Journal of Anxiety Disorders 100 (2023) 102789

Contents lists available at ScienceDirect

Journal of Anxiety Disorders

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

The free-viewing matrix task: A reliable measure of attention allocation in psychopathology

Dana Shamai-Leshem[a,b,*], Rany Abend[c], Gal Arad[b], Omer Azriel[b], Lyndsey Chong[d], Peter de Jong[e], Chelsea Dyan Gober Dykan[b], Greg Hajcak[d], Julia Klawohn[d,f], Alexandria Meyer[d], Klavdia Neophytou[g], Yuval Neria[h], Georgia Panayiotou[g], Franklin Schneier[h], Ali Soleymani[e,i], Noga Yair[a], Daniel S. Pine[c], Yair Bar-Haim[a,b,1], Amit Lazarov[b,1]

[a] Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel
[b] School of Psychological Sciences, Tel Aviv University, Tel Aviv, Israel
[c] Emotion and Development Branch, National Institute of Mental Health, National Institutes of Health, USA
[d] Department of Psychology, Florida State University, Tallahassee, FL, USA
[e] Department of Clinical Psychology & Experimental Psychopathology, University of Groningen, Groningen, the Netherlands
[f] Department of Medicine, MSB Medical School Berlin, Berlin, Germany
[g] Department of Psychology, University of Cyprus, Nicosia, Cyprus
[h] Department of Psychiatry, Columbia University Irving Medical Center and New York State Psychiatric Institute, New York, NY, USA
[i] Centre of Education and Learning, Technical University of Delft, Delft, the Netherlands

ARTICLE INFO

Keywords:
Attention bias
Eye-tracking
Gaze-tracking
Psychometrics
Test-retest reliability

ABSTRACT

Aberrant attention allocation has been implicated in the etiology and maintenance of a range of psychopathologies. However, three decades of research, relying primarily on manual response-time tasks, have been challenged on the grounds of poor reliability of its attention bias indices. Here, in a large, multisite, international study we provide reliability information for a new eye-tracking-based measure of attention allocation and its relation to psychopathology and age. Data from 1567 participants, across a wide range of psychiatric diagnoses and ages, were aggregated from nine sites around the world. Of these, 213 participants also provided retest data. Acceptable overall internal consistency and test-retest reliability were observed among adult participants (Cronbach's alpha = 0.86 and $r(213) = 0.89$, respectively), as well as across all examined psychopathologies. Youth demonstrated lower internal consistency scores (Cronbach's alpha = 0.65). Finally, the percent dwell time index derived from the task statistically differentiated between healthy participants and participants diagnosed with social anxiety disorder, major depression, and post-traumatic stress disorder. These results potentially address a long-standing reliability crisis in this research field. Aberrant attention allocation patterns in a variety of psychiatric disorders may be targeted with the hope of affecting symptoms. The attention allocation index derived from the matrix task offers reliable means to measure such cognitive target engagement in clinical contexts.

1. Introduction

Allocation of greater attentional resources to disorder-related stimuli over neutral stimuli has been implicated in the etiology and maintenance of various psychiatric disorders including anxiety, depression, addiction, and post-traumatic stress, as reflected in a vast literature (Bar-Haim et al., 2007; Gade et al., 2022; Klanecky Earl et al., 2020; Peckham et al., 2010; Shamai-Leshem et al., 2022). Various response-time (RT)-based paradigms have been utilized in the assessment of attention bias, including the dot probe, emotional Stroop, exogenous cueing task, and visual search. The most used approach to measure attention bias is the dot probe, which is based on the underlying

* Correspondence to: Tel Aviv University, Sagol School of Neuroscience, Tel Aviv 69978, Israel.
E-mail address: danasham@mail.tau.ac.il (D. Shamai-Leshem).
[1] This work reflects equal contribution of the last two authors.

https://doi.org/10.1016/j.janxdis.2023.102789
Received 8 November 2022; Received in revised form 25 September 2023; Accepted 12 October 2023
Available online 27 October 2023

assumption that people respond faster to probes presented in attended spatial locations relative to unattended locations. Despite its widespread use, the dot probe task, along with other RT-based tasks, had faced criticism for capturing only a snapshot of a single end point of the attention process, and hence were deemed limited in their ability to provide comprehensive information on underlying attentional mechanisms (Armstrong & Olatunji, 2012; Jiang & Vartanian, 2020; Mogg & Bradley, 1998). In addition, concerns about the reliability of RT-based attentional bias measures, which are typically characterized by poor psychometric properties (Lilienfeld & Strother, 2020; McNally, 2019; Rodebaugh et al., 2016; Streiner, 2016), have further stalled progress in this promising line of research (McNally, 2019).

In recent years, eye-tracking technology has been increasingly used as an alternative attentional assessment approach. Despite preliminary evidence for attention bias in various psychopathologies (Armstrong & Olatunji, 2012; Lazarov, Suarez-Jimenez, Tamman et al., 2019; Lisk et al., 2020), description of the reliability of attention eye-tracking tasks and measures has been rather limited. Extant studies that did explore the reliability of such tasks report encouraging results (e.g., Sanchez et al., 2017; Sears et al., 2019; Skinner et al., 2018; Waechter et al., 2014). For example, Sears et al. (2019) examined the internal consistency and test–retest reliability of eye-tracking-based attention indices derived from a free-viewing paradigm. Results showed moderate-to-excellent internal consistency for total dwell time across an 8-second presentation, which diminished for shorter time intervals. Moderate-to-high test-retest was also found when using naturalistic images as stimuli, and a lower test-retest was noted for face images. Skinner et al. (2018) tested the reliability of 12 eye-tracking-based attention bias indices reflecting overall attention, early attention, and late attention processes in a free-viewing task. Internal consistency was high across most of the attention bias indices (Cronbach's alpha ranging .70 to .99), but lower for first fixation duration. Waechter et al. (2014) examined the internal consistency of several eye-tracking-based attention indices while participants completed the dot-probe task, including the proportion of viewing time spent fixating on emotional stimuli (angry, disgust, and happy facial expressions) relative to neutral facial expression, as well as first fixation indices. The results of this study showed excellent reliability, with Cronbach's alphas exceeding .94 for the former and low reliability for the latter. While providing important insights regarding the reliability of eye-tracking-based attention allocation indices, extant studies have mostly used small non-clinical samples.

Here we use a large, multisite, international sample to examine the reliability of an attention allocation index derived from a free-viewing eye-tracking-based task – *the matrix task*. In this task, first described by Lazarov et al. (2016), and later applied to various psychopathologies (e.g., Klawohn et al., 2020; Lazarov et al., 2021; Soleymani et al., 2020), participants freely view matrices of 16 photographs of two types: eight with neutral or positive content and eight with disorder-related content (e.g., happy vs. dysphoric faces in depression, non-alcoholic vs. alcoholic beverages in alcohol-dependency). The proportion of dwell time on disorder-relevant content (relative to total dwell time, DT%), is used to index attentional allocation patterns.

In the current study, we pooled subject-level data from nine different sites world-wide to estimate the internal consistency and test-retest reliability of the DT% index, both overall and within specific psychopathologies and stimulus types. In addition, because the matrix task and the DT% index were developed in the context of psychopathology and clinical research, we leveraged the extant data to compare healthy participants and participants with various diagnosed psychopathologies. Finally, prior attentional research has identify differences in attention allocation across age (Mather & Carstensen, 2016; Mogoaşe et al., 2014; Price et al., 2016). Therefore, we also explored the developmental aspects of DT% by computing its association with age, and whether psychopathology (present, absent) further modifies the age-DT% association.

2. Method

2.1. Participants

Data of 1567 participants (M_{age} = 25.43 years, *SD* = 11.03, *range* = 6 – 73, 777 females) were aggregated from nine laboratories (see Table S1 for sample characteristics by site). Of those, 707 were unselected (i.e., without determined inclusion criteria), 140 verified as healthy via formal clinical interview, 193 were diagnosed with social anxiety disorder (SAD), 122 with major depressive disorder (MDD), 37 with post-traumatic stress disorder (PTSD), and 61 with a pediatric anxiety disorder (SAD, generalized anxiety disorder, or separation anxiety disorder, which are highly co-morbid in pediatric samples; see Table S2 for demographic characteristics by group). Two-hundred and thirteen participants from laboratories in Israel, the Netherlands, and the United States (M_{age} = 30.84 years, *SD* = 11.76, *range* = 19–73, 137 females) also provided one-week retest data. Institutional review board approval was obtained independently per site. Written informed consent was provided by all participants.

2.2. The free-viewing matrix task

Fig. 1 depicts an example of a single trial of the matrix task (Lazarov et al., 2016). Each trial begins with a centrally presented fixation cross shown until a fixation of 1000 ms is recorded, thereby confirming that each trial begins only when participant's gaze is fixated at the matrix's center. Then a matrix of 16 pictures is presented for 6000 ms, followed by an inter-trial-interval of 2000 ms. Participants are instructed to view each matrix in any way they choose.

Each matrix presents 16 chromatic pictures of two equally represented stimuli categories (e.g., eight disgusted and eight neutral faces; eight alcoholic and eight non-alcoholic beverages). Pictures' location within matrices is random, while ensuring that the four inner locations always contain two pictures of each type. For face matrices, two additional qualifications were kept: a) each actor appeared only once in each matrix; and b) each matrix contained eight male and eight female faces. The task typically presents one or two blocks of 30 trials each. [2]

For all current analyses, two areas of interest (AOIs) were defined for each matrix in each relevant contrast: one for the relatively negative picture category, and one for the relatively positive picture category (e.g., in a matrix containing dysphoric vs. neutral faces the neutral faces would be the relatively positive AOI. Conversely, in happy-neutral matrices, neutral faces would be the relatively negative AOI; Basel et al., 2022). In matrices containing pictures of objects, the stimuli towards which attention is biased in the context of the relevant disorder were considered relatively negative (e.g., in the context of alcohol abuse, pictures of alcoholic beverages were considered as the relatively negative stimuli compared to pictures of soft drinks). Fixations were defined as at least 100 ms of stable fixation within 1-degree visual angle. DT% on the relatively negative AOI was calculated as the proportion of time participants dwelled on that AOI relative to dwell time on both AOIs (i.e., dwell time on relatively negative AOI + relatively positive AOI). A score > 50 % reflects longer dwelling on the relatively negative AOI, whereas a score < 50 % reflects longer dwelling on the relatively positive AOI.

2.3. Data Analysis

Internal consistency and test-retest reliability were calculated for the full sample, for broad stimuli categories (i.e., faces, non-faces), and for

[2] Deviations from this practice were noted in two cases: (1) a task that contained 54 matrices and included pictures of alcoholic and non-alcoholic beverages, and (2) a task that contained 21 matrices and included pictures of high- and low-calorie food. See Table 1.

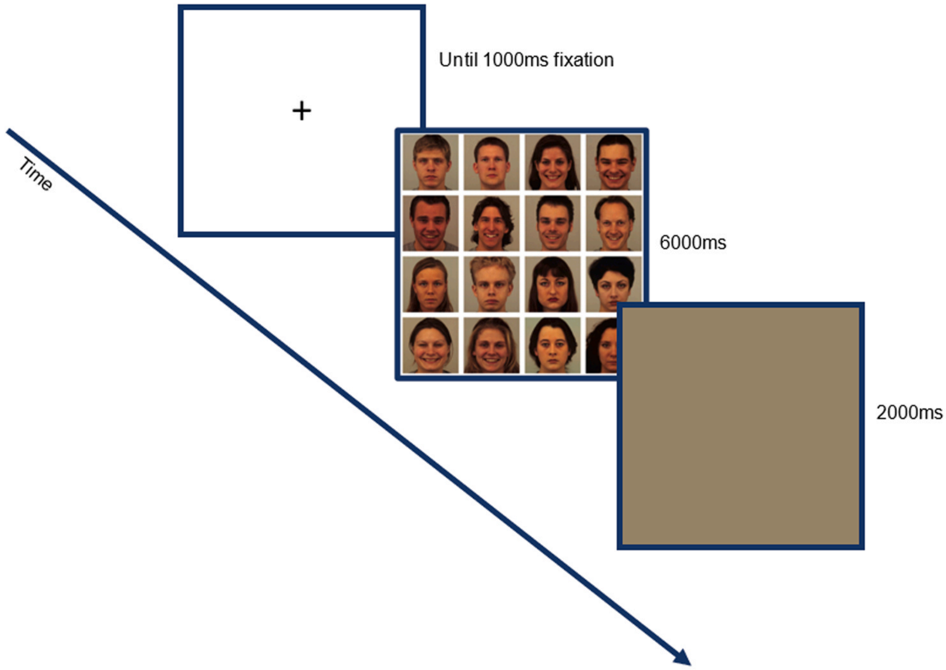

Fig. 1. An example of a single trial in the free-viewing matrix task. Face stimuli were taken from the Karolinska Directed Emotional Faces stimulus set.

each specific contrast type (e.g., angry-neutral faces; alcoholic-nonalcoholic beverages etc.). Internal consistency and test-retest reliability were also calculated per diagnostic group (e.g., Healthy, PTSD, MDD, SAD, and pediatric anxiety). Reliability by gender was also calculated. Internal consistency was calculated separately for one block (30 matrices) and two blocks (60 matrices) using Cronbach's alpha, considering each trial as a single item. In cases of missing data, internal consistency was calculated using list-wise deletion, excluding missing records from analysis. Following recommended values for cognitive tasks (Barch et al., 2008; Bland & Altman, 1997), we regarded Cronbach's alpha ≥ 0.7 as the acceptable range. Test-retest reliability was assessed using Pearson's correlation coefficient, with $r \geq 0.8$ considered as indicative of acceptable test-retest reliability (Kline, 2000; Polit, 2014). To examine differences in DT% scores between the test and retest sessions, overall and for the different stimuli and diagnoses (Table 2), we employed paired-samples t-tests.

To evaluate the relations between psychopathology and DT%, we examined differences in DT% between diagnostic groups and in relation to healthy control participants. Because dysphoric and threatening content were found to have different effects in different psychopathologies (e.g., Armstrong & Olatunji, 2012), DT% on threat content (i.e., angry, disgust, or fearful facial expressions vs. neutral/happy faces) and DT% on dysphoric content (dysphoric faces versus neutral/happy faces) were computed separately.[3] A one-way analysis of variance (ANOVA) was used to compare DT% across diagnostic groups. Significant results were followed-up with independent samples t-tests. Homogeneity of variance was assessed using Levene's test. When equal variance could not be assumed, t-test statistics were calculated using un-pooled variances and a correction to the degrees of freedom was employed.

To estimate the association between age and DT% a Pearson's correlation was used. To evaluate whether psychopathology further modifies the age-DT% association, we applied a moderation analysis procedure using the PROCESS macro in SPSS (IBM Corp., 2021), model 1 (Hayes, 2013). To further evaluate developmental aspects of DT%, we compared adults and youth on DT% using independent samples t-test. We then divided the youth sample into smaller age bins (6–9 year-olds, 10–18 year-olds) and conducted independent samples t-tests on DT%. In addition, to evaluate the influence of pediatric anxiety on the results, we divided each age bin into healthy youth and youth diagnosed with pediatric anxiety and compared the groups. To further evaluate whether age and pediatric anxiety predict DT% among youth, we performed a hierarchical regression analysis on DT% in the youth sample (i.e., age < 18). Standardized age in Z-scores and pediatric anxiety as a dummy categorical variable were entered into the model as predictors in Step 1, and their interaction term was entered into the model in Step 2.

Finally, because differences between youth and adults in internal consistency scores were evident (see Section 3.1.1 below), we conducted post-hoc analyses to test whether adults and youth also differed on DT% fluctuations (i.e., DT% variance) across matrices (see Macleod et al., 2019). To this end, we calculated the standard deviation of DT% across matrices in specific task-blocks for each participant. Using these data, we computed the mean standard deviation for the different groups. An independent-samples t-test was then used to examine differences in DT% variability between youths (6–18 years) and adults (18–73 years). To examine whether DT% variability among youths reflects difficulty in staying engaged with the task over time, we also examined the internal consistency scores among youth separately for the first and second halves of the task. Bonferroni correction for multiple comparisons was employed throughout data analysis. All analyses were conducted in SPSS, version 28 (IBM Corp., 2021).

3. Results

3.1. Reliability

3.1.1. Internal consistency

Overall internal consistency of DT% across the sample was Cronbach's alphas = 0.85 and 0.93 for 30 and 60 trials, respectively. Similar results were obtained for male and female participants for both 30 trials (Cronbach's alphas = 0.85 and 0.86, respectively) and 60 trials

[3] Since non-facial stimuli were assessed only among unselected participants, these were not included in this analysis.

Table 1
Internal consistency for percent dwell-time in the Free-Viewing Matrix Task by stimuli and diagnosis.

	30 Matrices		60 Matrices	
Stimuli contrast	Cronbach's alpha	n	Cronbach's alpha	n
Faces				
Angry-Neutral	0.78	569	0.71	65
Fear-Neutral	0.92	100		
Disgust-Neutral	0.86	371	0.95	141
Sad-Neutral	0.88	280		
Happy-Neutral	0.84	175		
Angry-Happy[a]	0.67[a]	28[a]		
Sad-Happy	0.84	263	0.93	138
Total Faces	0.85	1786	0.93	344
Non-Faces				
Alcoholic versus nonalcoholic beverages	0.86	98	0.91[b]	98[b]
High versus low calorie food	0.78[c]	111[c]		
Total non-faces	0.79[c]	209[c]	0.91[b]	98[b]
Diagnosis	Cronbach's alpha	*n*	Cronbach's alpha	*n*
Healthy				
Healthy adults				
Angry-Neutral	0.96	27		
Fear-Neutral	0.97	27		
Disgust-Neutral	0.94	19	0.97	19
Sad-Neutral	0.95	61		
Happy-Neutral	0.86	30		
Sad-Happy	0.93	37	0.96	37
Healthy children				
Disgust-Neutral	0.52	45		
Angry-Happy	0.69	15		
Sad-Happy	0.58	16		
Non-selected participants				
Angry-Neutral	0.69	459	0.71	65
Alcoholic-nonalcoholic	0.86	98	0.91[b]	98[b]
High-low calorie food	0.78[c]	111[c]		
PTSD				
Angry-Neutral	0.77	36		
Fear-Neutral	0.77	30		
Sad-Neutral	0.90	30		
SAD				
Disgust-Neutral	0.89	181	0.94	84
MDD				
Sad-Neutral	0.71	55		
Happy-Neutral	0.84	54		
Sad-Happy	0.85	65	0.89	63
Pediatric anxiety disorders				
Disgust-Neutral	0.68	48		
Angry-Happy	0.65	13		
Sad-Happy	0.54	14		

Note. PTSD = post-traumatic stress disorder, SAD = social anxiety disorder, MDD = major depressive disorder.
[a] Based on a single study sample with youth participants only.
[b] Calculated for 54 matrices.
[C] Calculated for 21 matrices.

Table 2
Test-retest for percent dwell-time on the relatively negative stimuli in The Free-Viewing Matrix Task by stimuli and by diagnosis.

Stimuli	$M_{DT\%}$-test (*SD*)	$M_{DT\%}$-retest (*SD*)	*r (p)*	*n*
Faces				
Angry-Neutral	50.71 (10.92)	49.04 (11.45)	0.86 (< 0.001)	101
Fear-Neutral	50.90 (10.79)	49.25 (11.38)	0.72 (< 0.001)	100
Disgust-Neutral	45.92 (8.92)	45.94 (8.70)	0.77 (< 0.001)	98
Sad-Neutral	49.8 (10.69)	49.53 (11.70)	0.84 (< 0.001)	101
Sad-Happy	46.76 (7.13)	46.86 (7.98)	0.83 (< 0.001)	41
Faces Overall	49.11 (10.27)	48.30 (10.71)	0.78 (< 0.001)	442
Non-Faces				
Alcoholic - nonalcoholic beverages	51.53 (7.42)	50.71 (8.04)	0.86 (< 0.001)	17
Diagnosis	$M_{DT\%}$-test (*SD*)	$M_{DT\%}$-retest (*SD*)	*r (p)*	*n*
Healthy	42.24 (13.62)	41.78 (13.58)	0.94 (< 0.001)	50
Non-selected	51.59 (7.32)	50.71 (8.04)	0.86 (< 0.001)	17
PTSD	55.64 (6.09)	53.99 (7.05)	0.81 (< 0.001)	37
SAD	47.81 (6.38)	47.18 (7.47)	0.62 (< 0.001)	32

Note. PTSD = post-traumatic stress disorder, SAD = social anxiety disorder. All *t*-tests examining the differences between $M_{DT\%}$-test and $M_{DT\%}$-retest yielded non-significant results (all p's > .05).

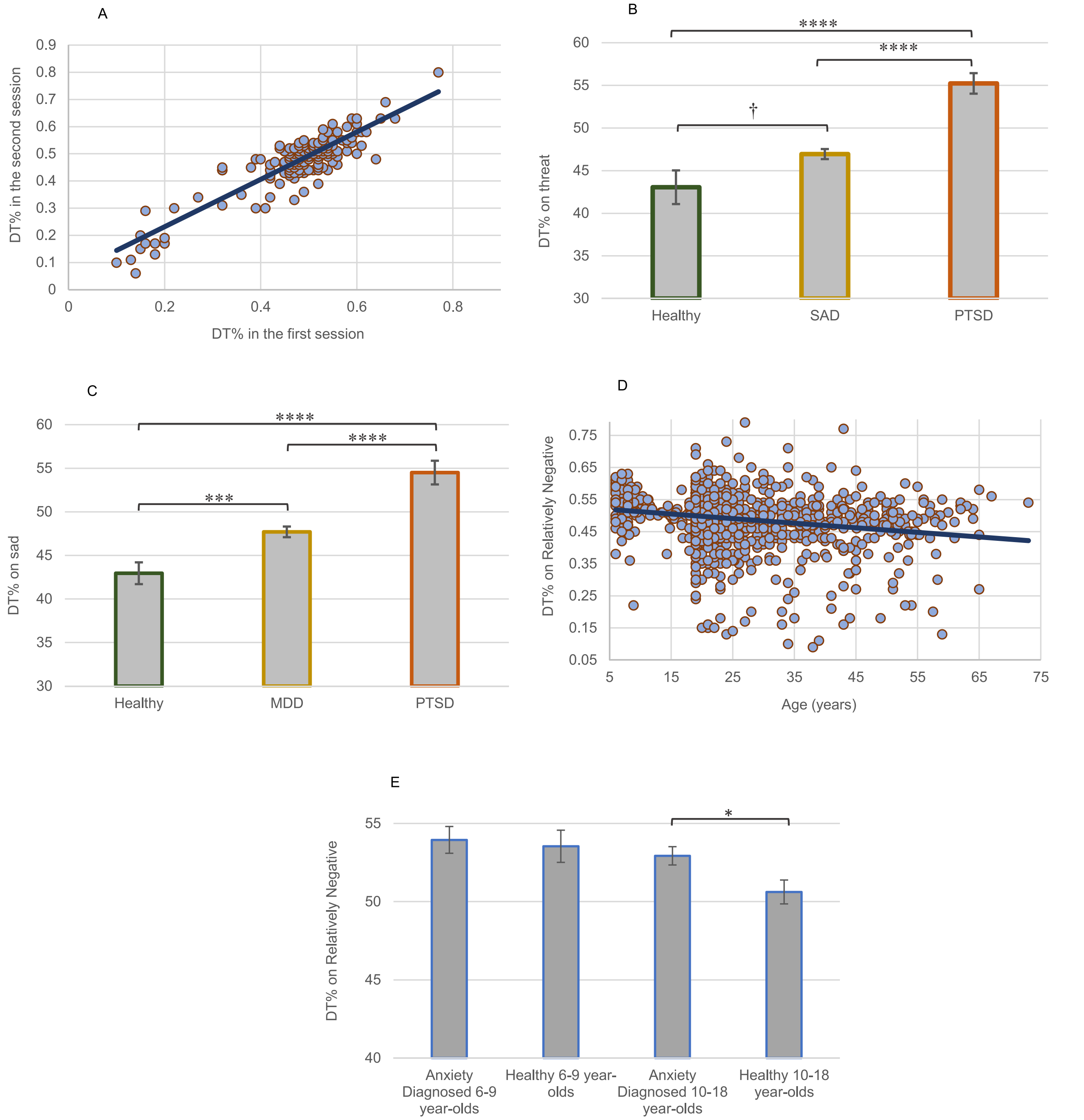

Fig. 2. Panel A: correlation between DT% in the first and second sessions for test-retest reliability; Panel B: DT% on threatening faces by diagnosis, error bars represent standard errors; Panel C: DT% on sad faces by diagnosis, error bars represent standard errors; Panel D: correlation between DT% on relatively negative stimuli and age; Panel E: DT% on relatively negative stimuli by age group, error bars represent standard errors. *Note.* DT% = percent dwell-time; PTSD = post-traumatic stress disorder. Threat = angry, fearful, or disgusted faces. Percent dwell-time was calculated as the proportion of dwell time on the relatively negative stimuli in the matrix relative to the sum of dwell-time on the negative and non-negative stimuli in the matrix. † = $p < .10$; * = $p < .05$; *** = $p < .001$; **** = $p < .0001$.

(Cronbach's alphas = 0.94 and 0.92, respectively). Internal consistency was 0.86 for adults and 0.65 for youth. Internal consistencies of DT% by stimuli contrasts and psychiatric diagnoses are presented in Table 1.

3.1.2. Test-retest reliability

Overall one-week test-retest reliability was $r(213) = 0.89$, $p < .001$, (see Fig. 2A), with similar results obtained for male and female participants ($r(76) = 0.89$ and $r(137) = 0.89$), respectively. No significant difference was noted between the mean DT% of the first ($M_{DT\%} = 48.10$, SD = 9.84) and second ($M_{DT\%} = 47.61$, SD = 9.71) measurement ($t(212) = -1.48$, $p = .14$, $d = 0.10$, CI = −0.03, 0.24). Test-retest results per stimuli contrasts and diagnoses, as well as the corresponding means and SDs of the test and retest sessions, are presented in Table 2.

3.2. Psychopathology and DT%

3.2.1. Percent dwell-time on threat content

Three groups provided data for this analysis: healthy participants ($n = 50$, $M_{DT\%} = 43.05$, SD = 13.96), patients with PTSD ($n = 37$, $M_{DT\%} = 55.23$, SD = 5.64), and patients with SAD ($n = 193$, $M_{DT\%} = 46.94$, SD = 8.20) – see Fig. 2B. The three groups differed significantly in DT% ($F(2) = 19.07$, $p < .001$, $\eta^2 = 0.12$, CI = .06,.19). Follow-up analyses indicated that DT% on threat faces was larger among participants with PTSD compared to healthy participants, $t(68) = -5.58$, $p < .001$, $d = -1.09$, CI = −1.54, − 0.63, and that DT% on threat faces was larger among participants with SAD compared to healthy participants at a trend-significance level, $t(58) = -3.47$, $p = .06$, $d = -0.40$, CI = −0.72, − 0.09. DT% on threat faces was larger among participants with PTSD compared to participants with SAD, $t(228) = 5.88$, $p < .001$, $d = 1.06$, CI = 0.69, 1.42.

3.2.2. Percent dwell-time on dysphoric content

Three groups provided data for this analysis: healthy participants ($n = 98$, $M_{DT\%} = 42.95$, SD = 12.37), participants with MDD ($n = 122$, $M_{DT\%} = 47.70$, SD = 6.77), and participants with PTSD ($n = 37$, $M_{DT\%} = 54.50$, SD = 8.26) – see Fig. 2C. A significant difference in DT% emerged between the three groups, $F(2) = 20.81$, $p < .001$, $\eta^2 = 0.14$, CI = .07,.22. Follow-up analyses indicated a significant difference between healthy participants and participants with MDD, $t(143) = -3.4$, $p < .001$, $d = -0.49$, CI = −0.76, − 0.22, and PTSD, $t(97) = -6.26$, $p < .001$, $d = -1.01$, CI = −1.41, − 0.61. DT% on dysphoric faces among participants with either MDD or PTSD was larger compared to healthy participants. Participants with PTSD had larger DT% on dysphoric faces compared to participants with MDD, $t(157) = -5.10$, $p < .001$, $d = -0.95$, CI = −1.33, − 0.57.

3.3. Age and DT%

Fig. 2D depicts a scatterplot of DT% on relatively negative stimuli as a function of age. Age inversely associated with DT%, $r(1442) = -0.21$, $p < .001$, such that as age increases, dwelling on negative stimuli decreases. Moderation analysis indicated that psychopathology moderated the effect of age on DT% ($b = 0.0005$, $SE = 0.0002$, $p = .003$). Participants without diagnosis exhibit stronger relation between age and DT% relative to those diagnosed with a psychopathology (b = - 0.002, SE = 0.0003, $p < .0001$ and $b = -0.001$, $SE = 0.003$, $p = .0001$, respectively).

Independent samples t-test indicated that DT% on relatively negative stimuli was larger among youth ($n = 118$, $M_{DT\%} = 52.99$, SD = 4.94) than among adults ($n = 1445$, $M_{DT\%} = 48.75$, SD = 7.79), $t(159) = 8.50$, $p < .001$, $d = 0.56$, 95 % CI = 0.31, 0.69. Dividing the youth sample into younger children (6–9 year-olds) and older children (10–18 year-olds) showed that DT% on relatively negative stimuli was larger among younger children ($n = 71$, $M_{DT\%} = 53.74$, SD = 5.62) than among older children ($n = 47$, $M_{DT\%} = 51.85$, SD = 3.42), $t(115) = 2.28$, $p = .02$, $d = 0.39$, 95 % CI = 0.02, 0.76. Dividing each of the youth groups into children with and without pediatric anxiety indicated no difference in DT% between healthy children and children with pediatric anxiety among 6–9 year-olds, and a difference among 10–18 year-olds, $t(45) = 2.43$, $p = .02$, $d = 0.71$, 95 % CI = 0.12, 1.30 (See Fig. 2E). Results of hierarchical regression analysis on DT% among youth with age and pediatric anxiety as predictors are presented in Table S3.

Additional analyses revealed that the variability in DT% in task-blocks conducted by youth ($M = 20.02$, SD = 5.31) was larger relative to this variability in adults, ($M = 15.85$, SD = 6.02), $t(2121) = 8.74$, $p < .001$, $d = 0.70$, CI = 0.54, 0.86. Internal consistency among youth was low for both halves of the task (Cronbach's alphas =.44 and.47 for the first and second halves, respectively).

4. Discussion

The current, large, multisite, international study examined the reliability of the DT% index derived from performance on the eye-tracking-based matrix task, and its association with a variety of psychopathologies and age. Results indicate acceptable internal consistency and test-retest reliability in adults across the analyzed contrasts and psychopathologies, but lower internal consistency scores in youth. Analyses further suggest that relative to healthy participants, participants diagnosed with different psychopathologies had larger DT% on relatively negative content. Finally, DT% on relatively negative stimuli inversely correlated with age.

For adult participants, the obtained overall internal consistency and test-retest coefficients for the DT% index reveal acceptable values for estimating population effect sizes and comparing groups in research (Bland & Altman, 1997; Bonett & Wright, 2015; Lohr, 2002; Nunnally & Bernstein, 1994). When using 60 matrices the Cronbach's alpha values surpass the recommended standard of 0.90 that is considered acceptable for individual classification (Bland & Altman, 1997; Bonett & Wright, 2015; Nunnally & Bernstein, 1994). Together, the current results suggest that the DT% index has an acceptable reliability for group (when using 30 or 60 matrices) and individual (when using 60 matrices) measurements. Studies evaluating the reliability of RT-based indices of attention allocation revealed mostly unacceptable internal consistency and test-retest coefficients (e.g., Brown et al., 2014; Eide et al., 2010; Evans et al., 2018; Macleod et al., 2019; Molloy & Anderson, 2020; Price et al., 2015; Schmukle, 2005; Spanakis et al., 2018; Staugaard, 2009). These limited psychometric properties of RT-based indices of attention allocation posed a critical barrier to advancement in research on attention mechanisms in psychopathology, and specifically on examination of change in attention allocation patterns following treatment. The acceptable reliability of the DT% index derived from the matrix task can help overcome this barrier and may also be used to evaluate the magnitude of change in attention allocation from pre- to post-treatment. Future studies could further establish the threshold for meaningful changes in DT% scores, indicating what is the minimal change that can be confidently attributed to treatment effect (Vaz et al., 2013).

The current study suggests lower internal consistency of the DT% index in youth than in adults. Although this internal consistency is still higher than that of most traditional RT-based indices of attention allocation among youth (Brown et al., 2014; Price et al., 2015), it is still below what would be deemed acceptable for both group means research and individual assessments (Bland & Altman, 1997; Bonett & Wright, 2015; Lohr, 2002; Nunnally & Bernstein, 1994). Importantly however, the matrix task was applied with youth participants only using 30 trials/matrices. The current analyses suggest that the observed lower reliability of the DT% index in youth cannot be attributed to reduced engagement or boredom with the task in this population. First, internal consistency was similar for the first and second halves of the task trials. Second, increasing the number of trials from 15 to 30 increased internal consistency from ~0.45 to ~0.65. As results in adults indicate that doubling the number of trials from 30 to 60 increases reliability from 0.86 to 0.93, future studies with youth could consider utilizing 60 matrices rather than 30. If a similar increase in internal consistency would be achieved, it could drive the internal consistency in youth into the acceptable range for group means research (Bland & Altman, 1997; Bonett & Wright, 2015; Nunnally & Bernstein, 1994). However, an increase in the number of trials should also be considered in light of the capacity of young participants to comply with elevated task demands. Our exploratory analyses also suggest that within-person DT% variability is greater among children than among adults, which may account, at least in part, for the observed difference in internal consistency between these age groups (Macleod et al., 2019). Considering the potentially higher variability of the DT% index in youth, future studies could evaluate the utility of dynamic measures of attention bias, such as attention bias variability (Alon et al., 2023; Todd et al., 2022, 2023), in measuring attention biases among youth.

The present study also revealed that as age increases DT% on relatively negative stimuli decreases. This result is consistent with the notion of a positivity bias among older adults (Isaacowitz et al., 2006; Reed et al., 2014), namely, an age-related increase in the tendency to prefer positive over negative information. It is also consistent with the 'moderation model' of attentional bias suggested by Field and Lester (2010), proposing that attention bias toward negative information characterizes all young children, and that this bias is reduced with age as a function of cognitive, emotional, and social development. These age-related changes in attention to negative stimuli carry implications for both measurement and treatment development for younger age groups.

DT% on threat- or dysphoric-related stimuli differed between healthy and psychiatrically diagnosed participants (PTSD, SAD, and MDD). The acceptable reliability of the DT% index point to a potential utility in generating population-wide distribution of attention bias scores. Such distribution norms could be used to identify patients suitable for attention-related intervention within the framework of precision medicine (Rodebaugh et al., 2016). Thus far, the poor reliability of RT measures hindered the possibility of creating pre-specified norms to select participants or patients based on their pre-treatment attention bias scores (e.g., Eldar et al., 2012; Hsu et al., 2021; Lazarov, Suarez-Jimenez, Abend et al., 2019). From this perspective, the acceptable reliability of the DT% index could generate interest in renewed attempts to use pre-treatment attention bias scores as a selection criterion. For example, population wide, age stratified, norms would allow offering ABM treatment only to patients with meaningful baseline attentional bias.

Some limitations of the current analyses should be considered. First, our test-retest data ($n = 213$) relies on an adult only sample. This calls for additional research on test-retest reliability of the DT% index in youth. Second, test-retest evaluation in the current study was examined over one week. Research using longer retest durations could be informative. Third, the number of participants in the PTSD group is relatively small. While this sample size is sufficient for internal consistency calculations (Bujang et al., 2018), additional research using larger samples of this population would be valuable. Fourth, the current results demonstrate that the DT% index reliably assesses the relative time participants spend on negative information (i.e., sustained attention). However, the DT% index itself does not clarify whether this sustained attention reflects deficits in engagement with or disengagement from negative stimuli. These more fine-grained processes could be better assessed exploring first-fixation measures from the Matrix task (i.e., first fixation latency, location, and dwell time) or by using other designated tasks such as engagement vs. disengagement tasks (Grafton & MacLeod, 2014; Sanchez et al., 2013). Finally, the current study focused on the reliability of the DT% attention index and its relation to psychopathology and age. Future studies could also test convergent and discriminant validity.

5. Conclusions

Basic research of attention-related processes in psychopathology and clinical treatments such as ABM, CBT, and SSRIs, which potentially target attentional processes, is dependent on reliable and stable measurement tools of disorder-relevant attention. The current evidence of the reliability of the DT% attention allocation index derived from the matrix task provides one such platform. The internal consistency and test-retest reliability of the DT% index suggest that how one attends to the presented stimuli in the matrix task may reflect a trait-like characteristic. The data further suggests that such trait-like characteristics differentiate participants with certain psychiatric disorders from healthy controls. Importantly, data from clinical trials further suggest that gaze-contingent attention bias modification protocols are effective in normalizing aberrant DT% patterns in patients (Arad et al., 2023; Lazarov et al., 2017; Shamai-Leshem et al., 2020), warranting additional clinical research and treatment development. Finally, the current results suggest that the free-viewing matrix task may provide a viable platform for population wide performance norms for specific stimuli contrasts. The creation of such norms could result in reliable criteria differentiating between normal attention allocation patterns and those characterizing specific psychopathologies, a key component toward personalized ABM interventions.

Funding information

This work was supported by the Israel Science Foundation (grant numbers 374/20 and 1811/17).

Declaration of interest

In the past 3 years Dr. Schneier has received consulting fees from Receptor Life Sciences; an honorarium from UpToDate; and research support from Otsuka Pharmaceuticals, National Institutes of Mental Health (NIMH) and VistaGen Therapeutics. Dr. Pine is supported by the NIMH Intramural Research Program. There are no conflicts of interest with the work described here. All the other authors report no conflict of interest.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

Data will be made available on request.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.janxdis.2023.102789.

References

Alon, Y., Bar-Haim, Y., Dykan, C. D. G., Suarez-Jimenez, B., Zhu, X., Neria, Y., & Lazarov, A. (2023). Eye-tracking indices of attention allocation and attention bias variability are differently related to trauma exposure and PTSD. *Journal of Anxiety Disorders, 96*, Article 102715. https://doi.org/10.1016/J.JANXDIS.2023.102715

Arad, G., Azriel, O., Pine, D. S., Lazarov, A., Sol, O., Weiser, M., Garber, E., Bloch, M., & Bar-Haim, Y. (2023). Attention bias modification treatment versus a selective serotonin reuptake inhibitor or waiting list control for social anxiety disorder: A randomized clinical trial. *The American Journal of Psychiatry, 180*(5), 357–366. https://doi.org/10.1176/APPI.AJP.20220533. 〈Https://Doi.Org/10.1176/Appi.Ajp.20220533〉

Armstrong, T., & Olatunji, B. O. (2012). Eye tracking of attention in the affective disorders: A meta-analytic review and synthesis. *Clinical Psychology Review, 32*(8), 704–723. https://doi.org/10.1016/j.cpr.2012.09.004

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: A meta-analytic study. *Psychological Bulletin, 133*(1), 1–24. https://doi.org/10.1037/0033-2909.133.1.1

Barch, D. M., Carter, C. S., & Committee, the C. E. (2008). Measurement issues in the use of cognitive neuroscience tasks in drug development for impaired cognition in schizophrenia: A report of the second consensus building conference of the CNTRICS initiative. *Schizophrenia Bulletin, 34*(4), 613–618. https://doi.org/10.1093/SCHBUL/SBN037

Basel, D., Aviram, T., & Lazarov, A. (2022). Lack of an attention bias away from relatively negative faces in dysphoria is not related to biased emotion identification. *Behavior Therapy, 53*(2), 182–195. https://doi.org/10.1016/J.BETH.2021.07.001

Bland, J. M., & Altman, D. G. (1997). Statistics notes: Cronbach's alpha. *BMJ, 314*(7080), 572. https://doi.org/10.1136/BMJ.314.7080.572

Bonett, D. G., & Wright, T. A. (2015). Cronbach's alpha reliability: Interval estimation, hypothesis testing, and sample size planning. *Journal of Organizational Behavior, 36*(1), 3–15. https://doi.org/10.1002/JOB.1960

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. *Journal of Anxiety Disorders, 28*(1), 97–107. https://doi.org/10.1016/j.janxdis.2013.11.004

Bujang, M. A., Omar, E. D., & Baharum, N. A. (2018). A review on sample size determination for cronbach's alpha test: A simple guide for researchers. *The Malaysian Journal of Medical Sciences: MJMS, 25*(6), 85. https://doi.org/10.21315/MJMS2018.25.6.9

Eide, P., Kemp, A., Silberstein, R. B., Nathan, P. J., & Stough, C. (2010). Test-retest reliability of the emotional stroop task: examining the paradox of measurement change. *The Journal of Psychology, 136*(5), 514–520. https://doi.org/10.1080/00223980209605547. ⟨Http://Dx.Doi.Org/10.1080/00223980209605547⟩

Eldar, S., Apter, A., Lotan, D., Edgar, K. P., Naim, R., Fox, N. A., Pine, D. S., & Bar-Haim, Y. (2012). Attention bias modification treatment for pediatric anxiety disorders: A randomized controlled trial. *American Journal of Psychiatry, 169*(2), 213–220. https://doi.org/10.1176/APPI.AJP.2011.11060886/ASSET/IMAGES/LARGE/AJP_169_2_213_F002.JPEG

Evans, T. C., Walukevich, K. A., Seager, I., & Britton, J. C. (2018). A psychometric comparison of anxiety-relevant attention measures. *Anxiety, Stress, & Coping, 31*(5), 539–554. https://doi.org/10.1080/10615806.2018.1489536. ⟨Https://Doi.Org/10.1080/10615806.2018.1489536⟩

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? *Clinical Child and Family Psychology Review, 13*(4), 315–332. https://doi.org/10.1007/S10567-010-0078-8/TABLES/1

Gade, P. R., Molloy, A., & Anderson, P. L. (2022). The use of trial-level bias scores to examine attention bias and attention bias variability among people with and without social anxiety disorder. *Journal of Clinical Psychology, 78*(5), 847–856. https://doi.org/10.1002/JCLP.23259

Grafton, B., & MacLeod, C. (2014). Enhanced probing of attentional bias: The independence of anxiety-linked selectivity in attentional engagement with and disengagement from negative information. *Cognition and Emotion, 28*(7), 1287–1302. https://doi.org/10.1080/02699931.2014.881326. ⟨Https://Doi.Org/10.1080/02699931.2014.881326⟩

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford. ⟨https://www.guilford.com/books/Introduction-to-Mediation-Moderation-and-Conditional-Process-Analysis/Andrew-Hayes/9781462549030⟩.

Hsu, K. J., Shumake, J., Caffey, K., Risom, S., Labrada, J., Smits, J. A. J., Schnyer, D. M., & Beevers, C. G. (2021). Efficacy of attention bias modification training for depressed adults: a randomized clinical trial. *Psychological Medicine*, 1–9. https://doi.org/10.1017/S0033291721000702

Isaacowitz, D. M., Wadlinger, H. A., Goren, D., & Wilson, H. R. (2006). Is there an age-related positivity effect in visual attention? A comparison of two methodologies. *Emotion, 6*(3), 511. https://doi.org/10.1037/1528-3542.6.3.511

Jiang, M. Y. W., & Vartanian, L. R. (2020). A review of existing measures of attentional biases in body image and eating disorders research. *Australian Journal of Psychology, 70*(1), 3–17. https://doi.org/10.1111/AJPY.12161. ⟨Https://Doi.Org/10.1111/Ajpy.12161⟩

Klanecky Earl, A. K., Robinson, A. M., Mills, M. S., Khanna, M. M., Bar-Haim, Y., & Badura-Brack, A. S. (2020). Attention bias variability and posttraumatic stress symptoms: The mediating role of emotion regulation difficulties. *Cognition and Emotion, 34*(6), 1300–1307. https://doi.org/10.1080/02699931.2020.1743235. ⟨Https://Doi.Org/10.1080/02699931.2020.1743235⟩

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. *Behaviour Research and Therapy, 135*, Article 103762. https://doi.org/10.1016/J.BRAT.2020.103762

Kline, P. (2000). In *Personality and individual differences. The Handbook of psychological testing*. Routledge. https://doi.org/10.1016/s0191-8869(96)90047-1.

Lazarov, A., Abend, R., & Bar-Haim, Y. (2016). Social anxiety is related to increased dwell time on socially threatening faces. *Journal of Affective Disorders, 193*, 282–288. https://doi.org/10.1016/j.jad.2016.01.007

Lazarov, A., Pine, D. S., & Bar-Haim, Y. (2017). Gaze-contingent music reward therapy for social anxiety disorder: A randomized controlled trial. *American Journal of Psychiatry, 174*(7), 649–656. https://doi.org/10.1176/appi.ajp.2016.16080894

Lazarov, A., Suarez-Jimenez, B., Abend, R., Naim, R., Shvil, E., Helpman, L., Zhu, X., Papini, S., Duroski, A., Rom, R., Schneier, F. R., Pine, D. S., Bar-Haim, Y., & Neria, Y. (2019). Bias-contingent attention bias modification and attention control training in treatment of PTSD: a randomized control trial. *Psychological Medicine, 49*(14), 2432–2440. https://doi.org/10.1017/S0033291718003367

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. *Psychological Medicine, 49*(5), 705–726. https://doi.org/10.1017/S0033291718002313

Lazarov, A., Suarez-Jimenez, B., Zhu, X., Pine, D. S., Bar-Haim, Y., & Neria, Y. (2021). Attention allocation in posttraumatic stress disorder: An eye-tracking study. *Psychological Medicine*. https://doi.org/10.1017/S0033291721000581

Lilienfeld, S. O., & Strother, A. N. (2020). Psychological measurement and the replication crisis: Four sacred cows. *Canadian Psychology/Psychologie Canadienne, 61*(4), 281. https://doi.org/10.1037/CAP0000236

Lisk, S., Vaswani, A., Linetzky, M., Bar-Haim, Y., & Lau, J. Y. F. (2020). Systematic review and meta-analysis: Eye-tracking of attention to threat in child and adolescent anxiety. *Journal of the American Academy of Child & Adolescent Psychiatry, 59*(1), 88–99.e1. https://doi.org/10.1016/J.JAAC.2019.06.006

Lohr, K. N. (2002). Assessing health status and quality-of-life instruments: Attributes and review criteria. *Quality of Life Research, 11*(3), 193–205. https://doi.org/10.1023/A:1015291021312/METRICS

Macleod, C., Grafton, B., & Notebaert, L. (2019). Anxiety-linked attentional bias: Is it reliable? *Annual Review of Clinical Psychology, 15*, 529–554. https://doi.org/10.1146/ANNUREV-CLINPSY-050718-095505. ⟨Https://Doi.Org/10.1146/Annurev-Clinpsy-050718-095505⟩

Mather, M., & Carstensen, L. L. (2016). Aging and attentional biases for emotional faces. *Psychological Science, 14*(5), 409–415. https://doi.org/10.1111/1467-9280.01455. ⟨Https://Doi.Org/10.1111/1467-9280.01455⟩

McNally, R. J. (2019). Attentional bias for threat: Crisis or opportunity? *Clinical Psychology Review, 69*, 4–13. https://doi.org/10.1016/J.CPR.2018.05.005

Mogg, K., & Bradley, B. P. (1998). A cognitive-motivational analysis of anxiety. *Behaviour Research and Therapy, 36*(9), 809–848. https://doi.org/10.1016/S0005-7967(98)00063-1

Mogoaşe, C., David, D., & Koster, E. H. W. (2014). Clinical efficacy of attentional bias modification procedures: An updated meta-analysis. *Journal of Clinical Psychology, 70*(12), 1133–1157. https://doi.org/10.1002/jclp.22081

Molloy, A., & Anderson, P. L. (2020). Evaluating the reliability of attention bias and attention bias variability measures in the dot-probe task among people with social anxiety disorder. *Psychological Assessment, 32*(9), 883. https://doi.org/10.1037/PAS0000912

Nunnally, J.C., & Bernstein, I.H. (1994). Psychometric theory / Jum C. Nunnally, Ira H. Bernstein. In *Psychometric theory*.

Peckham, A. D., McHugh, R. K., & Otto, M. W. (2010). A meta-analysis of the magnitude of biased attention in depression. *Depression and Anxiety, 27*(12), 1135–1142. https://doi.org/10.1002/da.20755

Polit, D. F. (2014). Getting serious about test-retest reliability: A critique of retest research and some recommendations. *Quality of Life Research, 23*(6), 1713–1720. https://doi.org/10.1007/S11136-014-0632-9/METRICS

Price, R. B., Kuckertz, J. M., Siegle, G. J., Ladouceur, C. D., Silk, J. S., Ryan, N. D., Dahl, R. E., & Amir, N. (2015). Empirical recommendations for improving the stability of the dot-probe task in clinical research. *Psychological Assessment, 27*(2), 365–376. https://doi.org/10.1037/pas0000036

Price, R. B., Wallace, M., Kuckertz, J. M., Amir, N., Graur, S., Cummings, L., Popa, P., Carlbring, P., & Bar-Haim, Y. (2016). Pooled patient-level meta-analysis of children and adults completing a computer-based anxiety intervention targeting attentional bias. In *, 50. Clinical psychology review* (pp. 37–49). Elsevier Inc.. https://doi.org/10.1016/j.cpr.2016.09.009

Reed, A. E., Chan, L., & Mikels, J. A. (2014). Meta-analysis of the age-related positivity effect: Age differences in preferences for positive over negative information. *Psychology and Aging, 29*(1), 1–15. https://doi.org/10.1037/a0035194

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. *Journal of Abnormal Psychology, 125*(6), 840. https://doi.org/10.1037/ABN0000184

Sanchez, A., Romero, N., & De Raedt, R. (2017). Depression-related difficulties disengaging from negative faces are associated with sustained attention to negative feedback during social evaluation and predict stress recovery. *PLoS One, 12*(3), Article e0175040. https://doi.org/10.1371/journal.pone.0175040

Sanchez, A., Vazquez, C., Marker, C., LeMoult, J., & Joormann, J. (2013). Attentional disengagement predicts stress recovery in depression: an eye-tracking study. *Journal of Abnormal Psychology, 122*(2), 303–313. https://doi.org/10.1037/a0031529

Schmukle, S. C. (2005). Unreliability of the dot probe task. *European Journal of Personality, 19*(7), 595–605. https://doi.org/10.1002/per.554

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. *Behavior Research Methods, 51*(6), 2748–2760. https://doi.org/10.3758/S13428-018-1147-Z/FIGURES/2

Shamai-Leshem, D., Linetzky, M., & Bar-Haim, Yair (2022). Attention biases in previously depressed individuals: A meta-analysis and implications for depression recurrence. *Cognitive Therapy and Research, 46*(6), 1033–1048. https://doi.org/10.1007/S10608-022-10331-Y, 2022 46:6.

Shamai-Leshem, D., Lazarov, A., Pine, D. S., & Bar-Haim, Y. (2020). A randomized controlled trial of gaze-contingent music reward therapy for major depressive disorder. *Depression and Anxiety*. , Article da.23089. https://doi.org/10.1002/da.23089

Skinner, I. W., Hübscher, M., Moseley, G. L., Lee, H., Wand, B. M., Traeger, A. C., Gustin, S. M., & McAuley, J. H. (2018). The reliability of eyetracking to assess attentional bias to threatening words in healthy individuals. *Behavior Research Methods, 50*(5), 1778–1792. https://doi.org/10.3758/S13428-017-0946-Y/FIGURES/1

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. *Addictive Behaviors, 100*, Article 106117. https://doi.org/10.1016/j.addbeh.2019.106117

Spanakis, P., Jones, A., Field, M., & Christiansen, P. (2018). A stroop in the hand is worth two on the laptop: Superior reliability of a smartphone based alcohol stroop in the real world. *Substance Use & Misuse, 54*(4), 692–698. https://doi.org/10.1080/10826084.2018.1536716. ⟨Https://Doi.Org/10.1080/10826084.2018.1536716⟩

Staugaard, S. R. (2009). Reliability of two versions of the dot-probe task using photographic faces. *Psychology Science Quarterly, 51*(Issue 3).

Streiner, D. L. (2016). Statistics commentary series: Commentary #16— Regression toward the mean. In *Journal of Clinical Psychopharmacology, 36* pp. 416–418). Lippincott Williams and Wilkins,. https://doi.org/10.1097/JCP.0000000000000551

Todd, J., Coutts-Bain, D., Wilson, E., & Clarke, P. (2023). Is attentional bias variability causally implicated in emotional vulnerability? A systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews, 146*, Article 105069. https://doi.org/10.1016/J.NEUBIOREV.2023.105069

Todd, J., Wilson, E., Coutts-Bain, D., & Clarke, P. J. F. (2022). Attentional bias variability and its association with psychological and health symptoms: A meta-analysis. *Neuroscience & Biobehavioral Reviews, 140*, Article 104812. https://doi.org/10.1016/J.NEUBIOREV.2022.104812

Vaz, S., Falkmer, T., Passmore, A. E., Parsons, R., & Andreou, P. (2013). The case for using the repeatability coefficient when calculating test–retest reliability. *PLoS One, 8*(9), Article e73990. https://doi.org/10.1371/JOURNAL.PONE.0073990

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. *Cognitive Therapy and Research, 38*(3), 313–333. https://doi.org/10.1007/s10608-013-9588-2