Journal of Affective Disorders 290 (2021) 169–177

Contents lists available at ScienceDirect

Journal of Affective Disorders

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

Research paper

Increased attention allocation to socially threatening faces in social anxiety disorder: A replication study

Amit Lazarov [a,b,*], Dana Basel [a], Sarah Dolan [c], Daniel G. Dillon [d], Diego A. Pizzagalli [e], Franklin R. Schneier [f]

[a] School of Psychological Sciences, Tel Aviv University, Tel Aviv 69978, Israel
[b] Department of Psychiatry, Columbia University Irving Medical Center, NY, USA
[c] New York State Psychiatric Institute, NY, USA
[d] Center for Depression, Anxiety and Stress Research at McLean Hospital, MA, USA; Harvard Medical School, MA, USA
[e] Center for Depression, Anxiety and Stress Research at McLean Hospital, Belmont, MA, USA, McLean Imaging Center, McLean Hospital, MA, USA, Harvard Medical School, MA, USA
[f] New York State Psychiatric Institute and Department of Psychiatry, Columbia University Irving Medical Center, NY, USA

ARTICLE INFO

Keywords:
Social anxiety disorder
Eye tracking, Attention allocation, Attention bias
Replicability
Reliability

ABSTRACT

Background: Threat-related attention bias has been implicated in the etiology and maintenance of social anxiety disorder (SAD), with attentional research increasingly using eye-tracking methodology to overcome the poor psychometric properties of response-time-based tasks and measures. Yet, extant eye-tracking research in social anxiety has mostly failed to report on psychometrics and attempts to replicate past results are rare. Therefore, we attempted to replicate a previously published eye-tracking study of gaze patterns in socially anxious and non-anxious participants as they viewed social threatening and neutral faces, while also exploring the psychometric properties of the attentional measures used.
Methods: Gaze was monitored as participants freely viewed 60 different matrices comprised of eight socially-threatening and eight neutral faces, presented for 6000 ms each. Gaze patterns directed at threat and neutral areas of interest (AOIs) were compared by group. Internal consistency and test-retest reliability were also evaluated.
Results: Relative to healthy controls, socially anxious patients dwelled significantly longer on threat faces, replicating prior findings with the same task. Internal consistency of total dwell time on threat and neutral AOIs was high, and two-week test-retest reliability was acceptable.
Limitations: Test-retest reliability was only examined for the control group, which had a small sample size.
Conclusion: Increased dwell time on socially threatening stimuli is a reliable, stable, and generalizable measure of attentional bias in adults with social anxiety.

1. Introduction

Cognitive models of social anxiety disorder (SAD) consistently implicate threat-related information processing biases, including biased attention, in the etiology, development and maintenance of the disorder (Clark and Wells, 1995; Rapee and Heimberg, 1997). Specifically, these models suggest that the attentional system of socially anxious individuals may be distinctively sensitive to, or biased toward, socially threatening stimuli in the environment, resulting in heightened fear and anxiety (Bogels and Mansell, 2004; Morrison and Heimberg, 2013). Three main theoretical accounts as to the causal relationship between attention biases and anxiety have been proposed (Field and Lester, 2010; Lazarov and Bar-Haim, 2020). The first views threat-related attention biases as innate emotional elements evident from early childhood only in anxious individuals. The second conceptualizes attention biases to threat as normative early in life. A failure of these early normative attention biases to diminish over maturation due to individual differences in, for example, fearfulness, and behaviorally inhibited temperament, eventually leads to anxiety disorders. Finally, the third considers threat-related attention biases to be caused by specific events during development and therefore the result of direct experiences. This latter proposition is in line with cognitive models of SAD suggesting social

* Corresponding author at: School of Psychological Sciences, Tel Aviv University, Tel Aviv 69978, Israel.
E-mail address: amitlaza@post.tau.ac.il (A. Lazarov).

https://doi.org/10.1016/j.jad.2021.04.063
Received 22 November 2020; Received in revised form 16 February 2021; Accepted 23 April 2021
Available online 1 May 2021
0165-0327/

anxiety to be related to negative social learning experiences, such as ongoing criticism, social rejection or social trauma (Kuo et al., 2011) and with research showing that trauma-exposure (Lazarov et al., 2019) and fear conditioning procedures (Mulckhuyse et al., 2013; Nissens et al., 2017; Preciado et al., 2017; Schmidt et al., 2015) can induce attention biases to the feared situations/stimuli. While extant research has yet to elucidate the exact mechanism(s) through which attention to threat is causally associated with anxiety (Chong and Meyer, 2020), threat-related attention biases have been established to be present in SAD (Bantin et al., 2016; Chen and Clarke, 2017), and have been suggested as targets for therapeutic interventions via computerized attention bias modification treatments (ABMT; for reviews see Heeren et al., 2015; Linetzky et al., 2015; Mogg et al., 2017).

Early research on biased attention in social anxiety mostly employed first-generation cognitive tasks, such as the dot-probe and Stroop tasks, inferring threat-related attention biases from differences in response time (RT) contingent on the valence of presented stimuli (for reviews see Bantin et al., 2016; Bogels and Mansell, 2004). While considerably advancing our knowledge, RT-based attentional tasks suffer from some inherent limitations. First, as RT-based attentional measures are derived from keypresses occurring at the very end of the information processing sequence, the examined attentional components, which take place earlier in the process, can only be indirectly inferred from facilitated or impaired performance. Thus, RT-based measures of attention are fundamentally limited in their ability to differentiate the various aspects of attention, including facilitated threat detection, difficulty disengaging attention from threat once detected, and ongoing threat-related attentional allocation (i.e., avoidance, sustained attention), which is critical for clarifying which aspects of attention might be affected, and hence should be targeted in treatment of adults with SAD (Lazarov et al., 2016; Lazarov et al., 2018; Lazarov et al., 2019). Second, the attentional mechanisms of interest and the key presses used to measure them are separated in time, with processes such as decision-making and motor preparation occurring between them. Hence, this distanced relationship makes RT an inherently coarse measure (Armstrong and Olatunji, 2012; In-Albon and Schneider, 2010; Lazarov et al., 2019; Price et al., 2016). Finally, RT-based tasks demonstrate poor reliability, with respect to both internal consistency and test-retest reliability (Brown et al., 2014; Rodebaugh et al., 2016; Schmukle, 2005; Staugaard, 2009; Waechter et al., 2014; Waechter and Stolz, 2015). Reliability, signifying the proportion of a measure's variance that reflects *true* score variance, as opposed to measurement *errors* (Waechter et al., 2014), is crucial for our ability to trust emerging results. Reliable, psychometrically sound attentional measures are imperative for enhancing our understanding of the attentional processes that might be implicated in psychopathology (Rodebaugh et al., 2016; Spiller et al., 2020). For these three reasons, we need new and improved paradigms to assess—and subsequently modify—attentional biases in SAD (Chen and Clarke, 2017; Lazarov et al., 2019; McNally, 2019; Van Bockstaele et al., 2014).

Eye-tracking, a non-invasive approach that continuously samples gaze data, can address the limitations of RT-based tasks and measures (Armstrong and Olatunji, 2012; Chen and Clarke, 2017; Suslow et al., 2020). By continuously measuring eye movements at speeds up to 2000 data-points per second (2000 Hz), eye-tracking offers a nearly instantaneous measure of visual attention. This greatly improves our ability to delineate the time course and different components of attentional processes, and it also removes potentially confounding elements as no motor responses, other than the eye movements themselves, are required (Armstrong and Olatunji, 2012; Lazarov et al., 2016; Lazarov et al., 2019). For example, in free-viewing tasks, one of the most widely used paradigms in attention research, participants freely view arrays of stimuli without any additional requirements or demands, while their eye-data are being continuously recorded. The recorded eye data are then processed and analyzed to infer different attentional components. Location and latency of first fixations (i.e., the initial fixations following stimulus onset) are used to index threat detection/vigilance, with a greater proportion of first fixations on threat compared with neutral stimuli, or shorter latencies to first fixate on threat compared with neutral stimuli, considered evidence of facilitated threat detection. First fixation dwell time is used to reflect difficulty in disengaging attention from threat, once detected, as indicated by increased dwell time on threat compared with neutral stimuli. Accumulating the durations (i.e., total dwell time) of all fixations made during stimulus presentation per specific stimuli type is used to reflect ongoing attention allocation, with increased attention allocation to threat over neutral stimuli reflecting sustained attention on threat, with the opposite pattern indicating attentional avoidance (Elias et al., 2021; Lazarov et al., 2019). Importantly, eye-tracking-based measures also demonstrate appropriate psychometric properties, including good internal consistency and high test-retest reliability, especially for measures assessed over longer time periods (i.e., beyond first fixations or the initial 500 ms time epoch; Armstrong and Olatunji, 2012), such as total fixation duration or number of fixations (Chong and Meyer, 2020; In-Albon and Schneider, 2010; Lazarov et al., 2016; Lazarov et al., 2018; Lazarov et al., 2019; Sears et al., 2019; Skinner et al., 2017; Waechter et al., 2014). However, while improved psychometrics have been noted for eye-tracking attentional research in general, similar research in SAD has rarely reported the psychometric properties of tasks and measures used. This is a key limitation, because measures must be reliable if the results are to inspire confidence (Lazarov et al., 2016; Parsons et al., 2018; Waechter et al., 2014).

To the best of our knowledge, only two studies have reported on the psychometric properties of eye-tracking tasks in social anxiety. In the first (Waechter et al., 2014), students with high and low levels of social anxiety viewed pairs of faces, one emotional (anger, disgust, happy) and one non-emotional (calm or neutral), for 5000 ms. Four indices commonly used in free viewing eye-tracking tasks were assessed. Two were based on first fixations—location and latency of first fixations—-with both showing low internal consistency. Two others were based on the entire presentation duration—proportion of fixation frequency and proportion of viewing time spent fixating on the emotional vs. neutral stimuli—and these showed good-to-excellent internal consistency. Yet, while these latter two indices were highly reliable, the two groups did not significantly differ on any of these measures, which is to be expected under the attention bias hypothesis. In addition, test-retest reliability was not assessed in this study.

In the second study (Lazarov et al., 2016), students with high and low levels of social anxiety, as well as treatment-seeking patients with SAD, freely viewed face matrices comprised of eight socially threatening faces (i.e., faces with disgusted expressions; Staugaard, 2010) and eight neutral faces, presented concurrently for 6000 ms (Session 1); the two student samples returned to the laboratory a week later to complete the task a second time (Session 2). Latency to first fixations, first fixation location, and first fixation dwell time were assessed, as was total dwell time on threat and neutral stimuli, respectively (i.e., the accumulative time spent fixating on each pre-defined area of interest; AOI). Internal consistency and one-week test-retest reliability were evaluated. Replicating Waechter et al. (2014), no group differences emerged for the first fixation measures. However, relative to controls, the socially anxious students and participants with SAD dwelled significantly longer on threat faces in both Sessions 1 and 2. Reliability analysis of both Session 1 and 2 echoed the results of Waechter et al. (2014). Internal consistency was low for first fixation measures, ranging from .37 to .67. By contrast, high internal consistency was noted for total dwell time on threat faces and neutral faces, and for the percentage of total dwell time on threat faces out of total dwell time on all faces, with Cronbach's alphas of .95, .95, and .91, respectively. Internal consistency remained high in Session 2, with Cronbach's alphas of .89, .92, and .94, respectively. However, within-group internal consistencies were not examined. Finally, one-week test-retest reliability was significant and acceptable for all three dwell times-based measures (i.e., .68, .62, and .63 for the three dwell-tome measures described above, respectively), while for first fixation measures test-retest was non-significant and low, ranging .06 to .26.

These two studies lend preliminary support for the claim that dwell-time-based measures are reliable markers of attention allocation, with potential to serve as a viable treatment target. However, two studies are certainly not enough to establish this. Furthermore, only one of these two studies examined test-retest reliability, which is the only reliability index testing replicability over time. Importantly, prioritizing replicability in psychological sciences has gained much traction over the last several years, especially in light of well-known problems of reproducibility in psychology research (Collaboration, 2015; Wiggins and Christopherson, 2019), with problematic measurement being specifically implicated as one factor contributing to this crisis. Indeed, it has been argued that researchers in psychological sciences, including in research on attention processes, do not devote the required attention to the psychometric properties of measures used in their studies, neglecting them or taking them for granted (Lilienfeld and Strother, 2020; McNally, 2019; Parsons et al., 2018).

Here, we replicate our previous attention study in SAD using the same task and outcome measures, but among English-speaking US participants as opposed to the Hebrew-speaking Israeli participants used in the original study (Lazarov et al., 2016). While modifications of the task have been successfully used by different research groups (e.g., Abend et al., 2020; Klawohn et al., 2020; Lazarov et al., 2018; Soleymani et al., 2020), this is the first attempt to replicate the original study using exactly the same task and primary outcome measure (i.e., total dwell time), and in participants with the same form of psychopathology. Accordingly, treatment-seeking socially anxious participants and non-anxious healthy control participants completed the attention allocation task described above (Lazarov et al., 2016). We expected to replicate the group differences in participants' allocation of visual attention to threatening and neutral stimuli. While in the original study no group differences emerged for first fixation measures (i.e., latency, location, and dwell time), we still incorporated them in current analyses to be as consistent as possible with the original study. Finally, we also assessed internal consistency and test-retest reliability, extending the retest period from one to two weeks, hoping to show once again adequate reliability.

2. Methods

2.1. Participants

Participants included 53 treatment-seeking individuals with a clinical diagnosis of SAD who were recruited as part of a randomized control trial (RCT) examining the clinical efficacy of a novel treatment for SAD. Twenty age-, sex-, and race-matched participants with no lifetime psychiatric disorders were recruited as a healthy control (HC) group. Demographic and psychopathological characteristics (see Measures below) by group are presented in Table 1.

Participants were recruited via online advertisement (e.g., Craigslist, RecruitMe, the Anxiety Disorders Clinic's website), local media, and community postings. Primary and co-morbid psychiatric diagnoses were assessed by a trained psychiatrist using the Mini-International Neuropsychiatric Interview (MINI version 7.0.2; Sheehan et al., 1998, see below), a well-validated structured interview for psychiatric diagnoses (Lecrubier et al., 1997; Sheehan et al., 1997). Severity of social anxiety and depression was further assessed using the Liebowitz Social Anxiety Scale (LSAS; Liebowitz, 1987) and Hamilton Rating Scale for Depression (HAM-D; Hamilton, 1960), respectively.

Inclusion criteria for the SAD group were: a) primary DSM-5 diagnosis of SAD; b) LSAS self-report version score ≥50; c) 18-60 years old; d) fluent English; and e) normal or corrected-to-normal vision, excluding multi-focal eye wear to prevent eye-tracking calibration difficulties. An LSAS cutoff score of 50 was used as this score permits identification of SAD with optimal balance between specificity and sensitivity (Amir and Taylor, 2012; Mennin et al., 2002), and it was also used in previous studies using the same paradigm (Lazarov et al., 2016; Lazarov et al., 2017). Exclusion criteria were: a) current severe depression, indicated by a HAM-D (Hamilton, 1960) score >20; b) clinically significant suicidal ideation or behavior; c) current or past psychosis; d) current or past diagnosis of PTSD, obsessive-compulsive disorder, bipolar disorder, manic episode, tic disorder, or attention deficit hyperactivity disorder (ADHD); e) severe alcohol or cannabis use disorder, and/or any severity of other substance use disorder (except nicotine use disorders); f) current unstable or untreated medical illness; g) current or past organic mental disorder, seizure disorder or brain injury; h) current unstable or untreated medical illness; i) use of any psychotropic medication in the past month (with the exception of serotonin reuptake inhibitors (SSRIs), serotonin-norepinephrine reuptake inhibitors (SNRIs), or zolpidem for sleep, taken at a stable dose for at least three months); j) concurrent cognitive behavioral therapy (CBT), or other psychotherapy that was initiated within the past three months; or k) eye-tracking calibration difficulties. Inclusion criteria for the HC group were: a) 18–60 years old; b) fluent English; c) normal or corrected-to-normal vision, excluding multi-focal eye wear to prevent eye-tracking calibration difficulties; and d) LSAS (self-report version) score <30. Exclusion criteria were: a) current or past history of any DSM-5 psychiatric disorder; b) current or past organic mental disorder, seizure or brain injury; and c) current unstable or untreated medical illness. Of the 53 participants with SAD included in the study, 18 also met criteria for major depressive disorder depressive (MDD), and five met for generalized anxiety disorder (GAD). Three participants were on stable SSRI and one was on stable SNRI medication, and of these medicated participants two had a current diagnosis of a major depressive episode (MDE) and two were in remission.

The study was conducted in accordance with ethical guidelines of the Declaration of Helsinki and approved by the New York State Psychiatric Institute Institutional Review Board, with participants providing written informed consent. No participant had prior experience with eye-tracking. Participants were paid for completing the study.

Table 1
Demographic and psychopathological characteristics by group.

	SAD group (*n* = 53)		HC group (*n* = 20)		Statistics	
	M (or %)	*SD*	*M (or %)*	*SD*	*Statistic*	*p-value*
Age (years)	29.19	6.64	28.65	6.18	0.31	.75
Gender (% Female)	58.49	-	60.00	-	0.01	.91
Education (years)	15.81	1.86	15.95	2.95	0.24	.81
Race (% White)	39.62	-	50.00	-	0.64	.42
LSAS						
Total score	85.40	14.69	7.50	6.98	22.69	<.001
Fear subscale	44.66	7.66	4.30	4.30	22.22	<.001
Avoidance subscale	40.77	8.73	3.20	3.24	18.69	<.001
HAM-D	5.98	4.48	0.70	1.22	5.22	<.001

Note. SAD = Social Anxiety Disorder; HC = Healthy Control; LSAS = Liebowitz Social Anxiety Scale; HAM-D = Hamilton Rating Scale for Depression.

2.2. Measures

2.2.1. Social anxiety

Social anxiety was measured using the clinician-rated LSAS (Liebowitz, 1987), which lists 24 socially relevant situations. Each situation is rated on two scales ranging 0-3, representing fear and avoidance of the described situation in the past week, yielding a total score and subtotal scores for fear and avoidance. Scores of 30 for non-generalized SAD and 60 for generalized SAD show the best balance between specificity and sensitivity using both the clinician-administered (Mennin et al., 2002) and self-report (Rytwinski et al., 2009) LSAS. The LSAS has strong psychometric properties, including high internal consistency, strong convergent and discriminant validity, and high test-retest reliability (Baker et al., 2002; Fresco et al., 2001; Heimberg et al., 1999). Cronbach's α for the clinician-rated LSAS in the present sample was .98.

2.2.2. Depression

Current depressive symptoms were evaluated using the clinician-rated HAM-D (Hamilton, 1960), a 17-item measure covering core symptoms of depression over the past week. The HAM-D has strong internal consistency and inter-rater and test-retest reliability (Trajkovic et al., 2011). The HAM-D was administered using the Structured Interview Guide for the Hamilton Rating Scale for Depression (SIGH-D; Williams, 1988), which has strong psychometric properties in clinical samples (Williams, 1988). Cronbach's α in the current sample was .80.

2.2.3. Primary and co-morbid diagnoses

Primary and co-morbid diagnoses were assessed using the DSM-5 version of the MINI (Sheehan et al., 1998), a structured diagnostic interview for psychiatric disorders. The MINI is a valid and time-efficient alternative to the Structured Clinical Interview for DSM Patients (SCID-P) and the Composite International Diagnostic Interview (Lecrubier et al., 1997; Sheehan et al., 1997).

2.3. Eye-tracking task

The eye-tracking task was identical to the one used in our previous study of attention allocation in SAD (Lazarov et al., 2016). Color photographs of eight male and eight female actors, each contributing a disgusted and a neutral emotional facial expression, were taken from the Karolinska Directed Emotional Faces database (KDEF; Lundqvist et al., 1998). During each trial a 4 × 4 matrix of 16 faces (one from each actor)—eight with a disgusted expression (i.e., socially threatening stimuli; Staugaard, 2010; Waechter et al., 2014) and eight with a neutral expression—was presented (see Fig. 1). Each face appeared randomly at each matrix position. Furthermore, each actor appeared only once per matrix, each matrix contained eight male and eight female faces, and the four inner faces always contained two disgusted and two neutral expressions.

Each trial began with a fixation-cross, shown until a mandatory fixation of 1000 ms was recorded; this verified that each trial began when the participant's gaze was fixated at the center of the to-be-presented matrix. Next, the matrix was presented for 6000 ms of free viewing, followed by an inter-trial interval of 2000 ms. The entire task included 60 different matrices, administered in two blocks of 30 with a one-minute break in-between. Each block was preceded by a 5-point calibration and 5-point validation of the participant's gaze.

2.4. Eye-tracking measures

Eye-tracking data were processed using EyeLink Data Viewer software (SR Research Ltd.; version 3.1.246). Fixations were defined as at least 100 ms of stable fixation within 1-degree visual angle. For each of the 60 matrices we defined two Areas of Interest (AOIs), one including the eight disgusted expressions (threat AOI) and one including the eight neutral expressions (neutral AOI). As in previous studies using the same task, total dwell time per AOI was calculated by averaging the total

Fig. 1. An example of a single matrix. The eight disgusted faces comprise the threat area of interest (AOI) and the eight neutral faces comprise the neutral AOI.

dwell time (in milliseconds) on each AOI in each matrix across the 60 matrices. First fixation latency was calculated by averaging the latency to first fixations, in milliseconds, for each AOI. First fixation location was measured by counting the number of times the first fixation was in each AOI. First fixation dwell time was computed by averaging first fixation duration, in milliseconds, for each AOI (Lazarov et al., 2016; Lazarov et al., 2018).

2.5. Apparatus

Eye movements were recorded using a remote high-speed Eyelink 1000+ eye tracker (SR Research Ltd., Mississauga, Ontario, Canada) with a sampling rate of 500 Hz. Operating distance to the eye-tracking monitor was about 60–65 cm. The stimuli were presented on a 24-inch monitor with a screen resolution of 1920 × 1080 pixels.

2.6. Procedure

Participants were tested individually in a quiet room at the Anxiety Disorders Clinic, New York State Psychiatric Institute. They were told that they were going to complete a task examining gaze patterns using an eye-tracking apparatus, during which they would be sequentially presented with different matrices of faces. They were also informed that they would need to fixate on a central cross to make each matrix appear, and this contingency was demonstrated. Participants were told to look freely at each matrix in any way they chose until it disappeared. After the task, HC participants were invited to take part in a second session, held approximately two weeks later, while participants with clinical SAD began the RCT as scheduled and were not retested. Session 2 followed the same protocol as described for Session 1, but using new matrices from the same set of actors.

2.7. Data analysis

Independent sample *t*-tests and *Chi-square* tests were used to compare the groups on demographic and clinical measures. To examine group differences in first fixation measures per AOI and total dwell time on the two AOIs, we performed separate mixed-model analyses of variance (ANOVA) with group (SAD, HC) as a between-subject factor and AOI (threat, neutral) as a within-subject factor. Follow-up analysis included simple effect analysis. As groups also differed on baseline depression levels, we conducted analysis of covariance (ANCOVA) for significant findings, entering depression scores from the HAM-D as a covariate.

Reliability was assessed for three variants of the total dwell time measure—namely, dwell time on socially threatening faces, dwell time on neutral faces, and the percentage of dwell time on socially threatening faces out of total dwell time (% dwell time = dwell time on threat stimuli/dwell time on threat + neutral stimuli). Internal consistency was examined for the overall sample ($N = 73$) and separately by group (SAD, HC), using *Cronbach's alpha* and treating each trial (i.e., each matrix) as a single item. For the HC group, test-retest reliability was computed for the three measures using the Pearson (Lazarov et al., 2016; Lazarov et al., 2018; Waechter et al., 2014) and the Spearman's Rank correlation coefficients.

All statistical tests were 2-sided, using α of .05. Effect sizes for significant findings are reported using η^2_p values for ANOVAs and Cohen's *d* for mean comparisons, including 90% effect size confidence interval (CI).

3. Results

3.1. Demographics and clinical characteristics

Demographic and clinical characteristics are described in Table 1. As expected, groups differenced significantly on social anxiety and depression severity but not on demographic variables. The SAD group was characterized by clinically severe social anxiety and mild levels of depression.

3.2. First fixation measures

Non-significant group-by-AOI interaction effects were noted for first fixation latency, $F(1, 71) = .80$, $p = .37$, first fixation location, $F(1, 71) = .04$, $p = .85$, or first fixation dwell time, $F(1, 71) = .47$, $p = .49$. There were no main effects of group or AOI for these measures either.

3.3. Attention allocation (Total dwell time)

Total mean dwell times in milliseconds, by group and AOI (threat, neural), are presented in Fig. 2. Replicating the findings of Lazarov et al. (2016), a significant group-by-AOI interaction effect emerged, $F(1,71) = 6.43$, $p = .01$, $\eta^2_p = .08$, $CI = .01$-$.19$. Follow-up simple effects analysis for the threat AOI revealed more fixation time in the SAD group ($M = 2230$, $SD = 413$) vs. the HC group ($M = 1874$, $SD = 574$), $F(1,71) = 8.63$, $p = .004$, $\eta^2_p = .11$ (*Cohen's d* = 0.71), $CI = .02$-$.23$. Conversely, for the neutral AOI fixation time did not differ between the SAD ($M = 2520$, $SD = 412$) and HC ($M = 2778$, $SD = 830$) groups, $F(1,71) = 3.12$, $p = .08$. Within-group analyses showed that the HC group spent significantly more time fixating on the neutral AOI compared with the threat AOI, $F(1,19) = 9.71$, $p = .006$, $\eta^2_p = .39$, (*Cohen's d* = 1.27), while the SAD group showed the opposite pattern, dwelling significantly longer on the threat AOI compared with the neutral AOI, $F(1,52) = 8.14$, $p = .006$, $\eta^2_p = .14$, (*Cohen's d* = 0.70). The group-by-AOI interaction effect remained significant after entering HAM-D depression scores as a covariate, $F(1, 70) = 4.26$, $p = .009$, $\eta^2_p = .06$, $CI = .001$-$.16$, as did the follow-up analysis for the threat AOI, $F(1,70) = 4.92$, $p = .03$, $\eta^2_p = .07$, $CI = .003$-$.17$.

3.4. Reliability (Internal consistency and test-retest)

Across the groups, internal consistency for total dwell time on threat faces, total dwell time on neutral faces, and the percentage of total dwell time on threat faces was high, with Cronbach's alphas of .99, .99, and .96, respectively. Within-group analysis showed Cronbach's alphas of .87 and .99 for total dwell time on threat faces for the HC and SAD groups, respectively. For total dwell time on neutral faces, Cronbach's alpha was .99 for both groups. Finally, for percentage of total dwell time on threat faces, Cronbach's alphas were .97 and .93, for the HC and SAD groups, respectively.

Two weeks test-retest reliability (see Fig. 3), as indicated by the Pearson correlation coefficient, was significant for all three dwell time measures, namely, total dwell time on threat faces ($r(20) = 0.87$), total dwell time on neutral faces ($r(20) = 0.93$), and the percentage of total dwell time on threat faces ($r(20) = 0.92$), all $ps<.001$. Using the Spearman's Rank correlation coefficient, the two-week test-retest reliability for all three measures remained significant, namely, total dwell time on threat faces ($r(20) = 0.78$), total dwell time on neutral faces ($r(20) = 0.84$), and the percentage of total dwell time on threat faces ($r(20) = 0.72$), all $ps<.001$.

4. Discussion

This study aimed to replicate a previous attention allocation eye-tracking study in social anxiety, using the same task and outcome measures. Participants' first fixations latency, location, and dwell time, as well as their total dwell time, on socially threatening and neutral stimuli was assessed, and the task's reliability—in terms of both internal consistency and two-week test-retest reliability—was examined. The attempt was successful, as results were consistent with the findings from the original study (Lazarov et al., 2016). First, relative to healthy controls the SAD patients demonstrated increased dwell time on socially threatening faces. Moreover, including depression severity as a

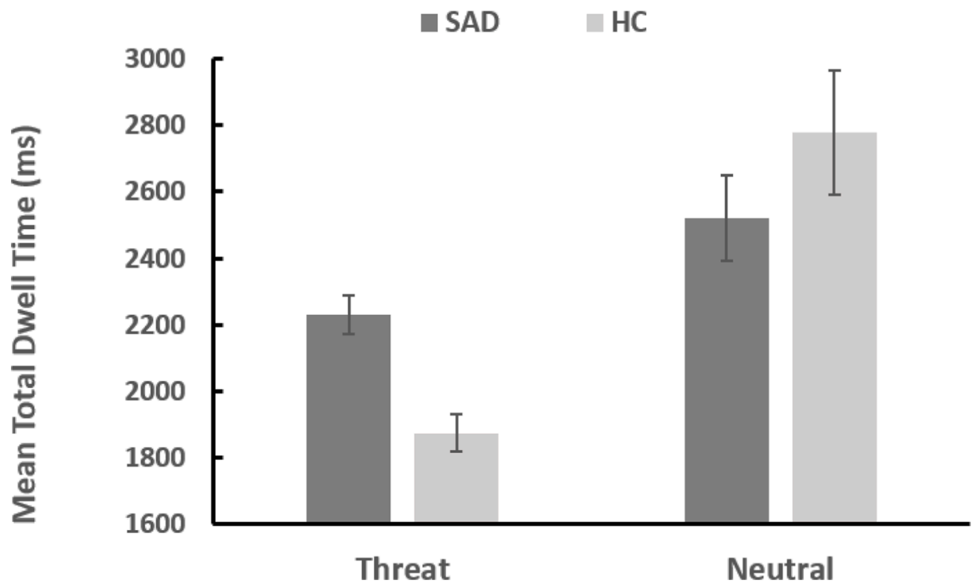

Fig. 2. Mean total dwell time per matrix by AOI and Group. Greater values indicate greater dwell time per matrix in milliseconds. Error bars denote standard error. Clinical social anxiety disorder (SAD) participants spent significantly more time fixating on the threat AOI in comparison to healthy control (HC) participants. There were no significant differences in dwell time on the neutral AOI.

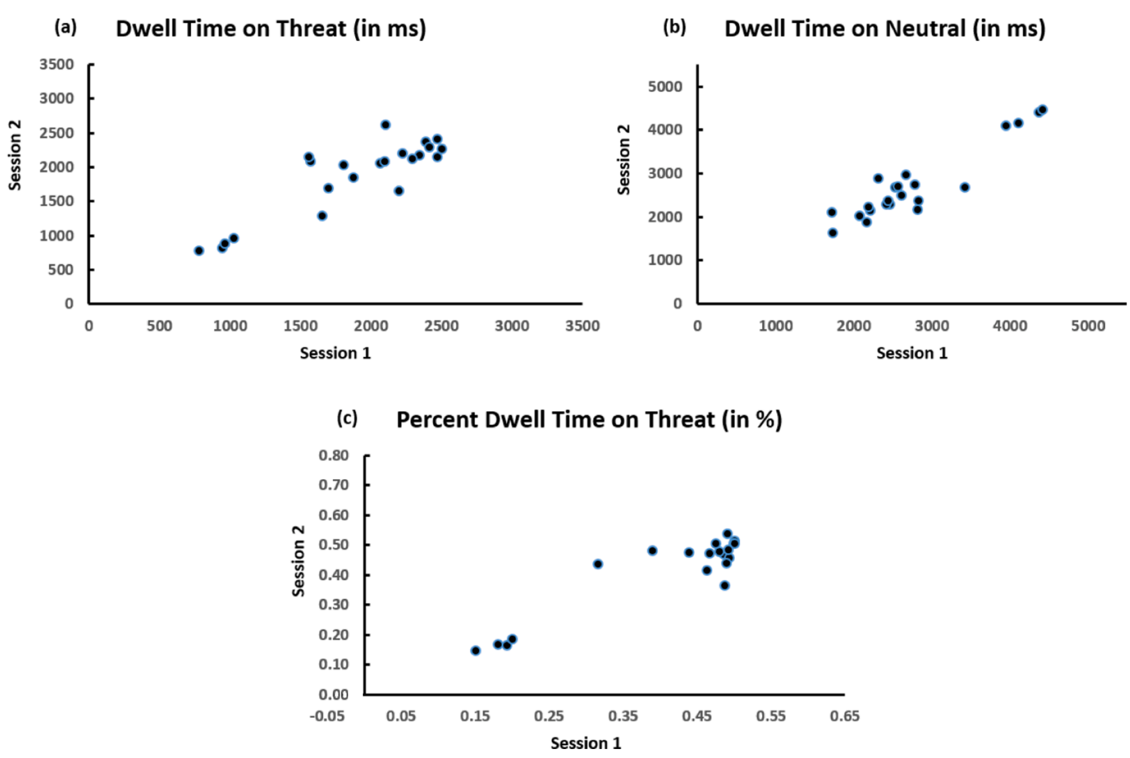

Fig. 3. Test-retest scatter plot (Session 1 on the *X* axis; Session 2 on the *Y* axis) for (a) Dwell time on threat faces (in milliseconds); (b) Dwell time on neutral faces (in milliseconds); and (c) Percent dwell time on threat faces (in percentage).

covariate did not affect this result, suggesting that the differences are mainly driven by social anxiety. Second, none of the first fixation measures differentiated the two groups. Third, the task exhibited acceptable-to-good internal consistency and test-retest reliability. Importantly, while these psychometric results are similar to those reported in the original study, here we extend on that prior work by also reporting acceptable within-group internal consistencies and by extending the time-interval between sessions from one to two weeks.

While the present study replicated the original one by Lazarov et al. (2016), a few critical differences in the study design support the generalizability of the findings. First, the present study was conducted at a US research clinic, recruiting English-speaking participants of varied ethnicity, as opposed to a homogeneous sample of Hebrew-speaking Israeli participants, which increases the generalizability of obtained results by showing the task to be valid across different cultures/nationalities. Second, unlike the original study, in which the non-socially anxious control participants were first-year psychology students who scored at the bottom of the sampling pool (i.e., undergraduate students screened for social anxiety using the LSAS at the beginning of the school year), here control participants were healthy

Journal of Affective Disorders 290 (2021) 169–177

adults who were clinically screened and matched on age, race, and gender distribution to the SAD group. This is important as student participants, especially those scoring extremely low on different measures of psychopathology, may be inherently different from the general population, thereby driving some of the observed group differences (Hanel and Vione, 2016). Finally, in the present study the clinical assessment of social anxiety and depression levels was based on clinician-administered scales (i.e., LSAS, HAM-D) rather than self-report measures, which is valuable because clinician-administered scales allow inquiries in case of ambiguous or vague answers (Trull et al., 2001), reduce the risk for response biases, and are less subject to wording-effects (Moskowitz, 1986; Schwarz, 1999). The consistency of the current and previous results is encouraging and suggests that dwell time on socially threatening stimuli, as measured by eye-tracking, may reflect a stable, replicable, and generalizable measure of attention allocation, with increased dwell time on threat being a reliable characteristic of social anxiety.

In light of replications and strong psychometrics properties, the here reported attentional bias to threat in social anxiety may serve as a target for treatment. Put differently, if the attentional system of patients with social anxiety, as compared with non-socially anxious healthy participants, consistently favors socially threatening cues over more benign ones (e.g., neutral faces), then modifying one's attention away from these stimuli may lead to a corresponding reduction in social anxiety symptoms. This assumption is in line with the disengagement difficulty model of social anxiety (Amir et al., 2003; Buckner et al., 2010; Rapee and Heimberg, 1997; Schofield et al., 2012) suggesting that if less attention is devoted to socially threatening stimuli then less anxiety and fear will ensue. This claim is supported by results from a randomized controlled trial that used a modified version of the current (assessment) task to reduce dwell time on threat faces by reinforcing dwell time on neutral faces with gaze-contingent music-reward. Specifically, SAD patients freely viewed different face matrices (similar to those used in the present study) while music of their choosing played in the background. In the active group (gaze-contingent music reward therapy; GC-MRT) the chosen music played only when fixating on any of the eight neutral faces. Fixating on any threat face stopped the music. Conversely, in the control group the chosen music played continuously, not contingent on participants' gaze patterns. Overall, training included eight 12-minute training sessions. Results showed greater reductions in SAD symptoms in the GC-MRT group, compared with the control group, on both clinician-rated and self-reported social anxiety measures, effects that were maintained at three-month follow-up. Importantly, and most relevant to the present study, the GC-MRT group, but not the control group, reduced dwell-time on threat faces across training sessions, which partially mediated the observed clinical effects. Finally, the GC-MRT group, but not control group, also showed reduced dwell-time on threat faces in the post-treatment assessment task (using faces that were not used in training) reflecting near-transfer generalization effects (Lazarov et al., 2017). Thus, gaze-contingent music reward therapy, or GC-MRT, holds potential as a novel treatment for social anxiety. While showing encouraging results, the aforementioned trial included a small sample size of only 20 patients in each group. Future research should examine the clinical effects of GC-MRT in a larger randomized control trial to further establish its clinical efficacy. Future work could also examine GC-MRT as an adjuvant treatment to cognitive behavioral therapy (CBT) for social anxiety, as prior research has shown a dot probe-based ABMT procedure to augment the effect of CBT for social anxiety (Lazarov et al., 2018).

Our findings are also broadly consistent with previous studies that used modified versions of the original task to assess participants with different forms of psychopathology, supporting the tasks' reliability and utility in assessing attention. First, in a version of the task that used matrices with sad and happy faces, both depressed students and patients with MDD were found to dwell longer on sad faces than non-depressed students, with the task showing good psychometric properties (Lazarov et al., 2018). A recent follow-up study used separate blocks of sad-neutral and happy-neutral matrices and showed that depressed individuals dwelled longer on sad faces than did never-depressed participants, with no difference for dwell times on happy faces (Klawohn et al., 2020). This study also reported good-to-excellent within-group internal consistencies for dwell-time on sad and neutral faces in the sad-neutral block, and for happy and neutral faces in the happy-neutral block. Second, another study presented participants with a modified version of the task that included matrices presenting alcoholic and non-alcoholic beverages (Soleymani et al., 2020). Participants who scored higher on craving and alcohol problems also dwelled longer on the alcoholic beverages, with the dwell time-based measure demonstrating excellent internal reliability and considerable stability as indexed by an eight-days test-retest reliability. Finally, using a study in pediatric anxiety showed that treatment-seeking, medication-free youths (age 8-to-18) with anxiety disorders dwelled longer on negative vs. non-negative faces compared with healthy youths, with the task demonstrating adequate reliability (Abend et al., 2020). Reliability was further established in a community sample of healthy children aged 6-to-9, with the task showing good psychometric properties, comparable to those noted for adults (Chong and Meyer, 2020).

The current results should be considered in light of a few limitations. First, while the present study examined reliability via internal consistency for the entire sample and within each group, test-retest reliability was only assessed for nonanxious healthy participants, as patients with SAD were scheduled to begin the RCT for which they were recruited as planned. As test-retest reliability may vary depending on the (sub) population being examined, the obtained test-retest measure of the HC group cannot be generalized with certainty to the SAD group (Hedge et al., 2018; Lebel and Paunonen, 2011; Parsons et al., 2018). Thus, future research should examine test-retest reliability of the total dwell time measure in socially anxious participants. Second, although significant group differences were observed in total dwell time on disgust faces, the HC group in the current study was small, potentially limiting our power to detect group differences for dwell time on neutral faces. Third, as our main goal was to replicate the original Lazarov et al. (2016) study, we opted to use the same emotional stimuli used in the original study (i.e., disgust vs neutral faces). Thus, we did not include negative emotional expressions other than disgust (e.g., anger, fear), nor did we examine attention allocation patterns to positive stimuli (e.g., using happy-neutral matrices), which limits the generalization of present results to different emotions. Future research should therefore explore whether reliability and group differences extend to other emotional expressions. Fourth, SAD participants exhibited some comorbidities (i.e., 18 participants with comorbid MDD and five with comorbid GAD) also differentiating the SAD group from the HC group. Thus, the presence of comorbidity confounds the ability to attribute group differences to SAD alone. Indeed, prior research has shown threat-related attention biases in depression (Lazarov et al., 2018; Suslow et al., 2020) and GAD (Goodwin et al., 2017). However, the fact that results remained significant after controlling for depression levels, and the small number of participants with GAD, increase our confidence in obtained results. Relatedly, as we did not include a measure of ADHD symptoms we could not control for this in our analyses. Future research should address these two limitations. Finally, while present results show replicability in SAD and specificity over co-morbid depressive symptoms, the same task may yield similar results also in other psychopathologies in which attention biases are implicated (e.g., MDD, PTSD, GAD). Hence, replicating the present study across other disorders may help in determining the specificity of findings to social anxiety.

This study strengthens our confidence in total dwell time on threat as a reliable, replicable, and generalizable measure of attention allocation in SAD, by showing that socially anxious individuals dwell longer on threatening social stimuli, and by demonstrating that this finding, when assessed with eye-tracking, is reliable both within sessions (in SAD and HC) and across sessions (although this was only tested in the HC group). Hence, this measure could serve as a viable index of enhanced attention allocation to threat in adults with social anxiety, suggesting that modifying attention allocation habits may be an effective treatment for social anxiety disorders.

Author Contributers

Amit Lazarov: Conceptualization, Investigation, Supervision, Formal analysis, Writing - original draft and revised version

Dana Basel: Formal analysis, Writing - review & editing – original and revised manuscript

Sarah Dolan: Data curation, Project administration

Daniel G. Dillon: Writing - review & editing – original and revised manuscript

Diego A. Pizzagalli: Writing - review & editing – original and revised manuscript

Franklin R. Schneier: Conceptualization, Funding acquisition, Investigation, Resources, Supervision, Writing - review & editing – original and revised manuscript

Authors declaration

We declare that this manuscript is original and that it has not been published before or has been posted on a web site and that it is not currently being considered for publication elsewhere.

Role of the Funding source

This research was supported by National Institute of Mental Health (NIMH), Grant number R61MH116089 to Franklin R. Schneier.

Daniel G. Dillon was partially supported by NIMH, Grant number R61MH116089.

Diego A. Pizzagalli was partially supported by NIMH, Grant number R37 MH068376.

Amit Lazarov was partially supported by NIMH, Grant number R61MH116089, and by the Israel Science Foundation, Grant number 374/20.

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The funding organization had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Institutional Board Review

The authors assert that all procedures contributing to this work comply with APA ethical standards and with the Helsinki Declaration of 1975, as revised in 2008. All procedures were approved by the New York State Psychiatric Institute Institutional Review Board

Declaration of Competing Interest

Dr. Schneier has received research support from Otsuka Pharmaceuticals and Compass Pathways. Over the past 3 years, Dr. Pizzagalli has received consulting fees from BlackThorn Therapeutics, Boehringer Ingelheim, Compass Pathway, Concert Pharmaceuticals, Engrail Therapeutics, Otsuka Pharmaceuticals, and Takeda Pharmaceuticals; one honorarium from Alkermes, and research funding from NIMH, Dana Foundation, Brain and Behavior Research Foundation, Millennium Pharmaceuticals. In addition, he has received stock options from BlackThorn Therapeutics. No funding from these entities was used to support the current work, and all views expressed are solely those of the authors. All other authors declare that they have no financial disclosures in association with this work. We wish to confirm that there are no other known potential conflicts of interest associated with this publication.

Acknowledgments

We greatly appreciate and thank the participants who participated in this study.

References

Abend, R., Bajaj, M.A., Matsumoto, C., Yetter, M., Harrewijn, A., Cardinale, E.M., Kircanski, K., Lebowitz, E.R., Silverman, W.K., Bar-Haim, Y., Lazarov, A., Leibenluft, E., Brotman, M., Pine, D.S., 2020. Converging multi-modal evidence for implicit threat-related bias in pediatric anxiety disorders. J. Abnorm. Child Psychol. 1–14. https://doi.org/10.1007/s10802-020-00712-w.

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, 1325–1335. https://doi.org/10.1016/S0005-7967(03)00039-1.

Amir, N., Taylor, C.T., 2012. Combining computerized home-based treatments for generalized anxiety disorder: an attention modification program and cognitive behavioral therapy. Behav. Ther. 43, 546–559. https://doi.org/10.1016/j.beth.2010.12.008.

Armstrong, T., Olatunji, B.O., 2012. Eye tracking of attention in the affective disorders: a meta-analytic review and synthesis. Clin. Psychol. Rev. 32, 704–723. https://doi.org/10.1016/j.cpr.2012.09.004.

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, 701–715. https://doi.org/10.1016/S0005-7967(01)00060-2.

Bantin, T., Stevens, S., Gerlach, A.L., Hermann, C., 2016. What does the facial dot-probe task tell us about attentional processes in social anxiety? A systematic review. J. Behav. Ther. Exp. Psychiatry 50, 40–51. https://doi.org/10.1016/j.jbtep.2015.04.009.

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

Brown, H.M., Eley, T.C., Broeren, S., MacLeod, C., Rinck, M., Hadwin, J.A., Lester, K.J., 2014. Psychometric properties of reaction time based experimental paradigms measuring anxiety-related information-processing biases in children. J. Anxiety Disord. 28, 97–107. https://doi.org/10.1016/j.janxdis.2013.11.004.

Buckner, J.D., Maner, J.K., Schmidt, N.B., 2010. Difficulty disengaging attention from social threat in social anxiety. Cognit. Ther. Res. 34, 99–105. https://doi.org/10.1007/s10608-008-9205-y.

Chen, N.T.M., Clarke, P.J.F., 2017. Gaze-based assessments of vigilance and avoidance in social anxiety: a review. Curr. Psychiatry Rep. 19 (9), 59. https://doi.org/10.1007/s11920-017-0808-4.

Chong, L., Meyer, A., 2020. Psychometric properties of threat-related attentional bias in young children using eye-tracking. Psychophysiology 57. https://doi.org/10.1002/dev.22053. S45-S45https://doi.org/.

Clark, D.M., Wells, A., 1995. A cognitive model of social phobia. In: Heimberg, R.G., Liebowitz, M.R., Hope, D.A., Schnier, F.R. (Eds.). Social Phobia: Diagnosis, Assessment and Treatment. Guildford Press, New York, pp. 69–93.

Elias, S., Massad, R., Lazarov, A., 2021. Visual attention patterns of socially anxious individuals when using facebook – an eye tracking study. Behav. Ther. https://doi.org/10.1016/j.beth.2020.12.007.

Field, A.P., Lester, K.J., 2010. Is there room for 'development' in developmental models of information processing biases to threat in children and adolescents? Clin. Child Fam. Psychol. Rev. 13, 315–332. https://doi.org/10.1007/s10567-010-0078-8.

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, 1025–1035. https://doi.org/10.1017/S0033291701004056.

Goodwin, H., Yiend, J., Hirsch, C.R., 2017. Generalized Anxiety Disorder, worry and attention to threat: a systematic review. Clin. Psychol. Rev. 54, 107–122. https://doi.org/10.1016/j.cpr.2017.03.006.

Hamilton, M., 1960. A rating scale for depression. J. Neurol. Neurosurg, Psychiatry 23, 56–62. https://doi.org/10.1136/jnnp.23.1.56 https://doi.org/.

Hanel, P.H.P., Vione, K.C., 2016. Do student samples provide an accurate estimate of the general public? PLoS One 11 (12). https://doi.org/10.1371/journal.pone.0168354 e0168354. https://doi.org/.

Hedge, C., Powell, G., Sumner, P., 2018. The reliability paradox: why robust cognitive tasks do not produce reliable individual differences. Behav. Res. Methods 50, 1166–1186. https://doi.org/10.3758/s13428-017-0935-1.

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

Heimberg, R.G., Horner, K.J., Juster, H.R., Safren, S.A., Brown, E.J., Schneier, F.R., Liebowitz, M.R., 1999. Psychometric properties of the Liebowitz social anxiety scale. Psychol. Med. 29, 199–212. https://doi.org/10.1017/S0033291798007879.

In-Albon, T., Schneider, S., 2010. Using eye tracking methodology in children with anxiety disorders. Hadwin, J.A., Field, A.P. (Eds.). Information Processing Biases and Anxiety: a Developmental Perspective. John Wiley & Sons Ltd, Chichester, UK, pp. 129–149.

Klawohn, J., Bruchnak, A., Burani, K., Meyer, A., Lazarov, A., Bar-Haim, Y., Hajcak, G., 2020. Aberrant attentional bias to sad faces in depression and the role of stressful life events: evidence from an eye-tracking paradigm. Behav. Res. Ther. 135, 103762 https://doi.org/10.1016/j.brat.2020.103762.

Kuo, J.R., Goldin, P.R., Werner, K., Heimberg, R.G., Gross, J.J., 2011. Childhood trauma and current psychological functioning in adults with social anxiety disorder. J. Anxiety Disord. 25, 467–473. https://doi.org/10.1016/j.janxdis.2010.11.011.

Lazarov, A., Abend, R., Bar-Haim, Y., 2016. Social anxiety is related to increased dwell time on socially threatening faces. J. Affect. Disord. 193, 282–288. https://doi.org/10.1016/j.jad.2016.01.007.

Lazarov, A., Bar-Haim, Y., 2020. Emerging domain-based treatments for pediatric anxiety disorders. Biol. Psychiatry. https://doi.org/10.1016/j.biopsych.2020.08.030.

Lazarov, A., Ben-Zion, Z., Shamai, D., Pine, D.S., Bar-Haim, Y., 2018. Free viewing of sad and happy faces in depression: a potential target for attention bias modification. J. Affect. Disord. 238, 94–100. https://doi.org/10.1016/j.jad.2018.05.047.

Lazarov, A., Marom, S., Yahalom, N., Pine, D.S., Hermesh, H., Bar-Haim, Y., 2018. Attention bias modification augments cognitive-behavioral group therapy for social anxiety disorder: a randomized controlled trial. Psychol. Med. 48 (13), 2177–2185. https://doi.org/10.1017/S003329171700366X.

Lazarov, A., Pine, D.S., Bar-Haim, Y., 2017. Gaze-Contingent music reward therapy for social anxiety disorder: a randomized controlled trial. Am. J. Psychiatry 174, 649–656. https://doi.org/10.1176/appi.ajp.2016.16080894.

Lazarov, A., Suarez-Jimenez, B., Tamman, A., Falzon, L., Zhu, X., Edmondson, D.E., Neria, Y., 2019. Attention to threat in posttraumatic stress disorder as indexed by eye-tracking indices: a systematic review. Psychol. Med. 49, 705–726. https://doi.org/10.1017/S0033291718002313.

Lebel, E.P., Paunonen, S.V., 2011. Sexy but often unreliable: the impact of unreliability on the replicability of experimental findings with implicit measures. Pers. Soc. Psychol. B 37, 570–583. https://doi.org/10.1177/0146167211400619.

Lecrubier, Y., Sheehan, D.V., Weiller, E., Amorim, P., Bonora, I., Sheehan, K.H., Janavs, J., Dunbar, G.C., 1997. The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. Eur. Psychiatry 12, 224–231. https://doi.org/10.1016/S0924-9338(97)83296-8.

Liebowitz, M.R., 1987. Social phobia. Modern Problems in Pharmacopsychiatry 22, 141–173. https://doi.org/ 10.1159/000414022.

Lilienfeld, S.O., Strother, A.N., 2020. Psychological measurement and the replication crisis: four sacred cows. Can. Psychol. https://doi.org/10.1037/cap0000236.

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. Depress. Anxiety 32, 383–391. https://doi.org/10.1002/da.22344.

Lundqvist, D., Flykt, A., Öhman, A., 1998. KDEF: the Karolinska Directed Emotional Faces [CD-ROM]. Karolinska Institutet, Department of Clinical Neuroscience, Psychology Section, Stockholm.

McNally, R.J., 2019. Attentional bias for threat: crisis or opportunity? Clin. Psychol. Rev. 69, 4–13. https://doi.org/10.1016/j.cpr.2018.05.005.

Mennin, D.S., Fresco, D.M., Heimberg, R.G., Schneier, F.R., Davies, S.O., Liebowitz, M.R., 2002. Screening for social anxiety disorder in the clinical setting: using the Liebowitz Social Anxiety Scale. J. Anxiety Disord. 16, 661–673. https://doi.org/10.1016/S0887-6185(02)00134-2.

Mogg, K., Waters, A.M., Bradley, B.P., 2017. Attention Bias Modification (ABM): review of effects of multisession abm training on anxiety and threat-related attention in high-anxious individuals. Clin. Psychol. Sci. 5, 698–717. https://doi.org/10.1177/2167702617696359.

Morrison, A.S., Heimberg, R.G., 2013. Social anxiety and social anxiety disorder. Annu. Rev. Clin. Psychol. 9, 249–274. https://doi.org/10.1146/annurev-clinpsy-050212-185631.

Moskowitz, D.S., 1986. Comparison of self-reports, reports by knowledgeable informants, and behavioral observation data. J. Pers. 54, 294–317. https://doi.org/10.1111/j.1467-6494.1986.tb00396.x.

Mulckhuyse, M., Crombez, G., Van der Stigchel, S., 2013. Conditioned fear modulates visual selection. Emotion 13, 529–536. https://doi.org/10.1037/a0031076.

Nissens, T., Failing, M., Theeuwes, J., 2017. People look at the object they fear: oculomotor capture by stimuli that signal threat. Cognit. Emot. 31 (8), 1707–1714. https://doi.org/10.1080/02699931.2016.1248905.

Open Science Collaboration, 2015. Estimating the reproducibility of psychological science. Science 349, 6215. https://doi.org/10.1126/science.aac4716 aac4716https://doi.org/.

Parsons, S., Kruijt, A., Fox, E., 2018. Psychological Science needs a standard practice of reporting the reliability of cognitive behavioural measurements. Adv. Methods Pract. Psychol. Sci. 2 (4), 378–395. https://doi.org/10.1177/2515245919879695.

Preciado, D., Munneke, J., Theeuwes, J., 2017. Was that a threat? Attentional biases by signals of threat. Emotion 17, 478–486. https://doi.org/10.1037/emo0000246.

Price, R.B., Greven, I.M., Siegle, G.J., Koster, E.H.W., De Raedt, R., 2016. A novel attention training paradigm based on operant conditioning of eye gaze: preliminary findings. Emotion 16, 110–116. https://doi.org/10.1037/emo0000093 https://doi.org/.

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

Rodebaugh, T.L., Scullin, R.B., Langer, J.K., Dixon, D.J., Huppert, J.D., Bernstein, A., Zvielli, A., Lenze, E.J., 2016. Unreliability as a threat to understanding psychopathology: the cautionary tale of attentional bias. J. Abnorm. Psychol. 125, 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., Stein, M.B., Hofmann, S.G., 2009. Screening for social anxiety disorder with the self-report version of the Liebowitz social anxiety scale. Depress. Anxiety 26, 34–38. https://doi.org/10.1002/da.20503.

Schmidt, L.J., Belopolsky, A.V., Theeuwes, J., 2015. Attentional capture by signals of threat. Cognit. Emot. 29, 687–694. https://doi.org/10.1080/02699931.2014.924484.

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

Schofield, C.A., Johnson, A.L., Inhoff, A.W., Coles, M.E., 2012. Social anxiety and difficulty disengaging threat: evidence from eye-tracking. Cognit. Emot. 26, 300–311. https://doi.org/10.1080/02699931.2011.602050.

Schwarz, N., 1999. Self-reports – how the questions shape the answers. Am. Psychol. 54, 93–105. https://doi.org/10.1037/0003-066X.54.2.93.

Sears, C., Quigley, L., Fernandez, A., Newman, K., Dobson, K., 2019. The reliability of attentional biases for emotional images measured using a free-viewing eye-tracking paradigm. Behav. Res. Methods 51, 2748–2760. https://doi.org/10.3758/s13428-018-1147-z.

Sheehan, D.V., Lecrubier, Y., Sheehan, K.H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., Dunbar, G.C., 1998. The Mini-International Neuropsychiatric Interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J. Clin. Psychiatry 59, 22–33.

Sheehan, D.V., Lecrubier, Y., Sheehan, K.H., Janavs, J., Weiller, E., Keskiner, A., Schinka, J., Knapp, E., Sheehan, M.F., Dunbar, G.C., 1997. The validity of the Mini International Neuropsychiatric Interview (MINI) according to the SCID-P and its reliability. Eur. Psychiatry 12, 232–241. https://doi.org/10.1016/S0924-9338(97)83297-X.

Skinner, I.W., Hubscher, M., Moseley, G.L., Lee, H., Wand, B.M., Traeger, A.C., Gustin, S. M., McAuley, J.H., 2017. The reliability of eyetracking to assess attentional bias to threatening words in healthy individuals. Behav. Res. Methods 1–15. https://doi.org/10.3758/s13428-017-0946-y.

Soleymani, A., Ivanov, Y., Mathot, S., de Jong, P.J., 2020. Free-viewing multi-stimulus eye tracking task to index attention bias for alcohol versus soda cues: Satisfactory reliability and criterion validity. Addict. Behav. 100, 106117 https://doi.org/10.1016/j.addbeh.2019.106117.

Spiller, T.R., Levi, O., Neria, Y., Suarez-Jimenez, B., Bar-Haim, Y., Lazarov, A., 2020. On the validity of the centrality hypothesis in cross-sectional between-subject networks of psychopathology. BMC Med. 18, 297. https://doi.org/10.1186/s12916-020-01740-5.

Staugaard, S.R., 2009. Reliability of two versions of the dot-probe task using photographic faces. Psychol. Sci. 51, 339–350.

Staugaard, S.R., 2010. Threatening faces and social anxiety: a literature review. Clin. Psychol. Rev. 30, 669–690. https://doi.org/10.1016/j.cpr.2010.05.001.

Suslow, T., Hußlack, A., Kersting, A., Bodenschatz, C.M., 2020. Attentional biases to emotional information in clinical depression: a systematic and meta-analytic review of eye tracking findings. J. Affect. Disord. 274, 632–642. https://doi.org/10.1016/j.jad.2020.05.140.

Trajkovic, G., Starcevic, V., Latas, M., Lestarevic, M., Ille, T., Bukumiric, Z., Marinkovic, J., 2011. Reliability of the Hamilton Rating Scale for Depression: a meta-analysis over a period of 49 years. Psychiatry Res. 189, 1–9. https://doi.org/10.1016/j.psychres.2010.12.007.

Trull, T.J., Widiger, T.A., Burr, R., 2001. A structured interview for the assessment of the Five-Factor Model of personality: Facet-level relations to the axis II personality disorders. J. Pers. 69, 175–198. https://doi.org/10.1111/1467-6494.00141.

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

Waechter, S., Nelson, A.L., Wright, C., Hyatt, A., Oakman, J., 2014. Measuring attentional bias to threat: Reliability of dot probe and eye movement indices. Cogni. Ther. Res. 38, 313–333. https://doi.org/10.1007/s10608-013-9588-2.

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

Wiggins, B.J., Chrisopherson, C.D., 2019. The replication crisis in psychology: an overview for theoretical and philosophical psychology. J. Theor. Philos. Psychol. 39, 202–217. https://doi.org/10.1037/teo0000137.

Williams, J.B.W., 1988. A structured interview guide for the Hamilton depression rating-scale. Arch. Gen. Psychiatry 45, 742–747. https://doi.org/10.1001/archpsyc.1988.01800320058007.