($F_{(3,91)} = 1.34$, $p = .267$). The result suggests that attentional

bias did not change after ABM using the PIM task, regardless of training contingency or stimuli used.

We further explored bias modification at the individual level by calculating the reliable change indices for each participant, using methods proposed by Jacobson and Truax (1991). First, the standard error of measurement (SE_M) was calculated based on the sample baseline standard deviation and split-half reliability of the pre-training dot-probe task. The Spearman-Brown corrected average reliability estimate of 5000 random splits (Parsons, 2018) served as a measure of internal reliability for bias index. The resulting estimate was very low ($r = 0.05$). A Standardised Difference Score (S_{diff}) was computed based on the standard error of measurement ($S_{diff} = \sqrt{(2 \ast SE_M{}^2)}$). If an individual's bias index was reduced by at least 1.96 times the S_{diff}, they were classified as showing a reliably improved bias (i.e. reduced attentional bias towards threat). If an individual's bias index increased by at least 1.96 times the S_{diff}, they were classified as showing a reliably deteriorated bias towards negative. If an individual's bias change fell within the range of 1.96 S_{diff}, they were classified as unchanged. The results showed that five participants had a reliable deterioration in bias, 88 participants showed no reliable change, and only two participants

Table 1
Participant demographics and characteristics at baseline.

		2D disgust (N = 22)	2D neutral (N = 23)	3D disgust (N = 25)	3D neutral (N = 25)	Between-groups comparison
Male	N	8	14	11	15	$\chi^2 = 4.07, p = .254$
	(%)	(36%)	(61%)	(44%)	(60%)	
Tertiary education	N	16	14	13	14	$\chi^2 = 2.33, p = .508$
	(%)	(73%)	(61%)	(52%)	(56%)	
Age	M	40.23	39.35	35.92	38.08	$F_{(3,91)} = 0.53, p = .660$
	(SD)	(13.53)	(12.40)	(10.88)	(13.08)	
Bias index (ms)	M	9.47	−23.39	9.11	−2.21	$F_{(3,91)} = 3.56, p = .017^*$
	(SD)	(37.29)	(48.28)	(37.72)	(32.80)	
Liebowitz Social Anxiety Scale, Self-reported (LSAS-SR)	M	63.91	62.22	65.28	67.72	$F_{(3,91)} = 0.25, p = .862$
	(SD)	(20.61)	(18.15)	(27.37)	(23.52)	
Patient Health Questionnaire (PHQ-9)	M	6.09	5.74	4.48	5.48	$F_{(3,91)} = 0.85, p = .470$
	(SD)	(3.60)	(3.43)	(3.50)	(4.10)	
Generalised Anxiety Disorder 7-item scale (GAD-7)	M	4.64	4.96	4.68	5.40	$F_{(3,91)} = 0.20, p = .898$
	(SD)	(4.11)	(2.87)	(4.07)	(4.32)	
Difficulties in Emotion Regulation Scale-16 (DERS-16)	M	37.50	31.74	36.00	39.76	$F_{(3,48.7)} = 2.12, p = .110$
	(SD)	(14.08)	(8.32)	(16.21)	(15.75)	
Brunnsviken Brief Quality of Life Inventory (BBQ)	M	52.23	43.61	54.40	46.84	$F_{(3,50.3)} = 1.50, p = .226$
	(SD)	(16.76)	(16.12)	(26.62)	(18.65)	

Note. ANOVAs were conducted for all other measures. Levene's test revealed that assumption of homogeneity of variances was violated for DERS-16 ($p = .02$) and BBQ ($p = .02$), thus these two ANOVAs were conducted without assumption of equal variances. For bias index, a Tukey's test revealed that the 2D neutral group had a significantly lower average bias index at baseline compared to the 2D disgust and 3D disgust groups.

* $p < .05$. For between-groups comparison, Pearson's chi-squared tests were conducted for sex and education.

achieved reliable improvement in their attentional bias after ABM training (Fig. 5).

3.3. Anxiety change

A mixed model approach was used to analyse how anxiety scores changed across the four groups over time. Two dummy-coded variables replaced the group variable to dissociate the effects of training condition (disgust = 0 vs. neutral = 1) and stimuli used (2D = 0 vs. 3D = 1). We used the nlme() package (Pinheiro and Bates, 2014) in R to compare different models on their fit to the data using the Akaike

information criterion (AIC), with values ranging from 2986.3 (effect of time only) to 3353.8 (null model with intercept only). For details of model comparison, see Supplementary Materials.

A linear mixed-effects model analysis was carried out using the full model. Time, condition, 2D/3D stimuli, and all 2-way and 3-way interactions were modelled as fixed effects. Random intercepts and random slopes for each participant were modelled as random effects. For main effects, only time was significant – on average, participants showed a reduction of 4.6 points in their LSAS-SR score at each successive assessment ($t_{(265)} = −3.45$, $p < .001$, Nagelkerke's pseudo $r^2 = 0.65$). No other main effects or interactions were significant
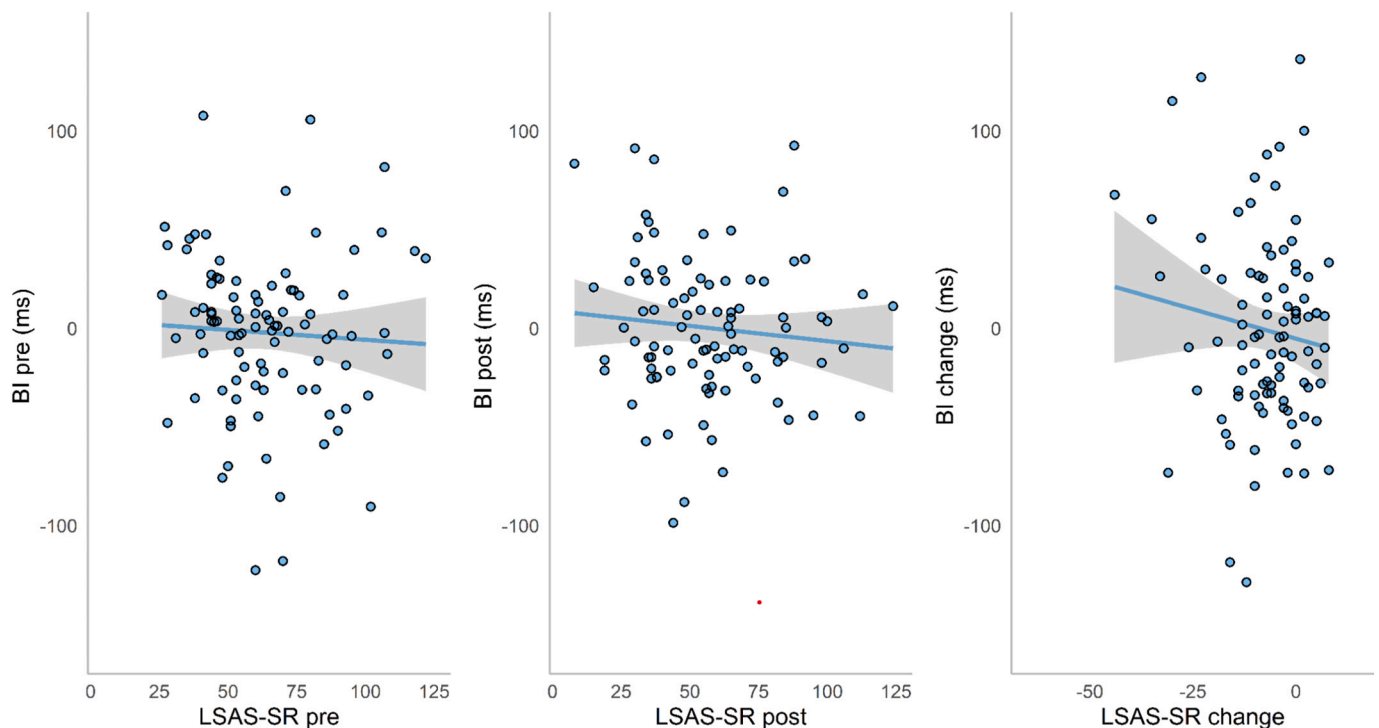

Fig. 4. Scatterplots with trend lines of bias index (BI) and anxiety scores (LSAS-SR) at pre-ABM, post-ABM, and their changes over time. Shaded regions indicate 95% CI.
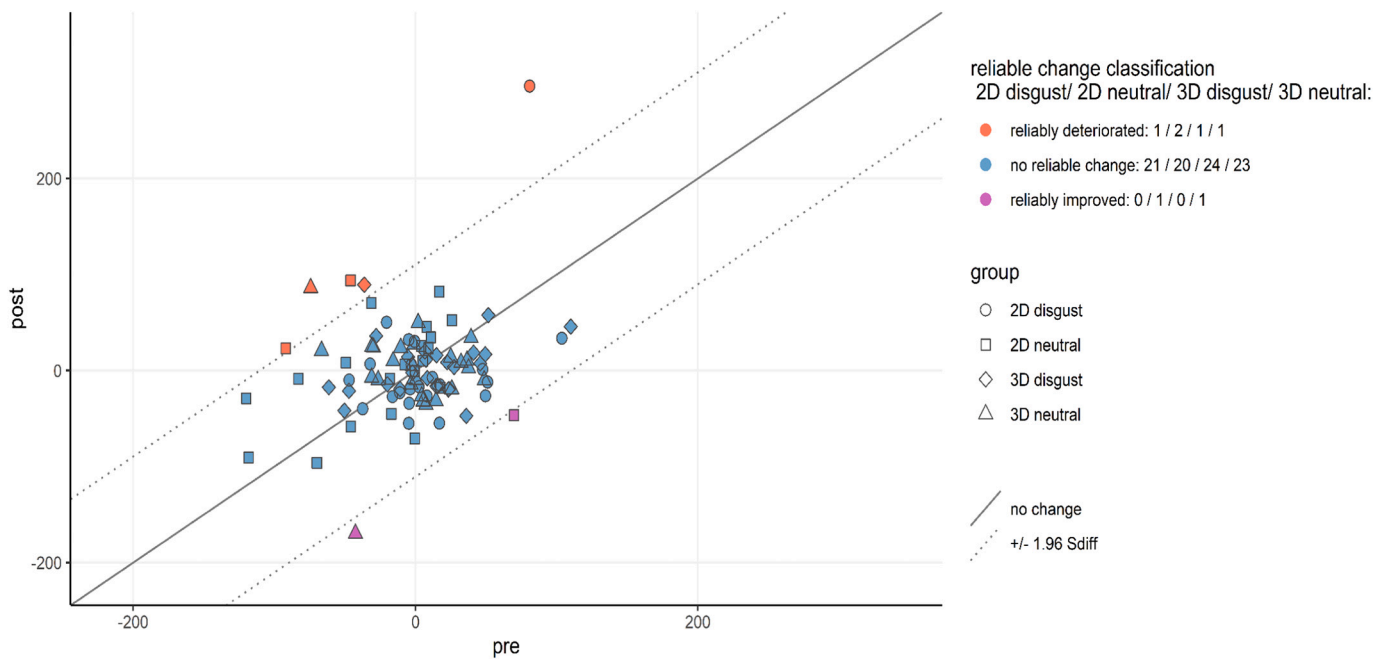
Fig. 5. Reliable change plot for individual participant's bias index.

Table 2
Fixed effects parameter estimates.

Effect	Estimate	SE	DF	t	p
Intercept	67.56	5.78	265	11.68	< 0.001***
Time	−4.57	1.10	265	−3.45	< 0.001***
Condition	−3.90	8.09	91	−0.48	0.63
2D/3D	1.66	7.93	91	0.21	0.83
Time x condition	1.14	1.85	265	0.61	0.54
Time x 2D/3D	−0.52	1.82	265	−0.29	0.77
Condition x 2D/3D	5.92	11.16	91	0.53	0.60
Time x condition x 2D/3D	−0.36	2.57	265	−0.14	0.89

*** $p < .001$.

(Table 2). Fig. 6 illustrates the LSAS-SR reduction over time, separated by groups.

Although a simpler model with only the main effect of time ($t_{(268)} = -6.82$, $p < .001$) had the lowest AIC value, directly comparing this model against the full model did not reveal a significantly better fit (likelihood-ratio = 1.75, $p = .941$).

3.4. Clinically significant change

Jacobson-Truax clinical change indices were computed for LSAS-SR scores. Here we applied the full clinical change index calculation (as opposed to just the reliable change calculation done for bias change). For each participant, reliable change was determined first (defined as change surpassing 1.96 S_{diff}), followed by application of the A criterion to determine clinical change. The A criterion was based on the sample baseline distribution of LSAS-SR scores – participants whose post-training scores were lower than the baseline group mean score minus 1.96 times the baseline standard deviation were classified as 're-covered', indicating that their post-ABM scores fall outside the 95% confidence interval of the sample's distribution at baseline.

For the calculation of the reliable change criterion, Cronbach's alpha was determined using the psych package (Revelle, 2018). The internal reliability of the LSAS-SR was found to be satisfactory (α = 0.95). Scores at post-training, 1-week, and 3-month follow-ups were all compared to baseline. At each time point, participants who showed reliable change (i.e. changed more than 1.96 S_{diff}) and a score below the

cut-off were classified as 'recovered'. Participants who scored below the cut-off but did not show reliable change were classified as 'non-reliably recovered'. Participants who showed reliable change but did not score below the cut-off points were classified as 'improved'. Participants who did not show reliable change were classified as 'unchanged'. Participants who showed reliable increase in LSAS-SR scores would have been classified as 'deteriorated'. At the 3-month follow-up, four participants were classified as 'recovered', two 'non-reliably recovered', 34 'improved', 44 'unchanged', and four 'deteriorated' (Fig. 7).

3.5. Secondary outcome measures

All secondary outcomes were analysed using linear mixed models. None of the analyses revealed any significant effect, indicating that the ABM training did not have an impact on depression, generalised anxiety, emotional regulation or quality of life.

4. Discussion

The current study investigated the efficacy of a VR-based person-identity-matching task in reducing attentional bias and social anxiety in participants with LSAS-SR scores comparable to a clinical population. After a single session of training, we did not observe any changes in attentional bias. Contrary to our hypotheses, all groups showed reduction in anxiety symptoms post-training ($p < .001$, Nagelkerke's pseudo $r^2 = 0.65$), regardless of group affiliation (neutral vs disgust PIM; 2D vs 3D stimuli). This reduction was maintained at the 1-week and 3-month follow-ups. At the 3-month follow-up, only 4% of participants met the 'recovered' criteria for clinically significant change, while anxiety scores were classified as 'unchanged' for 46% of the participants.

4.1. Failure to detect bias

The lack of evidence for attentional bias in socially anxious individuals has been reported in numerous recent studies (Boettcher et al., 2013; Heeren et al., 2015a; Ma et al., 2019; Miloff et al., 2015; for a recent meta-analysis of baseline bias in ABM RCTs, see Kruijt et al., 2019). In addition, we found no evidence of change in bias after ABM
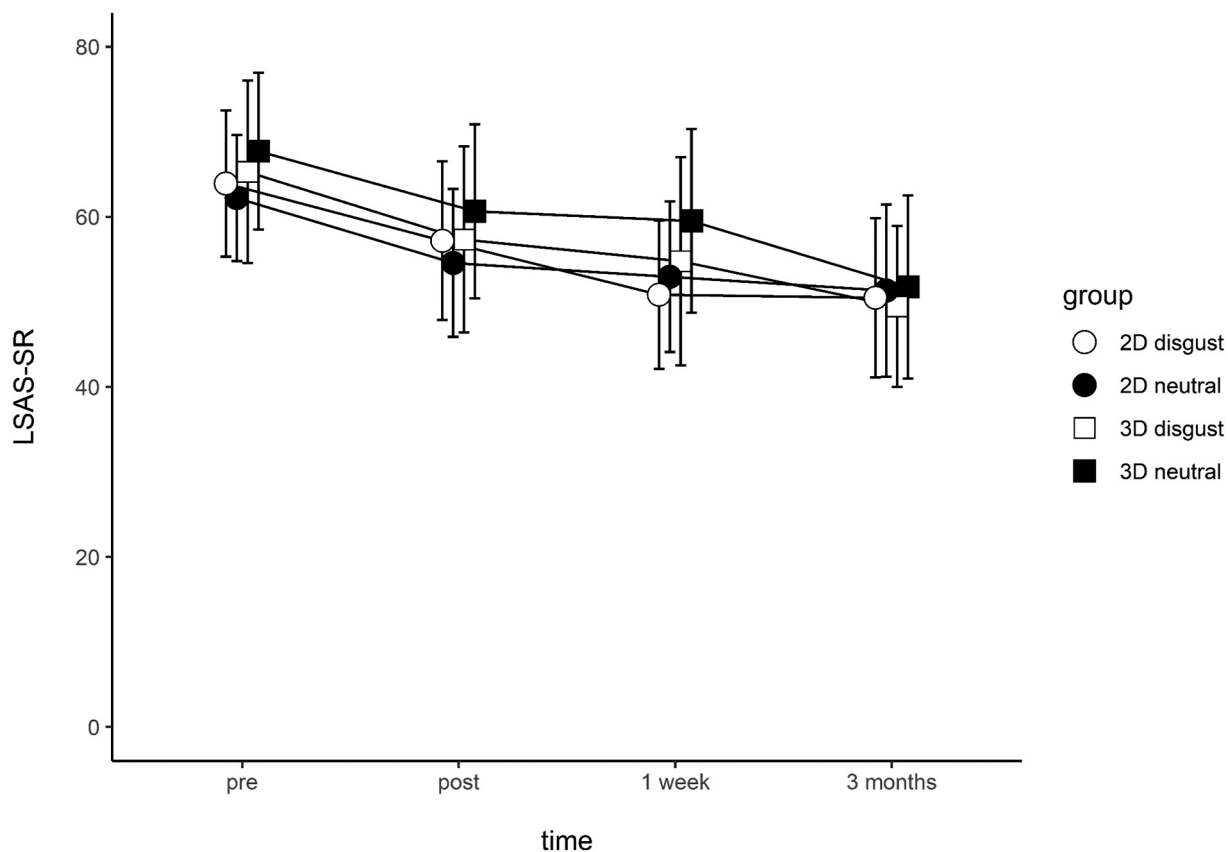
Fig. 6. Liebowitz Social Anxiety Scale, Self-reported (LSAS-SR) score change across time. Error bars represent 95% CI.

training using the PIM task (cf. Notebaert et al., 2015). One potential explanation for the failure to detect bias is the poor reliability of the dot-probe task (e.g. Chapman et al., 2017; Schmukle, 2005; Waechter and Stolz, 2015). The Spearman-Brown corrected split-half estimate for

internal reliability for the bias index was extremely low ($r = 0.05$). A similarly low reliability estimate ($r = -0.04$) for bias index obtained using the dot-probe task from another sample of 100 participants in a related study conducted by our research group has been reported
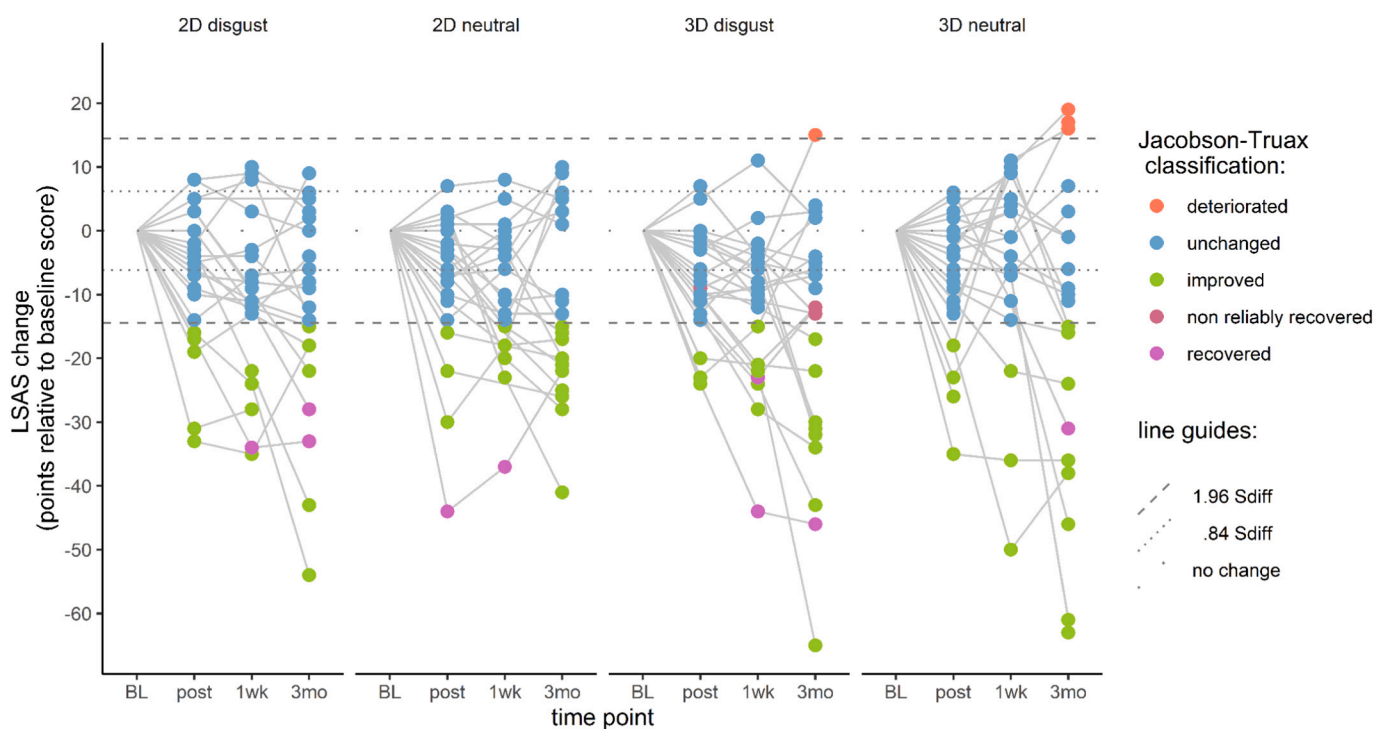

Fig. 7. Jacobson-Truax (Criterion A) classification of clinically significant change across time.

elsewhere (Ma et al., 2019). Therefore, developing reliable measures of attentional bias should be the top priority for ABM research (Huppert et al., 2018; Rodebaugh et al., 2016). Any attempt at achieving bias modification would be futile without a reliable way to measure bias.

4.2. Symptom reductions

There was an overall reduction in social anxiety scores post-ABM training across all participants, which was maintained at the 3-month follow-up. However, in terms of clinically significant change, only 4% (4 out of 95) of participants achieved the 'recovered' status. A number of studies have reported that both active ABM and mock ABM can induce similar levels of symptom reduction (e.g. Boettcher et al., 2013; Bunnell et al., 2013; Enock et al., 2014; McNally et al., 2013). These findings seem to suggest that even when ABM tasks failed to measure or change attentional bias, there might be components in the training procedures that produced therapeutic effect. The mechanism of this therapeutic effect is difficult to pinpoint. In our study, the lowered anxiety cannot be attributed to any specific training contingency or stimuli, because all groups exhibited similar levels of anxiety reduction. It has been proposed that nonspecific factors such as placebo effect resulted from participating in a study could also contribute to symptom reduction in ABM studies (Enock et al., 2014). Similarly, the exposure effect from viewing facial expressions throughout the VR session may have contributed to symptom reduction.

4.3. Limitations

The current study has a number of limitations that need to be taken into consideration when interpreting the results. Firstly, our finding of no bias change following PIM training is not in line with the original study by Notebaert et al. (2015). Using a novel task poses its own challenges when it comes to interpreting the results, and the comparison to the original study is further complicated by the change in facial expressions (i.e. neutral vs. disgusted instead of happy vs. angry), the lack of a stressor task, and the introduction of VR. An a priori power analysis performed during the planning stages of the study showed that to detect a medium effect size ($f = 0.25$) and an alpha of 0.5, a total sample size of 100 was sufficient to achieve 80% power. However, it has been suggested that effect sizes for ABM might be larger for stressor vulnerability than symptom reduction (Jones and Sharpe, 2017). Therefore, it is possible that the current study does not have a large enough sample size to detect ABM effects, especially in the absence of a stressor task. Secondly, all outcome measures except for attentional bias were self-reported. The participants in our study were not assessed by formal diagnostic criteria, thus cannot be considered a clinical population. However, the average LSAS-SR score at baseline was 65, which indicated probable SAD diagnosis (Rytwinski et al., 2009). Thirdly, the facial expressions used in the current study were not validated. Anecdotal reports from experimenters during data collection noted how some participants perceived the negative facial expressions as 'angry' instead of 'disgusted'. Fourthly, since the current study lacks a wait list control group, it is difficult to discern whether symptom reduction was ABM-specific, or due to factors such as spontaneous recovery or placebo effect stemming from taking part in a clinical study.

5. Conclusion

To summarise, a single-session, VR-based PIM task did not result in ABM-specific reduction in attentional bias or anxiety symptom. Both training towards neutral and training towards disgust achieved similar levels of anxiety reduction, which was maintained at 3-month follow-up. The anxiety reduction could not be attributed to changes in attentional bias, as we failed to detect bias at baseline, nor could we change bias with ABM training. More accurate, reliable, and precise measures of attentional bias are needed before we can properly assess the efficacy

of any ABM procedure.

Open science and pre-registration

We strive to adhere to the principles of Open Science. Unfortunately, the current study was not pre-registered before the commencement of data collection. In our effort to best compensate for the lack of pre-registration, all data used in the current study, as well as the complete R script used for data cleaning and analyses will be made openly accessible.

Declarations of competing interest

The authors declare no conflict of interest.

Acknowledgement

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.invent.2020.100334.

References

Amir, N., Elias, J., Klumpp, H., Przeworski, A., 2003. Attentional bias to threat in social phobia: facilitated processing of threat or difficulty disengaging attention from threat? Behav. Res. Ther. 41 (11), 1325–1335. https://doi.org/10.1016/S0005-7967(03)00039-1.

Amir, N., Weber, G., Beard, C., Bomyea, J., Taylor, C.T., 2008. The effect of a single-session attention modification program on response to a public-speaking challenge in socially anxious individuals. Journal of Abnormal 117 (4), 860–868. https://doi.org/10.1037/a0013445.

Andersson, G., Westöö, J., Johansson, L., Carlbring, P., 2006. Cognitive bias via the internet: a comparison of web-based and standard emotional Stroop tasks in social phobia. Cogn. Behav. Ther. 35 (1), 55–62. https://doi.org/10.1080/16506070500372469.

Baker, S.L., Heinrichs, N., Kim, H.-J., Hofmann, S.G., 2002. The Liebowitz social anxiety scale as a self-report instrument: a preliminary psychometric analysis. Behav. Res. Ther. 40 (6), 701–715. https://doi.org/10.1016/S0005-7967(01)00060-2.

Bar-Haim, Y., 2010. Research review: attention bias modification (ABM): a novel treatment for anxiety disorders. Journal of Child Psychology and Psychiatry and Allied Disciplines 51 (8), 859–870. https://doi.org/10.1111/j.1469-7610.2010.02251.x.

Barrera, T.L., Norton, P.J., 2009. Quality of life impairment in generalized anxiety disorder, social phobia, and panic disorder. Journal of Anxiety Disorders 23 (8), 1086–1090. https://doi.org/10.1016/j.janxdis.2009.07.011.

Beard, C., Sawyer, A.T., Hofmann, S.G., 2012. Efficacy of attention bias modification using threat and appetitive stimuli: a meta-analytic review. Behavior Therapy 43 (4), 724–740. https://doi.org/10.1016/j.beth.2012.01.002.

Bjureberg, J., Ljótsson, B., Tull, M.T., Hedman, E., Sahlin, H., Lundh, L.-G., ... Gratz, K.L., 2016. Development and Validation of a Brief Version of the Difficulties in Emotion Regulation Scale: The DERS-16. Journal of Psychopathology and Behavioral Assessment 38 (2), 284–296. https://doi.org/10.1007/s10862-015-9514-x.

Boettcher, J., Leek, L., Matson, L., Holmes, E.A., Browning, M., MacLeod, C., ... Carlbring, P., 2013. Internet-Based Attention Bias Modification for Social Anxiety: A Randomised Controlled Comparison of Training towards Negative and Training Towards Positive Cues. PLoS ONE 8 (9). https://doi.org/10.1371/journal.pone.0071760.

Bunnell, B.E., Beidel, D.C., Mesa, F., 2013. A randomized trial of attention training for generalized social phobia: does attention training change social behavior? Behav. Ther. 44 (4), 662–673. https://doi.org/10.1016/j.beth.2013.04.010.

Carlbring, P., Apelstrand, M., Sehlin, H., Amir, N., Rousseau, A., Hofmann, S.G., Andersson, G., 2012. Internet-delivered attention bias modification training in individuals with social anxiety disorder - a double blind randomized controlled trial. BMC Psychiatry 12 (1), 66. https://doi.org/10.1186/1471-244X-12-66.

Carpenter, J.K., Andrews, L.A., Witcraft, S.M., Powers, M.B., Smits, J.A.J., Hofmann, S.G., 2018. Cognitive behavioral therapy for anxiety and related disorders: a meta-analysis of randomized placebo-controlled trials. Depression and Anxiety 35 (6), 502–514.

https://doi.org/10.1002/da.22728.

Chapman, A., Devue, C., Grimshaw, G.M., 2017. Fleeting reliability in the dot-probe task. Psychol. Res. https://doi.org/10.1007/s00426-017-0947-6.

Cisler, J.M., Koster, E.H.W., 2010, March. Mechanisms of attentional biases towards threat in anxiety disorders: an integrative review. Clin. Psychol. Rev. https://doi.org/10.1016/j.cpr.2009.11.003.

Core Team, R., 2019. R: a language and environment for statistical computing. In: R Foundation for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria Retrieved from. http://www.r-project.org.

Cristea, I.A., Kok, R.N., Cuijpers, P., 2015. Efficacy of cognitive bias modification interventions in anxiety and depression: meta-analysis. Br. J. Psychiatry 206 (1), 7–16. https://doi.org/10.1192/bjp.bp.114.146761.

Cristea, I.A., Kok, R.N., Cuijpers, P., 2017. Invited commentary on ... Confusing procedures with process in cognitive bias modification research. Br. J. Psychiatry. https://doi.org/10.1192/bjp.bp.117.200972.

Dandeneau, S.D., Baldwin, M.W., Baccus, J.R., Sakellaropoulo, M., Pruessner, J.C., 2007. Cutting stress off at the pass: reducing vigilance and responsiveness to social threat by manipulating attention. J. Pers. Soc. Psychol. 93 (4), 651–666. https://doi.org/10.1037/0022-3514.93.4.651.

De Voogd, E.L., Wiers, R.W., Prins, P.J.M., Salemink, E., 2014. Visual search attentional bias modification reduced social phobia in adolescents. J. Behav. Ther. Exp. Psychiatry 45 (2), 252–259. https://doi.org/10.1016/j.jbtep.2013.11.006.

Enock, P.M., Hofmann, S.G., McNally, R.J., 2014. Attention bias modification training via smartphone to reduce social anxiety: a randomized, controlled multi-session experiment. Cogn. Ther. Res. 38 (2), 200–216. https://doi.org/10.1007/s10608-014-9606-z.

Fresco, D.M., Coles, M.E., Heimberg, R.G., Liebowitz, M.R., Hami, S., Stein, M.B., Goetz, D., 2001. The Liebowitz social anxiety scale: a comparison of the psychometric properties of self-report and clinician-administered formats. Psychol. Med. 31 (6), 1025–1035. https://doi.org/10.1017/S0033291701004056.

Grafton, B., Macleod, C., 2016. Engaging with the wrong people: the basis of selective attention to negative faces in social anxiety. Clin. Psychol. Sci. 4 (5), 793–804. https://doi.org/10.1177/2167702615616344.

Grafton, B., MacLeod, C., Rudaizky, D., Holmes, E.A., Salemink, E., Fox, E., Notebaert, L., 2017. Confusing procedures with process when appraising the impact of cognitive bias modification on emotional vulnerability. Br. J. Psychiatry 211 (5), 266–271. https://doi.org/10.1192/bjp.bp.115.176123.

Hakamata, Y., Lissek, S., Bar-Haim, Y., Britton, J.C., Fox, N.A., Leibenluft, E., ... Pine, D.S., 2010. Attention bias modification treatment: a meta-analysis toward the establishment of novel treatment for anxiety. Biological Psychiatry 68 (11), 982–990. https://doi.org/10.1016/j.biopsych.2010.07.021.

Hansson, M., Chotai, J., Nordstöm, A., Bodlund, O., 2009. Comparison of two self-rating scales to detect depression: HADS and PHQ-9. Br. J. Gen. Pract. 59 (566), 650–654. https://doi.org/10.3399/bjgp09X454070.

Hedman, E., Ljótsson, B., Rück, C., Furmark, T., Carlbring, P., Lindefors, N., Andersson, G., 2010. Internet administration of self-report measures commonly used in research on social anxiety disorder: a psychometric evaluation. Comput. Hum. Behav. 26 (4), 736–740. https://doi.org/10.1016/j.chb.2010.01.010.

Hedman, E., Andersson, E., Ljótsson, B., Axelsson, E., Lekander, M., 2016. Cost effectiveness of internet-based cognitive behaviour therapy and behavioural stress management for severe health anxiety. BMJ Open 6 (4), e009327. https://doi.org/10.1136/bmjopen-2015-009327.

Heeren, A., Mogoaşe, C., McNally, R.J., Schmitz, A., Philippot, P., 2015a. Does attention bias modification improve attentional control? A double-blind randomized experiment with individuals with social anxiety disorder. Journal of Anxiety Disorders 29, 35–42. https://doi.org/10.1016/j.janxdis.2014.10.007.

Heeren, A., Mogoaşe, C., Philippot, P., McNally, R.J., 2015b. Attention bias modification for social anxiety: a systematic review and meta-analysis. Clin. Psychol. Rev. 40, 76–90. https://doi.org/10.1016/j.cpr.2015.06.001.

Huppert, J.D., Kivity, Y., Cohen, L., Strauss, A.Y., Elizur, Y., Weiss, M., 2018. A pilot randomized clinical trial of cognitive behavioral therapy versus attentional bias modification for social anxiety disorder: an examination of outcomes and theory-based mechanisms. Journal of Anxiety Disorders 59, 1–9. https://doi.org/10.1016/j.janxdis.2018.08.002.

Jacobson, N.S., Truax, P., 1991. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. J. Consult. Clin. Psychol. 59 (1), 12–19. https://doi.org/10.1037/0022-006X.59.1.12.

Jakubovski, E., Johnson, J.A., Nasir, M., Müller-Vahl, K., Bloch, M.H., 2019. Systematic review and meta-analysis: dose-response curve of SSRIs and SNRIs in anxiety disorders. Depression and Anxiety 36 (3), 198–212. https://doi.org/10.1002/da.22854.

Johansson, R., Carlbring, P., Heedman, A., Paxling, B., Andersson, G., 2013. Depression, anxiety and their comorbidity in the Swedish general population: point prevalence and the effect on health-related quality of life. PeerJ 2013 (1). https://doi.org/10.7717/peerj.98.

Jones, E.B., Sharpe, L., 2017. Cognitive bias modification: a review of meta-analyses. J. Affect. Disord. 223, 175–183. https://doi.org/10.1016/j.jad.2017.07.034.

Koster, E.H., Fox, E., MacLeod, C., 2009. Introduction to the special section on cognitive bias modification in emotional disorders. J. Abnorm. Psychol. 118 (1), 1–4. https://doi.org/10.1037/a0014379.

Kroenke, K., Spitzer, R.L., Williams, J.B.W., 2001. The PHQ-9: validity of a brief depression severity measure. J. Gen. Intern. Med. 16 (9), 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x.

Kruijt, A.-W., Carlbring, P., 2018. Processing confusing procedures in the recent re-analysis of a cognitive bias modification meta-analysis. Br. J. Psychiatry. https://doi.org/10.1192/bjp.2018.41.

Kruijt, A.-W., Parsons, S., Fox, E., 2019. A meta-analysis of bias at baseline in RCTs of attention bias modification: no evidence for dot-probe bias towards threat in clinical anxiety and PTSD. J. Abnorm. Psychol. 128 (6), 563–573. https://doi.org/10.1037/abn0000406.

Lindner, P., Frykheden, O., Forsström, D., Andersson, E., Ljótsson, B., Hedman, E., ... Carlbring, P., 2016. The Brunnsviken Brief Quality of Life Scale (BBQ): Development and Psychometric Evaluation. Cognitive Behaviour Therapy 45 (3), 182–195. https://doi.org/10.1080/16506073.2016.1143526.

Lindner, P., Miloff, A., Hamilton, W., Reuterskiöld, L., Andersson, G., Powers, M.B., Carlbring, P., 2017. Creating state of the art, next-generation virtual reality exposure therapies for anxiety disorders using consumer hardware platforms: design considerations and future directions. Cogn. Behav. Ther. 46 (5), 404–420. https://doi.org/10.1080/16506073.2017.1280843.

Linetzky, M., Pergamin-Hight, L., Pine, D.S., Bar-Haim, Y., 2015. Quantitative evaluation of the clinical efficacy of attention bias modification treatment for anxiety disorders. Depression and Anxiety. https://doi.org/10.1002/da.22344.

Ma, L., Kruijt, A.-W., Nöjd, S., Zetterlund, E., Andersson, G., Carlbring, P., 2019. Attentional bias modification in virtual reality – a VR-based dot-probe task with 2D and 3D stimuli. Frontiers in Psychology 10. https://doi.org/10.3389/fpsyg.2019.02526.

MacLeod, C., Mathews, A., 2012. Cognitive bias modification approaches to anxiety. Annu. Rev. Clin. Psychol. 8 (1), 189–217. https://doi.org/10.1146/annurev-clinpsy-032511-143052.

MacLeod, C., Mathews, A., Tata, P., 1986. Attentional bias in emotional disorders. J. Abnorm. Psychol. 95 (1), 15–20. https://doi.org/10.1037/0021-843X.95.1.15.

MacLeod, C., Rutherford, E., Campbell, L., Ebsworthy, G., Holker, L., 2002. Selective attention and emotional vulnerability: assessing the causal basis of their association through the experimental manipulation of attentional bias. J. Abnorm. Psychol. 111 (1), 107–123. https://doi.org/10.1037/0021-843X.111.1.107.

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

McNally, R.J., 2018. Attentional bias for threat: crisis or opportunity? Clin. Psychol. Rev. https://doi.org/10.1016/j.cpr.2018.05.005.

McNally, R.J., Enock, P.M., Tsai, C., Tousian, M., 2013. Attention bias modification for reducing speech anxiety. Behav. Res. Ther. 51 (12), 882–888. https://doi.org/10.1016/j.brat.2013.10.001.

Miloff, A., Savva, A., Carlbring, P., 2015. Cognitive bias measurement and social anxiety disorder: correlating self-report data and attentional bias. Internet Interv. 2 (3), 227–234. https://doi.org/10.1016/j.invent.2015.03.006.

Mogg, K., Bradley, B.P., 2016. Anxiety and attention to threat: cognitive mechanisms and treatment with attention bias modification. Behav. Res. Ther. 87, 76–108. https://doi.org/10.1016/j.brat.2016.08.001.

Notebaert, L., Clarke, P.J.F., Grafton, B., MacLeod, C., 2015. Validation of a novel attentional bias modification task: the future may be in the cards. Behav. Res. Ther. 65, 93–100. https://doi.org/10.1016/j.brat.2014.12.007.

Parsons, S., 2018. splithalf: robust estimates of split half reliability. Version 3. figshare. https://doi.org/10.6084/m9.figshare.5559175.v2.

Phillips, M.L., Senior, C., Fahy, T., David, A.S., 1998. Disgust–the forgotten emotion of psychiatry. The British Journal of Psychiatry: The Journal of Mental Science 172, 373–375. https://doi.org/10.1192/bjp.172.5.373.

Pinheiro, J., Bates, D., 2014. Linear and nonlinear mixed effects models R-package nlme. Retrieved from. https://cran.r-project.org/package=nlme.

Revelle, W., 2018. Psych: procedures for personality, psychometric, and personality research. Retrieved from. https://cran.r-project.org/package=psych.

Rodebaugh, T.L., Scullin, R.B., Langer, J.K., Dixon, D.J., Huppert, J.D., Bernstein, A., ... Lenze, E.J., 2016. Unreliability as a Threat to understanding psychopathology: The cautionary tale of attentional bias. Journal of Abnormal Psychology 125 (6), 840–851. https://doi.org/10.1037/abn0000184.

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. https://doi.org/10.1002/da.20503.

Samuelsson, H., Jarnvik, K., Henningsson, H., Andersson, J., Carlbring, P., 2012. The umeå university database of facial expressions: a validation study. J. Med. Internet Res. 14 (5). https://doi.org/10.2196/jmir.2196.

Saris, I.M.J., Aghajani, M., van der Werff, S.J.A., van der Wee, N.J.A., Penninx, B.W.J.H., 2017. Social functioning in patients with depressive and anxiety disorders. Acta Psychiatr. Scand. 136 (4), 352–361. https://doi.org/10.1111/acps.12774.

Schmukle, S.C., 2005. Unreliability of the dot probe task. Eur. J. Personal. 19 (7), 595–605. https://doi.org/10.1002/per.554.

Spitzer, R.L., Kroenke, K., Williams, J.B.W., Löwe, B., 2006. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch. Intern. Med. 166 (10), 1092–1097. https://doi.org/10.1001/archinte.166.10.1092.

Urech, A., Krieger, T., Chesham, A., Mast, F.W., Berger, T., 2015. Virtual reality-based attention bias modification training for social anxiety: a feasibility and proof of concept study. Frontiers in Psychiatry 6 (OCT), 1–5. https://doi.org/10.3389/fpsyt.2015.00154.

Van Bockstaele, B., Verschuere, B., Tibboel, H., De Houwer, J., Crombez, G., Koster, E.H.W., 2014. A review of current evidence for the causal impact of attentional bias on fear and anxiety. Psychol. Bull. 140 (3), 682–721. https://doi.org/10.1037/a0034834.

Vlaescu, G., Alasjö, A., Miloff, A., Carlbring, P., Andersson, G., 2016. Features and functionality of the Iterapi platform for internet-based psychological treatment. Internet Interv. 6, 107–114. https://doi.org/10.1016/j.invent.2016.09.006.

Waechter, S., Stolz, J.A., 2015. Trait anxiety, state anxiety, and attentional bias to threat: assessing the psychometric properties of response time measures. Cogn. Ther. Res. 39 (4), 441–458. https://doi.org/10.1007/s10608-015-9670-z.

Zhang, X., Yin, L., Cohn, J.F., Canavan, S., Reale, M., Horowitz, A., ... Girard, J.M., 2014. BP4D-Spontaneous: A high-resolution spontaneous 3D dynamic facial expression database. Image and Vision Computing 32 (10), 692–706. https://doi.org/10.1016/j.imavis.2014.06.002.