Assessing the association between the error-related ERPs and trait anxiety using mass univariate statistics

by

Zora Chen

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Master of Arts

in

Psychology

Waterloo, Ontario, Canada, 2023

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Enhanced error monitoring, as reflected in increased amplitude of the error-related negativity (ERN) ERP component, has been suggested to be a vulnerability neuro-marker of anxiety disorders. However, the association between an enhanced ERN amplitude and anxiety levels in the nonclinical population have been inconsistent. In a sample of 82 adults, we examined the association between anxiety and the ERN with different analytical methods (mass-univariate statistics and conventional analyses), self-reported anxiety scales (STAI and STICSA), and trial numbers (all correct trials and equal numbers of correct and error trials). Both the conventional and mass-univariate analyses demonstrated a robust enhancement of the ERN and Pe relative to the correct-ERPs. However, the mass-univariate approach additionally unveiled a wider array of electrodes and a longer duration of involvement in this error enhancement. There was no consistent moderation of the findings by trial numbers, analyses, and anxiety scales. Across the analytic methods, the results showed a lack of consistent correlation between trait anxiety and error-related ERPs. The present results suggest a lack of enhancement of error monitoring by anxious traits in individuals with sub-clinical anxiety and those with clinical anxiety but without a clinical diagnosis. Importantly, the absence of such correlation questions the validity of the ERN as a neural marker for anxiety disorders. Future studies that investigate neuro-markers of anxiety may explore alternative neural signatures and task designs, and employ robust statistics to provide a more comprehensive understanding of anxiety vulnerability.

Acknowledgements

Table of Contents

List of Figures

List of Tables

Introduction

Adaptive behaviour requires the ability to monitor one's actions. Action monitoring is necessary for detecting mistakes and implementing compensatory behaviours (Botvinick et al., 2001; Rabbitt, 1966). As mistakes can coincide with physical, social, or financial costs, individuals tend to be highly motivated to avoid them. Error aversion appears to be amplified by anxious traits, and enhanced error monitoring has been correlated with various forms of clinical anxiety (Meyer & Gawlowska, 2017; Saunders & Inzlicht, 2020; Weinberg et al., 2010). However, it is unclear how error monitoring varies depending on individual differences in sub-clinical anxiety. Examining error monitoring in the neurotypical population may help elucidate the degree to which anxiety correlates with cognitive processes of performance.

The error-related negativity (ERN) and error positivity (Pe) are two electrophysiological measures used to examine error processing in relation to anxious traits. The ERN is a negative deflection occurring at fronto-central sites within 100 ms after an incorrect response (Gehring et al., 1993). A smaller negativity, known as the correct-related negativity (CRN), is observed following correct responses. The ERN and CRN are similar in time-course and topography (Vidal et al., 2000, 2003), although the ERN generally has a higher amplitude and is thought to reflect automatic error detection (Holroyd & Coles, 2002). The ERN and CRN are typically assessed with binary choice tasks like the Flanker task (Eriksen & Eriksen, 1974), in which participants make unambiguous correct/error responses within a limited time window. Modified versions of the paradigms may involve monetary incentives (Hajcak et al., 2005) and social context manipulations (Barker et al., 2015). For the current study, we discuss the ERN and CRN in the context of the basic paradigm (no laboratory manipulations). Numerous studies have shown that, compared to non-anxious participants, individuals diagnosed with anxiety-related disorders present with an enhanced ERN amplitude, an effect interpreted as reflecting enhanced error reactivity in clinical anxiety. This enhanced ERN has been found across a variety of clinical groups, including obsessive-compulsive disorder (e.g., Endrass et al., 2010, 2014; Riesel, 2019; Riesel et al.,

2017), social anxiety disorder (Endrass et al., 2014; Kujawa et al., 2016), and generalized anxiety disorder (Weinberg et al., 2010; Weinberg, Kotov, et al., 2015).

Interestingly, the association between anxious traits and error reactivity has not been consistently observed in neurotypical populations that display subclinical levels of anxiety. Associations of error processing with such subclinical anxiety traits are mixed in basic binary choice conditions (i.e., unambiguous responses with no incentives). Some evidence suggests that healthy individuals with high trait levels of anxiety (HTA) have enhanced error monitoring compared to those with low trait levels of anxiety (LTA), as reflected by heightened ERN amplitudes for the former group (Hajcak et al., 2003; Meyer et al., 2012). However, another study demonstrated the opposite pattern, where the HTA group showed reduced ERN amplitudes compared to the LTA group (Hsieh et al., 2021). Still, another study found that the HTA and LTA groups did not differ in ERN amplitudes (Aarts & Pourtois, 2010). Some studies using Pearson's correlations found no significant association between ERN amplitude and self-reported anxiety scores (Beste et al., 2013; Cavanagh & Allen, 2008; Saunders et al., 2015). Further, a recent meta-analysis reported that subclinical anxiety does not have a significant association with ERN amplitude (Saunders & Inzlicht, 2020). Thus, to date, the association between subclinical anxiety traits and ERN amplitude remains unclear.

The Pe is a positive deflection at parietal sites that occurs 200-400 ms after an incorrect response (Falkenstein et al., 1991). A smaller deflection, called the correct positivity (Pc), is also observed following a correct response with a similar timing and topography as the Pe. The Pe has been hypothesized to reflect the conscious awareness of errors (Endrass et al., 2005; Nieuwenhuis et al., 2001). Findings on the Pe association to individual differences in anxiety have been mixed, although most studies have not detected any significant modulation of Pe amplitude by anxiety (Endrass et al., 2014; Klawohn et al., 2014; Olvet & Hajcak, 2009b; Weinberg et al., 2010; Xiao et al., 2011).

One possibility underlying the mixed results between anxiety and the error-related ERPs is that the way trait anxiety is measured remains inconsistent across studies. In addition, depressive symptoms may moderate ERN amplitude (Chiu & Deldin, 2007; Holmes & Pizzagalli, 2008, 2010b; Weinberg et al., 2016; see Pasion & Barbosa, 2019 for a meta-analysis) and the anxiety-ERN association (Weinberg et al., 2012; Weinberg, Kotov, et al., 2015). Similarly, Pe amplitudes can be blunted by depression (Aarts et al., 2013; Olvet et al., 2010; Schrijvers et al., 2008). It is possible that non-specific anxiety scales that blend anxiety with depression may introduce further complexity when examining the anxiety-ERN association. Sub-clinical levels of anxiety trait have been often assessed by the State-Trait Anxiety Inventory (STAI; Spielberger, 1983; e.g., Aarts & Pourtois, 2010; Meyer & Gawlowska, 2017). However, the STAI has been criticized on its ability to discriminate between anxiety and depression symptoms (Bieling et al., 1998). Some STAI items (e.g., "I wish I could be as happy as others seem to be.") appear to reflect symptoms related to both anxiety and depression, which undermines its ability to assess anxiety as distinct from depression (Bieling et al., 1998; Caci et al., 2003). In light of the association between depressive symptoms and the Pe (Aarts et al., 2013; Olvet et al., 2010; Schrijvers et al., 2008) and ERN (Weinberg et al., 2012; Weinberg, Kotov, et al., 2015), using a specific scale that reflects pure anxiety may provide a clearer and more accurate picture of the correlation between anxiety and error-related ERPs.

Succinctly, the State-Trait Inventory for Cognitive and Somatic Anxiety (STICSA; Ree et al., 2008) improves upon the shortcomings of the STAI. The STICSA includes items that are relatively unique to anxiety and excludes items shared by depression (Grös et al., 2007). The STICSA demonstrates an improved ability to differentiate between symptoms of anxiety and depression compared to the STAI, thus offering a more precise assessment of anxiety (Grös et al., 2007). To the best of our knowledge, the STICSA scale has not been used in studies assessing performance monitoring with the ERN or Pe. One goal of the present study was to compare the association between ERN and Pe amplitudes with anxiety using both STAI and STICSA scales.

A second important factor that may be involved in the inconsistent anxiety association with error processing ERPs pertains to how the ERN and Pe are measured in the first place. The correct-related ERPs (i.e., the CRN and Pc) occur after every response and appear to reflect generic response monitoring. In order to isolate error-specific processes from generic response monitoring, subtraction-based difference waves between the CRN/ERN (ΔERN) and Pc/Pe (ΔPe) have been used (e.g., Meyer et al., 2015, 2018; Saunders et al., 2015; Weinberg et al., 2012; Weinberg 2010; see Luck, 2014 for a more thorough discussion). In addition, the ΔERN has shown more robust association with individual differences in anxiety than the ERN (Klawohn et al., 2020). Compared to the absolute error waveforms, the ΔERN and ΔPe are more error-specific measures as they remove stimulus-related activity and generic response monitoring that are shared between the error and correct responses (Klawohn et al., 2020).

The number of trials averaged to create the ERP is critical for the validity of any ERP quantification (Luck, 2014). Considering that the ΔERN and ΔPe calculations involve both error and correct trials, it is important that these conditions have comparable trial counts in the analysis. However, drastically different numbers of correct and erroneous trials have been included in previous calculations of the ΔERN and ΔPe. The correct trial numbers used for CRN and Pc calculations are typically much higher than the error trial numbers used for ERN and Pe calculations, in part due to the low ratio of errors to correct responses produced by the typical action monitoring paradigm (e.g., Flanker task). The imbalance of trials across conditions can result in lower signal-to-noise ratios of the error-related ERPs than the correct-related ERPs, which can be problematic for the examination of the difference wave. Equating trial numbers across conditions is thus necessary to avoid biased results. Bootstrapping (e.g., Buzzell et al., 2019) and permutation (e.g., Muir et al., 2020) have been used to combat this problem. However, to date, no study has directly tested the influence of trial counts by matching the number of correct and error trials for each individual. One goal of the present study was to fill this gap.

Finally, employing robust statistics can enhance the statistical examination of ERPs. Previous studies that examined error monitoring have utilized classic ERP analysis, which can inflate Type I error rate (Luck & Gaspelin, 2017). The classic practice (mean amplitude of pre-defined channels and time window) also limits the examination of brain activity to a specific time window and specific channels. Doing so decreases the chances of finding other potentially interesting neuro-markers related to error monitoring and its association to anxiety. Mass-Univariate (MU) analysis is a more robust way to study ERPs via a data-driven approach (Luck & Gaspelin, 2017). The MU approach allows for analyses across all electrodes and time points (Luck & Gaspelin, 2017), thereby reducing Type II error rate as compared to the limited range analysis required by the classic approach. With an appropriate correction for multiple comparisons, the MU approach can also control for overall Type I error rate (Luck & Gaspelin, 2017) while preserving statistical power (Pernet et al., 2011). Therefore, when examining the ERN and Pe, MU statistics may offer advantages compared to the classic approach.

In sum, the current study investigated error monitoring in relation to subclinical trait anxiety, while controlling for several confounding factors and utilizing robust statistics. The study aimed to examine whether an anxiety-specific scale (STICSA) may reveal a stronger correlation with error-related ERPs than a non-anxiety specific scale (STAI). To the best of our knowledge, the current study would be the first to use the STICSA scale in the error-monitoring literature. As an additional effort to account for shared variance between depression and anxiety, depression symptoms were assessed using Depression Anxiety Stress Scale-Depression subscale (Lovibond & Lovibond, 1995) and controlled for in the analyses. We hypothesized that the ERN amplitude would be closely correlated with anxiety, and that this correlation would be revealed by a pure anxiety measure. In light of the lack of significant associations between the Pe and anxiety in the literature, we did not have a-priori hypotheses for the Pe. The Pe analyses were exploratory.

A second goal of the study was to investigate the impact of methodological choices on the error-related ERPs and their associations with anxiety. First, we expected that the

difference in trial numbers between correct and error conditions would influence how the correct- and error-related ERPs compare. To this end, we created an additional ERP dataset that contained an equal number of correct and error trials per participant to match the signal-to-noise ratios of the correct- and error-related ERPs. With fewer trials, we expected a greater overlap of confidence intervals between the correct and error ERPs, and thus fewer time points of significant differences between the CRN and the ERN, and between the Pc and Pe. Second, we examined whether the analytical approach (conventional or robust statistics) would influence the findings. All ERP analyses were performed with MU statistics using the LIMO EEG toolbox (Pernet et al., 2011) as well as the conventional approach by creating average ERPs. For both analyses, we expected that the electrodes at which the ERN and Pe are usually maximal would show significant main effects of response type, such that error-related ERPs would be of higher amplitude than correct-related ERPs. We also expected stronger correlations of error-related ERPs with the STICSA than with the STAI. We had no a-priori hypotheses for ERPs other than the ERN and the Pe. Together, this study aimed to investigate the impact of methodological factors, including anxiety scales, trial numbers, and statistical analysis, on error-related ERPs and their associations with anxiety traits.

Methods

The study design, sample size, inclusion/exclusion criteria, as well as planned analyses were pre-registered on Open Science Framework (osf.io/d5amx).

Participants

A total of 87 participants were recruited from the Waterloo region through the University of Waterloo online system (SONA) and through local advertisements. Individuals between 17-29 years old were eligible to participate if they had no history of brain lesion, coma, or loss of consciousness for over 5 minutes. They reported no personal or familial history of epilepsy, seizures, or sensitivities to flashing lights. They also reported no history of neurological or psychiatric disease (e.g., major depressive disorder, generalized anxiety disorder, social anxiety disorder, autism spectrum disorder) and were not taking antidepressants or antipsychotic drugs at the time of testing. All participants had normal or corrected-to-normal vision. All study procedures were reviewed by and received ethic clearance through the Ethic Review Board of the University of Waterloo. Participants either received course credits or cash remuneration for their time ($15 CAN/hour).

Of the 87 participants tested, one was excluded due to low accuracy rate (< 50%). Four participants were excluded due to committing less than eight errors, the minimum number necessary for a reliable ERN amplitude (Foti et al., 2013; Olvet & Hajcak, 2009b). The final sample included 82 individuals (65% female) with a mean age of 21.21 years ($SD = 2.67$). Twenty-two percent (22%) of the participants identified as White/Caucasian, 29% were East Asian, 13% were South Asian, 16% Indian, 8% Middle Eastern, and 11% identified as other ethnic groups.

Self-report measures

STICSA-T. Trait anxiety was assessed using the trait scale of the State Trait Inventory for Cognitive and Somatic Anxiety (STICSA, Cronbach's $\alpha = .90$; Ree et al., 2008). Ten items (e.g., "I think the worst will happen") assessed cognitive aspects of anxiety,

and 11 items (e.g., "My throat feels dry") assessed somatic symptoms associated with anxiety. Participants indicated the degree to which each statement reflected how they felt in general (1 = *completely disagree*, 4 = *completely agree*). Average score was computed for the 21 items. Scores for 23 participants (28%) fell within the "clinical anxiety" range (score of 43 and above; Van Dam et al., 2013).

STAI-T. The State-Trait Anxiety Inventory (STAI, Cronbach's α = .92; Spielberger, 1983) includes 20 items (e.g., "I feel nervous and restless") assessing symptoms of anxiety. Participants reflected on how they generally felt and rated each statement from 1 (*completely disagree)* to 4 (*completely agree*). Average scores were computed for the 20 items. Scores for 44 participants (54%) were above the "clinical anxiety" threshold (score of 45 and above; Spielberger, 1983).

DASS-D. Trait levels of depressive symptoms were assessed using the depression subscale of the Depression Anxiety Stress Scale (DASS-D, Cronbach's α = .92; Lovibond & Lovibond, 1995). Participants used a 4-point scale (1 = *completely disagree*, 4 = *completely agree*) to indicate how well each of the 14 items described them in general (e.g., "I felt downhearted and blue"). Average scores were calculated for the 14 items. Scores for 1 participant (1%) was within the "no depression" range (0-9); scores for 41 participants (50%) fell within the "moderate depression" range (14-20), scores for 27 (33%) participants fell within the "severe depression" range (21-27), and scores for 13 participants (16%) were in the "extremely severe depression" range (28 and above; Lovibond & Lovibond, 1996).

Task and Materials

An arrowhead version of the flanker task (Eriksen & Eriksen, 1974) was presented using SR Research Experiment Builder version 2.3.38 (Figure 1).[1] Each trial started with a fixation cross presented at the centre of the screen for 250 ms. The fixation cross was followed by a blank screen that lasted for a varied interval from 300 to 600 ms. Five horizontally aligned arrowheads were then shown for 200 ms. Participants had up to 1000 ms

---

[1] Eye movements were not recorded.

to respond to the direction of the central arrow. After participant response, the screen would remain blank for an inter-trial interval that varied randomly from 1700 to 2000 ms. Half of the trials were congruent (">>>>>" or "<<<<<") and half were incongruent (">><>>" or "<<><<"). The order of congruent and incongruent trials was random. In addition, each arrowhead combination was presented for a maximum of three consecutive occurrences.

In the flanker task, participants pressed the up or down arrow key on a keyboard to indicate the direction of the middle arrow. Participants completed a practice block with six trials. They then completed six test blocks of 100 trials (600 trials total). The finger-arrow assignment was reversed after each block (up-arrow key for left central arrows and down-arrow key for right central arrows in one block; opposite in the next block etc.). The order of the key-arrow assignment for the first block was counterbalanced across participants. Participants were encouraged to respond as quickly and accurately as possible. Performance feedback was given at the end of each block. If the accuracy was below 75%, "Please be more accurate" was displayed on the monitor; if their accuracy was above 90%, "Please respond faster" was displayed; otherwise, "You're doing a good job" was displayed.

Figure 1

Study paradigm
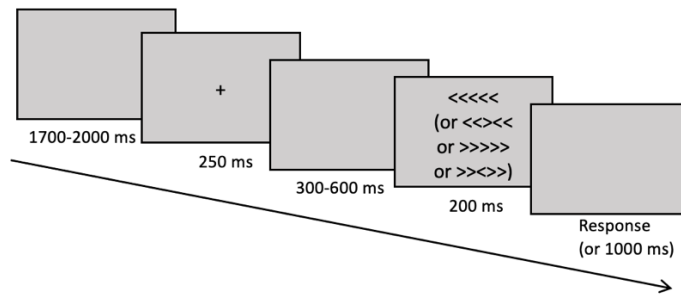
Note. The inter-trial interval had a random duration between 1700 to 2000 ms. On each trial, a fixation cross was presented for 250 ms, followed by a blank screen for a variable interval from 300-600 ms. Five arrows would then be presented for 200 ms. Each arrow combination was presented for 25% of trials for each block. Participants had up to 1000 ms to indicate the direction of the middle arrow using either the up or down arrow key on a keyboard.

Procedure

      Upon arriving at the laboratory, participants were told that the study examined people's neural activity during a computer task. They were given a description of the experiment. Participants provided informed written consent before the electroencephalogram (EEG) electrodes were attached. They then completed a demographic questionnaire about their sex, age, vision, and ethnicity. Detailed instructions of the flanker task were given. Participants were then seated in a chair with their head placed on a chin rest while they completed the flanker task. Participants then answered the anxiety and depression questionnaires. They were then debriefed and provided post-debriefing written consent.

EEG recording and data reduction

      EEG data were recorded at 512Hz using a BioSemi Active-Two system (Amsterdam, the Netherlands) at 66 scalp sites. Six additional electrodes were attached with stickers under the eyes, on the temples and the mastoids. Signals were referenced online to the Common Mode Sense (CMS) and Driven Right Leg (DRL) electrodes. Offline, EEG data were re-referenced to the common average reference. EEG data were processed offline with the EEGLab (version v2022.1; Delorme & Makeig, 2004) and the ERPLab toolboxes (version 9.00; Lopez-Calderon & Luck, 2014). Raw data were band-pass filtered (0.01 – 30 Hz). Epochs were extracted from -400 ms to +800 ms relative to response onset. Data were then baseline corrected using the -400 to -200 ms time window prior to response. Any frontal or non-ocular channels (i.e., excluding electrodes Fp1, Fpz, Fp2, AF3, AFz, AF4, AF8, AF7, IO1, IO2, LO1, LO2) that were consistently noisy were removed for interpolation using EEGLab spherical splines tool. A trial was rejected if there was a voltage change of more than 100 µV on any non-frontal and non-ocular channels. Data were then processed with Independent Component Analysis (ICA) to remove eye blinks. An average of 367 ($SD = 115$) correct trials and 53 ($SD = 37$) error trials across participants remained for the following analyses on electrophysiological data.

Data Analysis

Behavioural analyses

Behavioural results were analyzed using SPSS (Version 29). Greenhouse-Geisser corrected results were reported when the assumption of sphericity was violated. Two behavioural measures were analyzed: Post-error slowing (PES) and overall accuracy. PES was calculated as the mean RTs for correct trials following incorrect responses (correct-in) minus the mean RTs in correct trials following correct responses (correct-cor). We first performed a within-subjects one-way ANOVA with the factor Response Type (contrasting correct-in and correct-cor conditions) to validate the PES before proceeding with further analyses. To examine whether PES was correlated with anxiety, we performed linear hierarchical regressions with PES as the dependent variable, and STICSA and STAI scores as the predictor in two separate regression models. DASS-D scores were entered in the second step of each model to assess the variance of anxiety distinct from depression. The same regression analyses were run on accuracy.

ERP analyses

To examine whether the number of correct trials relative to error trials would influence the ERP results, we created a dataset that contained equal number of trials across the correct and error conditions.[2] Specifically, we used a MATLAB function to randomly select correct trials to match the number of error trials for each participant. The full and subset ERP datasets were subjected to the same analytical pipelines. That is, both datasets were analyzed with mass-univariate analysis as well as the conventional approach, described below.

---

[2] Two EEG datasets were created by randomly selecting a subset of correct trials to match the number of error trials. These datasets showed the same results, so only results of one of these datasets were reported here.

Mass-univariate analysis

Response-locked ERP activity was analyzed using the LIMO EEG toolbox (version 3.0; Pernet et al., 2011). The toolbox utilizes hierarchical linear modelling which accounts for both within- and between-subjects variance. The first level general linear model estimates parameters for each subject at each electrode and each time point independently using Weighted Least Squares. Then, the estimated parameters were grouped to perform second level analyses.

To investigate the difference in error and correct response processing, repeated-measures ANOVA was applied to the full epoch duration (-400 to 800 ms) with the within-subject factor response type (correct and error). For repeated-measured ANOVA, LIMO computes the Hotelling T^2 test to determine the significance at $p < .05$. Next, to examine the association between trait anxiety and error processing, we performed regression analysis with trait anxiety as the predictor, and the amplitude of correct- and error-related ERPs as the dependent variable. Separate regressions were run for STICSA and STAI scores as the predictor. Thus, four separate regression analyses were run. LIMO performs the Fisher F test for regression analyses at the significance level of $p < .05$. Threshold free cluster enhancement (TFCE) correction and bootstrap computation (1000 iterations) were applied in both the repeated-measure ANOVA and regression analyses to control for Type I family wise error rates (see Pernet et al., 2015 for simulation).

Conventional ERP analysis

Following the conventional approach, we created average ERPs for each condition and each participant. The ERN was calculated as the mean voltage amplitude between 0-100 ms (post-response, Barker et al., 2015; Meyer et al., 2015, 2017; Rabinak et al., 2013; Weinberg et al., 2012) at Fz, Cz, and FCz following an error response. The CRN was calculated on the same electrodes and duration following a correct response. The ΔERN was calculated by subtracting the CRN from the ERN. The Pe and Pc were calculated as the mean

amplitudes from 200-400 ms at Pz and CPz following an error and correct response, respectively. The ΔPe was defined by the difference between Pe and Pc amplitudes.

Conventional ERP analyses were conducted using SPSS (Version 29). Results were reported with Greenhouse-Geisser correction when the assumption of sphericity was violated. Bonferroni correction for multiple comparisons were applied for pairwise comparisons. A 2x3 repeated measure ANOVA was conducted on mean amplitudes, with within-subject factors of ERP type (ERN and CRN) and electrode (Cz, FCz, and Fz). Next, to examine whether anxiety correlated with error-specific processing, the association between ΔERN and anxiety was examined with logistic regressions with STICSA and STAI as predictors in separate models. Depression was adjusted for in the second step of each model to reveal the unique variance of anxiety.

Similar ANOVA and regression analyses were applied for the Pe, Pc, and ΔPe. Specifically, mean amplitudes were analyzed using a 2x2 repeated measures ANOVA with the within-subject factors of ERP type (Pe and Pc) and electrode (Pz and CPz). The association between ΔPe and anxiety was examined with logistic regressions with STICSA and STAI as predictors in separate models. Depression scores were adjusted for in the second step.

To examine whether the response-related ERPs reflect behaviour-related variables, we examined correlations between the ERPs (Pe, ΔPe, ERN and ΔERN) and behavioural measures of performance, including PES and overall accuracy.

Result

Behavioural Results

Please see Table 1 for descriptive statistics and inter-correlations of behavioural and self-reported measures. The overall accuracy was 85% (SD = 12.77). There was a significant main effect of Response Type on the RT in the next trial, $F(1, 81)$ = 21.02, $p < .001$, η^2 = 0.21. Specifically, participants took longer to make a correct response after error responses (M = 523.28, SD = 13.67) than after correct responses (M = 496.48, SD = 10.07). However, this Post Error Slowing was not predicted by STICSA (B = 0.02, p = .970) or STAI (B = 0.40, p = .525) scores (see Table 2).

Table 1

Descriptive statistics and Pearson correlations between self-reported and behavioural measures

	M (SD)	1	2	3	4	5
1. Overall RT (ms)	484.20 (95.64)	-	-	-	-	-
2. Accuracy (%)	84.70 (12.77)	0.12	-	-	-	-
3. PES (ms)	29.81 (58.87)	0.38**	0.08	-	-	-
4. STICSA	37.60 (10.01)	0.25*	-0.04	0.10	-	-
5. STAI	43.95 (10.50)	0.24*	0.14	0.07	0.76**	-
6. DASS-D	21.24 (6.73)	0.22*	0.12	0.16	0.64**	0.65**

Note. PES = Post-error slowing; STICSA = State-Trait Inventory for Cognitive and Somatic Anxiety; STAI = State-Trait Anxiety Inventory; DASS-D = Depression Anxiety Stress Scale-Depression subscale. PES was computed as the difference in average response times (RTs) between correct trials following error responses minus correct RTs following correct responses. Due to a significant correlation between STICSA and DASS-D scores, we opted for hierarchical regressions instead of Pearson correlations to account for depression-related variances in assessing the association between anxiety and ERN. * $p < .05$; ** $p < .001$

Table 2

Behavioural measures regression models

	Overall model	B	SE	p
STICSA models				
Model 1	$R^2 = 0.004, p = .970$			
STICSA		0.02	0.41	.970
Model 2	$R^2 = 0.165, p = .336$			
STICSA (step 1)		-0.17	0.43	.701
DASS-D (step 2)		1.51	1.01	.141
STAI models				
Model 1	$R^2 = 0.071, p = .525$			
STAI		0.40	0.63	.525
Model 2	$R^2 = 0.165, p = .337$			
STAI (step 1)		-0.31	0.82	.705
DASS-D (step 2)		1.71	1.27	.184

ERP Analyzed with LIMO

Main Effect of Response Types

Full dataset

The main effect of Response Type showed significance across the majority of channels (Figure 2A). The strongest significant differences were found over fronto-central, central, and central-parietal electrodes. Timeframes of significance extended from -209 ms to 664 ms with the maximal effect found at 332 ms on P4 ($F = 100.79, p = .001$, ΔError-Correct = 1.17 μV). In comparison, the maximal F-value in the ERN timeframe was found on FC1 at 31 ms, although the effect was smaller than the difference on P4 ($F = 57.48, p = .001$, ΔError-Correct = -1.90 μV). Single channel ERPs revealed that error-related waveforms had higher amplitudes than correct-related waveforms. In accordance with conventional analyses,

the Fz, FCz, and Cz electrodes showed significantly more negative ERN amplitudes than CRN amplitudes during the 0-80 ms timeframe (Figure 2B-D). Similarly, error responses elicited more positive Pe amplitudes relative to the Pc at CPz and Pz sites from 200-400 ms (Figure 2E-F). Contrary to our hypotheses, neither STAI nor STICSA scores correlated with correct or error ERPs after TFCE correction. To clearly demonstrate the lack of significant correlations, uncorrected F-values for the regressions are shown in Appendix A Figure A1.

Subset Dataset (i.e. same trial numbers per condition)

The pattern of Response Type main effect for the subset dataset did not appear to differ much from the full dataset (Figure 3A). The first timepoint of significant difference was observed at -217 ms, and the last significant timeframe was at 635 ms. The maximal effect occurred at 280 ms at P4 ($F = 100.79$, $p = .001$, ΔError-Correct = 1.17 µV). The maximal difference in the ERN timeframe was observed at 29 ms at FCz ($F = 66.10$, $p = .001$, ΔError-Correct = -2.56 µV). ERN amplitudes were more negative than CRN amplitudes at Fz, FCz, and Cz (Figure 3B-D). Similar error enhancement was also observed for Pe relative to Pc at CPz and Pz from 200-400 ms (Figure 3E-F). As seen for the full dataset, the regressions of STICSA and STAI scores predicting ERN and CRN amplitudes were not significant after TFCE correction (see uncorrected F-values in Appendix A Figure A2).

Figure 2

Response main effect with an alpha of 0.05 for the full dataset
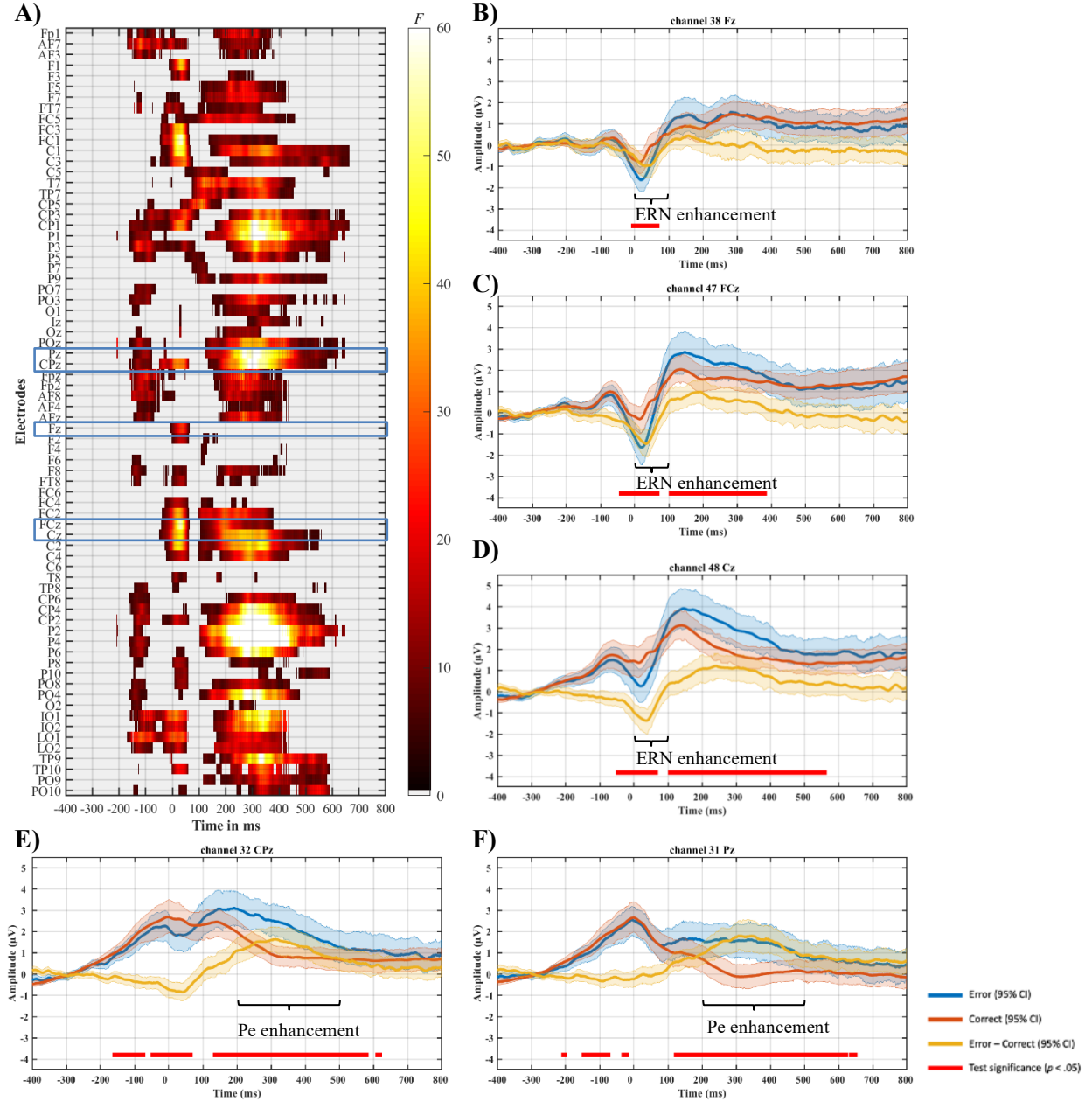
Note. A) *F*-values across the entire -400 to 800 ms epoch, with TFCE correction applied for multiple comparisons. Single channels are plotted for error, correct, and difference (error-correct) waves calculated with 20% trimmed means and 95% confidence intervals for the ERN electrodes at Fz (B), FCz (C), Cz (D), and for the Pe electrodes at CPz (E), and Pz (F).

17

Figure 3

Response main effect with an alpha of 0.05 for the dataset containing equal number of correct and error trials for each participant.
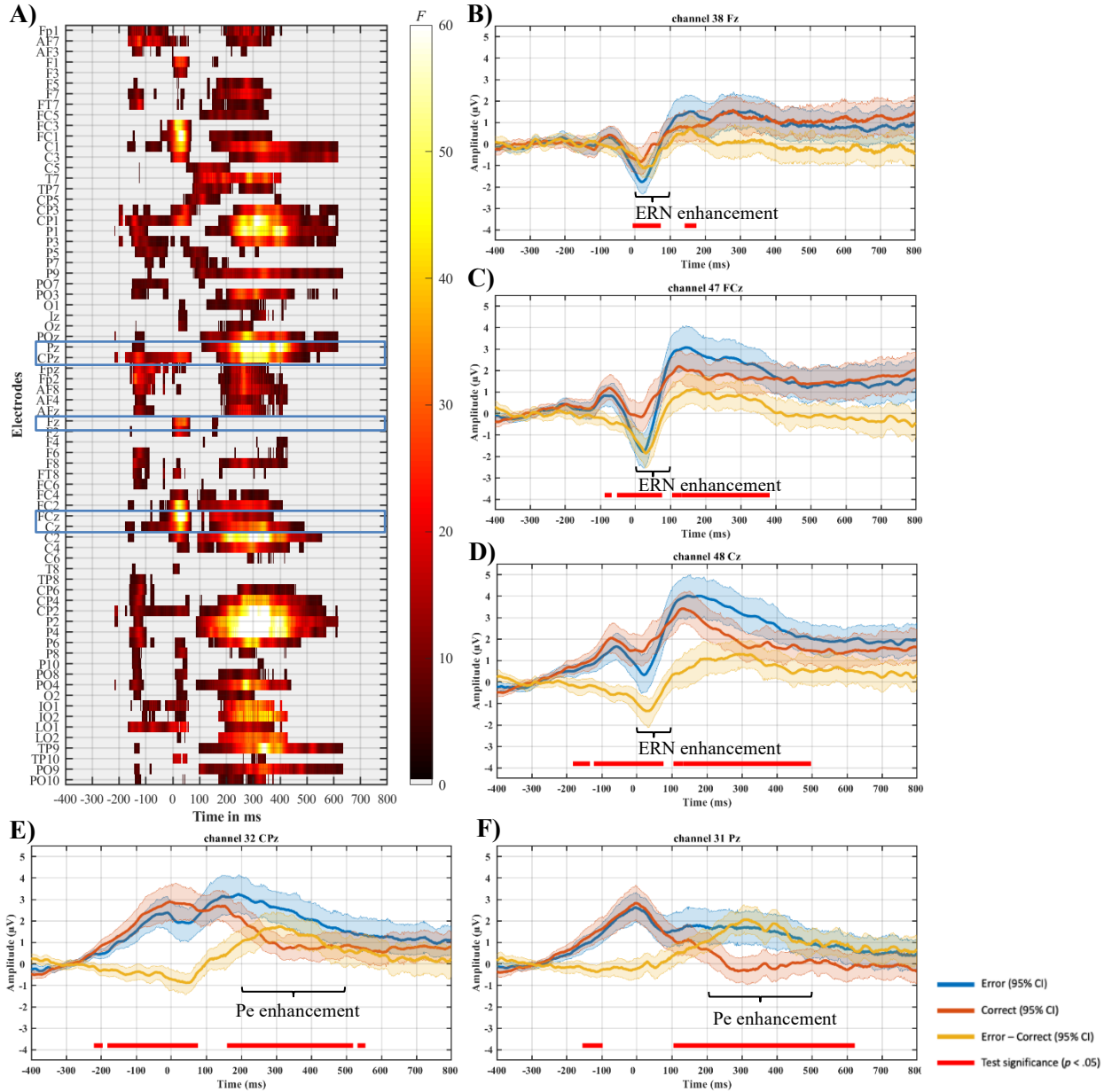
Note. A) *F*-values across the entire -400 to 800 ms epoch, with TFCE correction applied for multiple comparisons. Single channels are plotted for error, correct, and difference (error-correct) waves calculated with 20% trimmed means and 95% confidence intervals for the ERN electrodes at Fz (B), FCz (C), Cz (D), and for the Pe electrodes at CPz (E), and Pz (F).

ERP Analyzed with the Conventional Approach

Full Dataset

The response main effect, $F(1, 81) = 23.96$, $p < .001$, $\eta^2 = .23$, showed that the amplitude of the ERN ($M = 0.29$ µV , $SE = 0.32$) was more negative than the amplitude of the CRN ($M = 1.34$ µV, $SE = 0.29$). The channel main effect was significant, $F(1.2, 99.8) = 83.92$, $p < .001$, $\eta^2 = .51$, indicating that the overall ERN and CRN average amplitudes were the most positive at Cz, followed by FCz and Fz, all pairwise comparisons were significant ($ps < .001$). The interaction between channel and response was significant, $F(1.8, 143.4) = 3.20$, $p = .05$, $\eta^2 = .05$. Specifically, the ERN mean amplitude was the most negative at Cz, relative to FCz and Fz (pairwise $ps \leqslant .001$), and the CRN mean amplitude was the most positive at Cz, relative to FCz and Fz (pairwise $ps \leqslant .001$). Mean ERN amplitude was more negative than the CRN amplitude at all electrodes ($ps < .001$ at FCz, Fz, and Fz).

Likewise for the Pe timeframe, the main effect of response was significant, $F(1, 81) = 81.65$, $p < .001$, $\eta^2 = .50$, such that the Pe amplitude ($M = 3.59$ µV, $SE = 0.44$) was more positive than the Pc amplitude ($M = 1.04$ µV, $SE = 0.32$). The main effect of channel was significant, $F(1, 81) = 84.77$, $p < .001$, $\eta^2 = .51$, indicating that the overall amplitude across Pc and Pe was more positive at CPz than Pz ($p < .001$). The Pe and Pc amplitudes did not differ by channel (response x channel interaction $p = .132$).

We performed one regression model for each of the ΔERN electrodes (i.e., FCz, Cz, and Fz). For all channels, STICSA scores did not predict ΔERN amplitudes ($ps > .603$), even after controlling for depression ($ps > .545$).[3] Adding DASS-D scores to the regression models

---

[3] To stay consistent with analyses in previous studies, we also examined anxiety enhancement between high trait anxiety (HTA) and low trait anxiety (LTA) groups, divided based on STICSA or STAI scores relative to the sample mean. To investigate the ERN enhancement, a mixed-model ANOVA was then performed on the ΔERN with the between-subject factor of anxiety (HTA and LTA) and the within-subject factor of channel (FCz, Fz, and Cz). There was no main effect of Anxiety Group on ΔERN amplitudes (groups based on STICSA: $p = .973$; based on STAI: $p = .410$). A similar mixed ANOVA was performed on the ΔPe with the between-subject factor of trait anxiety (HTA and LTA) and the within-subject factor of channel (Pz and CPz). The group main effect was again non-significant ($p = .671$ for grouping based on STICSA; $p = .310$ for grouping based on STAI). To investigate whether depression reduced the group effects, we controlled for the depression scores as a covariate in an ANCOVA that included the between-subjects groups of anxiety (HTA and LTA) for

did not significantly improve the models, ps of $\Delta R^2 > .411$, showing that depression was not correlated with ΔERN amplitudes. The results were also non-significant for STAI scores at any of the channels (ps > .270; ps > .175 after adjusting for depression).

The regressions results were similar for ΔPe. Anxiety did not predict ΔPe amplitudes at Pz or CPz (for STICSA: ps > .588 and ps > .084 after adjusting for depression; for STAI: ps > .459 and ps > .945 after adjusting for depression). Although DASS-D scores significantly correlated with ΔPe model at CPz, when STICSA score was adjusted (ΔR^2 = 0.05; DASS-D: $B = 0.11$, $p = .041$), STICSA scores remained a non-significant predictor ($B = -0.06$, $p = .084$) and the overall model was also non-significant ($R^2 = 0.06$, $F(2, 81) = 2.32$, $p = .105$).

ERN amplitudes at Fz, FCz, or Cz were not associated with overall accuracy (ps >.239). However, ERN amplitudes were negatively correlated with PES (Fz: $B = -6.05$, $p = .007$; FCz: $B = -4.94$, $p = .011$; Cz: $B = -3.47$, $p = .050$). ΔERN amplitudes calculated from the full dataset showed that the difference wave was negatively associated with PES only at Fz (Fz: $B = -6.50$, $p = .040$; FCz and Cz: ps > .24). Accuracy correlations showed different results. ΔERN amplitude at Fz did not correlate with accuracy ($p = .436$), but the correlations were significant for ΔERN amplitudes at FCz ($B = -1.56$, $p = .010$) and Cz ($B = -1.57$, $p = .021$).

Pe amplitudes at either CPZ or Pz did not predict behavioural variables (PES: ps > .096; accuracy: ps > .384). ΔPe amplitudes calculated from the full dataset did not predict

---

the ΔERN and ΔPe. There were no significant group differences for any channel in either the ERN or Pe timeframe (ps for group main effects > .80).

To examine if error enhancement was evident only in individuals with clinical levels of anxiety, we conducted regression analyses in a subgroup containing only participants who met the clinical threshold of anxiety based on their anxiety scores (STICSA threshold = 43, n = 23; STAI threshold = 45, n = 45). Participants in the clinical subgroup did not show ΔERN modulation by symptom severity, even after adjusting for DASS-D (STICSA group: ps > .062 at Fz, FCz, and Cz; STAI threshold ps > .381).

Because the maximal response main effect was found at P4 within the Pe timeframe in the LIMO analysis, we examined whether the error and difference amplitudes at this electrode would correlate with anxiety scores. To this end, we performed a linear regression with anxiety scores predicting Pe and ΔPe at P4. The correlations were non-significant for both STICSA ($p = .681$ for Pe; $p = .203$ for ΔPe) and STAI scores ($p = .593$ for Pe; $p = .753$ for ΔPe).

post-error slowing ($p > .168$ at Pz; $p = .447$ at CPz) or accuracy ($p = .132$ at Pz), although ΔPe at CPz was associated with higher accuracy at trend level (B = 1.04, $p = .057$).

Subset Dataset (equal trial numbers per condition)

The results were replicated with the subset dataset. Specifically, the response main effect for the ERN and CRN was significant, $F(1, 81) = 27.37$, $p < .001$, $\eta^2 = .25$, such that the mean ERN amplitude ($M = -1.36\mu V$, $SE = 0.31$) was greater than the CRN ($M = -0.29$ μV, $SE = 0.32$) amplitude. The main effect of channel, $F(1.3, 104.3) = 81.72$, $p < .001$, $\eta^2 = .50$, revealed that Cz had an overall more positive amplitude ($M = 2.77$ μV, $SE = 0.38$) than FCz ($M = 0.41$ μV, $SE = 0.34$) and Fz ($M = -0.70$ μV, $SE = 0.28$), all pairwise $ps < .001$. The response by channel interaction was not significant, $p = .086$.

The main effect of response was also significant for the Pe timeframe, $F(1, 81) = 84.98$, $p < .001$, $\eta^2 = .51$, such that the mean Pe amplitude across channels ($M = 3.59$ μV, $SE = 0.44$) was greater than the mean Pc amplitude ($M = 1.04$ μV; $SE = 0.35$). The main effect of channel, $F(1, 81) = 83.01$, $p < .001$, $\eta^2 = .51$, indicated that the overall voltage at CPz ($M = 3.13$ μV, $SE = 0.39$) was more positive than the voltage at Pz ($M = 1.50$ μV, $SE = 0.38$). The interaction between channel and response was not significant, $p = .215$.

The regression results for the subset dataset slightly differed from the regression results of the full dataset. STICSA scores did not predict ΔERN amplitudes at FCz, Cz, or Fz ($ps > .517$; $ps > .305$ after controlling for depression), and STAI scores did not predict ΔERN amplitudes at FCz or Cz ($ps > .385$; $ps > .349$ after controlling for depression). Although STAI was not a significant predictor of ΔERN at Fz when depression was not controlled for ($B = -0.04$, $p = .097$), it predicted a smaller ΔERN when depression included in the model (DASS: $B = 0.08$, $p = .092$; STAI: $B = -0.07$, $p = .019$), and the overall model became marginally significant, $R^2 = .068$, $F(2, 81) = 2.90$, $p = .061$. STICSA and STAI scores did not predict ΔPe amplitudes at Pz or CPz, regardless of whether depression was controlled for ($ps > .278$), and depression was not a significant predictor ($ps > .122$).

ΔERN amplitudes did not predict PES at Fz, FCz, or Cz ($ps > .247$). For accuracy, ΔERN amplitudes at FCz ($B = -1.75$, $p = .005$) and Cz ($B = -1.76$, $p = .008$) predicted lower accuracy, whereas Fz did not correlate with accuracy ($p = .537$). ΔPe amplitudes calculated from the subset dataset did not predict post-error slowing at either Pz or CPz ($ps > .369$). ΔPe at Pz was not associated with accuracy ($p = .135$), whereas ΔPe at CPz approached significance for a positive correlation with accuracy ($B = 0.99$, $p = .057$).

Discussion

Prior work suggests enhanced error-monitoring is a reliable neural indicator of anxiety (Weinberg, Dieterich, et al., 2015). The ERN and Pe are two error-related ERPs that have been used to study error-monitoring and its association with anxiety (Falkenstein et al., 1991; Holroyd & Coles, 2002; Olvet & Hajcak, 2009a; Weinberg et al., 2010). Different methodological decisions of quantifying the ERN and the Pe can affect their psychometric properties (e.g., Klawohn et al., 2020; Sandre et al., 2020). However, it is unclear how methodological factors affect the response-related ERPs and their associations with anxiety. The current study sought to fill this gap by comparing different methodological approaches, including the choice of anxiety scales (STAI and STICSA) as well as the trial number that went into the quantification of correct-related ERPs with a relatively large sample and robust statistics. Results of the current study revealed a reliable increase in amplitudes of error-related ERPs (i.e., ERN and Pe) relative to correct-related ERPs (i.e., CRN and Pc). Importantly, we observed a general lack of ERN amplitude enhancement with increased anxiety. These findings corroborate previous evidence for the lack of enhanced error-monitoring ERPs in those without an anxiety-related psychiatric diagnosis (Saunders & Inzlicht, 2020). Below we discuss key observations from the analyses and the potential implications for ERPs quantification and their correlations with anxiety trait measures.

The present results replicated the error enhancement of response-related ERPs (compared to correct responses) using the conventional as well as the mass-univariate analyses. Importantly, not only did mass-univariate analysis confirm the error enhancement with more robust statistics than the conventional approach, for both ERN and Pe components, but it also revealed a longer timing of significant difference and a wider topographic distribution of the response effect. Significant difference between the error- and correct-related ERPs extended beyond the typical 200-400 ms time window of the Pe, ending as late as 600 ms. The conventional and mass-univariate analyses also converged on the finding that the error enhancement was stronger and more reliable in the Pe timeframe than

the ERN timeframe, reflected in stronger statistical significance in the Pe-Pc comparison relative to the ERN-CRN comparison. The mass-univariate analyses suggested channels displaying maximal differences different from those previously investigated. In accordance with previous literature showing a front-central distribution of the ERN (Moser et al., 2012), the maximal difference between correct and error-related within the ERN timeframe was found at FC1 and FCz, for the full and subset datasets, respectively. Notably, the maximal difference within the Pe timeframe was observed at P4, which is not one of the typical recording sites for the Pe. These results suggest that previous classic ERP analyses might have missed the electrodes and timings that showed the maximal effect. Moreover, the location of maximal response effect may differ depending on various factors, such as the EEG cap positioning (Sandre et al., 2020) and the reference used. Researchers may also favour the use of single-site measures rather than pooled sites for better internal consistency (Klawohn et al., 2020; Sandre et al., 2020). In light of the numerous time windows and electrode sites available to researchers, data-driven approaches such as mass-univariate analyses, are a good alternative to the classic ERP analyses that offers the advantage of performing more thorough temporal and topographic comparisons (i.e., reducing Type II statistical errors) while maintaining control of α error.

The present study tested the impact of trial numbers on the comparisons between correct- and error-related ERPs. We found that the number of correct trials that went into the quantification of the CRN and the Pc did not affect the statistical significance of the error enhancement, whether the number of correct trials was more than six times greater than, or equal to, the number of error trials. This surprising finding supports previous evidence for moderate to high within-subject reliability of the ERN, seen fairly reliably with less than 15 trials (e.g., Baldwin et al., 2015; Foti et al., 2013; Meyer et al., 2013; Olvet & Hajcak, 2009). Furthermore, the response effect is unbiased by the number of trials across the conditions, being similarly robust in each analysis (with all correct trials or just with the same correct and incorrect trial numbers). However, we should note that the classic analyses of the current study used mean amplitudes measures, which are more robust to the noise level than peak

amplitude measures (Luck, 2014). It is possible that the comparisons of response-related ERPs using peak amplitude measures, which are less robust to the noise level (Luck, 2014), may return different results. In addition, although few studies have tested the methodological properties of the Pc and Pe, the present study suggests that their comparison can be made reliably with relatively low trial numbers. More studies are needed to examine to what extent trial number imbalance and ERP amplitude measures play a role in the comparison between response-related ERPs, and specifically between the Pe and the Pc.

We also examined the impact of analytic decisions in relation to the association between the difference ERP waveforms and continuous anxiety symptoms. When mass-univariate statistics were used, no correlation with anxiety was seen at any time point for any electrode, regardless of the analysis (full dataset, subset dataset with equal trial numbers). When ERPs were analyzed in the classic way, ΔPe did not correlate with trait anxiety at Pz and CPz, replicating previous findings (Endrass et al., 2008, 2010, 2014; Riesel et al., 2012; Weinberg et al., 2010). However, in contrast to previous evidence for a positive association between ΔERN and trait-level anxiety tendency (Hajcak et al., 2003; Meyer et al., 2012), the present results suggest a lack of such association in general. The results also suggest that different analysis methods (classic approach and data-driven mass-univariate statistics) and anxiety scales (STICSA vs. STAI) have negligible impact on the ERN association with anxiety. Further, controlling for depression did not contribute to the predictive significance of anxiety. The only significant anxiety correlation was found in the dataset with equal correct and error trials, in which higher STAI scores were associated with smaller ΔERN at Fz after depression was controlled for (classic analyses). However, this correlation runs counter to the expected direction, and it was not replicated with the full dataset or the STICSA scale, raising questions about the validity of this effect. Rather, the finding seems like an outlier result found at a specific electrode, suggesting that the examination of only one electrode site likely inflates false positive rate.

A lack of ERN enhancement with anxiety in sub-clinical populations has been reported in a typical binary choice task (e.g., Aarts & Pourtois, 2010; Beste et al., 2013;

Cavanagh & Allen, 2008; Hsieh et al., 2021) and a recent meta-analysis concluded that sub-clinical anxiety does not enhance ERN amplitude (Saunders & Inzlicht, 2020). These meta-analytic results suggested that the ERN enhancement is seen only after a particular anxiety threshold is reached, such that only clinically significant anxiety would enhance this component (Saunders & Inzlicht, 2020). This notion was not supported by the present findings. Specifically, more than half of the current sample met the criteria for clinical anxiety, as indicated by their anxiety scores, but anxiety severity did not correlate with the ΔERN or ΔPe, even when only highly anxious individuals were analyzed separately (yielding a sample size on par with previous clinical sample studies, Riesel, 2019; Weinberg et al., 2010; Weinberg, Kotov, et al., 2015). Despite being the main tool used by past studies to assess traits, the self-report measures may reflect only limited aspects of trait anxiety. Future studies may explore the use of alternative physiological measures, such as the startle reflex (Riesel et al., 2013), in addition to self-reported measures to enhance reliability.

Alternatively, enhanced ERN in clinically anxious samples may be the collective result from multiple anxiety-related symptomology, beyond the cognitive and somatic symptoms evaluated through self-report measures. For example, participants' conformity to their clinical diagnosis of anxiety disorders may lead them to exert more cognitive effort in performance monitoring tasks, contributing to their ΔERN-anxiety correlation. Previous studies demonstrating ERN enhancement in clinical anxiety recruited participants with a formal diagnosis (e.g., Riesel, 2019; Weinberg et al., 2010; Weinberg, Kotov, et al., 2015). However, the current study failed to replicate the ERN enhancement in people with clinical levels of anxiety without a formal diagnosis. Moving forward, it will be critical for future studies to consider diagnostic status in their investigations of the ERN-anxiety association and examine its potential impact on the association. Because close to half of the sample met the criteria for clinical anxiety based on their anxiety scores, we cannot claim that sub-clinical anxiety (i.e., below clinical threshold) did not moderate error-related ERPs. However, we did not have a limited range of anxiety scores in the sample, which would have

limited our power to detect effects (Weinberg, Liu, et al., 2016). Rather, we had a large range of anxiety scores, which should have worked in our favour if this association was real.

The inconsistency observed between the current and previous findings on ERN enhancement in clinical anxiety may be attributed, in part, to variations in sample size. The number of trials, sample sizes, and characteristics of ERP components, such as the signal-to-noise ratio, are important factors in determining the statistical power of ERP studies (Boudewyn et al., 2018; Thigpen et al., 2017). The current finding indicates that trial numbers had no consistent impact on the anxiety-ERN association, suggesting that trial count was not a limiting factor of statistical power. Furthermore, mass-univariate and classic analyses converged on a non-significant association between ERN and anxiety, indicating that statistical methods had minimal influence on the results. Therefore, sample size emerges as the primary determinant of statistical power when investigating the association between ERN and anxiety. This notion aligns with prior work that emphasizes the necessity to increase sample size in investigating the association between the ERN and individual differences in anxiety (Saunders & Inzlicht, 2020).

Anxiety and ERN did not correlate with post-error slowing or accuracy, which is in accordance with prior findings (Endrass et al., 2008; Moser et al., 2013; Weinberg et al., 2016; see LoTemplio et al., 2023 for a review). Accuracy was positively related to ΔPe at CPz although this association was only a trend and was not seen at Pz, supporting the view that these classic analyses likely yield Type I errors and are inconsistent across neighbouring electrodes and thus unlikely trustworthy. More studies are needed to investigate how specific facets of error-related components may correlate with performance.

Mass univariate analyses (MUA) take into account the inter-trial variance for each condition and each individual, which is a different way of analyzing ERP data compared to the conventional methods where an average waveform is computed across trials for each condition and individual. For this reason, it is not possible to generate a difference wave between conditions with MUA and thus one cannot correlate the difference wave with trait

measures as done with conventional analyses. Regardless, the ERN-anxiety correlation was not replicated in the current study with MUS for either correct- or error-related ERPs. Future studies should consider examining ERP components using robust statistics. It may also be worthwhile for researchers to explore other powerful statistical techniques, such as multivariate approaches, to examine the neural signals of anxiety.

The current study demonstrated a lack of anxiety-ERN correlation in a basic binary choice paradigm without laboratory manipulations, but more studies would be needed to elucidate how the environment may modulate the ERN and its association with anxiety. Prior work has shown that monetary loss (Hajcak et al., 2005) and punishment for incorrect responses (Riesel et al., 2012) could increase the ERN. The ERN is also sensitive to the interaction between specific aspects of anxiety and the environment. For example, participants high in fear of negative evaluation showed greater ERN under social observation compared to those low in this trait (Barker et al., 2015). Future studies may target specific facets of anxiety via experimental paradigms to provide a more nuanced understanding of error monitoring.

It should be noted that the current study used a cross-sectional design with a young adult sample. Therefore, we could not infer the direction of effects between anxiety traits and the enhancement of error monitoring. More longitudinal studies in adult samples would be needed to understand the direction of effects between cognitive traits of anxiety and error monitoring. Further, the current results may not be generalizable to younger populations. Prior findings have shown that stressful life experiences during childhood or adolescence, such as punitive parenting (Brooker & Buss, 2014; Chong & Meyer, 2019; Lackner et al., 2018; Mehra et al., 2022; Meyer et al., 2015), adverse childhood experiences (Lackner et al., 2018), interpersonal stressors (Mehra et al., 2022), and natural disasters (Meyer et al., 2017), are associated with enhanced anxiety symptoms and elevated ERN. Alterations in ERN amplitudes in childhood and adolescence may have persistent effects. Specifically, enhanced ERN predicted risk for future development of anxiety disorders (Meyer, 2016, 2017, 2022). These longitudinal correlations were based on classic analyses, and the application of robust

statistics is fairly limited in the developmental literature of the ERN. Future investigations of the ERN in pediatric populations might benefit from employing robust statistics to gain a more thorough understanding of developmental implications of the ERN.

To summarize, the ERN has been proposed as a potential biomarker for vulnerability to anxiety disorders (Weinberg, Meyer, et al., 2016). However, the strongest enhancement of the ERN has been observed when comparing clinical and neurotypical populations, where the studies in general had relatively small group sizes (e.g., n < 40; Riesel, 2019; Weinberg et al., 2010, 2015). Furthermore, there have been inconsistencies in the results concerning the correlation between ERN and anxiety in nonclinical samples. The present study, which used robust statistics and a large sample size, aligns with a recent meta-analysis, which suggests that the ERN does not exhibit heightened levels in individuals without anxiety psychopathology (Saunders & Inzlicht, 2020). The growing body of evidence challenges the sensitivity of ERN to capture variations in anxiety among those without clinical anxiety, questioning the ability of the ERN to identify vulnerable individuals. In light of these concerns, it is imperative for researchers to explore alternative neural markers that can provide a more nuanced and accurate understanding of the physiological underpinning of anxiety. For example, the feedback-related negativity (Aarts & Pourtois, 2012; Gu et al., 2010; Simons, 2010) and P3 (Righi et al., 2009; Sehlmeyer et al., 2010) are ERPs that have also shown modulations by anxiety. To ensure the validity and reliability of these findings, it is also crucial for future neuroimaging studies that investigate anxiety markers to utilize robust statistical methods, ultimately providing a comprehensive and precise understanding of anxiety vulnerability.

References

Aarts, K., & Pourtois, G. (2010). Anxiety not only increases, but also alters early error-monitoring

    functions. *Cognitive, Affective, & Behavioral Neuroscience*, *10*(4), 479–492.

    https://doi.org/10.3758/CABN.10.4.479

Aarts, K., & Pourtois, G. (2012). Anxiety disrupts the evaluative component of performance

    monitoring: An ERP study. *Neuropsychologia*, *50*(7), 1286–1296.

    https://doi.org/10.1016/j.neuropsychologia.2012.02.012

Aarts, K., Vanderhasselt, M.-A., Otte, G., Baeken, C., & Pourtois, G. (2013). Electrical brain imaging

    reveals the expression and timing of altered error monitoring functions in major depression.

    Journal of Abnormal Psychology, *122*(4), 939–950. https://doi.org/10.1037/a0034616

Baldwin, S. A., Larson, M. J., & Clayson, P. E. (2015). The dependability of electrophysiological

    measurements of performance monitoring in a clinical sample: A generalizability and

    decision analysis of the ERN and Pe: EEG dependability. *Psychophysiology*, *52*(6), 790–800.

    https://doi.org/10.1111/psyp.12401

Barker, T. V., Troller-Renfree, S., Pine, D. S., & Fox, N. A. (2015). Individual differences in social

    anxiety affect the salience of errors in social contexts. *Cognitive, Affective, & Behavioral*

    Neuroscience, *15*(4), 723–735. https://doi.org/10.3758/s13415-015-0360-9

Beste, C., Konrad, C., Uhlmann, C., Arolt, V., Zwanzger, P., & Domschke, K. (2013). Neuropeptide

    S receptor (NPSR1) gene variation modulates response inhibition and error monitoring.

    NeuroImage, *71*, 1–9. https://doi.org/10.1016/j.neuroimage.2013.01.004

Bieling, P. J., Antony, M. M., & Swinson, R. P. (1998). The State--Trait Anxiety Inventory, Trait

    version: Structure and content re-examined. *Behaviour Research and Therapy*, *36*(7–8), 777–

    788. https://doi.org/10.1016/S0005-7967(98)00023-0

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict

    monitoring and cognitive control. *Psychological Review*, *108*(3), 624–652.

    https://doi.org/10.1037/0033-295X.108.3.624

Boudewyn, M. A., Luck, S. J., Farrens, J. L., & Kappenman, E. S. (2018). How many trials does it

    take to get a significant ERP effect? It depends. *Psychophysiology*, *55*(6), e13049.

    https://doi.org/10.1111/psyp.13049

Brooker, R. J., & Buss, K. A. (2014). Harsh parenting and fearfulness in toddlerhood interact to

    predict amplitudes of preschool error-related negativity. *Developmental Cognitive*

    Neuroscience, *9*, 148–159. https://doi.org/10.1016/j.dcn.2014.03.001

Buzzell, G. A., Barker, T. V., Troller-Renfree, S. V., Bernat, E. M., Bowers, M. E., Morales, S.,

    Bowman, L. C., Henderson, H. A., Pine, D. S., & Fox, N. A. (2019). Adolescent cognitive

    control, theta oscillations, and social observation. *NeuroImage*, *198*, 13–30.

    https://doi.org/10.1016/j.neuroimage.2019.04.077

Caci, H., Baylé, F. J., Dossios, C., Robert, P., & Boyer, P. (2003). The Spielberger trait anxiety

    inventory measures more than anxiety. *European Psychiatry*, *18*(8), 394–400.

    https://doi.org/10.1016/j.eurpsy.2003.05.003

Cavanagh, J. F., & Allen, J. J. B. (2008). Multiple aspects of the stress response under social

    evaluative threat: An electrophysiological investigation. *Psychoneuroendocrinology*, *33*(1),

    41–53. https://doi.org/10.1016/j.psyneuen.2007.09.007

Chiu, P. H., & Deldin, P. J. (2007). Neural Evidence for Enhanced Error Detection in Major

    Depressive Disorder. *Am J Psychiatry*, *164*(4), 608–161.

    https://doi.org/10.1176/ajp.2007.164.4.608

Chong, L. J., & Meyer, A. (2019). Understanding the Link between Anxiety and a Neural Marker of

    Anxiety (The Error-Related Negativity) in 5 to 7 Year-Old Children. *Developmental*

    Neuropsychology, *44*(1), 71–87. https://doi.org/10.1080/87565641.2018.1528264

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial

    EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*,

    134(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Endrass, T., Franke, C., & Kathmann, N. (2005). Error Awareness in a Saccade Countermanding

    Task. *Journal of Psychophysiology*, *19*(4), 275–280. https://doi.org/10.1027/0269-

    8803.19.4.275

Endrass, T., Klawohn, J., Schuster, F., & Kathmann, N. (2008). Overactive performance monitoring

    in obsessive-compulsive disorder: ERP evidence from correct and erroneous reactions.

    Neuropsychologia, *46*(7), 1877–1887.

    https://doi.org/10.1016/j.neuropsychologia.2007.12.001

Endrass, T., Riesel, A., Kathmann, N., & Buhlmann, U. (2014). Performance monitoring in

    obsessive–compulsive disorder and social anxiety disorder. *Journal of Abnormal Psychology*,

    123(4), 705–714. https://doi.org/10.1037/abn0000012

Endrass, T., Schuermann, B., Kaufmann, C., Spielberg, R., Kniesche, R., & Kathmann, N. (2010).

    Performance monitoring and error significance in patients with obsessive-compulsive

    disorder. *Biological Psychology*, *84*(2), 257–263.

    https://doi.org/10.1016/j.biopsycho.2010.02.002

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target

    letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143–149.

    https://doi.org/10.3758/BF03203267

Falkenstein, M., Hohnsbein, J., Hoormann, J., & Blanke, L. (1991). Effects of crossmodal divided
    attention on late ERP components. II. Error processing in choice reaction tasks.
    Electroencephalography and Clinical Neurophysiology, *78*(6), 447–455.
    https://doi.org/10.1016/0013-4694(91)90062-9

Foti, D., Kotov, R., & Hajcak, G. (2013). Psychometric considerations in using error-related brain
    activity as a biomarker in psychotic disorders. *Journal of Abnormal Psychology*, *122*(2), 520–
    531. https://doi.org/10.1037/a0032618

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A Neural System for
    Error Detection and Compensation. *Psychological Science*, *4*(6), 385–390.
    https://doi.org/10.1111/j.1467-9280.1993.tb00586.x

Grös, D. F., Antony, M. M., Simms, L. J., & McCabe, R. E. (2007). Psychometric properties of the
    State-Trait Inventory for Cognitive and Somatic Anxiety (STICSA): Comparison to the State-
    Trait Anxiety Inventory (STAI). *Psychological Assessment*, *19*(4), 369–381.
    https://doi.org/10.1037/1040-3590.19.4.369

Gu, R., Huang, Y.-X., & Luo, Y.-J. (2010). Anxiety and feedback negativity. *Psychophysiology*.
    https://doi.org/10.1111/j.1469-8986.2010.00997.x

Hajcak, G., McDonald, N., & Simons, R. F. (2003). Anxiety and error-related brain activity.
    Biological Psychology, *64*(1–2), 77–90. https://doi.org/10.1016/S0301-0511(03)00103-0

Hajcak, G., Moser, J. S., Yeung, N., & Simons, R. F. (2005). On the ERN and the significance of
    errors. *Psychophysiology*, *42*(2), 151–160. https://doi.org/10.1111/j.1469-8986.2005.00270.x

Holmes, A. J., & Pizzagalli, D. A. (2008). Spatiotemporal Dynamics of Error Processing
    Dysfunctions in Major Depressive Disorder. *Archives of General Psychiatry*, *65*(2), 179.
    https://doi.org/10.1001/archgenpsychiatry.2007.19

Holmes, A. J., & Pizzagalli, D. A. (2010). Effects of task-relevant incentives on the electrophysiological correlates of error processing in major depressive disorder. *Cognitive, Affective, & Behavioral Neuroscience*, *10*(1), 119–128. https://doi.org/10.3758/CABN.10.1.119

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*(4), 679–709. https://doi.org/10.1037/0033-295X.109.4.679

Hsieh, M.-T., Lu, H., Lin, C.-I., Sun, T.-H., Chen, Y.-R., & Cheng, C.-H. (2021). Effects of Trait Anxiety on Error Processing and Post-error Adjustments: An Event-Related Potential Study With Stop-Signal Task. *Frontiers in Human Neuroscience*, *15*, 650838. https://doi.org/10.3389/fnhum.2021.650838

Klawohn, J., Meyer, A., Weinberg, A., & Hajcak, G. (2020). Methodological choices in event-related potential (ERP) research and their impact on internal consistency reliability and individual differences: An examination of the error-related negativity (ERN) and anxiety. *Journal of Abnormal Psychology*, *129*(1), 29–37. https://doi.org/10.1037/abn0000458

Klawohn, J., Riesel, A., Grützmann, R., Kathmann, N., & Endrass, T. (2014). Performance monitoring in obsessive–compulsive disorder: A temporo-spatial principal component analysis. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(3), 983–995. https://doi.org/10.3758/s13415-014-0248-0

Kujawa, A., Weinberg, A., Bunford, N., Fitzgerald, K. D., Hanna, G. L., Monk, C. S., Kennedy, A. E., Klumpp, H., Hajcak, G., & Phan, K. L. (2016). Error-related brain activity in youth and young adults before and after treatment for generalized or social anxiety disorder. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *71*, 162–168. https://doi.org/10.1016/j.pnbp.2016.07.010

Lackner, C. L., Santesso, D. L., Dywan, J., O'Leary, D. D., Wade, T. J., & Segalowitz, S. J. (2018). Adverse childhood experiences are associated with self-regulation and the magnitude of the error-related negativity difference. *Biological Psychology*, *132*, 244–251. https://doi.org/10.1016/j.biopsycho.2018.01.006

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00213

LoTemplio, S. B., Lopes, C. L., McDonnell, A. S., Scott, E. E., Payne, B. R., & Strayer, D. L. (2023). Updating the relationship of the Ne/ERN to task-related behavior: A brief review and suggestions for future research. *Frontiers in Human Neuroscience*, *17*, 1150244. https://doi.org/10.3389/fnhum.2023.1150244

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy*, *33*(3), 335–343. https://doi.org/10.1016/0005-7967(94)00075-U

Lovibond, S. H., & Lovibond, P. F. (1996). *Manual for the depression anxiety stress scales* (2nd ed.). Psychology Foundation of Australia.

Luck, S. J. (2014). *An introduction to the event-related potential technique* (Second edition). The MIT Press.

Luck, S. J., & Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't): How to get significant effects. *Psychophysiology*, *54*(1), 146–157. https://doi.org/10.1111/psyp.12639

Mehra, L. M., Hajcak, G., & Meyer, A. (2022). The relationship between stressful life events and the

    error-related negativity in children and adolescents. *Developmental Cognitive Neuroscience*,

    55, 101110. https://doi.org/10.1016/j.dcn.2022.101110

Meyer, A. (2016). Developing Psychiatric Biomarkers: A Review Focusing on the Error-Related

    Negativity as a Biomarker for Anxiety. *Current Treatment Options in Psychiatry*, *3*(4), 356–

    364. https://doi.org/10.1007/s40501-016-0094-5

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. https://doi.org/10.1016/j.dcn.2017.08.001

Meyer, A. (2022). On the relationship between the error-related negativity and anxiety in children and

    adolescents: From a neural marker to a novel target for intervention. *Psychophysiology*,

    59(6), e14050. https://doi.org/10.1111/psyp.14050

Meyer, A., Carlton, C., Crisler, S., & Kallen, A. (2018). The development of the error-related

    negativity in large sample of adolescent females: Associations with anxiety symptoms.

    Biological Psychology, *138*, 96–103. https://doi.org/10.1016/j.biopsycho.2018.09.003

Meyer, A., Danielson, C. K., Danzig, A. P., Bhatia, V., Black, S. R., Bromet, E., Carlson, G., Hajcak,

    G., Kotov, R., & Klein, D. N. (2017). Neural Biomarker and Early Temperament Predict

    Increased Internalizing Symptoms After a Natural Disaster. *Journal of the American*

    Academy of Child & Adolescent Psychiatry, *56*(5), 410–416.

    https://doi.org/10.1016/j.jaac.2017.02.005

Meyer, A., & Gawlowska, M. (2017). Evidence for specificity of the impact of punishment on error-

    related brain activity in high versus low trait anxious individuals. *International Journal of*

    Psychophysiology, *120*, 157–163. https://doi.org/10.1016/j.ijpsycho.2017.08.001

Meyer, A., Hajcak, G., Torpey-Newman, D. C., Kujawa, A., & Klein, D. N. (2015). Enhanced error-related brain activity in children predicts the onset of anxiety disorders between the ages of 6 and 9. *Journal of Abnormal Psychology*, *124*(2), 266–274. https://doi.org/10.1037/abn0000044

Meyer, A., Riesel, A., & Proudfit, G. H. (2013). Reliability of the ERN across multiple tasks as a function of increasing errors. *Psychophysiology*, *50*(12), 1220–1225. https://doi.org/10.1111/psyp.12132

Meyer, A., Weinberg, A., Klein, D. N., & Hajcak, G. (2012). The development of the error-related negativity (ERN) and its relationship with anxiety: Evidence from 8 to 13 year-olds. *Developmental Cognitive Neuroscience*, *2*(1), 152–161. https://doi.org/10.1016/j.dcn.2011.09.005

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*(1), 3–10. https://doi.org/10.1111/j.1469-8986.2011.01279.x

Moser, J. S., Moran, T. P., Schroder, H. S., Donnellan, M. B., & Yeung, N. (2013). On the relationship between anxiety and error monitoring: A meta-analysis and conceptual framework. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00466

Muir, A. M., Hedges-Muncy, A., Clawson, A., Carbine, K. A., & Larson, M. J. (2020). Dimensions of anxiety and depression and neurophysiological indicators of error-monitoring: Relationship with delta and theta oscillatory power and error-related negativity amplitude. *Psychophysiology*, *57*(9), 603–624. https://doi.org/10.1111/psyp.13595

Nieuwenhuis, S., Ridderinkhof, K. R., Blom, J., Band, G. P. H., & Kok, A. (2001). Error-related brain potentials are differentially related to awareness of response errors: Evidence from an

antisaccade task. *Psychophysiology*, *38*(5), 752–760. https://doi.org/10.1111/1469-8986.3850752

Olvet, D. M., & Hajcak, G. (2009a). The effect of trial-to-trial feedback on the error-related negativity and its relationship with anxiety. *Cognitive, Affective, & Behavioral Neuroscience*, *9*(4), 427–433. https://doi.org/10.3758/CABN.9.4.427

Olvet, D. M., & Hajcak, G. (2009b). The stability of error-related brain activity with increasing trials. *Psychophysiology*, *46*(5), 957–961. https://doi.org/10.1111/j.1469-8986.2009.00848.x

Olvet, D. M., Klein, D. N., & Hajcak, G. (2010). Depression symptom severity and error-related brain activity. *Psychiatry Research*, *179*(1), 30–37. https://doi.org/10.1016/j.psychres.2010.06.008

Pasion, R., & Barbosa, F. (2019). ERN as a transdiagnostic marker of the internalizing-externalizing spectrum: A dissociable meta-analytic effect. *Neuroscience & Biobehavioral Reviews*, *103*, 133–149. https://doi.org/10.1016/j.neubiorev.2019.06.013

Pernet, C. R., Chauveau, N., Gaspar, C., & Rousselet, G. A. (2011). LIMO EEG: A Toolbox for Hierarchical LInear MOdeling of ElectroEncephaloGraphic Data. *Computational Intelligence and Neuroscience*, *2011*, 1–11. https://doi.org/10.1155/2011/831409

Pernet, C. R., Latinus, M., Nichols, T. E., & Rousselet, G. A. (2015). Cluster-based computational methods for mass univariate analyses of event-related brain potentials/fields: A simulation study. *Journal of Neuroscience Methods*, *250*, 85–93. https://doi.org/10.1016/j.jneumeth.2014.08.003

Rabbitt, P. M. (1966). Errors and error correction in choice-response tasks. *Journal of Experimental Psychology*, *71*(2), 264–272. https://doi.org/10.1037/h0022853

Rabinak, C. A., Holman, A., Angstadt, M., Kennedy, A. E., Hajcak, G., & Phan, K. L. (2013). Neural response to errors in combat-exposed returning veterans with and without post-traumatic

stress disorder: A preliminary event-related potential study. *Psychiatry Research: Neuroimaging*, *213*(1), 71–78. https://doi.org/10.1016/j.pscychresns.2013.01.002

Ree, M. J., French, D., MacLeod, C., & Locke, V. (2008). Distinguishing Cognitive and Somatic Dimensions of State and Trait Anxiety: Development and Validation of the State-Trait Inventory for Cognitive and Somatic Anxiety (STICSA). *Behavioural and Cognitive Psychotherapy*, *36*(03). https://doi.org/10.1017/S1352465808004232

Riesel, A. (2019). The erring brain: Error-related negativity as an endophenotype for OCD-A review and meta-analysis. *Psychophysiology*, *56*(4), e13348. https://doi.org/10.1111/psyp.13348

Riesel, A., Goldhahn, S., & Kathmann, N. (2017). Hyperactive performance monitoring as a transdiagnostic marker: Results from health anxiety in comparison to obsessive–compulsive disorder. *Neuropsychologia*, *96*, 1–8. https://doi.org/10.1016/j.neuropsychologia.2016.12.029

Riesel, A., Weinberg, A., Endrass, T., Kathmann, N., & Hajcak, G. (2012). Punishment has a lasting impact on error-related brain activity: Punishment modulates error monitoring. *Psychophysiology*, *49*(2), 239–247. https://doi.org/10.1111/j.1469-8986.2011.01298.x

Riesel, A., Weinberg, A., Moran, T., & Hajcak, G. (2013). Time Course of Error-Potentiated Startle and its Relationship to Error-Related Brain Activity. *Journal of Psychophysiology*, *27*(2), 51–59. https://doi.org/10.1027/0269-8803/a000093

Righi, S., Mecacci, L., & Viggiano, M. P. (2009). Anxiety, cognitive self-evaluation and performance: ERP correlates. *Journal of Anxiety Disorders*, *23*(8), 1132–1138. https://doi.org/10.1016/j.janxdis.2009.07.018

Sandre, A., Banica, I., Riesel, A., Flake, J., Klawohn, J., & Weinberg, A. (2020). Comparing the effects of different methodological decisions on the error-related negativity and its association with behaviour and gender. *International Journal of Psychophysiology*, *156*, 18–39. https://doi.org/10.1016/j.ijpsycho.2020.06.016

Saunders, B., & Inzlicht, M. (2020). Assessing and adjusting for publication bias in the relationship

    between anxiety and the error-related negativity. *International Journal of Psychophysiology*,

    155, 87–98. https://doi.org/10.1016/j.ijpsycho.2020.05.008

Saunders, B., Milyavskaya, M., & Inzlicht, M. (2015). What does cognitive control feel like?

    Effective and ineffective cognitive control is associated with divergent phenomenology:

    What does control feel like? *Psychophysiology*, *52*(9), 1205–1217.

    https://doi.org/10.1111/psyp.12454

Schrijvers, D., De Bruijn, E. R. A., Maas, Y., De Grave, C., Sabbe, B. G. C., & Hulstijn, W. (2008).

    Action monitoring in major depressive disorder with psychomotor retardation. *Cortex*, *44*(5),

    569–579. https://doi.org/10.1016/j.cortex.2007.08.014

Sehlmeyer, C., Konrad, C., Zwitserlood, P., Arolt, V., Falkenstein, M., & Beste, C. (2010). ERP

    indices for response inhibition are related to anxiety-related personality traits.

    Neuropsychologia, *48*(9), 2488–2495.

    https://doi.org/10.1016/j.neuropsychologia.2010.04.022

Simons, R. F. (2010). The way of our errors: Theme and variations. *Psychophysiology*, *47*(1), 1–14.

    https://doi.org/10.1111/j.1469-8986.2009.00929.x

Spielberger, C. D. (1983). *Manual for the State-Trait Anxiety Inventory*. Mind Garden.

Thigpen, N. N., Kappenman, E. S., & Keil, A. (2017). Assessing the internal consistency of the event-

    related potential: An example analysis. *Psychophysiology*, *54*(1), 123–138.

    https://doi.org/10.1111/psyp.12629

Van Dam, N. T., Gros, D. F., Earleywine, M., & Antony, M. M. (2013). Establishing a trait anxiety

    threshold that signals likelihood of anxiety disorders. *Anxiety, Stress & Coping*, *26*(1), 70–86.

    https://doi.org/10.1080/10615806.2011.631525

Vidal, F., Burle, B., Bonnet, M., Grapperon, J., & Hasbroucq, T. (2003). Error negativity on correct

trials: A reexamination of available data. *Biological Psychology*, *64*(3), 265–282.

https://doi.org/10.1016/S0301-0511(03)00097-8

Vidal, F., Hasbroucq, T., Grapperon, J., & Bonnet, M. (2000). Is the 'error negativity' specific to

errors? *Biological Psychology*, *51*(2–3), 109–128. https://doi.org/10.1016/S0301-

0511(99)00032-0

Weinberg, A., Dieterich, R., & Riesel, A. (2015). Error-related brain activity in the age of RDoC: A

review of the literature. *International Journal of Psychophysiology*, *98*(2), 276–299.

https://doi.org/10.1016/j.ijpsycho.2015.02.029

Weinberg, A., Klein, D. N., & Hajcak, G. (2012). Increased error-related brain activity distinguishes

generalized anxiety disorder with and without comorbid major depressive disorder. *Journal

of Abnormal Psychology*, *121*(4), 885–896. https://doi.org/10.1037/a0028270

Weinberg, A., Kotov, R., & Proudfit, G. H. (2015). Neural indicators of error processing in

generalized anxiety disorder, obsessive-compulsive disorder, and major depressive disorder.

Journal of Abnormal Psychology, *124*(1), 172–185. https://doi.org/10.1037/abn0000019

Weinberg, A., Liu, H., & Shankman, S. A. (2016a). Blunted neural response to errors as a trait marker

of melancholic depression. *Biological Psychology*, *113*, 100–107.

https://doi.org/10.1016/j.biopsycho.2015.11.012

Weinberg, A., Liu, H., & Shankman, S. A. (2016b). Blunted neural response to errors as a trait

marker of melancholic depression. *Biological Psychology*, *113*, 100–107.

https://doi.org/10.1016/j.biopsycho.2015.11.012

Weinberg, A., Meyer, A., Hale-Rude, E., Perlman, G., Kotov, R., Klein, D. N., & Hajcak, G. (2016).

Error-related negativity (ERN) and sustained threat: Conceptual framework and empirical

evaluation in an adolescent sample: ERN and sustained threat. *Psychophysiology*, *53*(3), 372–

385. https://doi.org/10.1111/psyp.12538

Weinberg, A., Olvet, D. M., & Hajcak, G. (2010). Increased error-related brain activity in generalized

anxiety disorder. *Biological Psychology*, *85*(3), 472–480.

https://doi.org/10.1016/j.biopsycho.2010.09.011

Xiao, Z., Wang, J., Zhang, M., Li, H., Tang, Y., Wang, Y., Fan, Q., & Fromson, J. A. (2011). Error-

related negativity abnormalities in generalized anxiety disorder and obsessive–compulsive

disorder. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *35*(1), 265–

272. https://doi.org/10.1016/j.pnpbp.2010.11.022

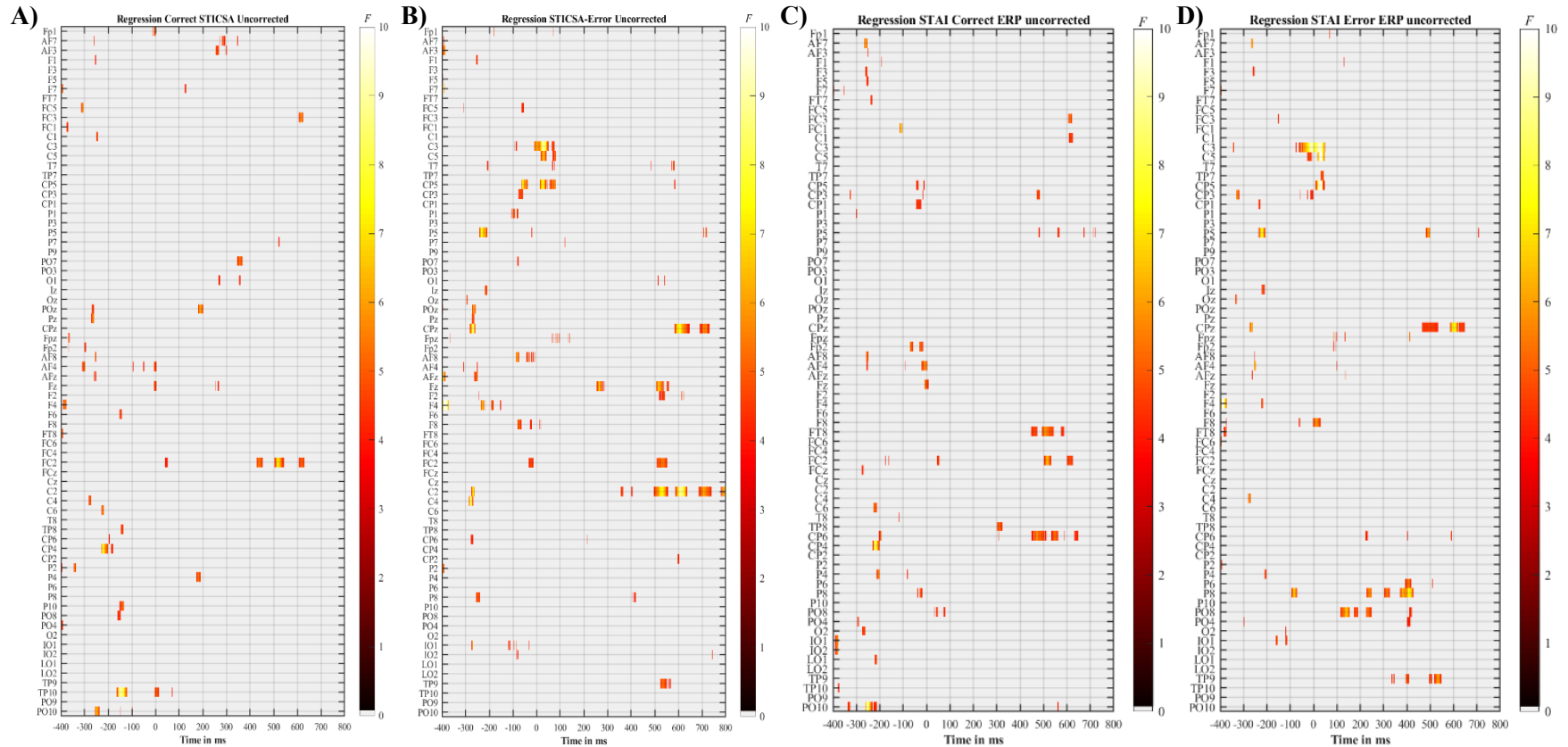# Regressions in LIMO for Anxiety Scores Predicting Response-ERPs

Figure A1. Uncorrected F-values of R^2 for the LIMO regression models between anxiety scores and ERPs with an alpha of 0.05, across the entire -400 to 800 ms epoch, on the full dataset. A) STICSA regression with correct-related ERP, B) STICSA regression with error-related ERP, C) STAI regression with correct-related ERP, and D) STAI regression with error-related ERP.
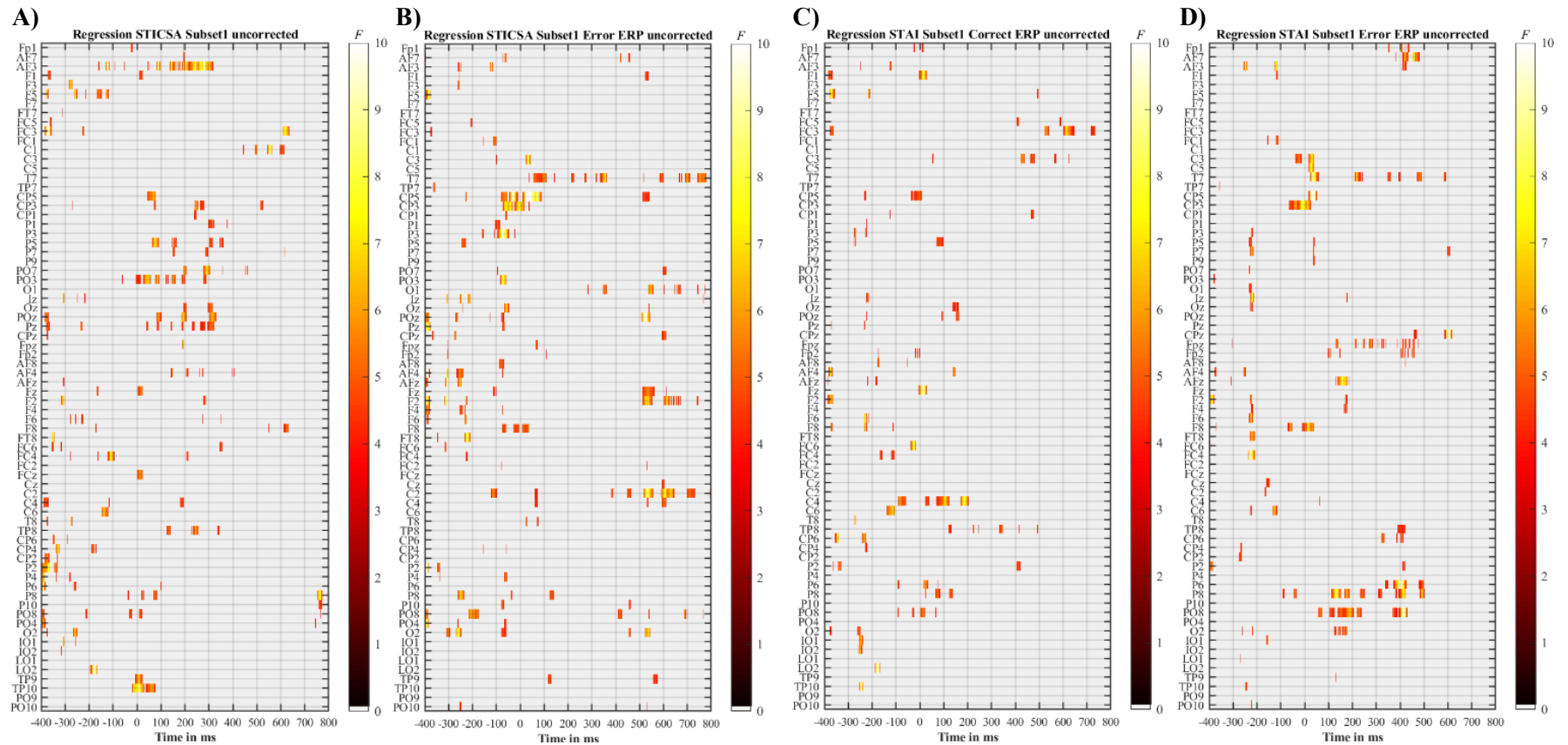
Figure A2. Uncorrected F-values of R^2 for the LIMO regression model between STICSA scores and ERPs with an alpha of 0.05, across the entire -400 to 800 ms epoch, on a dataset that contains equal numbers of correct and error trials for each participant. A) STICSA regression with correct-related ERP, B) STICSA regression with error-related ERP, C) STAI regression with correct-related ERP, and D) STAI regression with error-related ERP.