International Journal of Psychophysiology 155 (2020) 87–98

Contents lists available at ScienceDirect

International Journal of Psychophysiology

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

Assessing and adjusting for publication bias in the relationship between anxiety and the error-related negativity

Blair Saunders[a,*], Michael Inzlicht[b,c]

[a] Psychology, School of Social Sciences, University of Dundee, UK
[b] Department of Psychology, University of Toronto, Canada
[c] Rotman School of Management, University of Toronto, Canada

ARTICLE INFO

Keywords:
ERN
Anxiety
Meta-analysis
Publication bias

ABSTRACT

Many clinical neuroscience investigations have suggested that trait anxiety is associated with increased neural reactivity to mistakes in the form of an event-related potential called the error-related negativity (ERN). Several recent meta-analyses indicated that the anxiety-ERN association was of a small-to-medium effect size, however, these prior investigations did not comprehensively adjust effect sizes for publication bias. Here, in an updated meta-analysis ($k = 58$, $N = 3819$), we found support for an uncorrected effect size of $r = -0.19$, and applied a range of methods to test for and correct publication bias (trim-and-fill, PET, PEESE, Peters' test, three-parameter selection model). The majority of bias-correction methods suggested that the correlation between anxiety and the ERN is non-zero, but smaller than the uncorrected effect size (average adjusted effect size: $r = -0.12$, range: $r = -0.05$ to -0.18). Moderation analyses also revealed more robust effects for clinical anxiety and anxious samples characterised by worry, however, it should be noted that these larger effects were also associated with elevated indicators of publication bias relative to the overall analysis. Mixed anxiety and sub-clinical anxiety were not associated with the amplitude of the ERN. Our results suggest that the anxiety-ERN relationship survives multiple corrections for publication bias, albeit not among all sub-types and populations of anxiety. Nevertheless, only 50% of the studies included in our analysis reported significant results, indicating that future research exploring the anxiety-ERN relationship would benefit from increased statistical power.

1. Introduction

Safe and flexible behaviour depends on the ability to detect and compensate for mistakes. In addition to guiding adaptive behaviours, errors typically coincide with potential harm, including financial losses, embarrassment, and physical injury, meaning that errors are also motivationally significant (Hajcak, 2012). While mistakes are aversive for most people, error reactivity appears to be particularly high in clinical and subclinical forms of anxiety (Weinberg et al., 2012b). Considerable support for the relationship between anxiety and error-processing comes from clinical neuroscience investigations of an event-related potential called the error-related negativity (ERN; Gehring et al., 1993). Here, we applied bias-correction methods to explore the extent and impact of publication bias in the literature supporting the anxiety-ERN relationship.

1.1. Anxiety and the ERN

Anxiety is a multi-faceted construct that includes cognitive aspects (e.g., worry), as well as bodily symptoms (e.g., physiological arousal), and avoidant behaviours (e.g., avoiding socialising, work, etc.). The responses that characterize anxiety can be highly adaptive—stimulating increased attention and defensive action in response to threats (Nesse, 2005; Bergstrom and Meacham, 2016). However, more chronic forms of anxiety (e.g., generalized anxiety disorder; obsessive-compulsive disorder; social anxiety disorder) are also among the most prevalent and persistent mental health disorders world-wide (Baxter et al., 2014). This significant disease-burden has prompted psychopathological research exploring the neuro-cognitive underpinnings of anxiety, with one consistent finding indicating that enhanced neural reactivity to mistakes—in particular the amplitude of the ERN—is a biologically meaningful correlate of trait anxiety (Moser et al., 2013; Olvet and Hajcak, 2008; Weinberg, Dietrich, & Riesel, 2015).

The ERN is a negative-going response-locked ERP that peaks within

* Corresponding author at: Scrymgeour Building, University of Dundee, Dundee DD1 4HN, UK.
E-mail address: b.z.saunders@dundee.ac.uk (B. Saunders).

https://doi.org/10.1016/j.ijpsycho.2020.05.008
Received 19 August 2019; Received in revised form 24 April 2020; Accepted 25 May 2020
Available online 30 May 2020

100 ms of mistakes, and is putatively generated by the anterior midcingulate cortex (Dehaene et al., 1994). The ERN possesses psychometric properties supporting its use as a measure of individual difference: the component is internally consistent (Olvet and Hajcak, 2009); reliable over time (Weinberg and Hajcak, 2011); and is heritable (Anokhin et al., 2008). While a great deal of debate has been generated regarding the exact psychological mechanisms that give rise to the ERN, a considerable body of evidence suggests that the component is at least sensitive to the motivational significance of mistakes (Hajcak et al., 2005; Hajcak, 2012; for a recent review see Saunders et al., 2017). Most important for present concerns, enhanced ERN amplitudes have been reported across multiple forms of anxiety, including generalized anxiety disorder (Weinberg, Olvet, & Hajcak, 2010); obsessive-compulsive disorder (OCD, Gehring et al., 2000); and related individual differences in subclinical negative affectivity (Amodio et al., 2008). These findings are consistent with the idea that the ERN in part reflects the increased saliency and personal significance of mistakes among those with varying levels of anxiety.

The anxiety-ERN relationship persists after symptom remission (Stern et al., 2010), and enhanced ERN amplitudes are present in unaffected first-degree relatives of OCD patients (Riesel et al., 2011). This evidence suggests that the ERN may be a heritable, biologically-meaningful vulnerability marker for several forms of clinical anxiety (Olvet and Hajcak, 2008; Weinberg et al., 2015). Synthesising across many of these studies, two recent meta-analyses (Cavanagh and Shackman, 2015; Moser et al., 2013) supported negative small-to-medium sized correlations between trait anxiety and the ERN (Moser et al., 2013: $r = -0.25$; Cavanagh and Shackman, 2015: $r = -0.28$).

Anxiety is not a unitary construct, and it has been proposed that separable aspects of anxiety are differentially associated with ERN amplitude (Moser et al., 2013). One useful distinction in the present context is between cognitive aspects of anxiety comprising worry and rumination, often in response to ambiguous threats (i.e., *anxious apprehension*), and somatic and physiologic features of anxiety that typically arise in reaction to present threat (i.e., *anxious arousal*). It has recently been proposed that the anxiety-ERN relationship might be particularly related to anxious apprehension/worry (Moser et al., 2012; Zambrano-Vazquez and Allen, 2014; however, see Gorka et al., 2017). Supporting this suggestion, a meta-analysis by Moser et al. (2013) indicated that the anxiety-ERN link is largest in magnitude ($r = -0.35$) for *anxious apprehension* (comprising OCD, generalized anxiety disorder, social anxiety disorder, and behavioural inhibition) rather than *mixed anxiety* ($r = -0.09$), comprising a range of non-specific anxiety measures that to varying degrees combine anxious apprehension with anxious arousal and/or depression-related symptoms.

While these findings can be interpreted to support the view that the anxiety-ERN relationship reflects compensatory control mechanisms linked to worry and verbal rumination (Moser et al., 2013), other work points to a retained role for affective, visceral, and somatic contributions to anxiety related enhancement of the ERN (Hajcak et al., 2003; Gorka et al., 2017). Consequently, the exact dimensions of anxiety that underlie its relationship with ERN are still debated, representing an important avenue for ongoing research. Here, we retain the worry-mixed anxiety contrast to replicate an existing empirical moderation that exists within the meta-analytical literature on the ERN (i.e., Moser et al., 2013), while acknowledging that the exact psychological mechanisms underlying the anxiety-ERN relationship are outside the scope of the current review.

1.2. The current study

In the current study we applied a range of bias-correction tools to test the robustness of the anxiety-ERN relationship. While meta-analysis is a powerful tool to synthesise empirical data across studies, meta-analytic estimates are highly susceptible to common biases that exist in academic publishing (Rothstein et al., 2006). First, publication bias, also known as the file-drawer problem, occurs when statistically significant results enter the literature more frequently than null results. Second, questionable research practices (or p-hacking) refer to cases in which researchers analyse data flexibly with the goal to reduce the *p*-value below a given threshold (Simmons et al., 2011). Both practices mean that naïve[1] meta-analytical estimates can substantially overestimate the size of a given effect.

To combat this issue, several statistical tools have been developed to assess and adjust for bias in the published literature. Such a quantitative assessment of publication bias was never conducted on either of the prior meta-analyses of the anxiety-ERN relationship, leaving the evidential value of those meta-analyses in some doubt. It is important to note that one previous meta-analyses did visually inspect funnel plots for evidence of publication bias, and conducted fail-safe N analyses to estimate how many non-significant studies would be required to nullify the anxiety-ERN relationship (Cavanagh and Shackman, 2015). These methods do not quantitatively address publication bias for two reasons. First, neither method provides an estimate of the underlying effect size after adjusting for publication bias. As such, while they might suggest that an effect would remain significant after accounting for null results, they provide no indication of the levels of effect size inflation attributable to publication bias. Second, and perhaps more critically, both methods are widely considered to be insufficient adjustments for publication bias due to their relatively high false-positive rates (Becker, 2005; Carter et al., 2019). In light of these limitations, the primary goal of the current research was to use a range quantitative bias-correction measures with varying levels of conservativeness to estimate the magnitude of the anxiety-ERN link.

Since the completion of our analyses, two further meta-analyses have been published that also explored the anxiety-ERN relationship. Pasion and Barbosa (2019) conducted a meta-analysis on internalizing psychopathology and the ERN, and included 64 studies that spanned Anxiety (e.g., generalized anxiety disorder, PTSD, OCD) and depression. This analysis returned a small overall relationship between internalizing psychopathology and the ERN ($g = 0.18$),[2] with subsequent moderation analyses indicating larger associations for anxiety-related dimensions of internalizing psychopathology (anxiety related: $g = 0.24$; obsessive-compulsive related, $g = 0.34$). In one further meta-analysis, Riesel (2019) included 38 studies of the ERN specifically in OCD samples, finding a significant association of medium effect size in response conflict tasks ($g = 0.55$). As with Cavanagh and Shackman (2015), Riesel (2019) addressed publication-bias using methods that are considered insufficient correctives against publication bias (fail-safe N and visual inspection of funnel plots). However, the authors also found that effect size diminished with increasing sample sizes, indicating that small-study effects are potentially present in the OCD-ERN literature. Pasion and Barbosa (2019) provided Egger's regression test on the relationship between ERN amplitudes and internalizing psychopathology overall (including both anxiety and depression), rather than on the relationship between anxiety and the ERN specifically. As such, despite at least four existing meta-analyses on the anxiety-ERN relationship, the field currently lacks a rigorous assessment of the extent to which publication bias might influence this well-researched effect.

Given the wide interest in the anxiety-ERN relationship (as evidenced by the multiple existing meta-analyses), we think it is important to conduct a comprehensive assessment of potential bias in this literature. A skeptic would have good reason to suspect that publication bias might distort effect sizes in prior meta-analyses of the anxiety-ERN relationship (or any other uncorrected meta-analysis). Publication bias

[1] We use the term *naïve* to refer to a meta-analytical effect size that does not take account of publication bias, rather than as a pejorative term

[2] g = Hedge's g, an effect size used for differences between means. The magnitude of g can be interpreted using similar rule-of-thumb approaches as Cohen's d (0.2 = small; 0.5 = medium).

exists across scientific disciplines, and refers to a tendency of authors, reviewers, and editors to favour the publication of works based on their statistical significance and/or the direction of their effects, resulting in a literature that provides a skewed estimation of effects (Dickersin, 1990; Rothstein et al., 2006). Recent evidence suggests that such bias might is potentially prevalent in human neuroscience. One recent analysis of over 1119 cognitive neuroscience papers reported that studies typically obtained 12% power to detect small effects, and 44% power to detect medium-sized effects (Szucs and Ioannidis, 2017, see also Button et al., 2013). These power analyses suggest that human neuroscience investigations should return null results more often than not. However, most academic journals contain a majority of significant results. Such power considerations might be particularly pertinent in clinical neuroscience research where practical constraints (e.g., access to patient samples) often make it very difficult to achieve the large samples that are necessary to quantify individual differences.

Here, we used a range of quantitative methods to estimate the extent of publication bias in the literature supporting the relationship between trait anxiety and the ERN. In this process, we provided a range of estimates of effect size across bias-correction procedures with varying degrees of conservativeness.

2. Method

Our initial aim was to apply tests of publication bias to two prior meta-analyses of the anxiety-ERN relationship (Cavanagh and Shackman, 2015; Moser et al., 2013). However, we also sought published and unpublished studies produced since the production of these earlier meta-analyses, resulting in a meta-analysis that is substantially larger than the previous reports. Data files and analysis scripts, in addition to supplementary documentation including a log of analysis decisions and details of the literature search are available on our open-science framework (OSF) page (https://osf.io/r7dvc/).

2.1. Study selection criteria

Primary inclusion/exclusion and study selection criteria were largely pre-determined from the two previous meta-analyses (Cavanagh and Shackman, 2015; Moser et al., 2013) and are posted in detail on our OSF page (https://osf.io/5efxa/). The prior meta-analyses overlap considerably in terms of topic and included studies—46% of the studies in Cavanagh & Shackman also appeared in Moser et al. Both previous meta-analyses defined anxiety broadly, including measures of clinical and trait anxiety, as well as related dispositions (e.g., neuroticism; behavioural inhibition).

Studies were included in our meta-analysis if they at least included the measure of the ERN on error trials, and included an operationalisation of anxiety that could be related to the amplitude of the ERN. In order to maintain comparison with existing meta-analyses (Moser et al., 2013; Cavanagh and Shackman, 2015), we exclusively included studies that used the list of anxiety measurements included in these two prior reviews. This inclusion criteria allowed for a wide-range of anxiety measures, including studies where groups were defined based on clinical diagnosis (e.g., generalized anxiety disorder, social anxiety disorder, obsessive compulsive disorder, post-traumatic stress disorder), anxiety scales (e.g., Penn State Worry Questionnaire; Yale-Brown Obsessive Compulsive Scale; Anxiety Sensitivity Index), and a range of closely related individual differences (e.g., Big-5 Neuroticism, Behavioural Inhibition System Scale). Using only the measures used in these prior meta-analyses also meant that we closely match the existing analyses in scope, and, more importantly, we could operationalize key moderators—especially anxious apprehension/worry vs. mixed anxiety—in exactly the same manner as the prior investigations. Samples were not included if they focused on a clinical disorder besides anxiety. This is especially pertinent since co-morbid depression can mask the anxiety-ERN relationship (Weinberg et al., 2012a, 2012b).

For our overall analyses, we also had to reconcile some key differences between the two prior meta-analyses. The largest difference in inclusion criteria between the two previous meta-analyses we built on directly were that Moser et al. included participants with clinical diagnoses, whereas Cavanagh and Shackman focused on un-medicated, sub-clinical samples. In our overall analyses, we included both clinical and non-clinical samples, and included clinical status as a moderator in our analyses (see below). Moser et al. included only studies in which the ERN was assessed in a conflict paradigm (e.g., Stroop, Flanker, Go/No-go) with no motivational manipulation (e.g., monetary incentive), whereas Cavanagh and Shackman included motivational manipulations and other tasks (e.g., probabilistic learning).

As with Moser et al., we determined to focus solely on conflict paradigms with no motivational manipulations for the following reasons. First, it has previously been suggested that anxiety-ERN relationships are primarily present in conflict tasks rather than other tasks, such as probabilistic learning tasks (see Moser et al., 2013). Interestingly, this moderation was verified by a meta-analysis that was published subsequent to the current analyses (Riesel (2019)). Second, we were primarily interested in understanding how individual differences in ERN amplitude are associated with individual differences in anxiety. This is best achieved when the ERN is assessed in a relatively neutral context, without motivational manipulations. Indeed, it has previously been demonstrated that motivational manipulations might mask difference between groups, for example, when punishment elevates ERN amplitudes in control participants to similar levels as OCD groups (Endrass et al., 2010).

Data and syntax were obtained by request from James F. Cavanagh. The analyses from Moser et al. (2013) were replicated independently, and verified by comparison with the forest-plot in the original manuscript. Two effect sizes from the Moser et al. meta-analyses were unavailable in the original manuscripts and were estimated from the forest plot in the prior meta-analyses and denoted with “est.” in Fig. 1.

We aimed to locate and include relevant studies published in the years following the prior meta-analyses. First, we emailed all first and senior authors who had articles included in the prior meta-analyses asking for published and unpublished studies. This call for papers was also published publicly on our OSF page and shared on social media. Second, we identified further studies by searching first and senior author's websites, and reading recent systematic review articles (e.g., Weinberg et al., 2015). Third, we conducted a PubMed search using the anxiety-ERN relevant criteria published with the prior analyses. The PubMed search was conducted for a period starting in January 2012 and ending in June 2018. These searches identified 1171 potential manuscripts, leading to a detailed examination of 71 manuscripts, 24 of which were included in the subsequent meta-analyses (including 7 unpublished effect sizes). 47 of the full-text articles that were closely examined were excluded either because they included no eligible measure of anxiety (11), were re-analyses of old data (10), had no usable ERN-anxiety relationship (7), had insufficient statistics to be included (8), included a motivational manipulation (3), participants were primarily recruited for a clinical disorder besides anxiety (3), they did not use a conflict task (2), they were a prospective study (2), or because their statistics were ambiguous (1). Of the unpublished effects included there was one case of clinical anxiety with the rest being correlational studies with volunteer, non-clinical samples. A flow chart depicting our search process is available online (https://osf.io/wxzjk/).

Our primary results focused on an overall meta-analysis that incorporated the unique studies from Cavanagh and Shackman in addition to more recent studies into the earlier meta-analysis by Moser et al. Analyses focusing on the Cavanagh and Shackman and Moser et al. studies are included in tables identified as “Cavanagh” and “Moser”, respectively. When effect sizes extracted from any given study differed between the previous meta-analyses, differences were resolved by the first author and documented (https://osf.io/fw2pc/). The overall analyses included only effect sizes in which the anxiety-ERN association

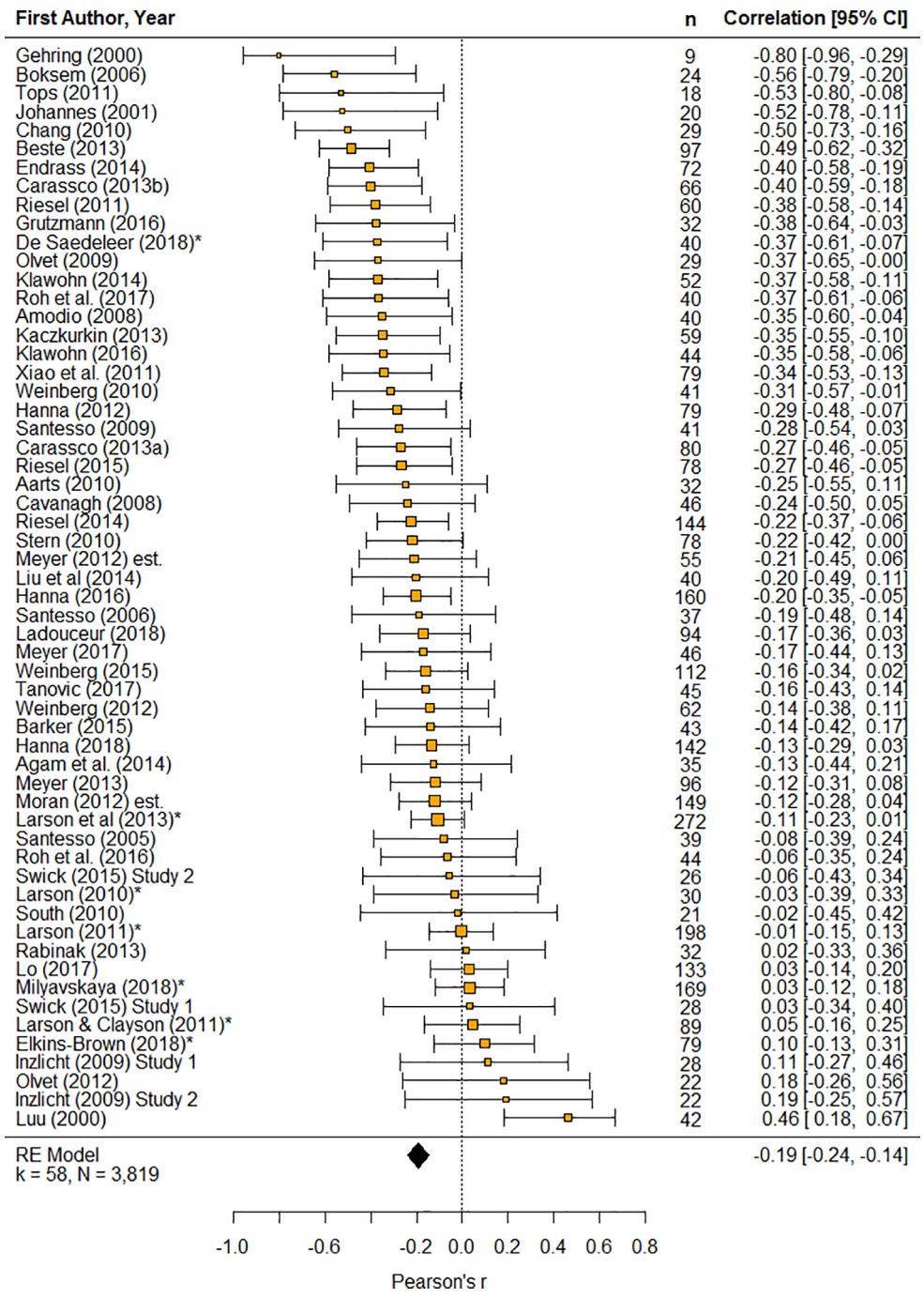

Fig. 1. Forest plot of correlation coefficients anxiety and ERN amplitude. Error-bars depict 95% confidence intervals. * = unpublished.

was measured in a conflict task without a motivational manipulation. Consequently, 15 effect sizes from Cavanagh and Shackman were excluded from the overall meta-analysis. This meant that the overall meta-analysis is most similar to that conducted by Moser et al., yet the bias-adjusted estimates for the Cavanagh and Shackman studies is presented in Tables 2 and 3.

Summary characteristics of the included studies are presented in Table 1, and details of effect size extraction available online (https://osf.io/thxws/). Half of the studies were clinical neuroscience investigations comparing the ERN in a clinically diagnosed anxious sample (OCD; generalized anxiety disorder; social anxiety disorder; post-traumatic stress disorder, PTSD) to healthy controls (k = 29). The rest comprised non-clinical volunteer samples using various trait scales. The effect size relating to a scale measure that was most closely related to worry/anxious apprehension (cf., Moser et al., 2013) was used when multiple effect sizes were available.

Pearson's r was extracted from each study. For group comparisons the effect size was extracted by calculating Cohen's d from the descriptive statistics, or from inferential statistics when descriptive statistics were unavailable, and converted into Pearson's r.

2.2. Meta-analyses

Analyses were conducted using the *metafor* package (Viechtbauer, 2010) in R. Random-effects meta-analyses were chosen because they are more appropriate for studies sampled from heterogeneous populations. Pearson's r is often not normally distributed, therefore, effect sizes were transformed onto the Fisher's Z-scale using *esclac*. The random-

Table 1
Features of studies included in the overall meta-analysis, adapted and updated from Moser et al. (2013) and Cavanagh and Shackman (2015). All samples are adult (i.e., > 18 y.o.) unless otherwise stated.

First author (year)	Population	Task	Measure	Type	ERN
Aarts (2010)[a,b]	Volunteer	Go/NoGo	STAI-T	M	Mean
Agam (2014)	Clinical	Antisaccade	SCID	AA	Mean
Amodio (2008)[a,b]	Volunteer	Go/NoGo	BIS	AA	Peak
Barker (2015)	Volunteer	Arrow flanker	LSAS-SR	AA	Mean
Beste (2013)[a]	Volunteer	Flanker/GoNogo	ASI	M	Mean
Boksem (2006)[a,b]	Volunteer	Letter flanker	BIS	AA	Mean
Carrasco (2013a)[a]	Clinical - pediatric OCD	Arrow flanker	K-SADS-PL	AA	Mean
Carrasco (2013b)[a]	Clinical - pediatric OCD	Arrow flanker	K-SADS-PL	AA	Mean
Cavanagh (2008)[a,b]	Volunteer	Letter flanker	BIS	AA	Peak
Chang (2010)[a,b]	Volunteer	Letter flanker	ASR	M	Peak
De Saedeleer (2018)*	Volunteer	Go/NoGo	PSWQ	AA	Mean
Elkins-Brown (2018)*	Volunteer	Go/NoGo	BIS	AA	Mean
Endrass (2014)	Clinical-OCD/SAD	Arrow flanker	SCID	AA	Peak
Gehring (2000)[a]	Clinical OCD	Colour Stroop	SCID	AA	Mean
Grutzmann (2016)	Clinical - OCD	Arrow flanker	SCID	AA	Peak
Hanna (2012)[a]	Clinical -pediatric OCD	Arrow flanker	K-SADS-PL	AA	Mean
Hanna (2016)	Clinical - pediatric OCD	Arrow flanker	Existing diagnosis	AA	Mean
Hanna (2018)	Clinical – adolescent OCD	Arrow flanker	Arrow flanker	AA	Mean
Inzlicht (2009), study 1[a]	Volunteer	Colour Stroop	BIS	AA	Mean
Inzlicht (2009), study 2[a]	Volunteer	Colour Stroop	BFI-N	M	Mean
Johannes (2001)[a]	Clinical OCD	Go/NoGo	SCID	AA	Mean
Kaczkurkin (2013)[a]	Volunteer	Letter flanker	OCI-R	AA	Peak
Klawohn (2014)	Clinical - OCD	Arrow flanker	SCID	AA	Peak
Klawohn (2016)	Clinical - OCD	Arrow flanker	SCID	AA	Peak
Ladouceur (2018)	Clinical – pediatric anxiety	Arrow flanker	K-SADS-PL	AA	Mean
Larson & Clayson (2011)*[a]	Volunteer	Arrow flanker	STAI-T	M	Mean
Larson (2010)*[a]	Volunteer	Colour Stroop	STAI-T	M	Mean
Larson (2011)*[a]	Volunteer	Arrow flanker	STAI-T	M	Mean
Larson (2013)*	Clinical – GAD	Letter flanker	SCID	AA	Mean
Liu (2014)	Clinical – pediatric OCD	Arrow flanker	K-SADS-PL	AA	Mean
Lo (2017)	Volunteer - child	Go/NoGo	RCADS-P	M	Mean
Luu (2000)[a,b]	Volunteer	Letter flanker	PANAS	M	Mean
Meyer (2012) est.[a,b]	Volunteer - child	Arrow flanker	Parent – SCARED	M	Mean
Meyer (2013)	Clinical - pediatric Anxiety	Go/NoGo	PAPA	AA	Mean
Meyer (2017)	Volunteer	Arrow flanker	STAI-T	M	Mean
Milyavskaya (2018)*	Volunteer	Arrow flanker	BIS	AA	Mean
Moran (2012) est.[a]	Volunteer	Letter flanker	PSWQ	AA	Mean
Olvet (2009)[a,b]	Volunteer	Letter flanker	DASS	M	Mean
Olvet (2012)[a]	Volunteer	Arrow flanker	BFI-N	M	Peak
Rabinak (2013)[a]	Clinical - PTSD	Arrow flanker	SCID	M	Mean
Riesel (2011)[a]	Clinical - OCD	Arrow flanker	SCID	AA	Peak
Riesel (2014)	Clinical - OCD	Arrow flanker	SCID	AA	Peak
Riesel (2015)	Clinical - OCD	Arrow flanker	SCID	AA	Peak
Roh (2016)	Clinical - OCD	Face flanker	SCID	AA	Mean
Roh (2017)	Clinical - OCD	Arrow flanker	SCID	AA	Mean
Santesso (2005)[b]	Volunteer – child	Letter flanker	JEPQR-S	M	Peak
Santesso (2006)[a]	Volunteer - child	Letter flanker	CBCL	AA	Peak
Santesso (2009)[b]	Volunteer - child	Letter flanker	Composite	M	Peak
South (2010)[b]	Volunteer – child	Arrow flanker	BIS	AA	Mean
Stern (2010)[a]	Clinical - OCD	Letter flanker	SCID	AA	Mean
Swick (2015) study 1	Clinical - PTSD	Arrow flanker	PCL-M	M	Mean
Swick (2015) study 2	Clinical - PTSD	Arrow flanker	PCL-M	M	Mean
Tanovic (2017)	Volunteer	Arrow flanker	PSWQ	AA	Mean
Tops (2011)[a]	Volunteer	Letter flanker	BIS	AA	Mean
Weinberg (2010)[a]	Clinical - GAD	Arrow flanker	SCID	AA	Mean
Weinberg (2012)[a]	Clinical - GAD	Arrow flanker	SCID	AA	Mean
Weinberg (2015)	Clinical	Arrow flanker	SCID	AA	Mean
Xiao et al. (2011)[a]	Clinical – GAD/OCD	Letter flanker	Chinese MINI	AA	Mean

Note: [a] = in Moser et al. (2013); [b] = in Cavanagh and Shackman (2015); * = unpublished effect size. GAD, Generalized Anxiety Disorder; OCD, Obsessive Compulsive Disorder; PTSD, Post-traumatic Stress Disorder. ASR, Achenbach Self-Report; BFI-N, Big Five Inventory-Neuroticism; BIS, Behavioral Inhibition System Scale; CBCL, Child Behavior Checklist; DASS, Depression Anxiety Stress Scale; JEPQR-S, Junior Eysenck Personality Questionnaire – Revised, Self-report; K-SADS-PL, Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version; LSAS-SR, Liebowitz Social Anxiety Scale – Self-Report; MINI, Mini-International Neuropsychiatric Interview; PANAS, Positive and Negative Affect Schedule; PCL-M, PTSD Checklist, Military version; PSWQ, Penn State Worry Questionnaire; RCADS-P, Revised Child Anxiety and Depression Scale—Parent Version; SCARED, Screen for Child Anxiety Related Disorders; SCID, Structured Clinical Interview for DSM Disorders; STAI-T, State and Trait Anxiety Inventory-Trait Version. AA, Anxious Apprehension; M, Mixed Anxiety. Mean, mean amplitude measure, Peak, peak detection method.

effects meta-analysis was conducted using a restricted maximum likelihood estimation. Z-scores were converted into Pearson's *r* for summary statistics/figures. Finally, we only included cases where effect sizes could be verified as arising from independent samples. As such, any cases where two or more effect sizes come from one publication (e.g., Swick et al., 2015), we were careful to ensure that the effect sizes

Table 2
Naïve estimates of effect-size from random-effects meta-analysis. Effect sizes for Moser and Cavanagh were replicated in our meta-analysis using the specific of studies/effect sizes that were included in these prior reviews.

Meta-analysis	k	N	r	Q	I^2
Overall	58	3819	**−0.19 [−0.24, −0.14]**	**121.63**	52.5% [36.1,74.5]
Worry	39	2792	**−0.22 [−0.27, −0.17]**	**60.45**	34.2% [5.5,68.4]
Mixed	19	1027	−0.10 [−0.21,0.02]	**51.74**	65.6% [36.1,84.6]
Clinical	29	2117	**−0.23 [−0.27, −0.18]**	32.52	8.6% [0.0,63.4]
Volunteer	29	1702	**−0.15 [−0.23, −0.06]**	**78.07**	65.8% [44.0,83.4]
Published	51	2942	**−0.22 [−0.27, −0.17]**	**90.31**	41.1% [23.1,71.8]
Unpublished	7	877	−0.03 [−0.11,0.05]	8.97	18.8% [0.00,92.3]
Moser	32	1706	**−0.24 [−0.32, −0.16]**	**84.64**	63.1% [44.7,83.8]
Cavanagh	28	942	**−0.29 [−0.38, −0.19]**	**59.48**	55.9% [30.8,80.3]

Note: k = number of studies; N = summed N across studies; r = Pearson's r; Q = test for heterogeneity, I^2 = variation across studies due to heterogeneity. Square brackets contain 95% CIs. Bold = $p < .05$.

came from different populations. In cases where two eligible clinical groups (e.g., obsessive compulsive disorder and social anxiety disorder; Endrass et al., 2014) shared a common control group, we pooled mean and standard-deviations to form one clinical group to compare with the control.

Meta-regression was used to assess moderation by the following study characteristics:

1. Anxiety-type. Moser et al. reported that the anxiety-ERN relationship is larger in anxious samples characterised by worry and verbal rumination about threat (i.e., anxious apprehension) compared to samples characterised by a mix of worry, physiological symptoms, and depressive symptoms (i.e., mixed anxiety). Consequently, we included anxiety-type (worry vs. mixed) as a moderator to replicate this prior investigation.
2. Clinical status. As half of the studies in our analyses included samples with a diagnosis of clinical anxiety, we included clinical status (clinical vs. volunteer) as a moderator. This moderator analysis also helps to account for a key difference in inclusion criteria between Moser et al. and Cavanagh and Shackman's prior meta-analyses.
3. Publication status. We tested for a difference in effect size between published and unpublished studies to assess publication bias.
4. Developmental stage. Since the ERN, and the frontal cortex more broadly, develops throughout the lifespan (cf., Tamnes et al., 2013), we included developmental stage (child vs. adult sample) as a categorical moderator of the anxiety-ERN relationship.
5. ERN quantification. We also included the method used to operationalize the ERN (peak amplitude vs. mean amplitude) as a categorical predictor of the anxiety-ERN relationship. This moderator analysis was conducted since differences in reliability might exist between studies using mean and peak measurements (Fischer et al., 2017).

2.3. Tests of publication bias

Small study effects occur when smaller studies over-estimate the size of an association or treatment effect. These effects often indicate publication bias, where only studies that overestimate the true underlying effect size generate "publishable" p-values. Multiple methods have been used to assess and correct for publication bias in existing literatures. While a great deal of debate exists concerning the most appropriate bias correction techniques—often concerning the amount or severity of adjustment—we opted to report multiple correction methods. This strategy was chosen to present a range of possible corrected effect sizes with varying levels of conservativeness, and also so that these adjusted effect sizes are available for future researchers interested in specific bias adjustments.

We first tested for small study effects using funnel plots (standard error against effect size). These plots should form a symmetrical funnel/triangle shape under conditions of low publication bias, with more variability in effect sizes for studies with high standard errors compared to more precise studies with low standard errors. Funnel-plot asymmetry is indicative of publication bias. We quantified asymmetry using Egger's tests (*regtest* in *metafor*). Trim-and-fill analyses were used to correct the meta-analytical estimate by imputing missing studies needed to make the funnel plot symmetrical (*trimfill* in *metafor*).

Trim-and-fill often under-corrects for small study effects (see Carter et al., 2019). Therefore, we additionally used a range of regression-based methods that provide more conservative bias corrections. These methods follow a simple logic in which a weighted-least squares regression is used to predict effect size from a measure of study precision. The intercept of this model is taken to be the most precise possible study, and, therefore, an unbiased assessment of the underlying effect. A significant slope indicates publication bias if increasing precision predicts decreasing effect size. In PET effect sizes are predicted from their standard errors, weighted by the inverse of the variances. PEESE is similar to PET, but instead uses a quadratic equation that allows for the assumption that studies with larger standard errors will systematically overestimate the underlying effect size more so than more precise studies. PET has a general tendency to over-correct small study effects (Inzlicht et al., 2015). Thus, a conditional logic has been suggested where PET is first used to establish if the corrected effect size is non-zero, followed by PEESE to provide an estimate of the underlying effect (Stanley and Doucouliagos, 2014). This conditional logic remains controversial, and here we report both measures, while acknowledging that PET is likely the more conservative estimate. The final—and perhaps most literal—test of small study effects used in this analysis was the Peter's test in which study effect sizes are predicted by the inverse of sample sizes, weighted by sample size (Peters et al., 2006).

Lastly, we used the three-parameter selection model. Selection models are weight functions that aim to explicitly model the process through which results are either published or suppressed due to publication bias. Recently endorsed by two simulation papers as a powerful method to correct for publication bias (McShane et al., 2016; Carter et al., 2019), the selection method we used has three parameters: the population effect size μ; heterogeneity of effect size τ^2; and the probability of a null result being published. We estimated this model using the *weightr* function and report the adjusted effect size and the likelihood ratio test which provides a χ^2 statistic comparing the unadjusted and adjusted effect-size estimates.

3. Results

3.1. Random effects meta-analysis

The overall meta-analysis returned a significant negative correlation

between anxiety and the ERN ($r = -0.19$, $p < .001$, see Fig. 1 and Table 2).[3] The *influence* function in *metafor* did not indicate that any individual study over-influenced the results. This analysis also showed significant heterogeneity, $Q(57) = 121.63$, $p < .001$, with $I^2 = 52.5\%$ suggesting a medium amount of variation between studies due to heterogeneity rather than chance (Pigott, 2012).

Anxiety-type (worry vs. mixed anxiety) significantly moderated the meta-analytical effect size, $Q(1) = 6.12$, $p = .013$, due to a significant relationship between worry and the ERN, $r = -0.22$, but not between mixed anxiety and the ERN, $r = -0.10$. Clinical status was also a significant moderator, $Q(1) = 4.50$, $p = .03$, due to larger effects in studies with clinical populations, $r = -0.23$, than in volunteers, $r = -0.15$. See Table 2.

As 68% of the effect sizes classified as worry were clinically anxious, we tested if the anxiety-type moderation holds when focusing separately on only clinical or volunteer samples. We found no significant moderation by anxiety-type when comparing worry ($k = 13$) with mixed anxiety ($k = 16$) in volunteers, $Q(1) = 0.65$, $p = .42$. However, worry did moderate the anxiety-ERN relationship when analyses were restricted to clinical samples, $Q(1) = 4.09$, $p = .04$, due to a larger effect size for clinical groups with worry, $r = -0.24$, $k = 26$, $p < .001$, than clinical groups with mixed anxiety, $r = 0.0005$, $k = 3$, $p = .997$. These subsample analyses suggest that the most robust Anxiety-ERN relationships are observed in clinical samples with worry (OCD, social anxiety disorder, generalized anxiety disorder). However, both the moderation by anxiety-type and clinical status should be interpreted with caution for a number of reasons, including their relatively high *p*-values, and the small number of clinical samples with mixed anxiety (all three were PTSD samples). In light of these moderation analyses, subsequent bias-corrections are provided for the overall meta-analysis, and for separate subsets split by anxiety-type and clinical status.

Finally, we found significant moderation by publication status, $Q(1) = 9.79$, $p = .0018$, due to larger effect sizes occurring in published, $r = -0.22$, relative to unpublished work that did not differ from zero, $r = -0.03$. It is important to note that the unpublished data included a higher proportion of volunteer samples and mixed anxiety than the overall analysis. As volunteer samples and mixed anxiety were both associated with smaller effect sizes, it is possible that the moderation by publication status is not fully explained by publication status. However, it is noteworthy that the one unpublished clinical study had a large sample ($n = 273$; Larson, 2013) yet did not show a significant relationship between anxiety and the ERN.

3.2. Quantitative corrections for publication bias

Egger's regression test did not support significant funnel-plot asymmetry, $Z = -1.72$, $p = .085$. Subsequent trim-and-fill analyses suggested four additional studies were required to make the funnel-plot symmetrical (see Fig. 2). These imputed values reduced the meta-analytical effect size, $r = -0.18$, $p < .001$, 95% CIs [−0.23, −0.13], however, it should be noted that this is only a small reduction with 95% CIs that include the uncorrected estimate (i.e., $r = -19$). Trim-and-fill returned a significant meta-analytical effect size for all sub-analyses except mixed anxiety (see Table 3 and Fig. 2). It might be suggested that conducting trim-and-fill without a significant Egger's test would be inappropriate in a case where significant funnel plot asymmetry was not evident. However, perhaps more importantly, if Egger's test is not significant, it is unlikely that trim-and-fill analyses will result in a substantial correction for publication bias (as was the case for our analyses). Thus, conducting trim-and-fill without a significant Egger's test would unlikely lead to inappropriate conclusions because the trim-and-fill analysis will provide little correction in cases where Egger's test does not detect bias.

A contour-enhanced funnel plot was used to explore the distribution of effects across different significance thresholds. Here, publication bias would occur if studies cluster around conventional significance thresholds (i.e., $p < .05 > .01$) with few null results. As Fig. 2 illustrates, 57% of the results included in overall meta-analysis were null results; this percentage decreased to 53% if unpublished effect sizes are removed. These numbers suggest that a decent proportion of null results on the relationship between anxiety and the ERN do enter the published literature, suggesting minimal effects for publication bias in this literature. Here, it should be noted that this analysis also includes the 7 unpublished studies that were solicited for this meta-analysis.

Given widespread concern that both funnel-plot asymmetry and trim-and-fill under-correct for publication bias (Carter et al., 2019; Idris, 2012; Simonsohn et al., 2014). We next conducted PET, Peters', and PEESE analyses that have lower false-positive rates than trim-and-fill (Carter et al., 2019). These analyses are summarised comprehensively in Table 3. First, it should be noted that the slopes for PET and Peters' tests were both significant for the overall analysis (PET, $b = -1.11$, *S.E.* = 0.53, $p = .04$; Peters', $b = -4.30$, *S.E.* = 2.09, $p = .04$), suggesting that effect sizes decreased as study precision increased, an effect that is consistent with publication bias. Largely similar results were found for PEESE, $b = -3.53$, *S.E.* = 1.78, $p = .05$. Together, these results are consistent with the suggestion that modest small study effects are present in the anxiety-ERN relationship, contrary to what is suggested in the funnel plots and the Egger's test, $Z = -1.72$ $p = .085$. While it is unfortunate that these metrics do not converge, it is consistent with global evaluations of these tools (Carter et al., 2019) and recommendations to use and consider a broad number of bias-corrective measures (Inzlicht et al., 2015).

We next turned our attention to the bias-corrected effect sizes obtained from the intercept value of these tests as an assessment of the attenuation of the Anxiety-ERN relationship. In the overall analysis, the intercept was significant for both the Peters', $b_0 = -0.11$, *S.E.* = 0.04, $p = .007$, 95% CIs [−0.19, −0.03], and PEESE, $b_0 = -0.12$, *S.E.* = 0.04, $p = .002$, 95% CIs [−0.19, −0.05] tests, suggesting a small, non-zero relationship between trait-anxiety and the ERN (see Fig. 2 and Table 3). The three parameter selection model also suggested that there was a small difference between bias-adjusted effect size, $r = -0.14$, $p < .001$, 95% CIs [−0.20, −0.06], and the unadjusted effect size ($r = -0.19$), $\chi^2(1) = 3.81$, $p = .05$. In contrast, the intercept value for PET was not statistically significant, $b_0 = -0.05$, *S.E.* = 0.07. $p = .51$, 95% CIs [−0.18, 0.09], meaning that the PET analysis was the only bias-correction test to suggest that the magnitude of the anxiety-ERN relationship was not distinguishable from zero.

In summary, with the exception of PET, each of these bias-correction effects returned a small but non-zero effect size for the anxiety-ERN relationship that ranged from $r = -0.11$ to $r = -0.14$. These point estimates are lower than previous uncorrected meta-analyses presented by Moser et al. (2013, $r = -0.25$, 95% CIs [−0.30, −0.20]) and Cavanagh and Shackman (2015; $r = -0.28$, 95%CIs [−0.19, −0.40]).

[3] A negative deflection is also observed on correct trials, the Correct Related Negativity (CRN). We found a small negative relationship between anxiety and the CRN, $r = -0.10$, 95% CIs [−0.14, −0.06], k = 40, $p < .001$. Moderation analyses also indicated that the CRN-Anxiety relationship was larger for peak-detection methods than mean amplitude, and that the CRN-Anxiety relationship was significant under trim-and-fill, PEESE, and the selection model (see https://osf.io/x8d3c/). These findings are consistent with the idea that anxiety might increase monitoring on both correct and incorrect trials, however, it is important to note that the CRN-Anxiety relationship appears to be smaller in magnitude than the anxiety-ERN relationship. In addition to the CRN, we also investigated the ΔERN (i.e., ERN *minus* CRN) that shows the extent to which performance monitoring differentiates between error and correct performance. In an overall analysis, we observed a small negative association between ΔERN and anxiety, $r = -0.109$, $p = .0029$, k = 20, $N = 1381$, however, this effect only remained significant after bias-correction using trim-and-fill, but not with any other method. Finally, as with the ERN, we also observed moderation of the anxiety-ΔERN relationship by anxiety-type, clinical status, and publication status (see https://osf.io/jahup/ for details).

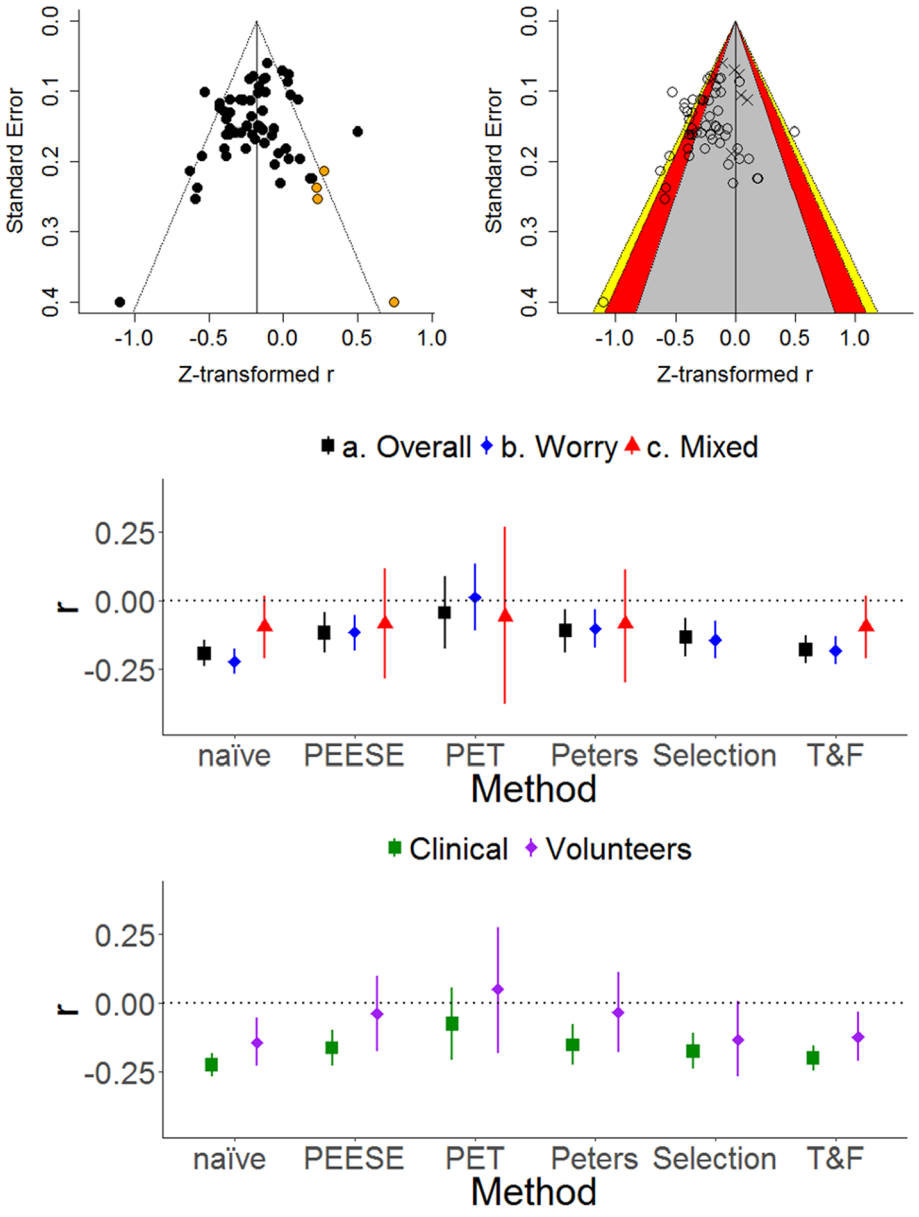

Fig. 2. Upper-left: funnel plot showing the original studies (black) and studies imputed by trim-and-fill analyses (orange). Upper-right: contour-enhanced funnel plot: studies in the gray area $p > .05$; studies in the red area have $p < .05$ and $> .01$; studies in the yellow area $p < .01$ and $> .005$; and studies in the white/background $p < .005$. Middle: meta-analytical effect sizes across the full range of estimation methods for the overall analyses, worry and mixed-anxiety subsamples. Lower: effect sizes across estimation methods for the clinical and volunteer subsamples. Error-bars depict 95% CIs. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3
Bias-correction statistics across methods and subsets. Adjusted *r* shows the bias-corrected meta-analytical effect size for each correction method.

	Overall	Worry	Mixed	Clinical	Volunteer	Moser	Cavanagh
Adjusted *r*							
Trim&Fill	**−0.18**	**−0.18**	−0.10	**−0.20**	**−0.12**	**−0.22**	**−0.29**
PET b_0	−0.05	0.01	−0.06	−0.08	0.05	−0.04	−0.05
PEESE b_0	**−0.12**	**−0.12**	−0.09	**−0.16**	−0.04	**−0.14**	−0.16
Peters' b_0	**−0.11**	**−0.10**	−0.09	**−0.15**	−0.03	−0.13	−0.14
Selection	**−0.14**	**−0.14**	n.a.	**−0.18**	−0.13	−0.08	**−0.15**
Mean adjusted *r*	−0.12	−0.11	−0.09	**−0.15**	−0.05	−0.12	−0.16
Bias-estimate							
Egger's test	−1.72	**−3.81**	0.24	**−2.42**	−1.10	−1.36	−1.48
PET b_1	**−1.11**	**−1.93**	−0.23	**−1.32**	−1.31	−1.32	−1.31
PEESE b_1	−3.53	**−6.03**	−0.19	**−4.21**	**−4.15**	−3.85	−3.47
Peters' b_1	**−4.30**	**−7.57**	−0.22	**−5.39**	−4.81	−4.92	−4.44
Selection χ^2	3.81	**7.05**		**4.23**	0.04	**10.84**	**5.24**

Note: bold = $p < .05$. Egger's test = *Z*-statistic. Selection = three-parameter selection model. Selection χ^2 = likelihood ratio test comparing unadjusted and adjusted effect size estimates.

However, the smallest estimate of effect size provided by the 95% CIs of each previous meta-analysis (~$r = -0.20$) overlap, if only slightly, with the largest effect size from each of our bias correction methods (range $r = -0.23$ to -0.18). Thus, our effect sizes are consistent with the range presented in previous analyses, albeit with most of the plausible effect sizes estimated from our analyses being lower in magnitude.

Finally, in relation to the interpretation of our overall results, it is important to note that a conditional logic has been applied to PET-PEESE analyses, where if the effect size (i.e., the intercept value) returned by PET is significant, it is suggested to use PEESE to estimate the effect size. Otherwise, it is recommended that the effect size is estimated using the effect size estimated by PET. Thus, the logic of PET-PEESE would suggest that our overall effect size and all sub-set analyses are not different from zero. However, as we have already noted, the use of PET-PEESE remains controversial.

3.3. Bias-correction in sub-groups revealed by moderation analyses

We next turned our attention to the significant moderation observed both by anxiety sub-type (i.e., worry vs. mixed anxiety) and clinical-status (i.e., clinical vs. non-clinical studies). Estimates from the selection method were not provided for the mixed analyses because the *p*-value interval contained too few values to create reliable estimates.

It is important to note that more significant indicators of publication bias/small study effects were observed in sub-groups associated with larger effect sizes (e.g., clinical groups, worry) than sub-groups of studies associated with smaller effect sizes (e.g., non-clinical groups, mixed anxiety). Indeed, every test of potential publication bias was statistically significant for both the worry and clinical sub-analyses. For the Worry sub-analyses, the Egger's test, and the slopes from PET, PEESE, and Peter's test all returned p-values < .001, and the likelihood ratio test from the three-parameter selection model was significant at the 1% level, $X^2(1) = 7.05$, $p = .008$. These findings comprise robust evidence for small study bias within the Worry sub-analysis. The same tests for the clinical sub-analysis were all significant at the 5% level (Egger's test, $p = .02$; PET, $p = .02$; PEESE, $p = .02$; Peters', $p = .02$; and three-parameter selection model $X^2(1)$, $p = .04$), again consistent with robust evidence of small-study effects.

It is interesting to consider the specificity of the relationships between worry and the ERN in light of these apparent small-study effects. In a previous meta-analysis (Moser et al., 2013), worry (i.e., anxious apprehension) showed a particularly strong association with ERN amplitude, r = −0.35, $k = 20$, 95% CIs [−0.40, −0.29]. The uncorrected effect size from our random effects meta-analysis for the worry studies was already lower than the previous estimate, r −0.22, $k = 39$, 95% CIs [−0.27, −0.17], providing initial evidence that the specific worry-ERN relationship might be smaller than previously thought. However, this smaller effect size was further reduced by each of our bias-correction methods (see Table 3). For example, the three-parameter selection method returned an effect size for the worry-ERN less than half the size of the earlier meta-analysis, $r = 0.14$, 95% CIs [−0.21, −0.08]. Furthermore, the size of the average worry-ERN relationship observed across all bias correction methods, $r = -0.11$, was largely similar to the same metric observed for the mixed anxiety-ERN relationship, $r = -0.09$. Thus, while our results replicated the moderation of the anxiety-ERN relationship by subtype of anxiety, our bias-corrected effect sizes suggested that this moderation is largely attenuated after correction for small study effects.

The anxiety-ERN relationship in clinical studies was reduced after correction for publication bias. However, the magnitude of effect size reduction during bias correction for clinical ($r = -0.23$ to $r = -0.15$, reduction of 0.08) and non-clinical ($r = -0.15$ to $r = -0.06$, reduction of 0.09) studies were similar, indicating that the gap in effect size between clinical and volunteer studies did not narrow after bias-correction in the same manner as the worry-mixed relationship.

4. Discussion

Our uncorrected meta-analysis indicated a small negative correlation between trait anxiety and the ERN ($r = -0.19$), with moderator analyses further indicating that the anxiety-ERN relationship was present in worry but not mixed anxiety (see also Moser et al., 2013), and was larger in samples identified by clinical diagnosis rather than in non-clinical volunteer groups. We also found evidence for publication bias. First, while published studies were associated with a statistically significant effect size ($r = -0.22$), the meta-analytical effect size for unpublished studies was not distinguishable from zero ($r = -0.03$). Second, 40% of the statistical tests of publication bias were significant in the overall meta-analysis (PET and Peters' test), and *every* bias test (PET, PEESE, Peters', Egger's test, selection model X^2) was significant for both the worry and clinical subsamples of studies. These findings are consistent with the presence of publication bias in the literature supporting a link between anxiety and the ERN.

We also provided a range of estimates of effect size after correction for potential bias. Effect sizes were generally smaller but non-zero after bias correction. However, the extent of effect size attenuation in the overall analysis ranged from relatively trivial (trim-and-fill, $r = -0.18$), to more modest reductions (three-parameter selection method, $r = -0.14$), to the complete abolition of the relationship between anxiety and the ERN (PET, $r = 0.05$). Here, it is noteworthy that the confidence intervals from the uncorrected meta-analytic effect overlapped with each corrected effect size, suggesting that, while effect sizes were attenuated by correction for publication bias, the range of plausible 'true' effect sizes were overlapping across analyses.

The same analyses on specific sub-groups of studies revealed that indicators of publication bias and effect size attenuation was particularly pronounced for the sub-selection of studies identified as anxious apprehension (i.e., worry), where both the uncorrected random effects meta-analysis ($r = -0.22$) and the average of the bias-corrected estimates ($r = -0.11$) were significantly smaller (i.e., 95% CIs did not overlap) than the corresponding effect size from one previous meta-analysis ($r = -0.35$; see Moser et al., 2013). Expressed in terms of variance explained, while previous estimates indicated that worry and the ERN share 12.3% variance, our uncorrected effect size reduces the same statistic to 4.8%, with the average of our bias-corrected effect sizes reducing shared variance to only 1.2%. These effects point to less close coupling between anxious apprehension and the ERN than previously thought, with implications for theory that we will return to in a subsequent section.

When regarding our results, a reader might worry which of our range of effect sizes they should take as most reflective of the anxiety-ERN relationship. If only one of these outcomes reflects the *true* underlying effect, which estimate should we trust? This is indeed a significant challenge that is yet to be fully addressed methodologically (see Carter et al., 2019 for discussion). Nevertheless, it is important to note that there is a considerable degree of overlap among our bias correction methods, with the exception of the PET analyses that supported a null association between anxiety and the ERN on every occasion. Instead of considering these tests as definitively isolating the true effect, these bias-correction tools are perhaps better viewed as providing a range of plausible values under different levels of conservativeness. Consequently, the trim-and-fill results might be considered the least conservative bias-corrected estimate, while PET is most conservative—perhaps to a fault (Inzlicht et al., 2015). Partly contradicting the conservativeness of PET, recent simulation studies have suggested that both PET-PEESE and the three-parameter selection model perform adequately well under conditions with moderate levels of publication bias and heterogeneity, particularly when, as in our case, k approaches 60 (Carter et al., 2019). However, in these simulations the three-parameter model tended to outperform PET-PEESE in the majority of cases, and, as such, the effect size estimates from the selection ($r = -0.14$) model might appropriately guide future theorising and empirical

investigation on the anxiety-ERN relationship.

One positive message from our analysis is that the majority of our results point to a significant, albeit small, relationship between anxiety and the ERN. These bias correction results stand in contrast to other bias-corrected meta-analyses in which effects became indistinguishable from zero after correction for publication bias (e.g., ego-depletion: Carter et al., 2015). This positive news for the field coheres with recent meta-scientific investigations which indicated that studies published in three major psychophysiology journals (Psychophysiology; International Journal of Psychophysiology; Journal of Psychophysiology) showed good evidential value and relatively low evidence for selective reporting (Carbine et al., 2019).

4.1. Theoretical implications

While our results suggest that the anxiety-ERN relationship is likely non-zero, the attenuation of effect sizes observed in our study relative to other investigations prompts deeper consideration of effect size in relation to theory. The ERN has been implicated in the pathogenesis of anxiety (Meyer, 2017), and has been identified as a candidate trait vulnerability marker for clinical anxiety (Olvet and Hajcak, 2008; Weinberg et al., 2015). According to these hypotheses, the ERN need not be continuously related to anxious symptomatology. These theories would predict that the ERN should be enhanced in clinical samples, and that this enhancement should make the individual more susceptible to future anxious psychopathology. In our view, the meaningfulness of the ERN as a trait-like vulnerability marker for clinically significant anxiety must be bounded by the strength of the relationship between the ERN and anxiety. The average of our corrected effect sizes, however, suggests that clinical anxiety explains 2.3% variance in ERN amplitude (i.e., $r = -0.15$). This small effect size limits the extent to which the ERN could play a role in the aetiology of anxious psychopathology, for example, by predicting the development of anxious symptoms over time. Here, it is important to note that vulnerability to heterogeneous disorders is likely multifaceted, and it could also be argued that explaining a small portion of a complex problem has utility. However, ongoing theoretical framing should closely consider effect size when linking the ERN to the development of clinical anxiety.

The significant moderations observed in our analyses are also of theoretic importance. First, we found no significant relationship between the amplitude of the ERN and studies classified as mixed anxiety. While previous meta-analyses found a small yet significant relationship between mixed anxiety and the ERN (Moser et al., 2013), this relationship was not even significant in our uncorrected meta-analysis ($r = -0.10$, $n = 1027$). One particularly salient explanation for the low association observed might come from the broader symptom and trait profile associated with measures that are used in studies classified as mixed anxiety. For example, the STAI-T includes items tapping anxiety "I feel nervous and restless" and "I am calm, cool, and collected" (reverse-coded), as well as depression "I feel like a failure", "I wish I could be as happy as others seem to be" (Spielberger, 1989). Similarly, Big-5 neuroticism includes characteristics such as "worries a lot", "can be tense", as well as "is depressed, blue", "can be moody" (cf., John & Srivastava, 1999). As previous studies have indicated that the presence of depressive symptoms can blunt the anxiety-ERN relationship (Weinberg et al., 2012a), it is possible that the relatively broad profile of mixed anxiety—blending worry and anxious arousal with depressive symptoms and general discontent—might provide a more confounded test of the anxiety-ERN relationship.

We also found that studies in which groups were identified by clinical diagnosis showed significantly larger effect sizes than groups with non-clinical volunteer samples. This finding could suggest a role for clinical status/symptom severity in the anxiety-ERN relationship. However, while clinical status and worry are *theoretically* dissociable, in our sample 66.7% of the effect sizes classified as worry were also clinical samples, while only 15.8% of the mixed anxiety effect sizes were from clinical samples (all PTSD). Consequently, the clinical-volunteer and worry-mixed contrasts are highly confounded. While we cannot rule out severity as a potential moderator of the anxiety-ERN relationship, it is also possible that the clinical diagnosis that are more prevalent in our analyses (e.g., OCD, generalized anxiety disorder) are more discrete examples of anxiety than the broader profile of negative affectivity (e.g., covering anxiety and depression) that is represented in the mixed anxiety sub-sample that made up the majority of the volunteer samples. Further speaking against a role of symptom severity per se in explaining the anxiety-ERN relationship, enhanced ERNs remain evident in clinically anxious samples after symptom remission (Stern et al., 2010) and are visible in asymptomatic relatives (Riesel et al., 2011). Thus, it is tentatively more likely that the anxiety-ERN relationship differs between qualitatively different dimensions that underlie anxiety, rather than as a function of symptom severity.

The moderation observed in our analyses must be interpreted with some caution given the significant indicators of publication bias observed for both the worry and clinical sub-analyses. For example, while the uncorrected effect sizes for the worry-mixed contrast differed meaningfully ($r = -0.22$ vs. $r = -0.10$, for worry and mixed anxiety, respectively), the average of the effect sizes after bias-correction was similar for worry ($r = -0.11$) and mixed anxiety ($r = -0.09$). The gap between average corrected effect sizes was slightly larger between clinical ($r = -0.15$) and volunteer ($r = -0.05$) samples. Thus, while the worry-ERN relationship was broadly non-zero after correction for publication bias—not the case for mixed anxiety—the magnitude of corrected effect sizes between subgroups was broadly similar. It is noteworthy that larger samples sizes contributed to the worry subset ($n = 2792$, k = 39) than mixed anxiety ($n = 1027$, k = 19), meaning that differences in statistical power could explain why small, yet significant effects were observed for one sample (i.e., worry) but not the other (i.e., mixed). In contrast to a view that the anxiety-ERN relationship depends on specific dimensions that cut across diagnoses (e.g., worry), the asymmetry in bias between worry and mixed anxiety, in addition to an asymmetry in statistical power, raises the possibility that the anxiety-ERN relationship is relatively general and non-specific, and that the moderation by worry arises from publication bias.

To summarise our theoretical points, our data point to a potential impasse in the literature. On the one hand, our results point to a small significant relationship between anxiety and the ERN; an effect that is most pronounced for clinical studies and worry. However, the higher evidence from publication bias and attenuation of effect sizes observed for worry and clinical subsamples make it difficult to tell if these studies have genuinely larger effects, or if these moderations are a relatively spurious result of an asymmetry in publication bias between the different categories of studies. To our knowledge, these questions cannot be addressed meta-analytically. Instead, new statistically powerful empirical investigations are required to disentangle the relationship between the ERN and various dimensions that underlie anxious psychopathology (e.g., Gorka et al., 2017).

4.2. Future directions and limitations

Our results can be interpreted to provide practical guidance for ongoing studies of the anxiety-ERN relationship. First, the range of effect sizes across subgroups of our analyses could be used to conduct *apriori* power analyses for ongoing investigations. At a practical empirical level, our results suggest that future studies would need to test at least 162 participants to achieve 80% power for a one-tailed of the uncorrected overall meta-analytical effect size, over 335 participants for the same test with the corrected effect size returned by the selection model ($r = -0.14$), and over 440 participants when using the average of the corrected effect sizes ($r = -0.12$). Similar large numbers of participants would be required to test the ERN-Anxiety relationship even for those sub-analyses that tended to return larger meta-analytical effect sizes (i.e., worry and clinical anxiety). The average sample size in

our analyses was 68.8; despite this sample size being perhaps larger than typical ERP investigations, our results suggest that ongoing research would benefit from increased sample sizes when studying the anxiety-ERN relationship.

Several limitations should be noted when interpreting our results. First, while we coded for a range of categorical moderators, future studies could advance the field by coding of continuous and categorical moderators of the anxiety-ERN relationship. The ERN-anxiety relationship is putatively moderated by sex (Moser et al., 2016) and comorbid depressive symptoms (Weinberg et al., 2012a, 2012b), and the ERN has also been implicated in the developmental trajectory of anxiety (Meyer, 2017). Consequently, future studies might benefit from taking more nuanced and statistically powerful approaches to coding moderators. For example, while we coded age as a binary categorical variable (child vs. adult), future studies could take a more continuous approach to coding age to investigate the anxiety-ERN relationship over the lifespan (from infancy to adulthood to older age). However, it should be noted that our reading of the literature suggests that there is a lack of studies assessing the anxiety-ERN relationship particularly into older age. Critically, then, addressing the lifespan trajectory of the anxiety-ERN relationship not only depends on changing the coding of a moderator, but also on further future studies plugging gaps in the literature.

Clinical and sub-clinical anxiety are highly comorbid with depressive symptoms (Brady and Kendall, 1992), and this comorbidity is undoubtedly present in our samples. As such, our conclusions are limited to anxiety without partialling out variance attributable to comorbid depressive symptoms—this was largely unavoidable due to the nature of anxiety measurement in most studies of the anxiety-ERN relationship. Here, it should be noted that while we suspect that comorbidity of depressive symptoms/negative affect is particularly problematic for the studies classified as mixed anxiety, it is nevertheless also true that studies classified as worry would also contain considerable comorbidity. Thus, while our analyses do reflect the broader anxiety-ERN literature, future meta-analyses might benefit from controlling for depressive symptoms if this becomes possible with sufficient reporting of comorbidity in ongoing investigations.

5. Conclusion

Our results suggested that there is a detectable relationship between anxiety and the ERN after many controls for publication bias, but that the magnitude of this effect is—perhaps considerably—smaller than suggested by previous estimates (average adjusted effect size: $r = -0.12$). Our results also suggest that the most reliable relationships between anxiety and the ERN are observed in clinical studies and anxious samples characterised by worry, while we observed no significant relationship between mixed anxiety and the ERN. However, it should be noted that the subsamples with the largest effect sizes (worry, clinical anxiety) also demonstrated the greatest evidence for publication bias, and resulting attenuations of effect sizes. These results suggest that ongoing research should increase statistical power—likely by increasing sample sizes—when conducting studies of the anxiety-ERN relationship.

References

Amodio, D.M., Master, S.L., Yee, C.M., Taylor, S.E., 2008. Neurocognitive components of the behavioral inhibition and activation systems: implications for theories of self-regulation. Psychophysiology 45, 11–19.

Anokhin, A.P., Golosheykin, S., Heath, A.C., 2008. Heritability of frontal brain function related to action monitoring. Psychophysiology 45, 524–534.

Baxter, A.J., Vos, T., Scott, K.M., Ferrari, A.J., Whiteford, H.A., 2014. The global burden of anxiety disorders in 2010. Psychol. Med. 44, 2363–2374.

Becker, B.J., 2005. Failsafe N or file-drawer number. In: Rothstein, H.R., Sutton, A.J., Borenstein, M. (Eds.), Publication Bias in Meta-analysis: Prevention, Assessment and Adjustments. John Wiley & Sons.

Bergstrom, C.T., Meacham, F., 2016. Depression and anxiety: maladaptive byproducts of adaptive mechanisms. Evolution, Medicine, and Public Health 2016, 214–218.

Brady, E.U., Kendall, P.C., 1992. Comorbidity of anxiety and depression in children and adolescents. Psychol. Bull. 111, 244.

Button, K.S., Ioannidis, J.P., Mokrysz, C., Nosek, B.A., Flint, J., Robinson, E.S., Munafò, M.R., 2013. Power failure: why small sample size undermines the reliability of neuroscience. Nat. Rev. Neurosci. 14, 365.

Carbine, K.A., Lindsey, H.M., Rodeback, R.E., Larson, M.J., 2019. Quantifying evidential value and selective reporting in recent and 10-year past psychophysiological literature: a pre-registered P-curve analysis. Int. J. Psychophysiol. 142, 33–49.

Carter, E.C., Kofler, L.M., Forster, D.E., McCullough, M.E., 2015. A series of meta-analytic tests of the depletion effect: self-control does not seem to rely on a limited resource. J. Exp. Psychol. Gen. 144, 796.

Carter, E., Schönbrodt, F., Gervais, W.M., Hilgard, J., 2019. Carter, E. C., Schönbrodt, F. D., Gervais, W. M., & Hilgard, J. (2019). Correcting for bias in psychology: a comparison of meta-analytic methods. Adv. Methods Pract. Psychol. Sci. 2, 115–144.

Cavanagh, J.F., Shackman, A.J., 2015. Frontal midline theta reflects anxiety and cognitive control: meta-analytic evidence. Journal of Physiology-Paris 109, 3–15.

Dehaene, S., Posner, M.I., Tucker, D.M., 1994. Localization of a neural system for error detection and compensation. Psychol. Sci. 5, 303–305.

Dickersin, K., 1990. The existence of publication bias and risk factors for its occurrence. JAMA 263, 1385–1389.

Endrass, T., Schuermann, B., Kaufmann, C., Spielberg, R., Kniesche, R., Kathmann, N., 2010. Performance monitoring and error significance in patients with obsessive-compulsive disorder. Biol. Psychol. 84, 257–263.

Endrass, T., Riesel, A., Kathmann, N., Buhlmann, U., 2014. Performance monitoring in obsessive–compulsive disorder and social anxiety disorder. J. Abnorm. Psychol. 123, 705.

Fischer, A.G., Klein, T.A., Ullsperger, M., 2017. Comparing the error-related negativity across groups: the impact of error-and trial-number differences. Psychophysiology 54, 998–1009.

Gehring, W.J., Goss, B., Coles, M.G., Meyer, D.E., Donchin, E., 1993. A neural system for error detection and compensation. Psychol. Sci. 4, 385–390.

Gehring, W.J., Himle, J., Nisenson, L.G., 2000. Action-monitoring dysfunction in obsessive-compulsive disorder. Psychol. Sci. 11, 1–6.

Gorka, S.M., Burkhouse, K.L., Afshar, K., Phan, K.L., 2017. Error-related brain activity and internalizing disorder symptom dimensions in depression and anxiety. Depress. Anxiety 34, 985–995.

Hajcak, G., 2012. What we've learned from mistakes: insights from error-related brain activity. Curr. Dir. Psychol. Sci. 21, 101–106.

Hajcak, G., McDonald, N., Simons, R.F., 2003. Anxiety and error-related brain activity. Biol. Psychol. 64, 77–90.

Hajcak, G., Moser, J.S., Yeung, N., Simons, R.F., 2005. On the ERN and the significance of errors. Psychophysiology 42, 151–160.

Idris, N.R.N., 2012. A comparison of methods to detect publication bias for meta-analysis of continuous data. J. Appl. Sci. 12, 1413–1417.

Inzlicht, M., Gervais, W., Berkman, E., 2015. Bias-correction techniques alone cannot determine whether ego depletion is different from zero: commentary on carter. Kofler, Forster, & McCullough 2015. Available at SSRN. https://ssrn.com/abstract=2659409 or. https://doi.org/10.2139/ssrn.2659409.

John, O.P., Srivastava, S., 1999. The Big Five trait taxonomy: History, measurement, and theoretical perspectives. Handbook of Personality: Theory and Research 2, 102–138.

McShane, B.B., Böckenholt, U., Hansen, K.T., 2016. Adjusting for publication bias in meta-analysis: an evaluation of selection methods and some cautionary notes. Perspect. Psychol. Sci. 11, 730–749.

Meyer, A., 2017. A biomarker of anxiety in children and adolescents: a review focusing on the error-related negativity (ERN) and anxiety across development. Developmental cognitive neuroscience 27, 58–68.

Moser, J.S., Moran, T.P., Jendrusina, A.A., 2012. Parsing relationships between dimensions of anxiety and action monitoring brain potentials in female undergraduates. Psychophysiology 49, 3–10.

Moser, J., Moran, T., Schroder, H., Donnellan, B., Yeung, N., 2013. On the relationship between anxiety and error monitoring: a meta-analysis and conceptual framework. Front. Hum. Neurosci. 7, 466.

Moser, J.S., Moran, T.P., Kneip, C., Schroder, H.S., Larson, M.J., 2016. Sex moderates the association between symptoms of anxiety, but not obsessive compulsive disorder, and error-monitoring brain activity: a meta-analytic review. Psychophysiology 53, 21–29.

Nesse, R.M., 2005. Natural selection and the regulation of defenses: a signal detection analysis of the smoke detector principle. Evol. Hum. Behav. 26, 88–105.

Olvet, D.M., Hajcak, G., 2008. The error-related negativity (ERN) and psychopathology: toward an endophenotype. Clin. Psychol. Rev. 28, 1343–1354.

Olvet, D.M., Hajcak, G., 2009. The stability of error-related brain activity with increasing trials. Psychophysiology 46, 957–961.

Pasion, R., Barbosa, F., 2019. ERN as a transdiagnostic marker of the internalizing-externalizing spectrum: A dissociable meta-analytic effect. Neuroscience Biobehavioral Reviews.

Peters, J.L., Sutton, A.J., Jones, D.R., Abrams, K.R., Rushton, L., 2006. Comparison of two methods to detect publication bias in meta-analysis. Jama 295 (6), 676–680.

Pigott, T., 2012. Advances in Meta-Analysis. Springer, New York, NY.

Riesel, A., Endrass, T., Kaufmann, C., Kathmann, N., 2011. Overactive error-related brain activity as a candidate endophenotype for obsessive-compulsive disorder: evidence from unaffected first-degree relatives. Am. J. Psychiatr. 168, 317–324.

Riesel, A., 2019. The erring brain: Error-related negativity as an endophenotype for OCD—A review and meta-analysis. Psychophysiology 56 (4), e13348.

Rothstein, H.R., Sutton, A.J., Borenstein, M. (Eds.), 2006. Publication Bias in Meta-analysis: Prevention, Assessment and Adjustments. John Wiley & Sons.

Saunders, B., Lin, H., Milyavskaya, M., Inzlicht, M., 2017. The emotive nature of conflict

monitoring in the medial prefrontal cortex. Int. J. Psychophysiol. 119, 31–40.
Simmons, J.P., Nelson, L.D., Simonsohn, U., 2011. False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. Psychol. Sci. 22, 1359–1366.
Simonsohn, U., Nelson, L.D., Simmons, J.P., 2014. P-curve and effect size: correcting for publication bias using only significant results. Perspect. Psychol. Sci. 9, 666–681.
Spielberger, C.D., 1989. State-trait Anxiety Inventory: Bibliography, 2nd ed. Consulting Psychologists Press, Palo Alto, CA.
Stanley, T.D., Doucouliagos, H., 2014. Meta-regression approximations to reduce publication selection bias. Res. Synth. Methods 5, 60–78.
Stern, E.R., Liu, Y., Gehring, W.J., Lister, J.J., Yin, G., Zhang, J., Taylor, S.F., 2010. Chronic medication does not affect hyperactive error responses in obsessive-compulsive disorder. Psychophysiology 47, 913–920.
Swick, D., Honzel, N., Turken, U., 2015. Intact error monitoring in combat veterans with post-traumatic stress disorder. Psychiatry Res. Neuroimaging 234, 227–238.
Szucs, D., Ioannidis, J.P., 2017. Empirical assessment of published effect sizes and power in the recent cognitive neuroscience and psychology literature. PLoS Biol. 15, e2000797.
Tamnes, C.K., Walhovd, K.B., Torstveit, M., Sells, V.T., Fjell, A.M., 2013. Performance monitoring in children and adolescents: a review of developmental changes in the error-related negativity and brain maturation. Developmental Cognitive Neuroscience 6, 1–13.
Viechtbauer, W., 2010. Conducting meta-analyses in R with the metafor package. J. Stat. Softw. 36, 1–48.
Weinberg, A., Olvet, D.M., Hajcak, G., 2010. Increased error-related brain activity in generalized anxiety disorder. Biological Psychology 85 (3), 472–480.
Weinberg, A., Hajcak, G., 2011. Longer term test–retest reliability of error-related brain activity. Psychophysiology 48, 1420–1425.
Weinberg, A., Klein, D.N., Hajcak, G., 2012a. Increased error-related brain activity distinguishes generalized anxiety disorder with and without comorbid major depressive disorder. J. Abnorm. Psychol. 121, 885.
Weinberg, A., Riesel, A., Hajcak, G., 2012b. Integrating multiple perspectives on error-related brain activity: the ERN as a neural indicator of trait defensive reactivity. Motiv. Emot. 36, 84–100.
Weinberg, A., Dieterich, R., Riesel, A., 2015. Error-related brain activity in the age of RDoC: a review of the literature. Int. J. Psychophysiol. 98 (2), 276–299.
Zambrano-Vazquez, L., Allen, J.J., 2014. Differential contributions of worry, anxiety, and obsessive compulsive symptoms to ERN amplitudes in response monitoring and reinforcement learning tasks. Neuropsychologia 61, 197–209.