Published in final edited form as:

Motiv Emot. 2020 October ; 44(5): 657–669. doi:10.1007/s11031-020-09834-6.

The stability and reliability of attentional bias measures in the dot-probe task: Evidence from both traditional mean bias scores and trial-level bias scores

Joshua M. Carlson, Lin Fang

Department of Psychological Science, Northern Michigan University, 1401 Presque Isle Avenue, Marquette, MI 49855

Abstract

A new method of calculating attentional bias from the dot-probe task measures fluctuations in bias towards and away from emotional stimuli over time using trial level bias score metrics. We assessed the stability and reliability of traditional attentional bias scores and trial level bias score measures of attentional bias across time in two five-block dot-probe task experiments in non-clinical samples. In experiments 1 and 2, both traditional attentional bias scores and trial level bias score measures of attentional bias did not habituate/decrease across time. In general, trial level bias score metrics (i.e., attention bias variability as well as the mean biases toward and away from threat) were more reliable than the traditional attention bias measure. This pattern was observed across both experiments. The traditional bias score, however, did improve in reliability in the later blocks of the fearful face dot-probe task. Although trial level bias score measures did not habituate and were more reliable across blocks, these measures did not correlate with state or trait anxiety. On the other hand, trial level bias score measures were strongly correlated with general reaction time variability—and after controlling for this effect no longer superior in reliability in comparison to the traditional attention bias measure. We conclude that general response variability should be removed from trial level bias score measures to ensure that they truly reflect attention bias variability.

Terms of use and reuse: academic research for non-commercial purposes, see here for full terms. <https://www.springer.com/aam-terms-v1>

Correspondence should be addressed to: Joshua M. Carlson, Ph.D., Department of Psychological Science, Northern Michigan University, 1401 Presque Isle Avenue, Marquette, MI 49855, joshcarl@nmu.edu.

Publisher's Disclaimer: This Author Accepted Manuscript is a PDF file of an unedited peer-reviewed manuscript that has been accepted for publication but has not been copyedited or corrected. The official version of record that is published in the journal is kept up to date and so may therefore differ from this version.

Conflict of Interest: Joshua Carlson declares that he has no conflict of interest. Lin Fang declares that she has no conflict of interest.

Ethical approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent: Informed consent was obtained from all individual participants included in the study.

Data Availability: The data associated with this manuscript has been shared on the Open Science Framework (OSF) in a project called “The stability and reliability of attentional bias measures in the dot-probe task: Evidence from both traditional mean bias scores and trial-level bias scores.”

Humans, and other species, reflexively orient attentional resources to salient stimuli within their environment (Kret, Jaasma, Bionda, & Wijnen, 2016; van Rooijen, Ploeger, & Kret, 2017). Visual signals of potential threat—including poisonous and predatory animals, weapons, violent acts, bodily harm, industrial pollution, threat-related words, threatening (fearful and angry) facial expressions and facial features, as well as other threat-related cues—are particularly salient and capture visual attention (Carlson, Fee, & Reinke, 2009; Carlson, Lehman, & Thompson, 2019; Carlson & Mujica-Parodi, 2015; Carlson & Reinke, 2014; Carlson, Torrence, & Vander Hyde, 2016; Cooper & Langton, 2006; Fox, 2002; Fox & Damjanovic, 2006; Koster, Crombez, Verschuere, & De Houwer, 2004; Macleod, Mathews, & Tata, 1986). The prioritized attentional processing of threat-related information is referred to as attentional bias. An attentional bias to threat is adaptive and can prepare an organism to ready resources for impending danger (Ohman & Mineka, 2001), but an exaggerated attentional bias to threat is linked to elevated anxiety (Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & van Ijzendoorn, 2007).

Within the laboratory, attentional bias to threat is most commonly measured with the dot-probe task (Macleod et al., 1986). During the task, two stimuli (one threat/emotion-relevant and the other emotionally neutral) are briefly displayed on opposite sides of a computer screen. These stimuli are followed by a target “dot” (or other target stimulus) occurring at one of the previously occupied locations. If attentional resources are biased to the location of the threat-related stimulus, participants should respond faster to targets occurring at the location previously occupied by the threatening stimulus (i.e., congruent trials) compared to targets replacing the neutral stimulus (i.e., incongruent trials). Traditionally, attention bias scores are calculated as the mean incongruent – congruent difference in reaction times (RTs) across all trials in the task. Although the dot-probe task is among the most widely used methods of measuring attention bias in the laboratory, the traditional quantification of attentional bias (i.e., mean incongruent – congruent differences in RTs) has come under scrutiny for having low internal and/or test–retest reliability (Price et al., 2015; Schmukle, 2005; Staugaard, 2009).

The poor reliability in the dot-probe task may in part be explained by the relatively recent understanding that attentional bias to threat is not static. Conversely, attentional bias to threat is dynamic with alternating periods of attentional focus towards and away from threat (Iacoviello et al., 2014; Zvielli, Bernstein, & Koster, 2014, 2015). The *trial level bias score*—or the difference between each congruent and incongruent trial and the closest opposing trial type—was introduced to quantify attentional bias on a moment-by-moment (i.e., trial-by-trial) basis (Zvielli et al., 2015). Trial level bias scores have been used to compute biases towards and away from threat at both mean and peak levels. In addition, trial level bias score variability can be calculated (Zvielli et al., 2015; see Method for quantification of the trial level bias score metrics). Prior research suggests that trial level bias score measures, relative to the traditional measure, better represent the fluctuating nature of attentional bias, and are uniquely associated with cognitive vulnerability to depression and anxiety (Davis et al., 2016; Zvielli et al., 2015; Zvielli, Vrijssen, Koster, & Bernstein, 2016).

The reliability of these trial level bias score metrics appears to be much higher than the traditional dot-probe bias score (Zvielli et al., 2016). However, most of the empirical support

for the improved reliability of these newer measures uses (sub)clinical samples, such as individuals with social anxiety disorder (Davis et al., 2016) and individuals in remitted depression (Zvielli et al., 2016). The questionable reliability of the dot-probe task has mostly been found in non-clinical studies (Broadbent & Broadbent, 1988; Mogg et al., 2000; Schmukle, 2005), in which the traditional attentional bias score not only shows low internal and/or test–retest reliability, but also an inconsistent relationship with anxiety. In addition, the validity of the trial level bias score metrics has been called into question with modeling data suggesting that general reaction time variability can influence measures of attention bias variability (Kruijt, Field, & Fox, 2016) and recent research indicating that these measures are indeed correlated (Clarke et al., 2020). Yet, the extent to which the reliability estimates of the newer dynamic attention bias measures are driven by general reaction time variability is untested. Thus, examining the reliability of newer attentional bias measures from the dot-probe task in non-clinical samples when general response time variability is controlled as well as the relationship between newer attention bias measures and anxiety will provide further empirical scrutiny into the reliability of attentional bias measures derived from the dot-probe task.

Beyond the abovementioned factors, previous research has shown that extended testing over multiple sessions (i.e., 1920 total trials) in the dot-probe task increases the test–retest reliability of the traditional attentional bias score (Aday & Carlson, 2019). It is unclear if prolonged testing increases the reliability of the trial level bias score metrics. Finally, research suggests that attentional bias to threat (as measured by the traditional bias score) does not habituate, or decrease, after repeated testing in the dot-probe task (Weber, Morrow, Rizer, Kangas, & Carlson, 2016). The stability of the trial level bias score metrics of attentional bias is still unknown.

To address these gaps in the literature, the current study measured traditional and newer (i.e., trial level bias score) attention bias measures in a dot-probe task with two experiments, one with fearful and neutral facial stimuli and the other one with threatening and neutral picture stimuli, across five blocks with 450 trials in total performed in each experiment. By including multiple blocks with a large number of trials in the dot-probe task, this allowed us to assess the degree to which the reliability of attention bias measures increases (or decreases) across time/block. We were also able to assess the stability of these measures across time/block. Based on previous findings using the traditional attention bias measure (Aday & Carlson, 2019; Weber et al., 2016), it was hypothesized that reliability would increase in later blocks and that these measures would be stable (i.e., not habituate) across block.

Experiment 1

Method

Participants—One-hundred and twenty-seven¹ adults (female = 92, right handed = 126) between 18 and 42 ($M = 22.17$, $SD = 4.92$) years of age participated in the study. These

¹128 participants were recruited, but one participant did not have sufficient data for block 1 of the dot-probe task and was therefore excluded.

participants completed a screening protocol for an ongoing clinical trial assessing the effects of attention bias modification on changes in brain structure (NCT03092609). All data included in this manuscript were collected during the screening session. Not all participants included here met inclusion criteria for the larger study. The determination of our sample size was based on the availability of this dataset. The study was approved by the Northern Michigan University Institutional Review Board. Participants received monetary compensation for their participation.

State-Trait Anxiety Inventory—State and trait anxiety were measured with the Spielberger State-Trait Anxiety Inventory (STAI; Spielberger, Gorsuch, & Lushene, 1970). The STAI-S consists of 20 items and yields a measure of state anxiety (how anxious one currently feels). The STAI-T also consists of 20 items and yields a measure of trait anxiety (how anxious one generally feels). Cronbach's alpha for the STAI-S and STAI-T items were 0.94 and 0.91, respectively.

Dot-Probe Task—The dot-probe task was programmed using E-Prime2 (Psychology Software Tools, Pittsburg, PA) and displayed on a 60 Hz 16" LCD computer monitor. Stimuli consisted of 20 fearful and neutral grayscale faces of 10 different actors (half female) (from Gur et al., 2002; Lundqvist, Flykt, & Öhman, 1998) that were cropped to exclude extraneous features and randomly presented². The ratings of the participants from the Experiment 2 showed that the fearful faces stimuli were perceived as more negative ($M = 3.83$, $SD = .30$) than neutral pictures ($M = 4.45$, $SD = .52$), $t(18) = 3.23$, $p = .005$. Each trial started with a white fixation cue (+) in the center of a black screen for 1000 ms. Two faces were presented simultaneously on the horizontal axis for 100 ms. Participants were seated 59 cm from the screen. Faces were approximately $5^\circ \times 7^\circ$ of the visual angle and were separated by 14° of the visual angle. The target dot appeared immediately after the faces disappeared. Responses were recorded with a Chronos E-Prime response box. Participants indicated left-sided targets by pressing the first, leftmost button using their right index finger and indicated right-sided targets by pressing the second button using their right middle finger. Participants were instructed to focus on the central fixation cue throughout the trial and respond to the targets as quickly as possible. The task included five blocks. At the end of each block, participants received feedback about their overall accuracy and reaction times to encourage accurate rapid responses.

The task included congruent trials (dot on the same side as the emotional face), incongruent trials (dot on the same side as the neutral face), and baseline trials (two neutral faces). Faster responses on congruent compared to incongruent trials was considered representative of attentional bias towards the respective emotion. The task consisted of five blocks with 450 trials in total. Each block contained and randomly presented 30 congruent, 30 incongruent, and 30 baseline trials.

²Fearful face stimuli were: 207F, 208F, 213F, 217F (Gur et al., 2002), AF14AFS, AF19AFS, AF22AFS, AM10AFS, AM22AFS, AM34AFS (Lundqvist, Flykt, & Öhman, 1998); Neutral face stimuli were: 207N, 213N, 217N (Gur et al., 2002), AF14NES, AF19NES, AF22NES, AM10NES, AM22NES, AM34NES (Lundqvist, Flykt, & Öhman, 1998).

Data preparation and analysis plan—Data were filtered to include correct responses between 150 and 750 ms post-target onset (96.72% of the data was included; Carlson, Aday, & Rubin, 2019) and the remaining 36851 trials were entered in the analyses. Similar to the steps outlined by Zvielli and colleagues (2015) and Carlson and colleagues (2019), trial level bias scores were computed by matching each congruent and incongruent trial to its closest opposing trial type (with a maximum distance of 5 trials backward or forward). The incongruent – congruent difference was then calculated for each pair. Using these trial level bias scores, several summary variables were derived including *Mean_{toward}* (Mean of trial level bias scores > 0 ms; i.e., mean bias towards emotional stimuli), *Mean_{away}* (Mean of trial level bias scores < 0 ms; i.e., mean bias away from emotional stimuli), *Peak_{toward}* (largest trial level bias scores > 0), *Peak_{away}* (largest trial level bias scores < 0), and variability. Variability was calculated as the summed distance between succeeding trial level bias scores divided by the total number of trial level bias scores. In addition, the traditional attention bias scores was calculated as the mean incongruent – congruent difference in reaction time. All attention bias measures were computed separately for each block. See Figure 1 for an example of trial level bias scores computed across blocks.

All the analyses were conducted in SPSS 26. To test for habituation (or potentiation) effects, one-way repeated measures analyses of Variance (ANOVAs) were conducted to assess the effect of block (1–5) on the following measures: *Mean_{toward}*, *Mean_{away}*, *Peak_{toward}*, *Peak_{away}*, trial level bias score variability, and the traditional attention bias score³. The Greenhouse-Geisser adjustment was applied to account for violations in the assumption of sphericity. To explore the trend across time, multilevel regression models were performed on traditional attentional bias (AB) score and trial-level bias score (TLBS) separately. For TLBS the fixed effects included block, whereas for traditional AB, the fixed effects contained trial type, and their interaction. For both models, random effects of subjects were calculated. Gamma regression were used when the distribution of the dependent variable was skewed. Pairwise comparisons were computed in post hoc analysis, with the Least Significant Difference (LSD) correction for multiple comparisons. The slope of the habituation effects was computed as the slope of the trend line with the index of the blocks as x-axis (i.e., from 1 to 5) and all the attentional measures as y-axis (separate for each measure index). To assess reliability in each attention bias measure across block, Pearson correlations were run with a one-tailed tests of significance for positive correlations across blocks. Pearson correlations were also used with one-tailed tests of significance for positive relationships between state and trait anxiety and attention bias measures as well as the slope. Fisher r-to-z transformation were applied to compare the difference between the average correlations of attentional bias indices. In addition, to control for the effect of the general response variability on the results, we also calculated the standard deviation of the RT from neutral-neutral trials and use the standard residual of the general RT variability and all the attentional bias indices as the un-confounded indices. We then examined the reliability of the indices again to see whether it is similar as using the original indices. Since *Mean_{toward}* and

³Although previous research suggests that there are sex differences in trial level bias score measures of attentional bias (Carlson, Aday, et al., 2019), supplementary analyses found no effects of sex in the current dataset (all $p > .32$). However, this may in part be due to the large percentage of female participants in the current dataset (72%). Thus, participant sex was not included in the analyses reported here.

$Peak_{toward}$ ($r = .87, p < .001$), $Mean_{away}$ and $Peak_{away}$ ($r = .85, p < .001$) were highly correlated, we only report the results of $Mean_{toward}$ and $Mean_{away}$ below to simplify the results.

Results

Stability of Attention Bias Measures across Blocks—The results of one-way repeated measures ANOVAs showed a significant main effect of block in traditional AB, $F(4, 504) = 10.78, p < .001, \eta_p^2 = .08$, TLBS, $F(4, 504) = 8.97, p < .001, \eta_p^2 = .07$, $Mean_{toward}$, $F(4, 504) = 14.06, p < .001, \eta_p^2 = .10$, and bias score variability, $F(4, 504) = 13.02, p < .001, \eta_p^2 = .09$, but not in $Mean_{away}$, $F(4, 504) = 1.39, p = .24, \eta_p^2 = .01$. As can be seen in Table 1, all bias indices increased linearly across blocks.

The results of multilevel regression models revealed that, for traditional AB, there was a significant main effect of block, $F(4, 36841) = 121.50, p < .001$, and trial type, $F(1, 36841) = 395.51, p < .001$. We also found a significant block by trial type interaction, $F(4, 36841) = 11.39, p < .001$. Post hoc (LSD) tests revealed that there was a trend that the traditional attention bias score was increased across blocks ($p < .001$). For TLBS, the main effect of block was significant, $F(4, 36405) = 23.52, p < .001$. Post hoc tests showed that, in general, there was a trend that TLBS increased linearly across blocks ($p < .05$), except that there was no significant difference between block 3 and 4 ($p = .47$).

Reliability of Attention Bias Measures across Blocks—The results of the comparisons between the average correlation of different attention bias indexes showed that $Mean_{away}$ (average $r = .39, z = 1.81, p = .07$), $Mean_{toward}$ (average $r = .50, z = 2.89, p < .01$) and bias score variability (average $r = .58, z = 3.78, p < .001$) were more reliable than the traditional attention bias score (average $r = .18$). $Mean_{away}$ ($z = 2.29, p < .05$), $Mean_{toward}$ ($z = 3.38, p < .001$), and TLBS variability ($z = 4.27, p < .001$) were also more reliable than the average TLBS (average $r = .12$). No other comparisons were significant or approached significance. See Table 2 for a correlation matrix of attention bias measures across blocks. However, after controlling the individual response variability, the differences between the traditional attention bias score and TLBS indices were not significant anymore. The reliability of the largest TLBS index TLBS variability (average $r = .26$) was similar to traditional attention bias (average $r = .18$), $z = 0.66, p = .51$, and TLBS (average $r = .11$), $z = 1.23, p = .22$. See Table 2 for more details.

Association between Anxiety and Attention Bias Measures across Blocks—

Levels of trait anxiety ranged from 27 to 79 ($M = 50.05, SD = 10.11$) and levels of state anxiety ranged from 22 to 80 ($M = 45.69, SD = 12.00$). State and trait anxiety were highly correlated, $r = .81, p < .001$. Across block, neither state nor trait anxiety were consistently correlated with any attention bias measure (See Table 4). In addition, habituation/potentiation slopes did not correlate with state or trait anxiety (See Table 4).

Discussion

In the Experiment 1, we assessed the stability and reliability of traditional attentional bias scores and (newer) trial level bias score measures of attentional bias across time in a five-

block dot-probe task with fearful faces as the stimuli. We consistently detected attentional bias to threat (i.e., fearful faces) across blocks in this experiment. In line with previous research using the traditional attention bias measure (Weber et al., 2016), we found that attentional bias to threat does not habituate/decrease across time. Our results extend this earlier research by suggesting that newer attention bias measures similarly do not habituate across time. In particular, we found that all the attentional bias indices linearly increased across time. Moreover, our findings suggest that trial level bias score measures—particularly attention bias variability, $Mean_{toward}$, and $Mean_{away}$ —are more reliable than the traditional attention bias measure and TLBS. To be noted, this superiority of the trial-level bias score measures were no longer significant after controlling for the general RT variability effect. Although trial level bias score measures did not habituate and were fairly reliable across time (i.e., block), these measures did not correlate with state or trait anxiety (in an unselected non-clinical sample). To determine if the findings in Experiment 1 using fearful faces generalize to other types of threat-related stimuli, we conducted the second experiment in which threatening pictures (i.e., IAPS images) were selected as stimuli.

Experiment 2

Method

Participants—Eighty two⁴ undergraduate students (female = 61, right handed = 74) between 18 and 30 ($M = 19.98$, $SD = 2.26$) years of age participated in experiment 2. The study was approved by the Northern Michigan University Institutional Review Board. Participants received course credit for their participation.

State-Trait Anxiety Inventory—As in Experiment 1, the Spielberger State-Trait Anxiety Inventory (STAI; Spielberger et al., 1970) was used. The Cronbach's alpha for the STAI-S and STAI-T items were 0.93 and 0.91, respectively.

Dot-Probe Task—The dot-probe task was similar to Experiment 1, except that a). The visual angle for the two stimuli that presented simultaneously was 12°. b). Stimuli consisted of 10 threatening and 10 neutral pictures⁵. c). Stimuli were presented for 500ms. The stimuli were selected from the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 2008). At the end of the dot-probe task, participants were asked to rate the IAPS pictures and the face stimuli used in the first study based on the standard rating procedure on a scale of 1 to 9 for unpleasant to pleasant (Lang et al., 2008). The ratings of the participants from the current experiment indicated that threatening pictures were perceived as more negative ($M = 3.05$, $SD = .53$) than neutral pictures ($M = 5.47$, $SD = .74$), $t(18) = 8.41$, $p < .001$.

Data preparation and analysis plan—The analysis plan was identical to Experiment 1. Data were filtered to include correct responses between 150 and 750 ms post-target onset

⁴86 participants were recruited, but one participant did not have sufficient data for block 1 and 2 of the dot-probe task and three did not have STAI score were therefore excluded.

⁵Threatening IAPS images were: 1030, 1080, 1201, 1300, 1930, 2120, 6200, 6260, 6510, 9440; Neutral IAPS images were: 7000, 7004, 2745.1, 7041, 7002, 2383, 5390, 7006, 7510, 7060.

(96.59% of the data was included; Carlson, Aday, et al., 2019) and the remaining 23761 trials were entered in the analyses. Similarly, since $Mean_{toward}$ and $Peak_{toward}$ ($r = .81$, $p < .001$), $Mean_{away}$ and $Peak_{away}$ ($r = .80$, $p < .001$) were highly correlated, we only reported the results of $Mean_{toward}$ and $Mean_{away}$ below.

Results

Stability of Attention Bias Measures across Blocks—In the one-way repeated measures ANOVAs, there were no effects of block for traditional AB, $F(4, 324) = 0.53$, $p = .99$, $\eta_p^2 = .001$, TLBS, $F(3.48, 281.63) = 0.29$, $p = .86$, $\eta_p^2 = .004$, $Mean_{toward}$, $F(3.34, 270.14) = 0.49$, $p = .71$, $\eta_p^2 = .01$, $Mean_{away}$, $F(4, 324) = 0.92$, $p = .46$, $\eta_p^2 = .01$, and bias score variability, $F(3.60, 291.23) = 0.77$, $p = .53$, $\eta_p^2 = .01$. The results suggested that both the traditional and trial level bias score attentional bias measures were stable across blocks.

In the multilevel regression model for traditional AB, we found a significant main effect of block, $F(4, 23751) = 138.29$, $p < .001$, and trial type, $F(1, 23751) = 20.80$, $p < .001$. However, the block by trial type interaction was not significant, $F(4, 23751) = 0.43$, $p = .79$. For TLBS, the main effect of block was not significant, $F(4, 23456) = 1.50$, $p = .20$.

Reliability of Attention Bias Measures across Blocks—The results of the comparisons between the average correlation of different attention bias indexes showed that $Mean_{away}$ (average $r = .48$, $z = 2.34$, $p < .05$), $Mean_{toward}$ (average $r = .41$, $z = 1.79$, $p = .07$) and bias score variability (average $r = .48$, $z = 2.34$, $p < .05$) were more reliable than the traditional attention bias score (average $r = .15$). $Mean_{away}$ ($z = 2.72$, $p < .05$), $Mean_{toward}$ ($z = 2.17$, $p < .05$), and TLBS variability ($z = 2.72$, $p < .05$) were also more reliable than the average TLBS (average $r = .09$). No other comparisons were significant or approached significance. See Table 3 for a correlation matrix of attention bias measures across blocks. Similar to Experiment 1, after considering the effect of the individual response variability, the differences between the traditional attention bias and TLBS indices were no longer significant. The reliability of the largest TLBS index $Mean_{away}$ (average $r = .17$) was the same as the traditional attention bias score (average $r = .17$), and not significantly larger than TLBS (average $r = .10$), $z = 0.45$, $p = .65$. See Table 3 for more details.

Association between Anxiety and Attention Bias Measures across Blocks—The level of trait anxiety ranged from 25 to 68 ($M = 44.34$, $SD = 9.98$) and the level of state anxiety ranged from 21 to 69 ($M = 37.98$, $SD = 11.26$). State and trait anxiety were highly correlated, $r = .59$, $p < .001$. Similar to Experiment 1, neither state nor trait anxiety were consistently correlated with any attention bias measure across blocks (See Table 4). Habituation/potentiation slopes did not correlate with state or trait anxiety as well (See Table 4).

Discussion

In the Experiment 2, we assessed the stability and reliability of traditional attention bias and (newer) trial level bias score measures of attentional bias across time in a five-block dot-probe task with neutral and threatening pictures as stimuli. We did not observe attentional bias to threat (i.e., threatening pictures) across blocks in this experiment. Similar to the

findings of the Experiment 1, attentional bias to threat does not habituate/decrease across time. Further, our findings suggest that trial level bias score measures—particularly attention bias variability, $Mean_{toward}$ and $Mean_{away}$ —are more reliable than the traditional attention bias measure and TLBS. However, in line with the Experiment 1, this superiority of the trial-level bias measures did not exist after controlling for general RT variability. The trial level bias score measures did not correlate with state or trait anxiety in another unselected non-clinical sample.

General Discussion

In the current study, we assessed the stability and reliability of traditional and (newer) trial level bias score measures of attentional bias across time in a five-block dot-probe task. Two experiments were conducted, one with neutral and fearful facial stimuli and another with neutral and threatening picture stimuli. Our findings consistently showed that trial-level bias score variability, $Mean_{away}$ and $Mean_{toward}$ are more reliable than traditional attention bias scores and average trial-level bias scores. We also found that attentional bias to threat does not habituate/decrease across time. However, the trend of attentional bias across time showed different patterns between the two experiments. In Experiment 1, we found that all attentional bias measurements linearly increased across blocks, whereas there were no significant differences across blocks in these measurements in Experiment 2. Interestingly, after controlling for the general variability in RTs, the un-confounded trial level bias score indices were no longer more reliable than the traditional attention bias score. In addition, these measures did not correlate with state or trait anxiety (in unselected non-clinical samples).

Stability of Attention Bias Measures across Blocks

We found a consistent capture of attention by fearful faces in the dot-probe across all five experimental blocks in Experiment 1. This finding is in accordance to earlier findings that threatening stimuli capture attention (Carlson et al., 2009; Carlson, Lehman, et al., 2019; Carlson & Mujica-Parodi, 2015; Carlson & Reinke, 2014; Carlson et al., 2016; Cooper & Langton, 2006; Fox, 2002; Fox & Damjanovic, 2006; Koster et al., 2004; Macleod et al., 1986; Pourtois, Grandjean, Sander, & Vuilleumier, 2004). However, it should be noted that several studies have failed to find attention bias effects in the dot-probe task (e.g., Puls & Rothermund, 2017). Indeed, in Experiment 2 using threatening IAPS images, we did not see a consistent capture of attention. Further research is needed to determine the task-relevant parameters that are optimal for detecting attentional bias in the dot-probe task. Cue to target timing has been shown to moderate attention bias effects in the dot-probe task (Cooper & Langton, 2006; Koster, Verschuere, Crombez, & Van Damme, 2005; Mogg & Bradley, 2006; Mogg, Bradley, de Bono, & Painter, 1997; Torrence, Wylie, & Carlson, 2017), but other factors such as size/visual angle of the cueing stimuli and the distance/visual angle between cueing stimuli as well as the type of task (i.e., detection vs discrimination) has received relatively less attention (although there is evidence for the detection task to be more effective; Salemink, van den Hout, & Kindt, 2007). Regardless of the differences in attentional capture between the two experiments, we found that attention bias measures do not decrease or habituate, across time/block. This finding is generally consistent with earlier

reports that attentional bias does not habituate (Herry et al., 2007; Lonsdorf, Juth, Rohde, Schalling, & Ohman, 2014; Weber et al., 2016).

The extent to which newer trial level bias score measures are moderated by time was until now unknown. Importantly, we found no evidence for habituation across any trial level bias score measure. From a practical perspective, this finding may be useful to researchers worried about potential practice or fatigue effects in dot-probe tasks with a larger number of trials. Our results suggest that across five blocks trial level bias scores do not weaken. However, in Experiment 1, all attention bias measures linearly increased across time. It is unclear why this pattern was observed in Experiment 1. It is possible that task familiarity increased these values. However, it should be noted that there was no task advantage to localizing the fearful or neutral face cues as the target dot occurred with equal probability at the location of each cue type. Regardless of the cause of this linear trend, it will be important for future studies comparing these trial level bias score measures across groups to ensure that all groups have (approximately) equal numbers of trials and excluded trials are not disproportionately occurring at earlier or later time points across groups. In short, we provide the first evidence that we are aware of that—similar to the traditional attention bias measure—trial level bias score measures of attentional bias do not habituate across time (i.e., block in the dot-probe task).

Reliability of Attention Bias Measures across Blocks

Consistent with previous research (Price et al., 2015; Schmukle, 2005; Staugaard, 2009), our results suggest that the traditional measure of attention bias had low reliability (i.e., correlations across blocks) in the dot-probe task with an average correlation of $r = .18$ in Experiment 1 and $r = .15$ in Experiment 2. To be clear, reliability (at the level of individual differences) is not just a concern for the dot-probe task, but for most robust and widely used cognitive tasks (Hedge, Powell, & Sumner, 2018). The results of Experiment 1 provide some evidence that extended testing in the dot-probe task can increase reliability (i.e., there were significant correlations between blocks 3 and 4 as well as 4 and 5). However, even these significant correlations were quite low ($r = .25$ and $r = .24$, respectively) and this trend was not observed in Experiment 2. Previous research has shown that extended testing in the dot-probe task can improve test–retest reliability with correlations of $r \approx .45 - .60$ (Aday & Carlson, 2019). However, it should be noted that this previous study assessed test–retest reliability in multiple sessions across a two-week timeframe rather than multiple blocks in a single session. This previous study also used self-relevant anxiety-related word stimuli rather than fearful faces or threatening images. It is possible that these factors resulted in relatively higher reliability relative to the current findings. Future research is needed to determine the feasibility of extended testing in the dot-probe task as a method improving reliability and what factors contribute to optimally increasing reliability. Nevertheless, our results are partially consistent with the notion that extended testing improves the reliability of the traditional attention bias measure.

Relative to the traditional attention bias measure, all trial level bias score measures showed higher reliability metrics in both Experiment 1 and Experiment 2 (see Table 2). Although a number of previous studies have shown that attention bias variability measures have

improved reliability (Naim et al., 2015; Price et al., 2015; Zvielli et al., 2016), this is the first evidence that we are aware of to demonstrate in unselected samples that all trial level bias score measures of attentional bias to fearful faces and threatening (IAPS) images have improved reliability. This adds to findings that split-half reliability is improved for trial level bias score measures of attentional bias to happy and sad faces in individuals with remitted depression (Zvielli et al., 2016). Taken together, it appears that trial level bias score measures of attentional bias to a variety of emotional stimuli have superior reliability compared to the traditional attention bias measure. Similar to the traditional attention bias measures in Experiment 1, reliability estimates of attention bias variability appeared to benefit from extended testing in the dot-probe task with the highest correlations occurring between blocks three to five. This pattern was not observed for other trial level bias score measures. In sum, we provide novel evidence that all trial level bias score measures of attentional bias to fearful faces have improved reliability and extended testing appears to further improve the reliability of the variability measure.

Importantly, however, we found across Experiments 1 and 2 that when general variability in RT was accounted for, the reliability of the trial level bias score measures was no longer superior to that of the traditional attention bias score. Thus, it appears that the superiority in reliability of the trial level bias score metrics is (at least partially) due to the reliability of RT in general (see Tables 2 & 3), rather than variability in attention bias per se. It should be noted that we only tested this in non-clinical samples. It is possible that variability in trial level bias scores in clinical samples is attributable to additional factors beyond the association with RT variability observed here. Attention bias variability measures have been shown to maintain their predictive value in some clinical samples when controlling for RT variability (Mazidi, Vig, Ranjbar, Ebrahimi, & Khatibi, 2019; Zvielli et al., 2016)⁶. Yet, other research points to a broader dysfunction in cognitive control linked to RT variability (Swick & Ashley, 2017). Although previous studies have shown that trial level bias score measures are more reliable in terms of test-retest reliability (Davis et al., 2016), whereas we assessed internal relation across block, rather than test sessions, it is likely that test-retest reliability will also be lower when controlling for general RT variability. Further research will be needed to assess the influence of RT variability as a potential confound in clinical samples and test-retest reliability. Given that there appears to be a link between trial level bias score measures and vulnerability for anxiety and depression (Davis et al., 2016; Zvielli et al., 2015; Zvielli et al., 2016), it will be of theoretical importance to determine whether this association is specifically related to attention bias variability—or variability in RT more broadly. Nevertheless, the results observed here should caution researchers before adopting these new trial level bias score metrics as reliable measures of attention bias per se, rather than reliable measures of RT variability more broadly. We encourage researchers to employ the strategy used here where general RT variability is removed from the trial level bias score measures to ensure that claims regarding attention bias, or attention bias variability, are unconfounded by general RT variability.

⁶It should be noted that RT variability in these studies was obtained from a very small number of practice trials, which may not be representative of RT variability during the task.

Based on these findings we cannot conclude that trial level bias score measures improve the reliability of the dot-probe task. Again, it should be pointed out that most cognitive tasks have poor reliability at the level of individual differences (Hedge et al., 2018). In particular, this appears to be driven by the poor reliability of RT based difference score measures (Draheim, Mashburn, Martin, & Engle, 2019; Goodhew & Edwards, 2019). Additionally, it should be noted that poor reliability for individual differences does not necessarily mean that the experimental effects associated with differences between conditions are unreliable (Draheim et al., 2019; Goodhew & Edwards, 2019; Hedge et al., 2018). Given the poor reliability generally associated with RT based difference scores, creating a new “novel” task of attentional bias would not be expected to ameliorate the concern in reliably measuring individual differences in attentional bias (at least using RT based difference scores). It has been suggested that accuracy based measures are a more reliable alternative (Draheim et al., 2019). However, accuracy in the dot-probe task is generally very high or at ceiling. Modifying the dot-probe task, or perhaps creating a new attention bias task, that allows for sufficient variability in accuracy may result in greater reliability for individual differences research. In addition, neuroimaging based measures such as event-related potentials (ERP), functional magnetic resonance imaging (fMRI), and even structural MRI based measures appear to be more reliable measures of individual differences in attention bias or changes in attention bias (Aday & Carlson, 2017; Kappenman, Farrens, Luck, & Proudfit, 2014; Reutter, Hewig, Wieser, & Osinsky, 2017; White et al., 2016). It will be important for future research to establish a suitable measure of individual differences in attentional bias; especially given the importance of attentional bias in clinical populations.

Association with Anxiety

Both the traditional attentional bias score and trial level bias score measures of attention bias were unrelated to state and trait anxiety across both experiments. It is important to note that our samples were unselected, which may in part explain the null relationship with anxiety. Similarly, a number of previous studies did not find an association between anxiety and attentional bias in the dot-probe task when using non-clinical samples (Broadbent & Broadbent, 1988; Mogg et al., 2000; Schmukle, 2005). Our findings provide new evidence that anxiety is also unrelated to temporal dynamics in attentional bias in non-clinical samples. Indeed, most observed relationships between anxiety and attention bias are often in studies that compared between low and high anxiety groups (e.g., Macleod et al., 1986; Mogg & Bradley, 1999, 2002). Yet, attention bias has previously been shown to correlate with anxiety in unselected samples—especially after extended testing in the dot-probe task (e.g., Aday & Carlson, 2019). The mixed findings, on the one hand, may suggest that the dot-probe task (behavioral measures) cannot provide a robust marker of anxiety-linked attentional bias in non-clinical samples (Mogg et al., 2010). On the other hand, accumulating evidence suggests that the relationship between anxiety and attention bias to threat is not as consistently observed across studies as previously thought (Heeren, Mogoase, Philippot, & McNally, 2015; Mogg, Waters, & Bradley, 2017).

Limitations and Conclusions

The findings of the current study addressed a research gap regarding the reliability of attentional bias measurements in dot-probe task. However, all measurements were collected

during the same session, which limits our conclusions concerning test-retest reliability. In addition, even though the extended number of trials and multiple block design in the present study help to unravel how attentional bias measures change across blocks, it would be interesting for future studies to systematically examine the effect of the length of the task on the reliability of attentional bias measures.

In summary, across two experiments, traditional and trial level bias score measures of attentional bias to threat did not weaken or habituate over time—suggesting that they are not susceptible to fatigue effects. Trial level bias score measures of attentional bias were more reliable than the traditional attention bias measures. However, the apparent superiority of the trial level bias score measures over the traditional attention bias measure were negligible after controlling for general RT variability—suggesting that the increased reliability of the trial level bias score measures is at least in part driven by the consistency of RT variability. Although attention bias measures were stable and were fairly reliable across time, these measures were unrelated to anxiety. Given that both traditional and trial level bias score measures were relatively stable and that the superior reliability of the trial level bias score measures was driven by general RT variability, these new measures do not appear to be more beneficial than the traditional attention bias measure (at least in non-clinical samples). Future research should remove general RT variability from trial level bias score measures to ensure that these measures reflect attention bias variability, rather than nonspecific RT variability.

Acknowledgements

Research reported in this publication was supported by the National Institute of Mental Health of the National Institutes of Health under Award Number R15MH110951. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We would like to thank Jeremy A. Andrzejewski and Taylor Susa, and all the other students in the Cognitive × Affective Behavior & Integrative Neuroscience (CABIN) Lab at Northern Michigan University for assisting in the collection of this data. In addition, we would like to thank three anonymous reviewers and the editor for their helpful feedback on an earlier draft of this manuscript. In particular, reviewer 3's suggestion to explore the potential influence of general reaction time variability on the reliability of attention bias indices.

Funding: This study was funded by the National Institute of Mental Health of the National Institutes of Health under Award Number R15MH110951

References

- Aday JS, & Carlson JM (2017). Structural MRI-based measures of neuroplasticity in an extended amygdala network as a target for attention bias modification treatment outcome. *Med Hypotheses*, 109, 6–16. doi:10.1016/j.mehy.2017.09.002 [PubMed: 29150296]
- Aday JS, & Carlson JM (2019). Extended testing with the dot-probe task increases test–retest reliability and validity. *Cogn Process*, 20, 65–72. doi:10.1007/s10339-018-0886-1 [PubMed: 30171401]
- Bar-Haim Y, Lamy D, Pergamin L, Bakermans-Kranenburg MJ, & van Ijzendoorn MH (2007). Threat-related attentional bias in anxious and nonanxious individuals: a meta-analytic study. *Psychol Bull*, 133(1), 1–24. doi:10.1037/0033-2909.133.1.1 [PubMed: 17201568]
- Broadbent DE, & Broadbent MHP (1988). Anxiety and attentional bias: state and trait. *Cognition & Emotion*, 2, 165–183.
- Carlson JM, Aday JS, & Rubin D (2019). Temporal dynamics in attention bias: effects of sex differences, task timing parameters, and stimulus valence. *Cogn Emot*, 33(6), 1271–1276. doi:10.1080/02699931.2018.1536648 [PubMed: 30336735]

- Carlson JM, Fee AL, & Reinke KS (2009). Backward Masked Snakes and Guns Modulate Spatial Attention. *Evolutionary Psychology*, 7(4), 527–537.
- Carlson JM, Lehman BR, & Thompson JL (2019). Climate change images produce an attentional bias associated with pro-environmental disposition. *Cogn Process*. doi:10.1007/s10339-019-00902-5
- Carlson JM, & Mujica-Parodi LR (2015). Facilitated Attentional Orienting and Delayed Disengagement to Conscious and Nonconscious Fearful Faces. *Journal of Nonverbal Behavior*, 39(1), 69–77. doi:DOI 10.1007/s10919-014-0185-1
- Carlson JM, & Reinke KS (2014). Attending to the fear in your eyes: Facilitated orienting and delayed disengagement. *Cognition and Emotion*, 28(8), 1398–1406. doi:10.1080/02699931.2014.885410 [PubMed: 24512521]
- Carlson JM, Torrence RD, & Vander Hyde MR (2016). Beware the eyes behind the mask: The capture and hold of selective attention by backward masked fearful eyes. *Motivation and Emotion*, 40(3), 498–505. doi:DOI 10.1007/s11031-016-9542-1
- Clarke P, Marinovic W, Todd J, Basanovic J, Chen NT, & Notebaert L (2020). What is attention bias variability? Examining the potential roles of attention control and response time variability in its relationship with anxiety.
- Cooper RM, & Langton SR (2006). Attentional bias to angry faces using the dot-probe task? It depends when you look for it. *Behav Res Ther*, 44(9), 1321–1329. [PubMed: 16321361]
- Davis ML, Rosenfield D, Bernstein A, Zvielli A, Reinecke A, Beevers CG, ... Smits JA (2016). Attention bias dynamics and symptom severity during and following CBT for social anxiety disorder. *J Consult Clin Psychol*, 84(9), 795–802. doi:10.1037/ccp0000125 [PubMed: 27362791]
- Draheim C, Mashburn CA, Martin JD, & Engle RW (2019). Reaction time in differential and developmental research: A review and commentary on the problems and alternatives. *Psychol Bull*, 145(5), 508–535. doi:10.1037/bul0000192 [PubMed: 30896187]
- Fox E (2002). Processing emotional facial expressions: the role of anxiety and awareness. *Cogn Affect Behav Neurosci*, 2(1), 52–63. [PubMed: 12452584]
- Fox E, & Damjanovic L (2006). The eyes are sufficient to produce a threat superiority effect. *Emotion*, 6(3), 534–539. [PubMed: 16938095]
- Goodhew SC, & Edwards M (2019). Translating experimental paradigms into individual-differences research: Contributions, challenges, and practical recommendations. *Conscious Cogn*, 69, 14–25. doi:10.1016/j.concog.2019.01.008 [PubMed: 30685513]
- Gur RC, Sara R, Hagendoorn M, Marom O, Hughett P, Macy L, ... Gur RE (2002). A method for obtaining 3-dimensional facial expressions and its standardization for use in neurocognitive studies. *J Neurosci Methods*, 115(2), 137–143. [PubMed: 11992665]
- Hedge C, Powell G, & Sumner P (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behav Res Methods*, 50(3), 1166–1186. doi:10.3758/s13428-017-0935-1 [PubMed: 28726177]
- Heeren A, Mogoase C, Philippot P, & McNally RJ (2015). Attention bias modification for social anxiety: A systematic review and meta-analysis. *Clinical Psychology Review*, 40, 76–90. doi:10.1016/j.cpr.2015.06.001 [PubMed: 26080314]
- Herry C, Bach DR, Esposito F, Di Salle F, Perrig WJ, Scheffler K, ... Seifritz E (2007). Processing of temporal unpredictability in human and animal amygdala. *J Neurosci*, 27(22), 5958–5966. doi:10.1523/JNEUROSCI.5218-06.2007 [PubMed: 17537966]
- Iacoviello BM, Wu G, Abend R, Murrough JW, Feder A, Fruchter E, ... Charney DS (2014). Attention bias variability and symptoms of posttraumatic stress disorder. *Journal of Traumatic Stress*, 27(2), 232–239. doi:10.1002/jts.21899 [PubMed: 24604631]
- Kappenman ES, Farrens JL, Luck SJ, & Proudfit GH (2014). Behavioral and ERP measures of attentional bias to threat in the dot-probe task: poor reliability and lack of correlation with anxiety. *Front Psychol*, 5, 1368. doi:10.3389/fpsyg.2014.01368 [PubMed: 25538644]
- Koster EH, Crombez G, Verschuere B, & De Houwer J (2004). Selective attention to threat in the dot probe paradigm: differentiating vigilance and difficulty to disengage. *Behav Res Ther*, 42(10), 1183–1192. [PubMed: 15350857]

- Koster EH, Verschuere B, Crombez G, & Van Damme S (2005). Time-course of attention for threatening pictures in high and low trait anxiety. *Behav Res Ther*, 43(8), 1087–1098. doi:10.1016/j.brat.2004.08.004 [PubMed: 15922291]
- Kret ME, Jaasma L, Bionda T, & Wijnen JG (2016). Bonobos (*Pan paniscus*) show an attentional bias toward conspecifics' emotions. *Proceedings of the National Academy of Sciences of the United States of America*. doi:doi/10.1073/pnas.1522060113
- Kruijt A-W, Field AP, & Fox E (2016). Capturing dynamics of biased attention: new attention bias variability measures the way forward? *PLoS One*, 11(11), e0166600. [PubMed: 27875536]
- Lang PJ, Bradley MM, & Cuthbert BN (2008). International affective picture system (IAPS): affective ratings of pictures and instruction manual. Retrieved from Gainesville, FL:
- Lonsdorf TB, Juth P, Rohde C, Schalling M, & Ohman A (2014). Attention biases and habituation of attention biases are associated with 5-HTTLPR and COMTval158met. *Cogn Affect Behav Neurosci*, 14(1), 354–363. doi:10.3758/s13415-013-0200-8 [PubMed: 23943515]
- Lundqvist D, Flykt A, & Öhman A (1998). The Karolinska directed emotional faces (KDEF). CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, 91–630.
- Macleod C, Mathews A, & Tata P (1986). Attentional Bias in Emotional Disorders. *Journal of Abnormal Psychology*, 95(1), 15–20. [PubMed: 3700842]
- Mazidi M, Vig K, Ranjbar S, Ebrahimi MR, & Khatibi A (2019). Attentional bias and its temporal dynamics among war veterans suffering from chronic pain: Investigating the contribution of post-traumatic stress symptoms. *J Anxiety Disord*, 66, 102115. doi:10.1016/j.janxdis.2019.102115 [PubMed: 31394483]
- Mogg K, Bradley B, Dixon C, Fisher S, Twelftree H, & McWilliams A (2000). Trait anxiety, defensiveness and selective processing of threat: an investigation using two measures of attentional bias. *Personality and Individual Differences*, 28, 1063–1077.
- Mogg K, & Bradley BP (1999). Some methodological issues in assessing attentional biases for threatening faces in anxiety: a replication study using a modified version of the probe detection task. *Behav Res Ther*, 37(6), 595–604. [PubMed: 10372472]
- Mogg K, & Bradley BP (2002). Selective orienting of attention to masked threat faces in social anxiety. *Behav Res Ther*, 40(12), 1403–1414. [PubMed: 12457635]
- Mogg K, & Bradley BP (2006). Time course of attentional bias for fear-relevant pictures in spider-fearful individuals. *Behav Res Ther*, 44(9), 1241–1250. doi:10.1016/j.brat.2006.05.003 [PubMed: 16870133]
- Mogg K, Bradley BP, de Bono J, & Painter M (1997). Time course of attentional bias for threat information in non-clinical anxiety. *Behav Res Ther*, 35(4), 297–303. [PubMed: 9134784]
- Mogg K, Holmes A, Garner M, & Bradley BP (2008). Effects of threat cues on attentional shifting, disengagement and response slowing in anxious individuals. *Behav Res Ther*, 46(5), 656–667. doi:10.1016/j.brat.2008.02.011 [PubMed: 18395185]
- Mogg K, Waters AM, & Bradley B (2017). Attention Bias Modification (ABM): Review of Effects of Multisession ABM Training on Anxiety and Threat-Related Attention in High-Anxious Individuals. *Clinical Psychological Science*. doi:10.1177/2167702617696359
- Naim R, Abend R, Wald I, Eldar S, Levi O, Fruchter E, ... Bar-Haim Y (2015). Threat-Related Attention Bias Variability and Posttraumatic Stress. *Am J Psychiatry*, appiajp201514121579. doi:10.1176/appi.ajp.2015.14121579
- Ohman A, & Mineka S (2001). Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychol Rev*, 108(3), 483–522. [PubMed: 11488376]
- Pourtois G, Grandjean D, Sander D, & Vuilleumier P (2004). Electrophysiological correlates of rapid spatial orienting towards fearful faces. *Cereb Cortex*, 14(6), 619–633. [PubMed: 15054077]
- Price RB, Kuckertz JM, Siegle GJ, Ladouceur CD, Silk JS, Ryan ND, ... Amir N (2015). Empirical recommendations for improving the stability of the dot-probe task in clinical research. *Psychol Assess*, 27(2), 365–376. doi:10.1037/pas0000036 [PubMed: 25419646]
- Puls S, & Rothermund K (2017). Attending to emotional expressions: no evidence for automatic capture in the dot-probe task. *Cogn Emot*, 1–14. doi:10.1080/02699931.2017.1314932 [PubMed: 27788634]

- Reutter M, Hewig J, Wieser MJ, & Osinsky R (2017). The N2pc component reliably captures attentional bias in social anxiety. *Psychophysiology*, 54(4), 519–527. doi:10.1111/psyp.12809 [PubMed: 28111770]
- Salemink E, van den Hout MA, & Kindt M (2007). Selective attention and threat: Quick orienting versus slow disengagement and two versions of the dot probe task. *Behav Res Ther*, 45(3), 607–615. [PubMed: 16769035]
- Schmukle SC (2005). Unreliability of the dot probe task. *European Journal of Personality: Published for the European Association of Personality Psychology*, 19(7), 595–605.
- Spielberger CD, Gorsuch RL, & Lushene RE (1970). *Manual for the State-Trait Anxiety Inventory (Self-Evaluation Questionnaire)*. Palo Alto, CA: Consulting Psychology Press.
- Staugaard SR (2009). Reliability of two versions of the dot-probe task using photographic faces. *Psychology Science Quarterly*, 51(3), 339–350.
- Swick D, & Ashley V (2017). Enhanced Attentional Bias Variability in Post-Traumatic Stress Disorder and its Relationship to More General Impairments in Cognitive Control. *Sci Rep*, 7(1), 14559. doi:10.1038/s41598-017-15226-7 [PubMed: 29109521]
- Torrence RD, Wylie E, & Carlson JM (2017). The time-course for the capture and hold of visuospatial attention by fearful and happy faces. *Journal of Nonverbal Behavior*. doi:doi:10.1007/s10919-016-0247-7
- van Rooijen R, Ploeger A, & Kret ME (2017). The dot-probe task to measure emotional attention: A suitable measure in comparative studies? *Psychon Bull Rev*. doi:10.3758/s13423-016-1224-1
- Weber MA, Morrow KA, Rizer WS, Kangas KJ, & Carlson JM (2016). Sustained, not habituated, activity in the human amygdala: A pilot fMRI dot-probe study of attentional bias to fearful faces. *Cogent Psychology*, 3(1), 1259881.
- White LK, Britton JC, Sequeira S, Ronkin EG, Chen G, Bar-Haim Y, ... Pine DS (2016). Behavioral and neural stability of attention bias to threat in healthy adolescents. *Neuroimage*, 136, 84–93. doi:10.1016/j.neuroimage.2016.04.058 [PubMed: 27129757]
- Zvielli A, Bernstein A, & Koster EH (2014). Dynamics of attentional bias to threat in anxious adults: bias towards and/or away? *PLoS One*, 9(8), e104025. doi:10.1371/journal.pone.0104025 [PubMed: 25093664]
- Zvielli A, Bernstein A, & Koster EH (2015). Temporal Dynamics of Attentional Bias. *Clinical Psychological Science*, 3(5), 772–788.
- Zvielli A, Vrijnsen JN, Koster EH, & Bernstein A (2016). Attentional bias temporal dynamics in remitted depression. *J Abnorm Psychol*, 125(6), 768–776. doi:10.1037/abn0000190 [PubMed: 27505407]

Figure 1. Example of trial level bias scores (TLBS), traditional attentional bias (AB), Mean Toward, Mean Away, TLBS variability of one block. For demonstration purpose, the indices used here were from a single subject.

Table 1.

Means (and SDs) in ms across blocks for attentional bias measures

Measure	Block1	Block 2	Block 3	Block 4	Block 5
<i>Experiment 1</i>					
Traditional AB	6.71 (16.01)	9.11 (17.51)	12.67 (16.52)	13.33 (18.52)	18.81 (18.78)
TLBS	5.99 (15.78)	9.37 (18.70)	13.66 (17.13)	12.78 (18.79)	17.75 (18.47)
Mean Toward	53.65 (20.81)	55.30 (21.54)	58.57 (21.69)	61.67 (22.06)	66.78 (24.95)
Mean Away	-49.70 (18.08)	-50.54 (21.22)	-50.45 (20.51)	-52.03 (20.64)	-53.91 (21.56)
TLBS Variability	41.37 (13.48)	42.74 (13.50)	44.32 (14.05)	46.13 (14.58)	48.97 (15.46)
<i>Experiment 2</i>					
Traditional AB	-4.11 (17.52)	-3.11 (19.97)	-3.94 (19.61)	-3.60 (17.21)	-3.16 (20.99)
TLBS	-3.90 (17.29)	-0.73 (23.22)	-3.29 (22.10)	-3.02 (21.81)	-2.25 (22.42)
Mean Toward	58.17 (19.05)	57.37 (21.83)	60.50 (32.79)	58.81 (21.58)	60.67 (22.54)
Mean Away	-60.34 (18.38)	-59.63 (22.47)	-63.24 (24.32)	-59.05 (23.04)	-59.29 (22.98)
TLBS Variability	47.36 (14.05)	46.97 (14.23)	48.57 (14.75)	47.32 (13.45)	49.41 (16.38)

Note: AB, attentional bias; TLBS, trial-level bias score.

*
 $p < .05$;

**
 $p < .01$,

 $p < .001$

Table 2.

Pearson correlations of attentional bias scores between blocks in Experiment 1 ($n = 127$)

Measure	Original							After controlling for the IRV							
	Block	1	2	3	4	5	Mean	IRV	Block	1	2	3	4	5	Mean
Traditional	1	1	.09	.21**	.13	.31***	.57***	.04	1	1	.09	.21**	.15*	.29***	.59***
AB															
.18 ^a	2		1	.22**	.07	.12	.50***	−.04	2		1	.22**	.07	.10	.50***
.18 ^b	3			1	.25**	.21**	.63***	.09	3			1	.25**	.19*	.63***
	4				1	.24**	.59***	−.06	4				1	.26**	.59***
	5					1	.66***	.21**	5					1	.62***
TLBS	1	1	.09	.13	.09	.18*	.51***	−.04	1	1	.09	.13	.09	.15*	.55***
.12 ^a	2		1	.09	.04	.08	.50***	−.04	2		1	.10	.04	.07	.43***
.11 ^b	3			1	.17*	.10	.54***	.06	3			1	.17*	.07	.55***
	4				1	.18*	.57***	−.05	4				1	.19*	.56***
	5					1	.58***	.21*	5					1	.55***
Mean Toward	1	1	.46***	.50***		.54***	.75***	.35***	1	1	.23**	.22**	.30***	.29**	.67***
.50 ^a	2		1	.57***	.33***	.44***	.73***	.43***	2		1	.30***	.04	.15	.56***
.22 ^b	3			1	.56***	.55***	.52***	.82***	3			1	.23**	.14	.62***
	4				1	.57***	.76***	.46***	4				1	.30***	.61***
	5					1	.31***	.57***	5					1	.61***
Mean Away	1	1	.49***	.49***	.33***	.33***	.72***	−.34***	1	1	.27**	.27**	.21*	.21*	.65***
.39 ^a	2		1	.38***	.40***	.29***	.72***	−.42***	2		1	.13	.14	.21**	.59***
.20 ^b	3			1	.40***	.39***	.74***	−.44***	3			1	.15*	.16*	.57***
	4				1	.43***	.72***	−.44***	4				1	.21**	.57***
	5					1	.69***	−.50***	5					1	.60***
TLBS	1	1	.59***	.55***	.46***	.50***	.76***	.41***	1	1	.31***	.25**	.26**	.31**	.66***
Variability															
.58 ^a	2		1	.64***	.58***	.55***	.82***	.52***	2		1	.27**	.17*	.31***	.64***
.26 ^b	3			1	.65***	.61***	.85***	.55***	3			1	.24**	.25**	.63***

Measure	Original								After controlling for the IRV						
	Block	1	2	3	4	5	Mean	IRV	Block	1	2	3	4	5	Mean
	4				1	.62 ***	.32 ***	.55 ***	4				1	.27**	.61 ***
	5					1	.32 ***	.61 ***	5					1	.67 ***

Note: AB, attentional bias; TLBS, trial-level bias score; IRV, individual reaction variability;

a. original average correlation coefficient;

b. average correlation coefficient after controlling for the IRV.

*
p < .05;

**
p < .01,

p < .001

Table 3.Pearson correlations of attentional bias scores between blocks in Experiment 2 ($n = 82$)

Measure	Original							After controlling for the IRV							
	Block	1	2	3	4	5	Mean	IRV	Block	1	2	3	4	5	Mean
Traditional AB	1	1	−.02	.02	.22 [*]	.20 [*]	.53 ^{***}	−.18	1	1	.03	.20 [*]	.21 [*]	.22 [*]	.57 ^{***}
.15 ^a	2		1	.33 ^{**}	.16	.10	.57 ^{***}	.17	2		1	.27 ^{**}	.20 [*]	.11	.56 ^{***}
.17 ^b	3			1	.15	.04	.60 ^{***}	.29 ^{**}	3			1	.21 [*]	.09	.61 ^{***}
	4				1	.18	.57 ^{***}	−.15	4				1	.16	.62 ^{***}
	5					1	.56 ^{***}	.04	5					1	.54 ^{***}
TLBS	1	1	−.08	.08	.29 ^{**}	.07	.45 ^{***}	−.10	1	1	−.04	.11	.29 ^{**}	.08	.54 ^{***}
.09 ^a	2		1	.39 ^{***}	.02	.13	.61 ^{***}	.29 ^{**}	2		1	.32 ^{**}	.07	.13	.56 ^{***}
.10 ^b	3			1	.15	−.11	.60 ^{***}	.30 ^{**}	3			1	.19 [*]	−.08	.58 ^{***}
	4				1	−.06	.53 ^{***}	−.09	4				1	−.07	.56 ^{***}
	5					1	.41 ^{***}	.10	5					1	.40 ^{***}
Mean Toward	1	1	.43 ^{***}	.31 ^{**}	.31 ^{**}	.35 ^{**}	.62 ^{***}	.41 ^{***}	1	1	.16	.04	.12	.07	.51 ^{***}
.41 ^a	2		1	.51 ^{***}	.46 ^{***}		.80 ^{***}	.49 ^{***}	2		1	.12	.34 ^{**}	.11	.64 ^{***}
.12 ^b	3			1	.34 ^{**}	.49 ^{***}	.80 ^{***}	.61 ^{***}	3			1	.08	.18	.52 ^{***}
	4				1	.33 ^{**}	.65 ^{***}	.30 ^{**}	4				1	−.02	.56 ^{***}
	5					1	.75 ^{***}	.67 ^{***}	5					1	.49 ^{***}
Mean Away	1	1	.55 ^{***}	.51 ^{***}	.48 ^{***}	.42 ^{***}	.75 ^{***}	−.50 ^{***}	1	1	.35 ^{**}	.17	.18	.26 ^{**}	.67 ^{***}
.48 ^a	2		1	.43 ^{***}	.38 ^{***}	.44 ^{***}	.73 ^{***}	−.49 ^{***}	2		1	.01	.06	.20 [*]	.56 ^{***}
.17 ^b	3			1	.55 ^{***}	.60 ^{***}	.82 ^{***}	−.59 ^{***}	3			1	.24 [*]	.34 ^{**}	.60 ^{***}
	4				1	.44 ^{***}	.75 ^{***}	−.53 ^{***}	4				1	−.07	.49 ^{***}
	5					1	.77 ^{***}	−.51 ^{***}	5					1	.59 ^{***}
TLBS Variability	1	1	.59 ^{***}	.42 ^{***}	.36 ^{***}	.42 ^{***}	.72 ^{***}	.58 ^{***}	1	1	.36 ^{***}	.06	.10	.05	.56 ^{***}
.48 ^a	2		1	.49 ^{***}	.48 ^{***}	.42 ^{***}	.77 ^{***}	.49 ^{***}	2		1	.06	.31 ^{**}	−.07	.60 ^{***}
.14 ^b	3			1	.45 ^{***}	.65 ^{***}	.79 ^{***}	.68 ^{***}	3			1	.21 [*]	.30 ^{**}	.58 ^{***}
	4				1	.51 ^{***}	.73 ^{***}	.40 ^{***}	4				1	.03	.59 ^{***}

Measure	Original							After controlling for the IRV							
	Block	1	2	3	4	5	Mean	IRV	Block	1	2	3	4	5	Mean
	5					1	.80 ***	.66 ***	5					1	.47 ***

Note: AB, attentional bias; TLBS, trial-level bias score; IRV, individual reaction variability;

a. original average correlation coefficient;

b. average correlation coefficient after controlling for the IRV.

* $p < .05$;

** $p < .01$,

*** $p < .001$

Table 4.

Pearson correlations of attentional bias scores between blocks

Measure	Anxiety	1	2	3	4	5	Mean	Slope
<i>Experiment 1</i>								
Traditional AB	Trait	.07	.12	.06	-.03	.12	.12	-.003
	State	-.04	.17	.07	-.04	.11	.10	.04
TLBS	Trait	.08	.03	.10	-.02	.05	.08	-.02
	State	-.02	.06	.03	-.01	.04	.04	.01
Mean Toward	Trait	.03	-.01	.05	-.01	.07	.03	.04
	State	-.03	.01	-.02	-.05	.005	-.02	.003
Mean Away	Trait	.08	-.07	-.03	.02	-.15	-.05	-.13
	State	.01	-.10	-.06	.01	-.09	-.07	-.03
TLBS Variability	Trait	-.01	.02	.06	-.01	.04	.02	.03
	State	-.004	.06	.03	-.04	-.01	.01	-.04
<i>Experiment 2</i>								
Traditional AB	Trait	-.10	.02	-.05	-.11	.03	-.07	.05
	State	-.16	.05	.11	-.13	-.06	-.05	-.004
TLBS	Trait	-.12	-.01	-.07	-.11	.02	-.11	.05
	State	-.13	.001	.10	-.15	-.05	-.07	-.01
Mean Toward	Trait	-.08	.07	.10	.13	.05	.09	.12
	State	-.12	.04	.20	.01	.05	.08	.12
Mean Away	Trait	-.13	-.17	-.11	-.17	-.01	-.16	.09
	State	-.09	-.15	-.10	-.15	-.09	-.15	-.01
TLBS Variability	Trait	-.07	.15	.18	.10	.09	.12	.10
	State	-.07	.04	.20	.04	.12	.09	.15

Note: AB, attentional bias; TLBS, trial-level bias score.