

**No evidence for attention bias towards threat in clinical anxiety: a meta-analysis of baseline bias in attention bias modification RCTs.**

Anne-Wil Kruijt<sup>1\*</sup>, Sam Parsons<sup>2</sup>, Elaine Fox<sup>2</sup>

<sup>1</sup> Department of Psychology, Stockholm University

<sup>2</sup> Department of Experimental Psychology, University of Oxford

\* corresponding author:

Anne-Wil.Kruijt@psychology.su.se  
mail@awkruijt.nl

**Background:** Considerable effort and funding are spent on developing Attention Bias Modification (ABM) as a treatment for anxiety disorders. ABM is theorized to exert therapeutic effects through reduction of an increased attentional bias towards threat. Yet, the available meta-analytical evidence for the common assertion that clinical anxiety is characterised by this treat-related attentional bias is thin: the largest meta-analysis to date included  $n=337$  clinically anxious individuals. We propose that baseline measurements in clinical ABM RCTs constitute a hitherto not assessed additional body of data on magnitude of biased attention in clinically anxious samples.

**Method:** We meta-analysed baseline dot-probe assessed bias for 1005 clinically anxious individuals enrolled in ABM RCTs.

**Results:** REML analysis indicated no evidence that mean bias index (BI) differs from zero ( $k=13$ ,  $n=1005$ , mean BI = 1.8 ms, SE = 1.26 ms,  $p = .144$ , 95% CI [-0.6 - 4.3]. Additional Bayes factor analyses also supported the traditional point-nil hypothesis ( $BF_{10} = .23$ ), whereas additional interval-based analysis indicated it unlikely that mean bias in clinical anxiety extends beyond the 0 to 5 ms interval.

**Discussion:** We discuss our findings with respect to strengths (larger samples, possible bypassing of publication bias), limitations (lack of control comparison, repurposing data), and theoretical and practical context. We suggest that it may be prudent to no longer classify anxious samples as being characterized by bias.

**Conclusion:** Clinically anxious individuals enrolled in RCTs for Attention Bias Modification are not characterized by attentional bias towards threat at baseline.

**Keywords:** attention bias, clinical anxiety, meta-analysis, cognitive bias modification, translational research

**GSS:** it is widely believed that anxiety is characterized by a tendency to orient attention specifically towards threatening information, and that this tendency (called attention bias) can be measured using a computer task called ‘dot-probe task’. In the past decade, many studies have tested whether a training version of this task can be used to modify bias, which might then be used as a new treatment (Attention Bias Modification). We analysed levels of attention bias measured before participants started the modification training in 13 studies enrolling 1005 diagnosed anxious patients. We found no evidence that clinically anxious people are characterized by attention bias towards threat (as measured by the dot-probe task.).

Preferential orienting of attention towards threatening information is commonly theorized to play a role in the aetiology of (clinical) anxiety. Consequently, it is also regarded a putative treatment target in anxiety disorders. Attention Bias Modification (ABM) techniques were initially developed to experimentally verify that a change in attention bias, in either direction, would lead to a change in anxiety vulnerability. These early findings provided (indirect) experimental evidence for the cognitive theory derived notion that biased information processing is involved in the aetiology and maintenance of anxiety disorders (MacLeod, Rutherford, Campbell, Ebsworthy, & Holker, 2002; Mathews & MacLeod, 2002, yet see: Harris & Menzies, 1998). However, it was not until 2009 that further ABM studies were published, and several of these took the form of clinical RCTs (Amir, Beard, Taylor, et al., 2009; Amir, Beard, Burns, & Bomyea, 2009; Hazen, Vasey, & Schmidt, 2009; Schmidt, Richey, Buckner, & Timpano, 2009). ABM was unambiguously identified as a potential treatment for anxiety disorders in a 2010 meta-analysis (Hakamata et al., 2010). Thus, focus shifted away from assessing attention bias's putative role in aetiology and maintenance of anxiety vulnerability, towards establishing the efficacy of ABM as a treatment for clinical anxiety disorders. It has been commonly asserted that anxiety disorders are indeed characterised by threat bias ever since. Yet, there exists less (meta-analytical) evidence of dot-probe assessed bias towards threat in diagnosed/clinical samples, than might be expected.

The first, and largest, meta-analysis on anxiety-related biased attention was published eleven years ago by Bar-Haim and colleagues and has been cited over 1500 times according to Scopus (Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & van IJzendoorn, 2007). This MA includes 172 studies (5869 individual participants) comparing healthy control, analogue low/high anxious, and/or clinically anxious samples on biased processing of negative information assessed with the emotional Stroop ( $k=77$ ), dot-probe ( $k=44$ ), or Posner/single

cueing task ( $k=4$ ). Relatively consistent evidence for the existence of anxiety related biased processing of threatening stimuli was obtained across tasks and participant categories with medium effect sizes  $d \approx .45$ . Yet given our interest in the evidence supporting ABM's proposed clinical target, and because ABM research relies almost exclusively on (variations of) the dot-probe task to measure and modify bias, we look at specifically dot-probe assessed bias observed in clinically anxious samples. From table 3 provided by Bar-Haim and colleagues (page 13), we learn that 17 clinically anxious samples differed significantly from healthy control groups on dot-probe assessed bias ( $d = .40$ , 95% CI [.29 - .60]<sup>1</sup>). This effect's size and direction are in line with the results for other tasks and non-clinical/highly anxious samples. Take note, however, that the average sample size is  $n = 20$ , for a total of  $n = 337$  clinically anxious individuals (compared to  $n = 341$  control participants) in 17 studies.

To the best of our knowledge, only two other meta-analysis assessing anxiety related attentional bias were published since 2007. A 2015 meta-analysis assessed whether anxious analogue or clinical samples showed greater bias for specifically disorder congruent versus disorder incongruent stimuli assessed with either Stroop or dot-probe tasks (Pergamin-Hight, Naim, Bakermans-Kranenburg, van IJzendoorn, & Bar-Haim, 2015). Six dot-probe studies with a total  $n$  of 115 clinically anxious individuals (PTSD, PD, SAD, or OCD) were included and for this subset of studies no meta-analytical evidence of disorder specific threat bias was obtained ( $d = .12$ ,  $p = .41$ ). The average sample size of the clinical groups was again  $\sim 20$ , and three of these six studies were also included in the 2007 meta-analysis. A 2016 meta-analysis of anxiety related bias focused on dot-probe assessed threat bias for facial stimuli associated with specifically social anxiety disorder (Bantin, Stevens, Gerlach, & Hermann, 2016). Four

---

<sup>1</sup> Throughout this manuscript, the findings of the meta-analysis by Bar-Haim and colleagues are given with 95% confidence intervals calculated from the 85% confidence intervals reported in the original paper.

out of ten included studies compared  $n = 109$  clinically anxious individuals to  $n = 129$  healthy controls. Three of these studies assessed bias for negative over neutral faces and for these a small to medium effect size was found ( $g = .38$ ) on the clinically anxious versus healthy controls contrast ( $k = 3$ ,  $n = 89$  versus  $n = 109$ ). Two of these three studies were also included in the 2007 meta-analysis by Bar-Haim and colleagues. The third is a study with  $n = 35$  generalized social phobia patients (Gotlib et al., 2004), which appears to be the largest clinically anxious sample assessed to date. Thus, from three meta-analyses a picture emerges that the, commonly assumed, phenomenon of dot-probe assessed threat bias may have been observed in circa 20-25 mostly rather small clinically anxious samples.

Perhaps the most major development in psychology research over the past decade has been the increasing awareness of statistical power and the need to assess sufficiently large samples. The 2007 meta-analysis indicated an estimated effect size  $d = 0.4$  for the comparison of clinical and control samples. To detect an effect of this size using a simple between-subjects t-test, a total  $n$  of 200 (100 per group) would be required to achieve 80% power, and a total  $n$  of 328 for 95% power (Faul, Erdfelder, Lang, & Buchner, 2007). After a decade, the absence of any large-scale studies following the 2007 meta-analysis is surprising. It is sobering to think that threat bias as measured by the dot-probe task has not been verified in a single study that has sufficient statistical power. This is of particular concern given the level of resources allocated to treatment-oriented research based on the premise that negative attention bias is a core feature of anxiety disorders.

Therefore, we present a meta-analysis of an hitherto not meta-analysed body of data on dot-probe threat bias, assessed in relatively large (up to  $n = 134$  in Rapee et al., 2013) and relatively well defined clinical samples, namely those enrolled in RCTs evaluating the effects

of ABM. We meta-analyse the baseline measures of biased attention to verify the existence of dot-probe assessed threat bias preceding attempted modification thereof.

In the preceding discussion of meta-analytical evidence, we focussed on the statistical contrasts assessing whether bias differs between clinically diagnosed and healthy control samples. Yet in order to utilize the (much larger) body of data that is available from the ABM RCTs, we shift to the question whether clinical samples exhibit attention bias. This is possible because the traditionally used bias index (BI) is measured on a bidirectional scale with an inherently meaningful 'zero value'. For many measures in clinical psychology we find ourselves limited to the question of whether a construct's effect size is larger or smaller in one population compared to another. For instance, for many symptom measures a zero-symptoms score is rare even in healthy/non-clinical populations, and we intuitively understand the need to compare a clinical sample's scores to a non-clinical sample's score. The traditional dot-probe derived estimate Bias Index (BI), on the other hand, can be understood as an inter-individual contrast comparing the individual's average response time between two sets of trials. To answer the question whether bias is present within a sample, we may employ a one-sample t-test, testing whether mean BI differs from zero. This test answers the question whether individuals in the sample on average responded faster (bias towards) or slower (bias away) on trials in which a response cue appeared in the location previously taken by an emotional stimulus (congruent trials), compared trials in which it appeared in the location of a neutral stimulus (incongruent trials - see: MacLeod, Mathews, & Tata, 1986; MacLeod et al., 2002). If the incongruent-congruent response time difference (BI) is not different from zero, the null-hypothesis that no bias is present cannot be rejected. Two of the previously discussed meta-analyses also assessed this 'within-group' BI effect. Their findings for this contrast in the subsets of dot-probe assessed clinical samples were as

follows: Bar-Haim et al (2007):  $k = 16$ ,  $n = 302$ ,  $d = .34$ , 95% CI[.18 - .50], and Bantini et al (2016):  $k = 3$ ,  $n = 89$ ,  $g = .48$ , 95% CI[.17 - .79].

One benefit of our approach is that it may partly bypass publication bias effects: in a typical RCT design, the baseline measure is not the outcome of interest and bias not being observed at baseline does not necessarily reduce the entire study to a difficult-to-publish null finding. Furthermore, several of the included RCTs were preregistered studies which strengthens the idea that this specific subset of attentional bias literature may be less affected by publication bias compared to the available literature directly comparing clinical and control groups.

Thus, we meta-analyse whether dot-probe measured attentional bias for threat is present in clinically anxious individuals enrolled in ABM RCTs, as evidenced by a baseline bias index (BI)  $> 0$ . In line with prevailing theory, our formal hypothesis is that attention bias towards negative information will be observed for the pooled clinically anxious samples<sup>2</sup>. In follow-up analyses we employ Bayesian methods to assess the relative strength of evidence for various BI effect sizes in milliseconds (ms).

## Methods

The Prisma checklist for this manuscript can be found in available as supplemental file x (Moher, Liberati, Tetzlaff, Altman, & PRISMA Group, 2009). Although no formal review protocol was prepared, a custom-built review app was built using R, as detailed below. The code is included as supplemental file x.

---

<sup>2</sup> Yet, from being involved with this field, we also know that a clear bias towards threat is often not observed at baseline in ABM RCTs. We disclose that our personal expectations run counter to the formal theory-derived hypothesis.

## **Record selection and data extraction**

The selection of records was done in several stages. First a search string was developed with the aim of retrieving as many English-written dot-probe based studies as possible from the Scopus database ([www.scopus.com](http://www.scopus.com)). The search string used was:

```
(TITLE-ABS-KEY (( "dot-probe" OR "probe detection task" OR "visual probe" OR  
"attentional probe task" OR "probe classification task" OR "atten* bias modification" ) OR ( "atten* retrain*" AND "probe" )) AND TITLE-ABS-KEY ( "bias*" OR "atten*" )) AND  
LANGUAGE ("English" )
```

The last update to our dataset was done on 20-03-2018, when a search in Scopus using the above string returned 1181 records, which were downloaded and imported into R for further processing. In R, the 1181 records were subjected to a filter which selected those records for which at least one the terms "RCT", "randomized controlled", "randomised controlled", "intervention" or "program", plus at least one of the terms "anx\*", " SAD ", " GAD ", " OCD", " phobi\*", " PTSD ", " panic", and at least one of the terms "patient\*", "diagnos\*", "clinic\*" were found across each record's title, abstract, and (index and author supplied) keywords.

This filter returned 333 records for manual assessment. These were subsequently loaded into a purpose built 'Shiny app' to aid the two assessors (SP & AWK) in the process of record selection and data extraction. The app guides the assessor through a two-stage record selection and data extraction procedure. For the first stage, each record's title, abstract, and keywords are shown and the assessor is asked to fill out their assessment for the following inclusion criteria:



- study aims to evaluate effects of a bias modification procedure (ABM / CBM / other)?
- assesses attention allocation bias to threatening information?
- participants are adults?
- clinical/diagnosed anxiety?

For each of the above questions the answer options were yes, no, and possibly.

If a 'no' answer was entered for either of the above four questions, the answer to the final question ('select for stage II') was automatically toggled from '?' to 'no' and vice-versa if the answer was changed again to yes or possibly. When all four criteria had an answer 'yes' or 'possibly', 'select for stage II' was toggled to 'yes'. The assessor manually submitted the information for each record before moving on to the next record.

For stage II, the assessor is again presented with a list of records to assess, now with an additional DOI based hyperlink to retrieve the paper and answer the remaining questions.

In stage II, the four inclusion criteria above had to be confirmed, and in addition the assessor was asked to indicate the primary diagnosis (choice of: GAD, OCD, Panic Disorder, PTSD, SP/SAD, simple phobia) and the diagnostic instrument used. Assessors were also asked to indicate whether individuals with comorbid mood disorders were excluded, the number of groups in the study, and various aspects of the bias assessment task used: type (dot-probe or Posner task), stimulus latency (< 500, 500 - <1000, > 1000, or 'other/mixed'), and stimulus type (words, faces, scenes, or other/mixed). If available in the paper, the assessors could enter for each group the number of participants as well as the mean and SD for the bias index (based on trials with negative and neutral stimuli) obtained at baseline. From these values, an overall mean bias index and SD were calculated, pooling over the two (or more) clinical samples. Mean bias was calculated as the n-weighted mean bias ( $\text{sum}(\text{mean} * n) / \text{sum}(n)$ ) and

pooled SD as  $\sqrt{\sum(n-1 * sd^2) / \sum (n-1)}$ . Thus, we obtained bias index mean and SD per study, collapsing the two (or more) clinical experimental and clinical control groups within each study. The assessor could enter comments, indicate if they felt additional data should be requested, and create additional records if a second study was presented in the same paper. Finally: the assessor had to check a checkbox to indicate their recommendation for inclusion in the MA and press submit.

A total of 380 records were 'stage 1 assessed' by each of the two assessors, who selected xx and 36 records for stage II assessment respectively. Following their individual stage II assessments, all data was gathered and the assessors used these records to reach consensus on the final set of studies to include. For most records the required data could be extracted from the published papers. It was verified that both assessors extracted the same values for each of these records, resulting in the discovery of one typing error which was subsequently corrected. As a final check, one of the assessors (AWK) manually compared the resulting selection to three recent meta-analysis assessing effects of ABM (Cristea, Kok, & Cuijpers, 2015; Linetzky, Pergamin-Hight, Pine, & Bar-Haim, 2015; Price et al., 2016) to verify that all relevant studies included in these MAs were also selected for the current MA. This resulted in identification of one additional study eligible for inclusion.

For 7 records, authors were contacted with a request to provide additional data. Most contacted authors kindly provided us with the requested data, and one group informed us that the bias data for their study was regrettably lost. One corresponding author did not respond to three emails sent over a nine-month period. For 1 study, mean and sd of the baseline bias index was inferred from a plot showing the baseline mean BI plus/min 1 sd on the x-axis

(Kuckertz et al., 2014, fig.4), while two other studies selected by the assessors could not be included in the final meta-analysis ( $n = 22$  and  $29$  – also see table 1).

### **Exclusion of Posner task assessed bias**

During our initial assessment of records, we also selected ABM RCTs assessing pre-training bias with the Posner/single cueing task with the intention of either reaching agreement with the involved authors on how to calculate an index of its four trial types that is similar enough to dot probe's (two trial types based) index, or performing a separate analysis of these studies. Three RCTs eligible for inclusion were identified (Amir, Beard, Taylor, et al., 2009; Amir, Weber, Beard, Bomyea, & Taylor, 2008; Boettcher, Berger, & Renneberg, 2011). We became aware, however, that while Posner tasks' four trial types can be combined into a single index (Mogg, Holmes, Garner, & Bradley, 2008), the interpretation of this index is not straightforward: depending on which of two validity effects occurred, opposing scores (more negative as well more positive) can be interpreted as indicative of more bias in the sense of more influence of the emotional stimulus on the eventual response time. It is perhaps for this reason that some ABM studies focussed on a contrast based on threat stimulus trials only, yet this index also has a different interpretation from dot probe BI (i.e. it does not correct for attention capturing by any stimulus regardless of emotional content). In addition, the corresponding author for two of these three papers did not respond to our requests. For these reasons we dropped the remaining record with the Posner task completely from our analysis (authors of this RCT assessed all four trial types using ANOVA and concluded that “participants in both groups showed a biased attention away from threat at pre- and at post-assessment” (Boettcher et al., 2011, p.530)).

### Statistical analysis:

The NHST meta-analysis was performed in *R*, using the `RMA()` function in the `metafor` package fitting a DerSimonian and Laird random-effects model (Viechtbauer, 2010). Inputs were mean BI values in ms for the effect size, and sampling variance calculated as  $(sdi^2/ni)$ . `Metafor` functions were also used to assess the model fit, to perform influence tests and the Duval and Tweedie trim and fill procedure, and to create funnel and forest plots.

Bayesian meta-analyses were performed using the `meta.ttestBF` function in the `BayesFactor` package (Morey & Rouder, 2015). The required *t*-values to input for each record (*i*) were calculated as  $(BI_i / \sqrt{sdi^2/ni})$ . Our primary Bayes factor analysis tests the relative likelihood of a point zero hypothesis. It differs from the REML analysis in that `meta.ttestBF()` assesses strength of evidence for a ‘singular underlying true effect’, and is thus essentially a fixed effect analysis. As a secondary Bayesian (and tertiary overall) analysis, we developed effect size interval analyses. To enable these, an overall ‘sigma’ value is calculated as the ‘*n*-weighted mean bias index divided by the mean delta’, in which mean delta is the ‘*n*-weighted average of effect size *d*’, which is computed as ‘ $ti / (\sqrt{ni})$ ’. This value is used to define null-intervals in milliseconds, in order to obtain Bayes factors expressing relative support for BI falling within each of a series of 1 ms-wide-intervals. We’ll introduce the interval-not\_interval Bayes factors further in the results section.

The full analysis script is available as supplemental file xx.

### Results

The final selection consists of  $k = 13$  studies, with a total  $n$  of 1005 clinically anxious individuals (range:  $n = 7$  to  $n = 134$ , table 1).

Table 1: overview of studies selected for inclusion

	study	primary diagnosis	diagnostic instrument	stimulus latency	stimulus type	additional data?	included	<i>n</i> total
1	Boettcher, Leek, Matson, Holmes, Browning, MacLeod, Andersson & Carlbring, 2013	SP / SAD	SCID	500 - <1000	words or words & faces		yes	129
2	Neubauer, von Auer, Murray, Petermann, Helbig-Lang, & Gerlach, 2013	SP / SAD	SCID	500 - <1000	faces	Authors provided values for baseline BI	yes	56
3	Rapee, MacLeod, Carpenter, Gaston, Frei, Peters, & Baillie, 2013	SP / SAD	ADIS	500 - <1000	words	Authors provided values for baseline S1BI	yes	134
4	Schoorl, Putman, & van der Does, 2013	PTSD	CAPS	500 - <1000	scenes		yes	102
5	Boettcher, Hasselrot, Sund, Andersson, & Carlbring, 2014	SP / SAD	SCID	500 - <1000	words & faces		yes	133
6	Kuckertz, Amir, Boffa, Warren, Rindt, Norman, Ram, Ziajko, Webb-Murphy, & McLay, 2014	PTSD	clinician	500 - <1000	words	No response. Mean & SD(BI) inferred from figure 4.	yes	29
7	Badura-Brack, Naim, Ryan, Levy Abend, Khanna, McDermott, Pine, & Bar-Haim, 2015 - S1	PTSD	CAPS	500 - <1000	words		yes	52
8	Badura-Brack, Naim, Ryan, Levy Abend, Khanna, McDermott, Pine, & Bar-Haim, 2015 – S2	PTSD	CAPS	500 - <1000	faces		yes	46
9	Carleton, Teale Sapach, Oriet, Duranceau, Lix, Thibodeau, Horswill, Ubbens, & Asmundson, 2015	SP / SAD	SCID	500 - <1000	words	Authors provided values for baseline BI	yes	82
10	Beard, Fuchs, Asnaani, Schulson, Schofield, Clerkin, & Weisberg, 2016	Panic Disorder	SCID	500 - <1000	faces		yes	7
11	Carleton, Teale Sapach, Oriet, & LeBouthillier, 2016	SP / SAD	SCID	500 - <1000	words	Authors provided values for	yes	90

						baseline BI		
12	Lazarov, Marom, Yahalom, Pine, Hermesh, Bar-Haim, 2017	SP / SAD	LSAS	500 - <1000	faces		yes	50
13	Naim, Kivity., Bar-Haim, Huppert, 2018	SP / SAD	MINI	500 - <1000	faces		yes	95
	Amir, Beard, Burns, & Bomyea, 2009	GAD	SCID-IV	500 - <1000	words	No response	no	29
	Fang, Sawyer, Aderka, & Hofmann, 2013	SP / SAD	ADIS	500 - <1000	faces	Authors kindly informed us that BI data is lost	no	22

## REML analysis

The REML model indicated no evidence that the mean observed BI differs significantly from zero ( $k=13$ ,  $n=1005$ , mean BI = 1.8, SE = 1.26,  $p = .144$ , 95% CI[-.6 - 4.3]. Given SE = 1.26 and  $n = 1005$ , an average bias of 1.8 milliseconds corresponds to a standardized effect size  $d = .05$ .

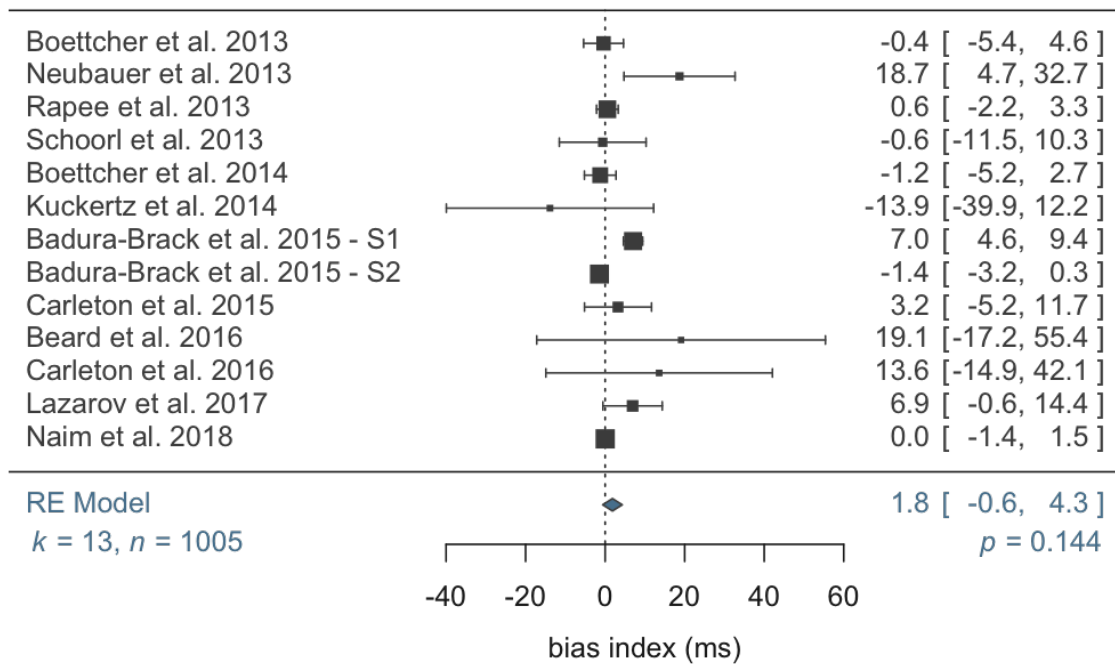


Figure 1. REML analysis forest plot. Estimates are in milliseconds bias.

The test for heterogeneity returned significant ( $Q(12) = 46.3, p < .001$ ), and influence tests indicated that the first study by Badura-Brack and colleagues forms an outlier in this set of studies (studentized residual = 5.2, Cook's distance = 2.5, dfBeta = 4.1). Excluding this record from the analysis, the Q-test no longer indicates heterogeneity ( $Q(11) = 16.8, p = .115$ ). As would be expected from eyeballing the forest plot (fig 1), excluding this record (Badura-Brack et al. 2015 – S1), does not change the result of no support for the hypothesis that the mean BI differs from point zero ( $k=12, n = 953, \text{mean BI} = -.16, \text{SE} = .52, p = .767, 95\% \text{ CI}[-1.8, .9]$ ). Yet, given that it is unclear what causes this record to be an outlier in this collection, we keep it included for the remainder of the analyses unless indicated otherwise. The reader may keep in mind that for any analysis, exclusion of this record would lower the estimated mean average bias index. Two of the included studies (Boettcher et al 2013, and Boettcher et al 2014, assessed bias for more negative information with a mixture of negative-neutral, negative-positive, and positive-negative trials. Excluding these two records (totalling 262 participants) does not meaningfully alter the results ( $k= 11, n= 743, \text{mean BI} = 2.7, \text{SE} = 1.58, p = .084, 95\% \text{ CI}[-.4 - 5.8]$ ). In this subset, the Q-test for heterogeneity is significant, with the first study by Badura-Brack being marked as an outlier. Removing this study in addition to the two mixed-trials studies yields again a homogenous set for which REML analysis returns: ( $k= 10, n= 691, \text{mean BI} = .3, \text{SE} = .78, p = .34, 95\% \text{ CI}[-1.3 - 1.8]$ ).

Figure 2 shows the forest plot for the complete ( $k=13$ ) dataset. The Duval and Tweedie trim and fill procedure gave no indication of publication bias based on this outcome (baseline BI), estimating only one possibly missing small-sized study in the lower left quadrant.

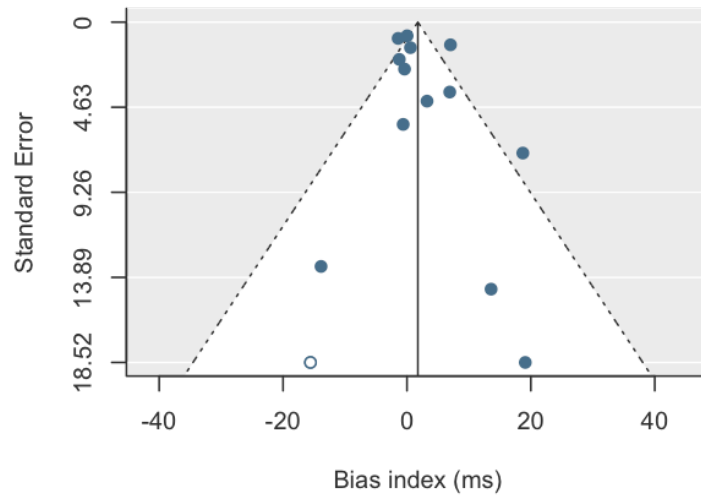


Figure 2. Funnel plot for the 13 included records and one Duval & Tweedie Trim and Fill procedure estimated potentially missing record (white). Estimates are in milliseconds (bias).

### REML subset analyses:

Repeating the REML analysis for subgroups of studies enrolling SP/SAD or PTSD patients or for studies employing only word or face stimuli did not lead to different or additional insights (see table 2).

Table 2: REML analysis for subsets by diagnosis or stimulus type

	<i>k</i>	<i>n</i>	<i>BI</i>	<i>SE</i>	<i>p</i>	<i>95% CI</i>	<i>Duval &amp; Tweedie estimated n missing</i>
SP/SAD	9	769	.4	.59	.531	[ -.8, 1.5]	3
PTSD	4	229	1.3	3.13	.670	[-4.8, 7.5]	1
<i>outlier removed*</i>	3	177	-1.4	.87	.095	[-3.2, .3]	0
Words only	5	387	3.5	2.28	.128	[-1.0, 8.0]	0
<i>outlier removed*</i>	4	335	.78	1.31	.552	[-1.7, 3.3]	0



Faces only	5	254	4.2	3.47	.221	[-2.6, 11.0]	1
<i>outlier removed*</i>	4	198	-.23	.88	.801	[-2.0, 1.5]	1

---

*\*When Q-tests indicated heterogeneity, the most influential studies were removed (one at a time) until the resulting Q-test is no longer significant.*

### **Introduction (interval-not-interval) Bayes factors:**

This part of the results starts with a brief intro on Bayes factors, to aid those less familiar with this outcome measure indicating relative support for one hypothesis over another. In traditional null-hypothesis statistical testing (NHST), we obtain the probability of the observed data given a null hypothesis (often ‘no difference between groups/conditions’, or ‘no difference from point zero’). If the probability of the observed data under the null-hypothesis falls below a certain threshold (typically  $p = .05$ ), the null hypothesis is rejected and, consequently, the alternative hypothesis accepted. Notice that there is no actual testing of the alternative hypothesis involved: in NHST it is only possible to reject or not reject the null hypothesis. With Bayesian analysis things are slightly different. Rather than the probability of the data given the null-hypothesis, we can obtain (and compare) the likelihood of competing hypotheses, given the observed data. Importantly, a low likelihood for one hypothesis does not automatically result in acceptance of an alternative: it is possible to conclude that the available data is insufficient to determine which hypothesis is most likely (also see: Dienes, 2014). The likelihood for one hypothesis over another can be expressed in a ratio called Bayes factor. A BF10 represents the likelihood of a 1-hypothesis over (divided by) the likelihood of a 0-hypothesis: a BF10 with value  $x$  indicates that the 1-hypothesis is  $x$  times as likely as the 0-hypothesis. Its inverse, the BF01, represents the likelihood of the 0-hypothesis over the 1-hypothesis: if  $\text{BF}_{10} = .33$ ,  $\text{BF}_{01} = 3$  (suggesting that the null-hypothesis is three times as likely as the alternative hypothesis). Bayes factors take value 1

when both hypothesis are equally likely given the data, leading to the conclusion of insufficient information/data. Finally, if we know Bayes factors for two alternative hypotheses relative to the same 0-hypothesis, we can divide one by the other to obtain a Bayes factor estimating the evidence in favour of one alternative hypothesis over the other:  $BF_{ab} = BF_{a0} / BF_{b0}$ . We use this ‘trick’ in our analysis to assess the likelihood that the mean bias index falls within a specified interval (iv) rather than outside this same interval (niv). Using the Bayesfactor package’s null-interval option, we first obtain two Bayes factors expressing the relative evidence for the hypotheses that the mean falls within a specified interval (iv) and ‘not in the interval’ (n\_iv), both relative to the null-hypothesis:  $BF_{iv\_0}$  and  $BF_{niv\_0}$ . Next, we divide these two Bayes factors to arrive at the BF for interval over not-interval ( $BF_{iv\_niv} = BF_{iv\_0} / BF_{niv\_0}$ ). This Bayes factor expresses the likelihood that the mean BI falls inside the specified interval relative to the likelihood that it falls outside the interval.

### **Bayesian point nil analysis:**

Returning to our data, the  $BF_{10}$  for ‘standard’ hypotheses  $H_0$ : ‘mean BI is point nil’ relative to  $H_1$ : ‘mean BI is not point nil’ is the Bayes factor equivalent to the NHST test of null-hypothesis ‘mean BI is 0’ assessed above. Using a standard Cauchy prior ( $r = .707$ ),  $BF_{10} = .23$  indicating substantial evidence for the  $H_0$  over the  $H_1$  (Wetzels & Wagenmakers, 2012). In other words: it is about 4.4 times as likely that the mean bias index is point nil than that the mean bias index is not point nil ( $BF_{01} = 1/BF_{10} = 4.4$ )

### **Bayesian ms interval analyses:**

Yet the point nil hypothesis is a very unlikely hypothesis: it tests the likelihood that the estimated mean is exactly 0. For this reason, the authors of the Bayesfactor package

implemented a null-interval option, which can be used to define an interval around zero indicating effect sizes that are considered too small to be of interest (Morey & Rouder, 2011). In the context of dot-probe derived bias this could be a minimum average BI for which consensus exists that such a small difference is likely not meaningful. However, the dot probe literature does not provide many clues as to what minimum BI size would be broadly accepted as being inconsequentially small. Therefore, we opted for a practical rather than a theoretical threshold and consider 1 ms as the smallest possible meaningful unit: millisecond precision of measurement is the absolute best we can hope to achieve with our current hardware and software, even if in practice this will often not be achieved.

For null-interval [-1:1], the  $BF_{iv\_0} = .78$ , indicating ‘anecdotal’ support for the point-zero hypothesis over the hypothesis that the mean falls within an interval of 1 ms around (and including) 0. When we assess the relative support for the hypothesis that mean BI falls outside of the [-1:1] ms interval, the  $BF_{niv\_0}$  is .21, indicating moderate support for the point-zero hypothesis over the hypothesis that BI is larger than 1 ms (in either direction). Finally, by dividing the  $BF_{iv\_0}$  by the  $BF_{niv\_0}$ , we ‘remove’ the point-nil hypothesis from the equation and obtain the BF for the competing hypotheses that the mean BI falls outside an interval of 1 ms to either side of 0, versus that the mean BI falls within this interval. The resulting  $BF_{iv\_niv} = 3.6$ , indicating that it is 3.6 times as likely that mean BI falls inside the [-1 :1] ms interval as that it falls outside this interval.

Finally, we take this analysis-format several steps further by assessing  $BF_{iv\_niv}$  for a series of 14 1-ms-wide intervals that are not centered on zero but ‘move’ along the range from -4 to +10 ms. The results are plotted in figure 4. This figure allows the reader to assess that the strongest (yet not decisive) support is obtained for BI to fall in the [2: 3] ms interval

( $BF_{iv\_niv} = 39.3$ ). It can also be observed that it's highly unlikely for BI to be larger than 2 ms away from threat, or 8 ms towards threat ( $BFs < 1/100$ ). Finally, it can be seen that BI will most likely fall in the 0 to 5 ms range of intervals ( $BFs > 3$ ). Indeed, the  $BF_{iv\_niv}$  for interval  $[0:5] = 369.1$ , which is interpreted as decisive evidence<sup>3</sup>.

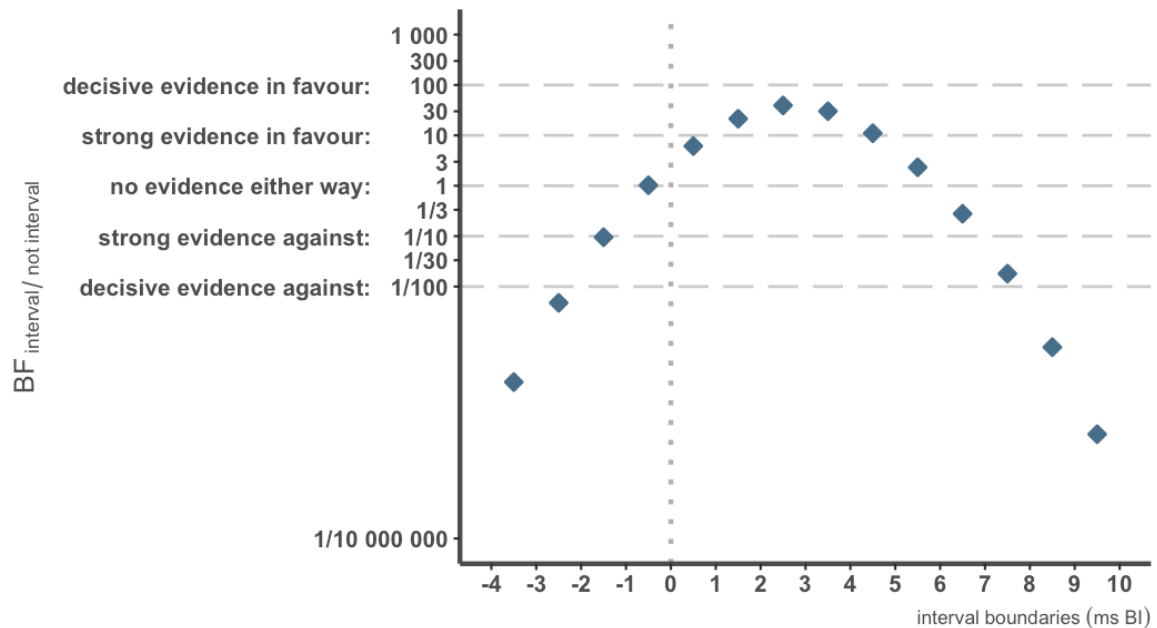


Figure 3. Interval/not interval Bayes factors plot. Intervals are defined in milliseconds bias. Bayes factor evidence labels as defined by Wetzels & Wagenmakers, 2012.

## Discussion

Meta-analysing dot-probe assessed indices of bias towards threat obtained at baseline for 1005 clinically diagnosed patients enrolled in RCTs for ABM, we obtained no evidence for the existence of such a bias using a NHST REML model. The estimated mean bias was 1.8 milliseconds, corresponding to an effect size  $d = .05$ . Furthermore, analysis using Bayes

<sup>3</sup> Note that the denominators (not interval) for the  $BF_{iv\_niv}$  are equally sized but not identical (as they 'move with' the interval defined), and therefore these BF's cannot be used to compute further BF's (for instance summation of the 1 ms intervals to a wider interval).

factors suggests that if a bias exists in these clinically anxious samples, it will most likely fall in the 0 to 5 ms range, which we consider to be likely inconsequentially small.

By meta-analysing baseline measurements from ABM RCTs, we add new meta-analytical evidence based on clinical samples ranging from  $n=7$  to  $n=134$ , with the median sample size ( $n=82$ ) being four times as large as the commonly used sample size of  $n=20$  in extant studies comparing clinical and control groups. The thirteen included studies enrolled a total of 1005 clinically anxious individuals, which is almost three times as many as were included in the ‘dot probe assessed in clinical samples’ sub analysis of the largest meta-analysis of attentional bias to date (Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & van IJzendoorn, 2007).

### **Data repurposing**

We meta-analysed ABM RCT baseline measures, which were not collected to be analysed in order to answer the question we sought to answer. We consider this to be a strength of our design because a) it is less likely that publication bias has affected this body of literature in general (as several studies were preregistered) and specifically not for this measure (precisely because the baseline bias is typically not an outcome of interest for these studies), and b) this is a clearly bounded and much larger body of data than is available from studies aimed at assessing the presences of bias in clinical versus control samples.

### **Specificity of findings to dot-probe assessed bias**

We stress that the current finding is specific to dot-probe assessed attentional bias. We consider dot-probe task assessed bias to be a different bias from, for instance, Stroop task assessed bias, although results obtained with either task were pooled for a major meta-

analysis published in 2007 (Bar-Haim et al., 2007). Dot probe bias reflects additional time required to spatially re-allocate attention if it were to be drawn to a specific position on a display, whereas Stroop bias reflects additional time required to resolve internal response selection. It's worth noting that the dot probe method was designed to overcome shortcomings of the Stroop task (most specifically, the difference in response time that may ensue from the mere presence of negative (versus neutral) stimuli, which is better controlled in the dot probe task which has stimuli of both valences on all trials (MacLeod, Mathews, & Tata, 1986b)). We also note that, back then, there were more, and seemingly more consistent, findings available using the Stroop than the dot-probe methodology ( $k = 77$  versus  $k = 44$  in Bar-Haim et al., 2007). Yet, since then focus has shifted considerably to the dot-probe, which we have previously suggested is at least partly due to the development (and perceived benefits) of dot-probe based ABM procedures (Kruijt, Field, & Fox, 2016). Importantly, the current findings do not rule out that other types of processing bias, or even biased attention allocation assessed with other tasks, play a role in anxiety.

### **implications regarding bias in healthy controls**

Our meta-analysis suggests clinically anxious individuals are not characterized by attention bias towards threat. Thus, if clinically anxious and 'healthy' populations were to differ in their attention orientation tendencies, this is only possible if healthy control samples display an attentional bias that is absent in clinically anxious samples (and theory would suggest that this would be a bias *away* from threat). Pending further research, it seems pertinent to change our manner of speech: if a pattern of reduced threat-avoiding were to be confirmed in sufficiently large clinical samples contrasted with sufficiently large control samples, and using sufficiently reliable measures, we could start characterising anxiety disorders as being associating with a lack of bias rather than implying that bias is

characteristic of anxious individuals. Yet, in absence of even a single qualifying study, we urge everyone involved with this field to reconsider any statements implying that clinical anxiety is shown to be associated with either biased attention towards threat or a lack of bias away from threat.

### **Lack of task reliability and adequate samples**

While we conclude that clinical anxiety is not characterized by dot probe assessed bias, the simplest explanation for our finding is that the dot-probe itself does not reliably measure what it is supposed to measure (Kappenman, Farrens, Luck, & Proudfit, 2014; Kruijt et al., 2016; McNally, 2018; Rodebaugh et al., 2016; Roy, Dennis, & Warner, 2015; Schmukle, 2005; Sigurjónsdóttir, Sigurðardóttir, Björnsson, & Kristjánsson, 2015; Staugaard, 2009; Waechter, Nelson, Wright, Hyatt, & Oakman, 2013; Waechter & Stolz, 2015). In addition, we note that over-reliance and over-generalisation from small analogue samples appears to have been ubiquitous practice in this field (and throughout experimental psychology) for decades. While we are very aware that the question of when to move from analogue to clinical samples is very complex and multi-faceted, it seems safe to say that we have missed a mark in delineating evidence of bias obtained in analogue samples from evidence in clinical samples. As a result, funding has been obtained to perform and enrol over a thousand patients in RCTs for ABM. In hindsight, it may have been more prudent to have first engaged larger patient samples in, less demanding, studies aimed at establishing that their information processing can be reliably observed to differ from healthy controls. Therefore, it seems paramount to refocus our efforts towards enabling verification of the existence of (various forms of) biased information processing in sufficiently large clinically

anxious samples (separately for various disorders), paired with adequate control samples, using demonstrably sufficiently reliable tasks<sup>4</sup>.

### **Implications for cognitive models and development of ABM as a treatment**

The existence of information processing biases is integral to cognitive behavioural theory, which is an important theoretical framework for clinical practice. Information processing-based theories of emotional disorders (e.g. Beck & Clark, 1997; Cisler & Koster, 2010; Dalgleish & Watts, 1990; Heeren, De Raedt, Koster, & Philippot, 2013; Mathews & Mackintosh, 1998; Mogg & Bradley, 1998; Rapee & Heimberg, 1997; Wells & Matthews, 1996; Williams, Watts, MacLeod, & Mathews, 1988) in particular, are founded on the notion that selective processing biases are an important characteristic of emotional disorders (but see: McNally, 2018; Mogg & Bradley, 2018; Roy et al., 2015). The conclusion that such biased processing is not at play has important implications reaching beyond the mere question of whether dot probe assessed bias is a suitable treatment target. There are a couple of things to consider, however. Firstly, various theories differ from each other mainly in the putative mechanisms that would give rise to a bias that might be observed as a higher dot probe bias index in anxious individuals (and for quite a few of the extant theories, the proposed automatic processes may be expected to have played out before the 500 ms timepoint at which most ABM dot probes aim to assess (and modify) attentional bias). Secondly, one ought to consider to what extent the theories predicted dot probe findings and to what extent they have been developed to incorporate and explain the findings of early dot probe literature. Therefore, our results mainly challenge the wisdom of continuing propagation of dot probe based attention bias modification as a treatment. Moreover, in a

---

<sup>4</sup> also see our call for implementation of a standard practice of reporting task reliability in experimental psychology (Parsons, Kruijt, & Fox, 2018).



forthcoming manuscript it is demonstrated that contingency based CBM may be rendered ineffective on a task-mechanical level by the absence of baseline bias (Kruijt & Carlbring, in prep). We conclude that we need to re-evaluate fundamental assumptions of ABM, and that therefore the most valuable approach we can take is to bring attention bias modification research back to its experimental origins. It is imperative to develop valid and reliable methods to assess putative biases and test (and revise) cognitive models accordingly.

## **Conclusion**

Clinically anxious individuals enrolled in RCTs for Attention Bias Modification are not characterized by attention bias towards threat at the start of their trials. This finding casts considerable doubt on the often-repeated assumption that dot probe assessed threat bias plays a role in aetiology and maintenance of clinical anxiety disorders. There also are implications for the on-going efforts to develop procedures to modify this bias with a view to implement these as clinical tools. Because we focussed on a body of data that typically does not include comparison to non-clinical groups, our meta-analysis leaves open the possibility that so-called healthy control or vulnerable samples are characterized by an attentional bias that is absent in clinically anxious individuals. Yet in order to conclude so, such an absence of bias (as well as other processing biases assumed to be associated with mood disorders) should first be verified in studies comparing sufficiently large clinical and control groups using sufficiently reliable bias measures. As a field, we should endeavour to set the record straight on this phenomenon that is commonly declared to characterize clinically anxious individuals.

## Disclosure

Nothing to disclose.

## Acknowledgements

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no. [324176].

## References

- Amir, N., Beard, C., Burns, M., & Bomyea, J. (2009). Attention modification program in individuals with generalized anxiety disorder. *Journal of Abnormal Psychology, 118*(1), 28–33. <https://doi.org/10.1037/a0012589>
- Amir, N., Beard, C., Taylor, C. T., Klumpp, H., Elias, J., Burns, M., & Chen, X. (2009). Attention training in individuals with generalized social phobia: A randomized controlled trial. *Journal of Consulting and Clinical Psychology, 77*(5), 961–73. <https://doi.org/10.1037/a0016685>
- Amir, N., Weber, G., Beard, C., Bomyea, J., & Taylor, C. T. (2008). The effect of a single-session attention modification program on response to a public-speaking challenge in socially anxious individuals. *Journal of Abnormal Psychology, 117*(4), 860–8. <https://doi.org/10.1037/a0013445>
- Badura-Brack, A. S., Naim, R., Ryan, T. J., Levy, O., Abend, R., Khanna, M. M., ... Bar-Haim, Y. (2015). Effect of Attention Training on Attention Bias Variability and PTSD Symptoms: Randomized Controlled Trials in Israeli and U.S. Combat Veterans. *American Journal of Psychiatry*. <https://doi.org/appi.ajp.2015.14121578>
- Bantin, T., Stevens, S., Gerlach, A. L., & Hermann, C. (2016). What does the facial dot-

probe task tell us about attentional processes in social anxiety? A systematic review.

*Journal of Behavior Therapy and Experimental Psychiatry*, 50, 40–51.

<https://doi.org/10.1016/j.jbtep.2015.04.009>

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-](https://doi.org/10.1037/0033-2909.133.1.1)

2909.133.1.1

Beard, C., Fuchs, C., Asnaani, A., Schulson, M., Schofield, C. A., Clerkin, E. M., &

Weisberg, R. B. (2016). A Pilot Open Trial of Cognitive Bias Modification for Panic Disorder. *Cognitive Therapy and Research*, 40(6), 792–798.

<https://doi.org/10.1007/s10608-016-9790-0>

Beck, A. T., & Clark, D. A. (1997). An information processing model of anxiety: Automatic and strategic processes. *Behaviour Research and Therapy*, 35(1), 49–58.

[https://doi.org/10.1016/S0005-7967\(96\)00069-1](https://doi.org/10.1016/S0005-7967(96)00069-1)

Boettcher, J., Berger, T., & Renneberg, B. (2011). Internet-Based Attention Training for

Social Anxiety: A Randomized Controlled Trial. *Cognitive Therapy and Research*, 36(5), 522–536. <https://doi.org/10.1007/s10608-011-9374-y>

Boettcher, J., Hasselrot, J., Sund, E., Andersson, G., & Carlbring, P. (2014). Combining

Attention Training with Internet-Based Cognitive-Behavioural Self-Help for Social Anxiety: A Randomised Controlled Trial. *Cognitive Behaviour Therapy*, 43(1), 34–48.

<https://doi.org/10.1080/16506073.2013.809141>

Boettcher, J., Leek, L., Matson, L., Holmes, E. A., Browning, M., MacLeod, C., ...

Carlbring, P. (2013). Internet-Based Attention Bias Modification for Social Anxiety: A Randomised Controlled Comparison of Training towards Negative and Training

Towards Positive Cues. *PLoS ONE*, 8(9), e71760.

<https://doi.org/10.1371/journal.pone.0071760>

Carleton, R. N., Teale Sapach, M. J. N., Oriet, C., Duranceau, S., Lix, L. M., Thibodeau, M.

A., ... Asmundson, G. J. G. (2015). A randomized controlled trial of attention modification for social anxiety disorder. *Journal of Anxiety Disorders*, 33, 35–44.

<https://doi.org/10.1016/j.janxdis.2015.03.011>

Carleton, R. N., Teale Sapach, M. J. N., Oriet, C., & LeBouthillier, D. M. (2017). Online

attention modification for social anxiety disorder: replication of a randomized controlled trial. *Cognitive Behaviour Therapy*, 46(1), 44–59.

<https://doi.org/10.1080/16506073.2016.1214173>

Cisler, J. M., & Koster, E. H. W. (2010). Mechanisms of attentional biases towards threat in anxiety disorders: An integrative review. *Clinical Psychology Review*, 30(2), 203–16.

<https://doi.org/10.1016/j.cpr.2009.11.003>

Cristea, I. A., Kok, R. N., & Cuijpers, P. (2015). Efficacy of cognitive bias modification interventions in anxiety and depression: meta-analysis. *The British Journal of Psychiatry*, 206(1), 7–16.

<https://doi.org/10.1192/bjp.bp.114.146761>

Daggleish, T., & Watts, F. N. (1990). Biases of attention and memory in disorders of anxiety and depression. *Clinical Psychology Review*, 10(5), 589–604.

[https://doi.org/10.1016/0272-7358\(90\)90098-U](https://doi.org/10.1016/0272-7358(90)90098-U)

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, 5, 781. <https://doi.org/10.3389/fpsyg.2014.00781>

Fang, A., Sawyer, A. T., Aderka, I. M., & Hofmann, S. G. (2013). Psychological treatment of social anxiety disorder improves body dysmorphic concerns. *Journal of Anxiety Disorders*, 27(7), 684–91.

<https://doi.org/10.1016/j.janxdis.2013.07.005>

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior*

*Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/BF03193146>

Gotlib, I. H., Kasch, K. L., Traill, S., Joormann, J., Arnow, B. a, & Johnson, S. L. (2004).

Coherence and specificity of information-processing biases in depression and social phobia. *Journal of Abnormal Psychology*, 113(3), 386–98. <https://doi.org/10.1037/0021-843X.113.3.386>

Hakamata, Y., Lissek, S., Bar-Haim, Y., Britton, J. C., Fox, N. a, Leibenluft, E., ... Pine, D.

S. (2010). Attention bias modification treatment: a meta-analysis toward the establishment of novel treatment for anxiety. *Biological Psychiatry*, 68(11), 982–90. <https://doi.org/10.1016/j.biopsych.2010.07.021>

Harris, L. M., & Menzies, R. G. (1998). Changing attentional bias: Can it effect self-reported anxiety? *Anxiety, Stress & Coping*, 11(2), 167–179.

<https://doi.org/10.1080/10615809808248310>

Hazen, R. a, Vasey, M. W., & Schmidt, N. B. (2009). Attentional retraining: a randomized clinical trial for pathological worry. *Journal of Psychiatric Research*, 43(6), 627–33.

<https://doi.org/10.1016/j.jpsychires.2008.07.004>

Heeren, A., De Raedt, R., Koster, E. H. W., & Philippot, P. (2013). The (neuro)cognitive mechanisms behind attention bias modification in anxiety: proposals based on theoretical accounts of attentional bias. *Frontiers in Human Neuroscience*, 7, 119.

<https://doi.org/10.3389/fnhum.2013.00119>

Kappenman, E. S., Farrens, J. L., Luck, S. J., & Proudfit, G. H. (2014). Behavioral and ERP measures of attentional bias to threat in the dot-probe task: poor reliability and lack of correlation with anxiety. *Frontiers in Psychology*, 5, article 1368: 1-9.

<https://doi.org/10.3389/fpsyg.2014.01368>

Kruijt, A.-W., Field, A. P., & Fox, E. (2016). Capturing dynamics of biased attention: Are new attention variability measures the way forward. *PLoS ONE*, 11(11), e0166600.

<https://doi.org/10.1371/journal.pone.0166600>

Kuckertz, J. M., Amir, N., Boffa, J. W., Warren, C. K., Rindt, S. E. M., Norman, S., ...

McLay, R. (2014). The effectiveness of an attention bias modification program as an adjunctive treatment for Post-Traumatic Stress Disorder. *Behaviour Research and Therapy*, 63, 25–35. <https://doi.org/10.1016/j.brat.2014.09.002>

Lazarov, A., Marom, S., Yahalom, N., Pine, D. S., Hermesh, H., & Bar-Haim, Y. (2017).

Attention bias modification augments cognitive-behavioral group therapy for social anxiety disorder: a randomized controlled trial. *Psychological Medicine*, 1–9.

<https://doi.org/10.1017/S003329171700366X>

Linetzky, M., Pergamin-Hight, L., Pine, D. S., & Bar-Haim, Y. (2015). Quantitative

evaluation of the clinical efficacy of Attention Bias Modification treatment for anxiety disorders. *Depression and Anxiety*, 32(6), 383–391. <https://doi.org/10.1002/da.22344>

MacLeod, C., Mathews, A., & Tata, P. (1986a). Attentional bias in emotional disorders.

*Journal of Abnormal Psychology*, 95(1), 15–20. <https://doi.org/10.1037/0021-843X.95.1.15>

MacLeod, C., Mathews, A., & Tata, P. (1986b). Attentional bias in emotional disorders.

*Journal of Abnormal Psychology*, 95(1), 15–20. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/3700842>

MacLeod, C., Rutherford, E., Campbell, L., Ebsworthy, G., & Holker, L. (2002). Selective

attention and emotional vulnerability: Assessing the causal basis of their association through the experimental manipulation of attentional bias. *Journal of Abnormal*

*Psychology*, 111(1), 107–123. <https://doi.org/10.1037/0021-843X.111.1.107>

Mathews, A., & Mackintosh, B. (1998). A Cognitive Model of Selective Processing in

Anxiety. *Cognitive Therapy and Research*, 22(6), 539–560.

<https://doi.org/10.1023/A:1018738019346>

- Mathews, A., & MacLeod, C. (2002). Induced processing biases have causal effects on anxiety. *Cognition & Emotion*, *16*(3), 331–354.  
<https://doi.org/10.1080/02699930143000518>
- McNally, R. J. (2018). Attentional bias for threat: Crisis or opportunity? *Clinical Psychology Review*. <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](https://doi.org/10.1016/S0005-7967(98)00063-1)
- Mogg, K., & Bradley, B. P. (2018). Anxiety and Threat-Related Attention: Cognitive-Motivational Framework and Treatment. *Trends in Cognitive Sciences*, *22*(3), 225–240.  
<https://doi.org/10.1016/J.TICS.2018.01.001>
- Mogg, K., Holmes, A., Garner, M., & Bradley, B. P. (2008). Effects of threat cues on attentional shifting, disengagement and response slowing in anxious individuals. *Behaviour Research and Therapy*, *46*(5), 656–667.  
<https://doi.org/10.1016/J.BRAT.2008.02.011>
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine*, *6*(7), e1000097. <https://doi.org/10.1371/journal.pmed.1000097>
- Morey, R. D., & Rouder, J. N. (2011). Bayes factor approaches for testing interval null hypotheses. *Psychological Methods*, *16*(4), 406–19. <https://doi.org/10.1037/a0024377>
- Morey, R. D., & Rouder, J. N. (2015). BayesFactor: Computation of Bayes Factors for Common Designs. Retrieved from <http://cran.r-project.org/package=BayesFactor>
- Naim, R., Kivity, Y., Bar-Haim, Y., & Huppert, J. D. (2018). Attention and interpretation bias modification treatment for social anxiety disorder: A randomized clinical trial of efficacy and synergy. *Journal of Behavior Therapy and Experimental Psychiatry*, *59*, 19–30. <https://doi.org/10.1016/j.jbtep.2017.10.006>

- Neubauer, K., von Auer, M., Murray, E., Petermann, F., Helbig-Lang, S., & Gerlach, A. L. (2013). Internet-delivered attention modification training as a treatment for social phobia: A randomized controlled trial. *Behaviour Research and Therapy*, 51(2), 87–97. <https://doi.org/10.1016/j.brat.2012.10.006>
- Parsons, S., Kruijt, A.-W., & Fox, E. (2018). Psychological Science needs a standard practice of reporting the reliability of cognitive behavioural measurements. <https://doi.org/10.17605/OSF.IO/6KA9Z>
- Pergamin-Hight, L., Naim, R., Bakermans-Kranenburg, M. J., van IJzendoorn, M. H., & Bar-Haim, Y. (2015). Content specificity of attention bias to threat in anxiety disorders: A meta-analysis. *Clinical Psychology Review*, 35, 10–18. <https://doi.org/10.1016/j.cpr.2014.10.005>
- Price, R. B., Wallace, M., Kuckertz, J. M., Amir, N., Graur, S., Cummings, L., ... Bar-Haim, Y. (2016, September). Pooled patient-level meta-analysis of children and adults completing a computer-based anxiety intervention targeting attentional bias. *Clinical Psychology Review*. <https://doi.org/10.1016/j.cpr.2016.09.009>
- Rapee, R. M., & Heimberg, R. G. (1997). A cognitive-behavioral model of anxiety in social phobia. *Behaviour Research and Therapy*, 35(8), 741–756. [https://doi.org/10.1016/S0005-7967\(97\)00022-3](https://doi.org/10.1016/S0005-7967(97)00022-3)
- Rapee, R. M., MacLeod, C., Carpenter, L., Gaston, J. E., Frei, J., Peters, L., & Baillie, A. J. (2013). Integrating cognitive bias modification into a standard cognitive behavioural treatment package for social phobia: a randomized controlled trial. *Behaviour Research and Therapy*, 51(4–5), 207–15. <https://doi.org/10.1016/j.brat.2013.01.005>
- Rodebaugh, T. L., Scullin, R. B., Langer, J. K., Dixon, D. J., Huppert, J. D., Bernstein, A., ... Lenze, A. (2016). Unreliability as a Threat to Understanding Psychopathology: The Cautionary Tale of Attentional Bias. *Journal of Abnormal Psychology*, 125(6), 840–851.



<https://doi.org/10.1037/abn0000184>

- Roy, A. K., Dennis, T. A., & Warner, C. M. (2015). A Critical Review of Attentional Threat Bias and Its Role in the Treatment of Pediatric Anxiety Disorders. *Journal of Cognitive Psychotherapy*, 29(3), 171–184. <https://doi.org/10.1891/0889-8391.29.3.171>
- Schmidt, N. B., Richey, J. A., Buckner, J. D., & Timpano, K. R. (2009). Attention training for generalized social anxiety disorder. *Journal of Abnormal Psychology*, 118(1), 5–14. <https://doi.org/10.1037/a0013643>
- 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>
- Schoorl, M., Putman, P., & Does, W. Van Der. (2012). Attentional bias modification in posttraumatic stress disorder: a randomized controlled trial. *Psychotherapy and ....* Retrieved from <http://www.karger.com/Article/FullText/341920>
- Sigurjónsdóttir, Ó., Sigurðardóttir, S., Björnsson, A. S., & Kristjánsson, Á. (2015). Barking up the wrong tree in attentional bias modification? Comparing the sensitivity of four tasks to attentional biases. *Journal of Behavior Therapy and Experimental Psychiatry*, 48, 9–16. <https://doi.org/10.1016/j.jbtep.2015.01.005>
- Staugaard, S. R. (2009). Reliability of two versions of the dot-probe task using photographic faces. *Psychology Science Quarterly*, 51(3), 339–350.
- Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the metafor Package. *Journal of Statistical Software*, 36(3), 1–48. <https://doi.org/10.1103/PhysRevB.91.121108>
- Waechter, S., Nelson, A. L., Wright, C., Hyatt, A., & Oakman, J. (2013). 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>
- Waechter, S., & Stolz, J. A. (2015). Trait Anxiety, State Anxiety, and Attentional Bias to

- Threat: Assessing the Psychometric Properties of Response Time Measures. *Cognitive Therapy and Research*, 39, 441–458. <https://doi.org/10.1007/s10608-015-9670-z>
- Wells, A., & Matthews, G. (1996). Modelling cognition in emotional disorder: The S-REF model. *Behaviour Research and Therapy*, 34(11–12), 881–888. [https://doi.org/10.1016/S0005-7967\(96\)00050-2](https://doi.org/10.1016/S0005-7967(96)00050-2)
- Wetzels, R., & Wagenmakers, E.-J. (2012). A default Bayesian hypothesis test for correlations and partial correlations. <https://doi.org/10.3758/s13423-012-0295-x>
- Williams, J. M. G., Watts, F. N., MacLeod, C., & Mathews, A. (1988). *Cognitive psychology and emotional disorders*. Chichester: Wiley.