Preprint of final manuscript available: psyarxiv.com/rfjup

A meta-analysis of bias at baseline in RCTs of attention bias modification: no evidence for dot-probe bias towards threat in clinical anxiety and PTSD.

Anne-Wil Kruijt[1*], Sam Parsons[2], Elaine Fox[2]

[1] Department of Psychology, Stockholm University
[2] Department of Experimental Psychology, University of Oxford

* corresponding author: dr. Anne-Wil Kruijt

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

Department of Psychology
Stockholm University
SE-106 91 Stockholm
Sweden

Ethics statement: this manuscript describes a meta-analytical study. Because all data originates from previous clinical trials, this study was deemed exempt from ethics committee approval.

***Background*:** Considerable effort and funding have been spent on developing Attention Bias Modification (ABM) as a treatment for anxiety disorders, theorized to exert therapeutic effects through reduction of a tendency to orient attention towards threat. However, meta-analytical evidence that clinical anxiety is characterized by threat-related attention bias is thin. The largest meta-analysis to date included dot-probe data for n=337 clinically anxious individuals. Baseline measures of biased attention obtained in ABM RCTs form an additional body of data that has not previously been meta-analyzed.

***Method*:** This paper presents a meta-analysis of threat-related dot-probe bias measured at baseline for 1005 clinically anxious individuals enrolled in 13 ABM RCTs.

Results: Random-effects meta-analysis indicated no evidence that the mean bias index (BI) differed 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 point-zero hypothesis (BF10 = .23), whereas interval-based analysis indicated that mean bias in clinical anxiety is unlikely to extend beyond the 0 to 5 ms interval.

Discussion: Findings are discussed with respect to strengths (relatively large samples, possible bypassing of publication bias), limitations (lack of control comparison, repurposing data, specificity to dot-probe data), and theoretical and practical context. We suggest that it should no longer be assumed that clinically anxious individuals are characterized by selective attention towards threat.

Conclusion: Clinically anxious individuals enrolled in RCTs for Attention Bias Modification are not characterized by threat-related attention bias at baseline.

Preprint available: psyarxiv.com/rfjup

Keywords: attention bias, clinical anxiety, meta-analysis, attention 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 the ‘dot-probe task’. Over the past decade, 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 analyzed 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

Preferential orienting of attention towards threatening information is theorized to play a role in the etiology and maintenance of (clinical) anxiety[1]. Consequently, it is also considered a putative treatment target in anxiety disorders. Attention Bias Modification (ABM) procedures were initially developed to test whether experimentally inducing or reducing threat-related attention bias results in concomitant changes in anxiety vulnerability. Early findings provided (indirect) experimental evidence for the cognitive theory-derived notion that biased information processing is involved in the etiology and maintenance of anxiety disorders (MacLeod, Rutherford, Campbell, Ebsworthy, & Holker, 2002; Mathews & MacLeod, 2002, yet see: Harris & Menzies, 1998). It was not until 2009, however, that further ABM studies were published, several of which were clinical randomized controlled trials (Amir, Beard, Taylor, et al., 2009; Amir, Beard, Burns, & Bomyea, 2009; Hazen, Vasey, & Schmidt, 2009; Schmidt, Richey, Buckner, & Timpano, 2009). These studies were followed within a year by the first meta-analysis evaluating ABM as a treatment (Hakamata et al., 2010). Thus, focus shifted away from studying attention bias's hypothesized role in the etiology and maintenance of anxiety vulnerability, towards establishing the efficacy of ABM as a treatment for clinical anxiety disorders. Since then, it has been assumed that anxiety disorders are indeed characterized by biased attention towards threat. Yet, the meta-analytical evidence of dot-probe bias towards threat in diagnosed anxious samples is not as strong as might be expected.

[1] In DSM-5 (American Psychiatric Association, 2013), PTSD was moved to a newly defined class of 'Trauma- and stressor-related disorders'. Before that time, it was considered an anxiety disorder. This is reflected in the attention bias and ABM literature, where PTSD is historically understood to be routinely included in the term "anxiety" (Bar-Haim et al., 2007). Throughout this manuscript, we adhere to the DSM-IV definitions and understand the term anxiety disorders to include PTSD.

The first, and largest, meta-analysis on anxiety-related biased attention was published eleven years ago by Bar-Haim and colleagues (Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & van IJzendoorn, 2007). Their meta-analysis includes 172 studies totaling 5869 individual participants and has been cited over 1600 times. Meta-analytical estimates were assessed for bias within healthy control, analog high anxious, and clinically anxious samples, as well as for differences in bias between sample types. Studies were included that measured biased processing of negative information with the emotional Stroop (k = 70/77 within/between comparisons), dot-probe (k = 35/44), or Posner/single cueing task (k = 7/4). Relatively consistent evidence of anxiety-related biased processing was found, with medium effect sizes $d \approx .45$. Given our interest in the evidence supporting ABM's proposed treatment target, and because ABM research relies almost exclusively on (variations of) the dot-probe task, we look at the estimates for specifically dot-probe bias in clinically anxious samples (diagnosed with either generalized anxiety disorder [GAD], obsessive compulsive disorder [OCD], post-traumatic stress disorder [PTSD], panic disorder, social phobia/social anxiety disorder [SP/ SAD], or simple phobia). From table 2 provided by Bar-Haim and colleagues (2007), it can be seen that dot-probe bias differed significantly from 0 (d = .34, 95% CI [.18 - .50]) for 302 clinically anxious participants enrolled in 16 studies. In addition, clinically anxious samples (n = 337 in k = 17) differed significantly from healthy control groups in magnitude of dot-probe bias (table 3: d = .40, 95% CI [.29 - .60][2]). These effects are consistent with those reported for other tasks and anxious analog samples (Bar-Haim et al., 2007).

[2] 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.

However, a serious limitation on these estimates are the small sample sizes. The clinically anxious groups consisted of ~ 20 participants on average.

To the best of our knowledge, only two other meta-analyses assessing anxiety-related attention bias have been published since 2007. In a 2015 meta-analysis (Pergamin-Hight, Naim, Bakermans-Kranenburg, van IJzendoorn, & Bar-Haim, 2015), no evidence for disorder-specific threat bias was found in a subset of six dot-probe studies enrolling 115 clinically anxious individuals (PTSD, PD, SAD, or OCD; d = .12, p = .41). The average sample size of the clinical groups was again ~ 20, and three of these six studies were also included in the 2007 meta-analysis. A 2016 meta-analysis focused on social anxiety related dot-probe bias for negative facial expressions (Bantin, Stevens, Gerlach, & Hermann, 2016). Three out of ten studies compared clinically anxious (n = 89) to healthy control (n = 129) participants on bias for negative faces. Bias was found to differ significantly from 0 across these three clinical samples (g = .48, 95% CI [.17 - .79]). In addition, magnitude of bias differed between the clinically anxious and healthy control samples (g = .38, 95% CI [.10, 0.66]). 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 in the 'measuring and comparing dot-probe bias' literature to date. Thus, from three meta-analyses a picture emerges that the, commonly assumed, phenomenon of dot-probe bias towards threatening information has only been documented for 20-25 small clinically anxious samples.

Increased awareness of statistical power, and the necessity of assessing sufficiently large samples, form a major development in psychology research over the past decade. The meta-analysis by Bar-Haim and colleagues (2007) indicated an estimated medium effect size (d = 0.40) for the comparison of clinical and control samples on threat-related dot-probe bias. For a single study to detect a between-subjects effect of this size, a total n of 198 (99 per group) would be required to achieve 80% power (and a total n of 328 for 95% power). However, more than a decade after the 2007 meta-analysis, no study has compared biased attention between clinical and control groups even approximately this size. Thus, the assumed association between clinical anxiety and dot-probe bias towards threat has not been verified in a single study with sufficient statistical power.

Yet, there exists an additional source of data on threat-related attention bias in clinical anxiety. In several ABM RCTs, threat-related attention bias was assessed before and after the intervention in relatively large (up to n = 134 in Rapee et al., 2013) and relatively well-defined clinical samples. Baseline measures of bias obtained in these RCTs have not previously been meta-analyzed. The previously discussed meta-analyses did not select ABM RCTs either because the meta-analysis predates publication of ABM RCTs (Bar-Haim et al., 2007); because RCTs tend not to allow comparison of bias for disorder-congruent and disorder-incongruent threat stimuli (Pergamin-Hight et al., 2015), or because RCTs tend not to include healthy control groups in addition to clinical or analog groups (Bantin et al., 2016). Yet, recent literature on biased attention in clinical anxiety consists almost entirely of bias modification studies, i.e., studies evaluating ABM interventions (often active versus placebo training) in groups that are not expected to differ at baseline (e.g., all

individuals are clinically anxious or vulnerable at baseline and randomized to treatment conditions). Although the question whether magnitude of bias differs between clinical and control samples is important, the assumption that clinical anxiety is characterized by threat-related bias can be verified using data from ABM RCTs, even if these enrolled only clinically anxious individuals. The dot-probe derived Bias Index (BI) is measured on a bidirectional scale with an inherently meaningful zero value which enables us to test whether the average bias within clinical samples differs from 0 (also see the within-group analyses in Bar-Haim et al., 2007). A one-sample test of mean BI against zero assesses whether individuals in the sample responded, on average, faster (bias towards) or slower (bias away) on trials in which a response cue appeared in the location previously taken by a threat stimulus (congruent trials), compared to trials in which the response cue appeared in the location of a neutral stimulus (incongruent trials; MacLeod, Mathews, & Tata, 1986; MacLeod et al., 2002). If mean BI does not significantly differ from zero, the null hypothesis that no bias is present cannot be rejected.

Apart from adding information based on a body of data that has not yet been meta-analyzed for this question, an additional benefit of meta-analyzing data from ABM RCTs is that it may partly bypass publication bias effects. This is because, in a typical RCT design, the baseline measure is not the outcome of interest. Therefore, bias not being observed at baseline does not necessarily reduce the entire study to a difficult-to-publish null finding. Several published ABM RCTs have been preregistered studies, which strengthens the idea that this subset of the attention bias literature might be less affected by publication bias than the wider literature comparing bias between clinical and control groups.

Thus, we present a meta-analysis of biased attention obtained at baseline in ABM RCTs enrolling clinically anxious samples. The aim of this meta-analysis is to verify the presence of threat-related biased attention, preceding attempted modification thereof. In line with prevailing theory, our formal hypothesis is that biased attention towards threatening information will be observed for the pooled clinically anxious samples[3]. 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 supplemental file S1 (Moher, Liberati, Tetzlaff, Altman, & PRISMA Group, 2009). Although no formal review protocol was prepared, a custom-built review app was built using *R* package shiny (Chang, Cheng, Allaire, Xie, & Jonathan, 2017) and detailed below. The code and required data are included as supplemental files s2a and s2b.

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 ABM studies as possible from the Scopus database (www.scopus.com). The last update to our dataset was done on 20-03-2018 when a search in Scopus using the above string returned 1181 records. The search string used in scopus was:

[3] 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 (also see: Mogg, Waters, & Bradley, 2017). We disclose that our personal expectations run counter to the formal theory-derived hypothesis, although we hope that a meta-analysis may uncover what is not clearly visible in separate studies.

(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")

Once imported in *R*, records were subjected to a filter selecting those records for which at least one of 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. Resulting records were subsequently loaded into a purpose-built 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, answer options were yes, no, and possibly. If a 'no' answer was entered for any 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 reconfirmed. 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 also had to indicate whether individuals with comorbid mood disorder 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). These options were adapted from the procedure described by Bar-Haim and colleagues (2007). 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 BI (bias index) obtained at baseline. From these, mean BI and SD were calculated, collapsing the two (or more) clinical groups within each study. Mean bias was calculated as the n-weighted mean BI (*sum(mean*n)/sum(n)*) and pooled SD as *sqrt(sum(n-1 *SD^2)/sum (n-1))*. 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 indicate their recommendation for inclusion in the meta-analysis.

A total of 394 records were 'stage I assessed' by each of the two assessors, who selected 36 and 37 records for stage II assessment respectively (29 records in common). Following their individual stage II assessments, all data was gathered and used 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 identical values for each of these records. This resulted in the discovery of one typing error and one mix-up of values, which were 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 meta-analyses were also selected for the current meta-analysis. 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 one 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),

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 employing the Posner task were identified as eligible for inclusion (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 (cue facilitation or inhibition of return), opposing scores can be interpreted as indicative of more bias in the sense of more influence of the emotional stimulus on the response time. It is perhaps for this reason that some ABM studies focused on an index based on threat stimulus trials only, yet this contrast does not correct for attention capturing (or inhibition of return) invoked 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 main meta-analysis was performed in *R,* using the RMA() function in the metafor package to a fit a restricted maximum-likelihood model (REML; Viechtbauer, 2010). Inputs were mean BI values in ms for the effect size, and sampling variance calculated as (SD*i*^2/n*i*). 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 M_i / sqrt($SD^{\wedge 2}_i$ / n_i). 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 therefore 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 was calculated as the 'n-weighted mean bias index divided by mean delta', in which mean delta is the 'n-weighted average of effect size *d*', and *d* is computed as t_i / sqrt(n_i). The overall sigma value was 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 ms-wide intervals. We will introduce the interval-not_interval Bayes factors further in the results section.

The full analysis script and data are available as supplemental files S3a and S3b.

Results

The final selection consisted of k = 13 studies, with a total *n* of 1005 clinically anxious individuals (range: *n* = 7 to *n* = 134; see table 1 for details).

Table 1: overview of studies selected for inclusion

	study	primary diagnosis	diagnostic instrument	stimulus latency	stimulus type	additional data?	included	*n* 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,	PTSD	CAPS	500 - <1000	faces		yes	46

	McDermott, Pine, & Bar-Haim, 2015 – S2							
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 baseline BI	yes	90
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	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 that mean BI does not differ 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. The forest plot for the complete (k = 13) dataset is presented in figure 1.

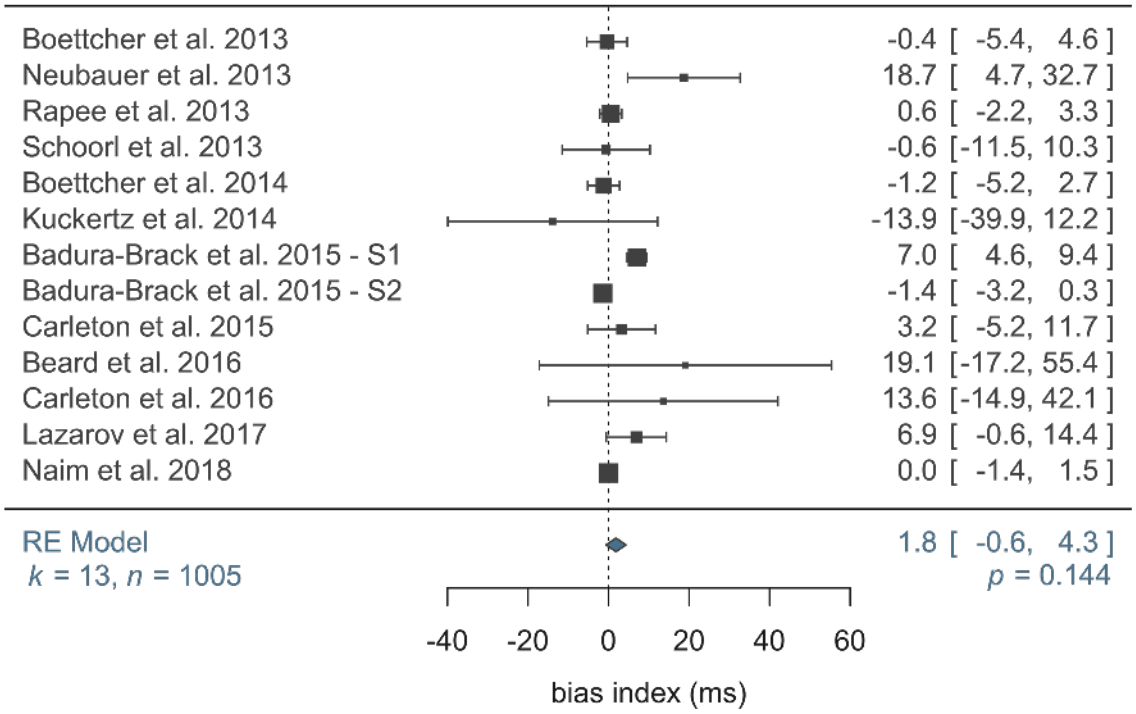

Figure 1. REML analysis forest plot. Estimates are in milliseconds bias.

The Q-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). When 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 (figure 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 ($k = 12$, $n = 953$, mean BI = -.16, $SE = .52$, $p = .767$, 95% CI [-1.8, .9]). Yet, given that it is unclear what caused this record to be an outlier in this collection, the record was retained 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.

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).

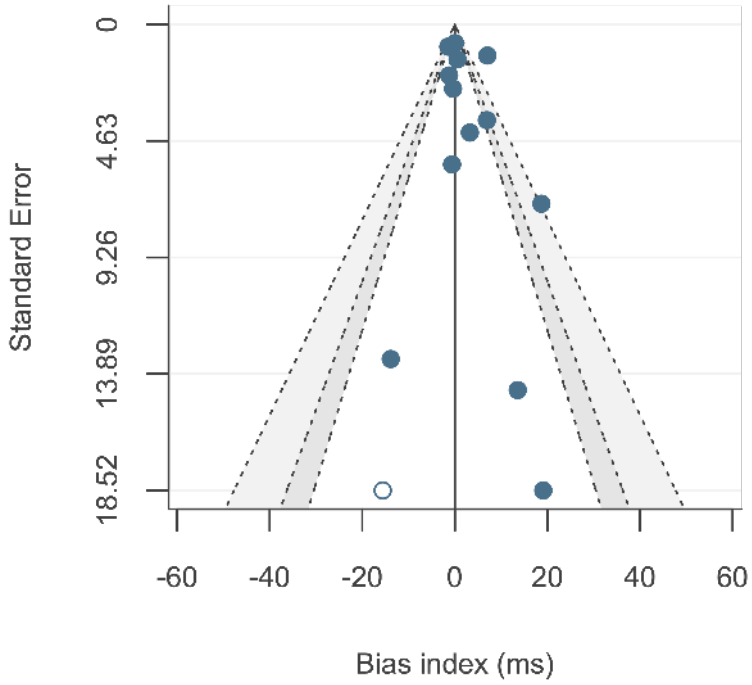

Figure 2. *Contour-enhanced funnel plot for the 13 included records and one Duval & Tweedie trim and fill procedure estimated potentially missing record (white). The contour lines indicate (from inside to outside) the boundaries for* p = *.10, .05, and .01. Estimates are in milliseconds (bias).*

REML subset analyses:

Repeating the REML analysis for subgroups of studies enrolling SP/SAD or PTSD patients and for subgroups of studies employing (only) word or face stimuli did not lead to different or additional insights (see table 2). It should be noted, however, that the funnel plot for the SP/SAD subset shows substantial asymmetry, which may indicate publication bias. Duval & Tweedie trim and fill procedure suggest that in this set three studies are missing on the left side (see supplemental file S4).

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 (totaling 262 participants) does not meaningfully alter the results (*k* = 11, *n* = 743, mean BI = 2.7, *SE* = 1.58, *p* = .084, 95% 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, mean BI = .3, *SE* = .78, *p* = .34, 95% CI [-1.3 - 1.8]). The funnel plot for this set also shows asymmetry (see supplemental file S4).

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

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

Bayes factor analyses:

Bayes factors are indices of relative support for one hypothesis over another. In traditional null hypothesis statistical testing (NHST), the probability of the observed data given a null hypothesis (e.g., 'no difference from zero') is calculated. 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 is accepted. Notice that there is no actual testing of the alternative hypothesis involved: in NHST it is only possible to reject or to not reject the null hypothesis. Rather than the probability of the data given the null hypothesis, Bayesian analysis provides 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 of one hypothesis over another can be expressed in a ratio called Bayes factor. A BF10 represents the likelihood of an alternative hypothesis over (divided by) the likelihood of a null hypothesis: a BF10 with value x indicates that the alternative hypothesis (H_1) is x times as likely as the null hypothesis (H_0). Its inverse, the BF01, represents the likelihood of the null hypothesis over the alternative hypothesis: if BF10 = .33, BF01 = 3 (indicating that, based on the available data, the null hypothesis is three times as likely as the alternative hypothesis). Bayes factors take value 1 when both hypotheses are equally likely given the data, leading to the conclusion of insufficient information/data. Finally, when we know Bayes factors for two alternative hypotheses relative to the same null hypothesis, we can divide one by the other to obtain a Bayes factor estimating the evidence in favor of one alternative hypothesis

over the other: BFab = BFa0 / BFb0. We use this method in our analyses to assess the likelihood that the mean bias index falls within a specified interval rather than outside this same interval. Using the Bayesfactor package 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' (niv), both relative to the null hypothesis: BF_iv_0 and BF_niv_0. Next, we divide these two Bayes factors to arrive at the Bayes factor 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 zero analysis:

The Bayes factor most similar to the NHST assessed one-sample test of null hypothesis 'mean BI is zero', is a BF10 comparing the hypotheses 'H_0: mean BI is zero' and 'H_1: mean BI is not zero'. Using a standard Cauchy prior (r = .707), BF10 = .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 zero than that the mean bias index is not point zero (BF01 = 1/BF10 = 4.4).

Bayesian ms-wide interval analyses:

Yet the point zero 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 mean 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 took 1 ms to be the smallest possible meaningful unit: millisecond precision of measurement is the absolute best we can hope to achieve with our current hard- 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). Next, we 'remove' the point-zero hypothesis from the equation (by dividing the BF_iv_0 by the BF_niv_0) 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 3. This figure allows the reader to assess that strong (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

is highly unlikely for BI to be larger than 2 ms away from threat, or 8 ms towards threat (BFs < 1/100). In addition, 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 the 5-ms-wide interval [0:5] = 369.1, which is a BF value that is typically interpreted as decisive evidence[4].

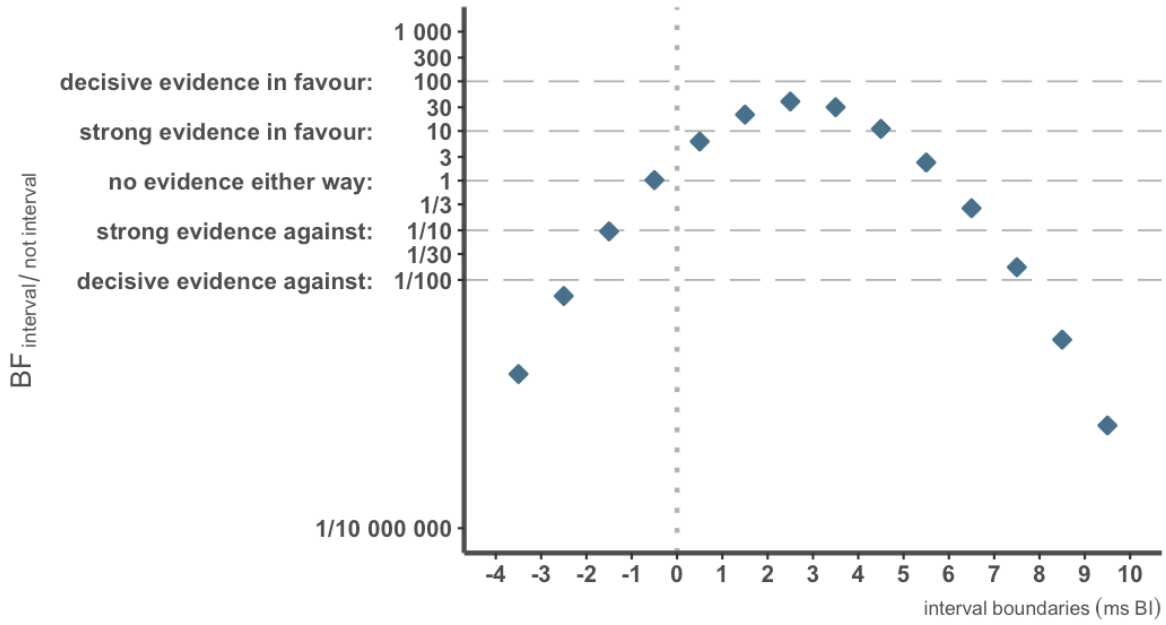

Figure 3. *Interval/not interval Bayes factors plot. Intervals are defined in milliseconds bias. Bayes factor evidence labels as defined by Wetzels & Wagenmakers, 2012.*

Discussion

The current meta-analysis found no evidence for threat-related attention bias in clinically anxious individuals. Data consisted of mean threat-related dot-probe indices obtained from 13 RCTs for ABM, representing a total of 1005 clinically

[4] 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 BFs cannot be used to compute further BFs (for instance summation of the 1 ms intervals to a wider interval).

diagnosed patients. REML estimated mean bias was 1.8 milliseconds, corresponding to a standardized effect size *d* = .05. Secondary analysis using Bayes factors suggested that if attention bias exists in these clinically anxious samples it most likely falls within the 0 to 5 ms range, which we consider to be inconsequentially small.

This meta-analysis adds evidence based on studies not previously included in meta-analyses of biased attention in clinical anxiety. The included clinical samples ranged in size 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 measuring and comparing bias between clinical and control groups. The thirteen included studies enrolled a total of 1005 clinically anxious individuals, which is about three times as many as were included in the 'dot-probe bias in clinically anxious samples' sub-analyses of the largest meta-analysis of attentional bias to date (Bar-Haim et al., 2007).

Data repurposing

We meta-analyzed baseline measures from ABM RCTs, which were not collected to be analyzed 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 since baseline bias is typically not an outcome of interest for these studies, and b) this is a larger body of data than is available from studies designed to measure and compare bias, enrolling often relatively well-defined samples of clinically anxious patients.

Specificity of findings to dot-probe bias

This meta-analysis is specific to threat-related attention bias as measured with the dot-probe. We did not include studies measuring bias by means of other tasks such as the emotional Stroop task, though results obtained with either task have been pooled in the past (Bar-Haim et al., 2007). The main reason is that ABM procedures are predominantly dot-probe based. Further, dot-probe and emotional Stroop tasks likely measure different cognitive processes in the sense that dot-probe bias is assumed to reflect the additional time required to spatially re-allocate attention when it has been drawn to a specific position on a display, whereas emotional Stroop bias is thought to reflect the additional time required to resolve a potential conflict in internal response selection. In 2007, considerably more published studies had employed the emotional Stroop than the dot-probe task (k = 77 versus k = 44 in Bar-Haim et al., 2007), yet since then focus has shifted to the dot-probe task. Although the dot-probe task was specifically designed to overcome shortcomings of the emotional Stroop task (MacLeod et al., 1986), we have previously suggested that the near-exclusive focus on dot-probe over the past decade may be at least partly due to the development of dot-probe based ABM procedures (Kruijt, Field, & Fox, 2016). Nonetheless, the current findings do not rule out the possibility that other types of processing bias, measured with other tasks, play a role in clinical anxiety. We are, however, hesitant to imply that the emotional Stroop, or any other currently existing measure of attention bias, might provide a more suitable basis for ABM. We discuss problems to do with task reliability and reliance on analog samples further below and suggest that these may apply also to other, currently existing, bias assessment methods.

Implications regarding attention bias in healthy controls

The current meta-analytical results indicate that clinically anxious individuals are not characterized by threat-related attention bias. Yet, our methods did not permit comparison of bias magnitude between clinically anxious and healthy control samples. It is, therefore, possible that healthy control samples display a dot-probe bias that is absent in clinically anxious samples. Non-anxious controls might, for instance, demonstrate a bias *away* from threat, which we found to be absent in clinically anxious samples. This would be a rather different pattern, however, from the pattern implied by the common statement that clinically anxious individuals are characterized by threat-related attention bias. Future research could investigate whether there is evidence for such differential threat-related biased attention by contrasting sufficiently large clinical and control samples on sufficiently reliable measures Yet given the current results, and in lieu of even a single qualifying study, we propose reconsidering any statements implying that clinical anxiety is shown to be associated with threat-related biased attention.

Lack of task reliability and adequate samples

While we conclude that clinical anxiety is not characterized by biased attention assessed with the dot-probe task, the simplest explanation for our finding is that the dot-probe task does not reliably assess biased attention. Several recent studies assessed internal reliability of the dot-probe bias index and found it to be unacceptably low (e.g., reliability estimates varying between -.70 and .59 were reported in the following studies: Bar-Haim et al., 2010; Brown, Eley, & Broeren, 2014; Enock, Hofmann, & McNally, 2014; Kappenman, Farrens, Luck, & Proudfit, 2014; Price et al., 2014; Rodebaugh et al., 2016; Schmukle, 2005; Staugaard, 2009;

Waechter, Nelson, Wright, Hyatt, & Oakman, 2013; Waechter & Stolz, 2015; in addition see: Kruijt et al., 2016; McNally, 2018; Mogg & Bradley, 2018; Parsons, Kruijt, & Fox, 2018; Roy, Dennis, & Warner, 2015; Sigurjónsdóttir, Sigurðardóttir, Björnsson, & Kristjánsson, 2015; Jones, Christiansen, & Field, 2018). This raises two important questions. First, whether any findings ever reported based on the dot-probe task have been reliable, and secondly why the notion that clinical anxiety is characterized by dot-probe related threat-bias became so well-established that more than a thousand anxiety patients have been enrolled in RCTs for dot-probe based ABM. Although a reasonably large number of studies assessing dot-probe bias in small clinically anxious samples have been published, a substantial part of the data indicating biased attention in anxiety-related information processing has been obtained from analog samples. Over-reliance on, and over-generalization from, analog samples has been common practice in the biased information-processing field (and indeed throughout experimental psychology) for decades. Similarly, awareness of the problems associated with small sample sizes and the probability of publication bias have long remained low (Tackett, Brandes, King, & Markon, n.d.). These four factors (low task reliability, over-reliance on analog samples, low sample sizes, and potential publication bias) may have contributed to the field being firmly under the impression that that anxiety-related biased attention was well established when bias modification methods were first developed, when the evidence base was, in fact, not as strong as was assumed. Two quotes from the seminal paper on ABM (MacLeod et al., 2002) may serve to illustrate this. In the introduction, it is stated that "*Although the existence of this association between anxiety vulnerability and negative attentional bias now stands beyond contention, no compelling evidence yet has served to establish the causal nature of the relationship*" (MacLeod et al., 2002,

p. 108). The paper then details two studies (each enrolling *n*=64 students) in which it was found that engaging in different versions of a training task, now known as dot-probe ABM, resulted in differential reactivity to a laboratory stressor procedure. Macleod and colleagues end their manuscript expressing that "*we hope that this research may signal the commencement of a new chapter within this literature, characterized by a collective endeavor to exploit the therapeutic potential of novel cognitive–experimental procedures, designed to directly modify the patterns of distorted information processing known to be associated with emotional pathology*" (MacLeod et al., 2002, p. 121). Five years later, the meta-analysis reported by Bar-Haim and colleagues (Bar-Haim et al., 2007) confirmed the assumption that the available literature indicated a consistent threat-related attention bias for clinical as well as analog groups, and ABM RCTs were published from 2009 onwards. In hindsight, the field (ourselves included) missed the fact that the very large 2007 meta-analysis included data on dot-probe bias for only a small number of clinically anxious individuals (*n* = 302/337 for the within/between analyses). In retrospect, it might have been better if larger patient samples had first been engaged in, relatively less demanding, studies aimed at establishing whether their information processing tendencies can be reliably observed to differ from healthy controls. By meta-analyzing the RCT baseline bias measures, the current meta-analysis provides part of the information that could have been obtained from such studies.

Implications for cognitive models and development of ABM as a treatment

The existence of information processing biases is integral to cognitive behavioral theory, which is an important framework for clinical practice. Information processing-based theories of emotional disorders (e.g. Beck & Clark, 1997; Cisler & Koster,

2010; Eysenck, Derakshan, Santos, & Calvo, 2007; Mathews & Mackintosh, 1998; Mogg & Bradley, 1998; Öhman, 1993; Rapee & Heimberg, 1997; Wells & Matthews, 1996; Williams, Watts, MacLeod, & Mathews, 1988), in particular, rely on the notion that biased information processing is characteristic of emotional disorders. A conclusion that clinical anxiety is not characterized by dot-probe bias towards threat has implications beyond the question of whether dot-probe bias forms a suitable treatment target. Yet, implications of the current meta-analysis will vary for various existing theories. It will be to the wider field to parse our findings with theory and future research. It is, for instance, worth noting that the currently analyzed bias indices all derived from dot-probe tasks using a stimulus duration of 500 ms. While this represents the standard task used in ABM research, several theoretical accounts may be interpreted to predict that bias is optimally measured at earlier timeframes. Moreover, some theories assume involvement of additional factors that are typically not assessed in ABM RCTs and were also not assessed in this meta-analysis. Cognitive-motivational models of anxiety, for instance, propose that multiple motivational and cognitive control processes interact and that this may result in anxiety-related attention being biased *away* from threat as well as *towards* threat (Mogg & Bradley, 2018). Verification of such variable bias will require newly developed tasks to provide a reliable estimation of individual bias. The current result of average dot-probe derived BI not differing from zero, however, could be interpreted to suggest that bias may be balanced, i.e., that at any given time-point a roughly equal proportion of clinically anxious individuals tends to orient towards and away from threat. With respect to the ongoing development of ABM, our results challenge the assumption that reducing threat-related biased attention will reduce anxiety vulnerability. All the more so because ABM may be rendered ineffective, on a

task-mechanical level, by the absence of baseline bias (Kruijt & Carlbring, 2018). Thus, the current results suggest that fundamental assumptions of ABM need to be re-evaluated and provide relevant information for theoretical revisions.

Conclusion

Clinically anxious individuals enrolled in RCTs for Attention Bias Modification are not characterized by threat-related attention bias at the start of their trials. The field should endeavor to set the record straight on this phenomenon that is commonly declared to characterize clinically anxious individuals. We propose that it will be important to a) develop better and more reliable ways of assessing information processing biases, and b) explore theoretical approaches that do not specifically predict preferential orienting towards threat to constitute a central feature of clinical anxiety disorders.

Disclosure

Nothing to disclose.

Acknowledgments

The research reported here was funded by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no. [324176]. Preliminary analyses on a smaller subset of studies were presented at the 9th Swedish Congress on internet interventions (SWEsrii) in November 2017. A preprint version of this manuscript is available at https://psyarxiv.com/rfjup.

References

References marked with an asterisk indicate studies included in the meta-analysis

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders : DSM-5.* (DSM-5 Task Force, Ed.). Arlington: American Psychiatric Publishing.

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), 961–973. 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), 860–868. 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*.

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., Holoshitz, Y., Eldar, S., Frenkel, T. I., Muller, D., Charney, D. S., … Wald, I. (2010). Life-Threatening Danger and Suppression of Attention Bias to Threat. *American Journal of Psychiatry*, *167*(6), 694–698. https://doi.org/10.1176/appi.ajp.2009.09070956

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

* 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

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–e71760. https://doi.org/10.1371/journal.pone.0071760

Brown, H., Eley, T., & Broeren, S. (2014). Psychometric properties of reaction time based experimental paradigms measuring anxiety-related information-processing biases in children. *Journal of Anxiety Disorders*, *28*, 97–107. https://doi.org/10.1016/j.janxdis.2013.11.004

* 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

Chang, W., Cheng, J., Allaire, J., Xie, Y., & Jonathan, M. (2017). shiny: Web Application Framework for R.

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), 203–216. 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

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

Enock, P. M., Hofmann, S. G., & McNally, R. J. (2014). Attention Bias Modification Training Via Smartphone to Reduce Social Anxiety: A Randomized, Controlled Multi-Session Experiment. *Cognitive Therapy and Research*, *38*(2), 200–216. https://doi.org/10.1007/s10608-014-9606-z

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion*, *7*(2), 336–353. https://doi.org/10.1037/1528-3542.7.2.336

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), 386–398. 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), 982–990. 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), 627–633. https://doi.org/10.1016/j.jpsychires.2008.07.004

Jones, A., Christiansen, P., & Field, M. (2018). Failed attempts to improve the reliability of the alcohol visual probe task following empirical recommendations. *Psychology of Addictive Behaviors*. https://doi.org/10.1037/adb0000414

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., & Carlbring, P. (2018). Mechanics of contingency-based Cognitive Bias Modification: pre-existing bias affects potency of active training but not placebo conditions. *PsyArXiv*. https://doi.org/10.31234/osf.io/scqf3

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. (1986). Attentional bias in emotional disorders. *Journal of Abnormal Psychology*, *95*(1), 15–20.

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

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

Mogg, K., Waters, A. M., & Bradley, B. P. (2017). Attention Bias Modification (ABM): Review of Effects of Multisession ABM Training on Anxiety and Threat-

Related Attention in High-Anxious Individuals. *Clinical Psychological Science*, *5*(4), 698–717. https://doi.org/10.1177/2167702617696359

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), 406–419. https://doi.org/10.1037/a0024377

Morey, R. D., & Rouder, J. N. (2015). BayesFactor: Computation of Bayes Factors for Common Designs.

* 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

Öhman, A. (1993). Fear and anxiety as emotional phenomena: Clinical phenomenology, evolutionary perspectives, and information-processing mechanisms. In *Handbook of emotions* (pp. 511–536). New York, NY, US: Guilford Press.

Parsons, S., Kruijt, A.-W., & Fox, E. (2018). Psychological Science needs a standard practice of reporting the reliability of cognitive behavioural measurements. *Psyarxiv*. 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., Kuckertz, J. M., Siegle, G. J., Ladouceur, C. D., Silk, J. S., Ryan, N. D., … Amir, N. (2014). Empirical Recommendations for Improving the Stability of the Dot-Probe Task in Clinical Research. *Psychological Assessment*, *27*(2), 365–376. https://doi.org/10.1037/pas0000036

Price, R. B., Wallace, M., Kuckertz, J. M., Amir, N., Graur, S., Cummings, L., … 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

* 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), 207–215. 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. V. D. (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.

Tackett, J. L., Brandes, C. M., King, K. M., & Markon, K. E. (n.d.). Psychology's Replication Crisis and Clinical Psychological Science. https://doi.org/10.31234/osf.io/kc8xt

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

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.