Robust single-trial event-related potentials differentiate between Distress and Fear disorders

Martin Randau[1], Nina Reinholt[1], Cyril Pernet[1], Bob Oranje[1], Belinda Rasmussen[1], and Sidse Arnfred[1]

[1]Copenhagen University Hospital

July 21, 2023

Abstract

Recent evidence indicates that measures of brain functioning as indexed by event-related potentials (ERP) on the electroencephalogram aligns more closely to transdiagnostic measures of psychopathology than to categorical taxonomies. The Hierarchical Taxonomy of Psychopathology (HiTOP) is a transdiagnostic, dimensional framework aiming to solve issues of comorbidity, symptom heterogeneity and arbitrary diagnostic boundaries. Based on shared features, the emotional disorders are allocated into subfactors Distress and Fear. Evidence indicate that disorders which are close in the HiTOP hierarchy share etiology, symptom profiles and treatment outcome. However, further studies testing the biological underpinnings of the HiTOP are called for. In this study, we assessed differences between Distress and Fear in a range of well-studied ERP components. Fifty-one patients with emotional disorders were divided into two groups (Distress, N = 26; Fear, N = 25) according to HiTOP criteria and compared against 37 healthy comparison subjects (HC). Addressing issues in traditional ERP preprocessing and analysis methods, we applied robust single-trial analysis as implemented in the EEGLAB toolbox LIMO EEG. Several ERP components were found to differ between the groups. Surprisingly, we found no difference between Fear and HC for any of the ERPs. This suggests that some well-established results from the literature, e.g., increased error-related negativity in OCD, is not a shared neurobiological correlate of the Fear subfactor. Conversely, for Distress, we found reductions compared to Fear and HC in several ERP components across paradigms. Future studies could utilize HiTOP-validated psychopathology measures to more precisely define subfactor groups.

[1] Research Unit for Psychotherapy & Psychopathology, Mental Health Service West, Copenhagen University Hospital – Psychiatry Region Zealand, Slagelse, Denmark[2] Psychiatric Research Unit, Copenhagen University Hospital – Psychiatry Region Zealand, Slagelse, Denmark[3] Deptartment of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark[4] Neurobiology Research Unit, Copenhagen University Hospital, Copenhagen, Denmark[5] Center for Neuropsychiatric Schizophrenia Research (CNSR), Center for Clinical Intervention and Neuropsychiatric Schizophrenia Research (CINS), Copenhagen University Hospital, Copenhagen, Denmark

Correspondence: Martin Randau <>

Introduction

The Hierarchical Taxonomy of Psychopathology (HiTOP) is a transdiagnostic, dimensional framework aiming to cover the full range of psychopathology, from healthy to ill, from symptom to spectrum (Kotov et al., 2017). Testing of this new proposal for neurobiological support is called for and would increase our understanding of the underlying mechanisms of mental disorders (DeYoung et al., 2023). Recent evidence

suggests that measures of brain functioning as indexed by event-related potentials (ERPs) on the electroencephalogram (EEG) aligns more closely to transdiagnostic measures of psychopathology than to categorical taxonomies such as ICD-10 and DSM-5, which impose arbitrary diagnostic thresholds, comorbidity and symptom heterogeneity within disorders (Donaldson et al., 2020; Granros, 2021; Macedo et al., 2021; Riesel et al., 2022). However, issues inherent in these traditional taxonomies as well as in ERP preprocessing and analysis methods contribute to inconsistent results (Feuerriegel & Bode, 2022). In this study, we apply novel robust single-trial ERP analysis methods to a range of well-studied ERPs in order to test the neurobiological underpinnings of the HiTOP subfactors 'Distress' and 'Fear' covering the emotional disorders including anxiety disorders, obsessive-compulsive disorder (OCD) and depression.

Established categorical taxonomies ICD-10 and DSM-5 posit that mental disorders are discrete entities with specific and unique symptoms and clear-cut boundaries between the mentally healthy and ill (DSM-IV-TR., 2000; WHO, 2004). While constituting the foundation of psychiatry, these taxonomies come with several problems for clinical practice and research (Clark et al., 2017). First, categorical taxonomies introduce comorbidity, the presence of more than one diagnosis in the same patient (González-Robles et al., 2018). For example, among patients with an anxiety disorder, 63% have a current and 81% a lifetime depressive disorder (Lamers et al., 2011). Second, symptom heterogeneity within disorders means that patients with the same diagnosis often present with vastly different symptom profiles or might have no overlapping symptoms at all (Clark et al., 2017; Kotov et al., 2017). Lastly, arbitrary boundaries between diagnoses and between healthy and ill pertains to loss-of-information in that the full range of psychopathology goes unaccounted for. Taken together, these issues limit the clinical and research usability of the categorical taxonomies (Perkins et al., 2020).

The HiTOP was developed to resolve the issues of categorical taxonomies and to facilitate basic research (Conway et al., 2022). The HiTOP is an evidence-based, data-driven, and atheoretical proposal for a taxonomy of psychopathology based solely on individual symptoms and maladaptive traits as the basic building blocks of mental disorders (Kotov et al., 2021; Lahey et al., 2021; Pianowski et al., 2019). Clustering of these building blocks according to observed covariation results in a hierarchical structure with at least four levels (Kotov et al., 2021; Ringwald et al., 2021). The lowest level consists of individual symptoms, such as anxious apprehension and compulsive thoughts, and maladaptive personality traits, such as negative affect and neuroticism. Correlated symptoms and traits are clustered into syndromes, roughly corresponding to diagnoses as we know them. However, HiTOP syndromes are dimensions, not categories, and do not necessarily map directly unto specific diagnoses. At the next level, syndromes are clustered into subfactors, examples of which are *Fear* and *Distress*containing between them all of the emotional disorders. As a result, syndromes corresponding to the diagnoses of OCD and social anxiety disorder (SAD), located in the Fear subfactor, have more symptoms and traits in common with each other than with either generalized anxiety disorder (GAD) or major depressive disorder (MDD), which are located in the Distress subfactor (Watson et al., 2022). At the next level, subfactors are clustered into one of six Spectra which represent the broadest possible division of mental disorders. Examples are the*Internalizing* , *Externalizing* and *Thought disorder*spectra. Recently, an all-encompassing general p-factor has been added at the top of the hierarchy. This level accounts for the positive correlation among all psychiatric symptoms, in essence signifying that a patient is more likely than a healthy comparison subject to experience any psychiatric symptom (Carragher et al., 2021; Forbes, 2020; Levin-Aspenson et al., 2021).

Accordingly, in the HiTOP, heterogeneity within disorders is modeled by clustering related symptoms together and unrelated symptoms elsewhere in the hierarchy. Comorbidity is accounted for in that higher levels in the hierarchy denote correlations among lower levels. For example, comorbidity within the emotional disorders is modeled by their common placement in the Internalizing spectrum.

In research, the HiTOP provides a refined perspective of the interface between neurobiology and psychopathology by allowing for testing at several levels of the hierarchy (Conway et al., 2019; Michelini et al., 2021; Perkins et al., 2019). Researchers can investigate whether a given ERP measure is a marker of an individual symptom or trait, of a subfactor or of an entire spectrum. Compelling evidence indicates that

closeness in the hierarchy implies similarities in etiology and genetic underpinnings as well as in treatment response and risk factors (Kotov et al., 2022; Waszczuk et al., 2020). While such correlations with biological measures have primarily been explored at the spectrum level, studies extending the scope to lower levels of the hierarchy are called for (Conway et al., 2022; Perkins et al., 2019). Two recent major reviews on potential biological markers of the HiTOP Internalizing and Thought disorder spectra indicate that EEG measures are implicated at several levels of the hierarchy, including spectrum, subfactor and syndrome levels (Kotov et al., 2020; Watson et al., 2022).

However, multiple issues complicate the use of past studies to validate the HiTOP framework. First, results are necessarily extrapolated from studies conducted on categorical taxonomies affected by aforementioned issues. Second, the field of clinical EEG, like other areas of neuroscience, is burdened by contradictory results and publication bias (Clayson et al., 2013, 2019, 2021; Clayson, 2020; Fields & Kuperberg, 2020; Klawohn, Meyer, et al., 2020; Saunders & Inzlicht, 2020). At least part of these inconsistencies stem from technicalities such as methods of baseline subtraction, hi-pass filtering and ERP amplitude extraction, which differ between studies (Alday, 2019; Feuerriegel & Bode, 2022; Luck & Gaspelin, 2017; Sandre et al., 2020; Šoškić et al., 2022; Tanner et al., 2016).

Single-trial analysis of EEG data as implemented in the EEGLAB toolbox LIMO EEG is a richer and more robust alternative to traditional methods and at least partly remedies preprocessing and amplitude extraction issues (Fields & Kuperberg, 2020); Delorme & Makeig (2004); C. R. Pernet et al. (2011); C. R. Pernet et al. (2021)]. LIMO EEG is a bootstrap-based approach to mass univariate analysis of single-trial EEG data and builds on statistical parametric mapping (SPM) widely used in the analysis of fMRI data (Kiebel & Friston, 2004, 2004). In contrast to traditional methods which require the *á priori* selection of channels and time windows to analyze, LIMO EEG analyzes all channels and time samples concurrently with equal or even greater statistical power (Fields & Kuperberg, 2020). In the method, the epoched or continuous EEG data is modeled in a two-level hierarchical generalized linear model (GLM). The first-level, the subject-level, models within-subject trial-to-trial variability due to noise and response variability. The second level, group-level, integrates subject-level estimates and models between-subject variability, e.g., differences between conditions or groups. A wealth of robust statistical tests can hereby be conducted, e.g., t-tests, ANOVA/ANCOVA and regression analyses with behavioral measures. False positives from multiple testing of $channels * timeframes$ is controlled with established and validated methods, e.g., maximum statistics, spatiotemporal clustering and threshold-free cluster enhancement (Maris & Oostenveld, 2007; C. R. Pernet et al., 2011; C. R. Pernet, 2015; Smith & Nichols, 2009). Another advantage of robust single-trial analysis concerns the preprocessing of ERP data, which amounts to removing as much artifacts and noise as possible without affecting the true brain signal. However, with no ground truth available, the effects of arbitrary preprocessing choices is often unknown (Feuerriegel & Bode, 2022; Robbins et al., 2020). It has been shown that robust single-trial analysis methods are less affected by preprocessing parameters and handles noisy data better than traditional methods, allowing for less strict cleaning and thereby more preserved brain activity (Alday & van Paridon, 2021; Bailey et al., 2022). Crucially, while single-trial analysis involves changes in in ERP methods, the units of analysis, i.e., peak amplitude and latency, remain unchanged, allowing for comparison with the literature.

In this study, we applied robust single-trial analysis to a range of well-studied ERPs with the aim to test for the biological underpinnings of the HiTOP classification of the emotional disorders into subfactors Distress and Fear under the Internalizing spectrum. Based on HiTOP criteria in terms of primary categorical ICD-10 diagnosis, 51 patients with emotional disorders about to start psychotherapy were divided into either a Distress or Fear subfactor group. These two groups were then compared against each other and against a group of 37 healthy comparison subjects (HC). All subjects were evaluated with three classic ERP paradigms eliciting a range of ERPs, including the error-related negativity (ERN), error positivity (Pe), correct-related negativity (CRN), correct positivity (Pc), P2, P3a, P3b, N1, two versions of the N2, as well as three versions of auditory mismatch negativity (MMN) and the corresponding difference wave P3a, dP3a (Luck & Kappenman, 2011).

To elucidate which ERPs, if any, align with the HiTOP would contribute to a greater understanding of the

neurobiological underpinnings of the mental disorders. While we did not expect all ERPs to align consistently with levels of the HiTOP hierarchy, we hypothesized that some ERPs can differentiate between the three groups. Specifically, we hypothesized that ERN is enhanced (a more negative amplitude) for the Fear group and reduced for the Distress group compared to each other and to the HC group (Macedo et al., 2021). P3b was likewise hypothesized to be enhanced for the Fear group and reduced for the Distress group compared to each other and to the HC group (Botelho et al., 2023; Gohle et al., 2008; Klawohn, Santopetro, et al., 2020). We also hypothesized that some ERPs might be abnormal in both subfactor groups compared to HC, indicating abnormalities in brain functioning at the spectrum or p-factor level.

Results will be presented at established regions of interest even though the analysis method evaluates the whole ERP time course at all channels simultaneously. Our study might therefore reveal significant group effects at time intervals and channels not previously described in the literature (Fields & Kuperberg, 2020).

Methods

Participants

Patients (N=51) of both sexes aged 18 to 59 with a primary ICD-10 diagnosis of either agoraphobia, depression, generalized anxiety disorder (GAD), obsessive-compulsive disorder (OCD), social anxiety disorder (SAD) or panic disorder (PD), with or without comorbidities including another emotional disorder, attention-deficit hyperactivity disorder and personality disorder about to start Unified Protocol (UP) transdiagnostic group cognitive behavior therapy were recruited from three tertiary free-of-charge public Mental Health Service outpatient clinics in Denmark, as described in detail in (Reinholt et al., 2021). Patients were referred to the clinic after two previous failed treatment attempts in primary care.

Exclusion followed the exclusion criteria for receiving treatment in the participating outpatient clinics: an ICD-10 F20 diagnosis, bipolar disorder or autism, alcohol- or substance use disorder, increased risk of suicide, recent ($<$ 4 weeks) onset or alteration of psychotropic medication, previous traumatic brain injury or organic brain disorder as assessed by journal look-up, and normal mental capabilities as estimated by having completed primary school. Healthy comparison subjects (HC, N=37), matched with the patient group as a whole in age and sex, were recruited from the local community through posters and online advertisement. Exclusion criteria were the same as for patients but also included no prior or present psychiatric diagnosis or psychotropic medication. All participants had normal or corrected-to-normal eye vision.

Based on a recent meta-analysis of the structure of the HiTOP hierarchy, patients were allocated according to primary diagnosis into either the Distress (depression and GAD) or the Fear (agoraphobia, OCD, PD and SAD) group (Ringwald et al., 2021).

Measures

Current medication status was extracted from the electronic health record including type of medication, dosage, treatment duration and changes hereof. Information on handedness and hearing status (normal/impaired) was interview-based. All participants were assessed with the the Mini-International Neuropsychiatric Interview version 7 (M.I.N.I.) diagnostic interview psychiatry specialist-trainee (MR) (Sheehan et al., 1998). For patients, a primary diagnosis within the emotional disorder spectrum was confirmed, and up to three concurrent secondary diagnoses were noted. Healthy comparison subjects were likewise screened for the absence of symptoms fulfilling criteria for any psychiatric diagnosis.

Patients were additionally administered self-report questionnaires relating to their primary diagnosis. Accordingly, patients with a primary ICD-10 diagnosis of agoraphobia were administered the Mobile Inventory (MI) (Chambless et al., 1985), depression the Patient Health Questionnaire (PHQ-9) (Kroenke & Spitzer, 2002), GAD the GAD-7 (Spitzer et al., 2006), OCD the Brief Obsessive-Compulsive Scale (BOCS) (Bejerot

et al., 2014), PD the self-report version of the Panic Disorder Severity Scale (PDSS) (Houck et al., 2002), and SAD the self-report version of the Liebowitz Social Anxiety Scale (LSAS) (Rytwinski et al., 2009). Questionnaire data were obtained using the online survey and database management web application REDCap licensed to Region Zealand, Denmark (Harris et al., 2019). Patients were instructed to answer +/- 1 week from the EEG recording.

Procedures

EEG laboratory setup

EEG was recorded at two psychiatric hospitals in Region Zealand, Denmark. Each session took place either in the morning or early afternoon and lasted approximately three hours including information, electrode and cap application, EEG recording (~1 hour), M.I.N.I diagnostic interview and breaks. Participants were instructed to show up rested and to avoid coffee and nicotine intake 2 hours before. Patients were also instructed to avoid, if possible, medication prescribed "as needed" on the night before and day of recording. During EEG recording, participants were seated in a comfortable armchair in a secluded room and instructed to sit as still as possible. Visual stimuli were presented on a 17" LCD monitor situated 1.5 meters from the participant. Audio stimuli were presented with airtube stereo insert earphones (C and H Distributors Inc., Milwaukee, WI, USA, 2021). Similar room luminosity at the two sites was ensured with blackout curtains but was not objectively measured.

EEG recording

EEG at the two sites was recorded with identical Biosemi ActiveTwo Mark 2 systems with 64 Ag/AgCl pin-type active electrodes attached to a cap according to the 10/20 system (BioSemi, Amsterdam, 2021). The signal was recorded reference-free with common mode sense (CMS) and driven right leg (DRL) electrodes as "ground" placed centrally close to POz. The signal was digitized at 24-bit resolution with a sampling rate of 2048 Hz. Electrode offset was kept below 40 μV.

EEG paradigms

All paradigms were presented using Presentation® software version 23.0 (Neurobehavioral Systems, 2021). The paradigms, presented in this order for all participants, were:

Attended oddball (AO)

Auditory stimuli delivered binaurally in a pseudorandom order: 10% target tones (1100 Hz, 50 ms duration), 6% distractor tones (a 50 ms bell sound) and 84% standard tones (1000 Hz, 50 ms). 50 dB sound intensity and 10 ms rise/fall for all stimuli. Participants were instructed to fixate on a white cross on a black background on the monitor and to press the left mouse button with their index finger when hearing the target tone while ignoring distractor stimuli. Participants started with a 30 stimuli test round.

Flanker

The flanker task was a modified version of the Eriksen Flanker (Eriksen & Eriksen, 1974) commonly used in the literature, e.g., (Riesel et al., 2022; Seow et al., 2020). Five horizontal arrows were presented in white on a black background on the monitor. Trials could be either congruent (<<<<< or >>>>>) or incongruent (<<><< or >><>>) and were presented for 200 ms. Trials were 50% congruent and 50% incongruent trials presented in random order. Participants were instructed to respond as quickly and accurately as possible by pressing either the left or right mouse button indicating the direction of the central arrow. Participants had 1050 ms to respond. Feedback was delivered on the monitor at the end of each block: if >90% correct responses or <25% missed trials ("Try to respond faster!") and if accuracy <75% ("Try to respond more accurately". Otherwise feedback was "Good job!". If participants had less than 17 errors in total, up to two extra blocks were administered in order to ensure internal consistency of the ERN (Clayson, 2020). Participants at first completed a test round consisting of 12 trials to ensure instructions were understood.

Unattended oddball (UO)

Participants watched a muted nature documentary while auditory stimuli were delivered binaurally at 50 dB in a pseudorandom order: 6% frequency deviant tones (1100 Hz, 50 ms duration), 6% duration deviant tones (1000 Hz, 100 ms duration), 6% combined frequency and duration deviant tones (1100 Hz, 100 ms duration) and 82% standard tones (1000 Hz, 50 ms). 50 dB sound intensity and 10 ms rise/fall for all stimuli. Participants were instructed to ignore all auditory stimuli and focus on the monitor.

EEG preprocessing

EEG data were processed offline in EEGLAB 2023.1 on MATLAB R2021b (Delorme & Makeig, 2004; Mathworks, 2022). Cleaning of artifacts and noise was with the The Reduction of Electroencephalographic Artifacts (RELAX) preprocessing pipeline, a novel pipeline based on an empirical assessment of established cleaning methods (Bailey et al., 2022, 2023). We applied the default RELAX pipeline, RELAX_MWF_wICA, which utilizes methods from the following published toolboxes: fieldtrip, the MWF toolbox, wICA (Castellanos & Makarov, 2006), ICLabel (Pion-Tonachini et al., 2019), PREP (Bigdely-Shamlo et al., 2015) and Zapline-plus (Klug & Kloosterman, 2021). Given that single-trial analysis handles noisy data better and in order to remove as little brain activity and obtain as many trials as possible, especially error trials in the Flanker paradigm, we applied RELAX with less-stringent settings than default for our main analysis (specified below).

Prior to processing in RELAX, the raw Biosemi EEG data were imported into EEGLAB reference-free and down-sampled to 250 Hz. In initial preprocessing steps, RELAX removed line-noise at 50 Hz with the Zapline-plus toolbox and referenced data to common average with the PREP toolbox after the automatic removal of extremely noisy or flat channels. Data were hi-pass filtered at 0.25 Hz and low-pass filtered at 80 Hz using the default RELAX Butterworth filter, which is suggested to perform better than EEGLAB's pop_eegfiltnew (Bailey et al., 2023). Note that RELAX applies a 0.25 Hz hi-pass filter by default instead of the commonly used 1 Hz, a trade-off which somewhat decreases the quality of the subsequent independent component analysis (ICA) decomposition but does not distort the ERP time course (Bailey et al., 2023; Luck, 2014; Tanner et al., 2016; Winkler et al., 2015).

Next, artifact reduction based on multiple wiener filtering (MWF) with a delay period of 30 and wavelet-enhanced ICA (wICA) with the extended infomax ICA algorithm proceeded with less strict than default RELAX cleaning parameters. Specifically, muscle slope threshold was -0.31 (default -0.59), no channels were deleted due to muscle artifacts (default: channels with 5% or more muscle artifacts deleted) and only channels with 15% or more extreme artifacts were deleted (default 5%). Other settings remained on default, including at most 20% removed channels. On average 59.2 channels remained for the AO, 59.4 for the Flanker and 59.9 for the UO paradigm. There was no difference between groups in number of removed channels.

After interpolating removed channels, the preprocessed data were epoched and baseline-corrected according to parameters predetermined for each of the three paradigms (see Table 1). Note that RELAX applies a regression-based baseline correction method instead of the traditional subtraction which has been shown to distort the ERP waveform (Alday, 2019). For response-locked ERPs in the Flanker paradigm, baseline regression correction was with one factor with two levels: correct and error response. For stimulus-locked ERPs from the Flanker paradigm and ERPs from the other two paradigms, regression was with zero factors. Next, epochs with an absolute voltage amplitude threshold exceeding 100 μV (default 60 μV) or a kurtosis/improbable data limit exceeding 3 standard deviations (SD)/median absolute deviation (MAD) overall or 5 SD/MAD at any channel were rejected.

On average 1069 epochs of all trial types remained for the AO, 418 for the Flanker paradigm and 993 for the UO paradigm. For AO, there was a significant difference between groups in number of remaining epochs (HC: 1095, Distress: 1046, Fear: 1066; $F(84,87) = 6.79$, $p = 0.002$). Follow-up 1-way ANOVAs revealed a significant difference in number of remaining standard stimulus trials (HC: 896, Distress: 879, Fear: 866; $F(84,87) = 6.79$, $p = 0.002$). This small difference of ˜3% between groups was determined to be of no

consequence for the main analysis. There were no further differences between groups in number of remaining epochs for any of the other paradigms. Finally, the preprocessed data were converted to BIDS format to facilitate the sharing of data with the community (C. R. Pernet et al., 2019).

Table 1 shows an overview of paradigm and ERP variables.

>> Table 1 here <<

ERP analysis and statistics

All demographic and behavioral statistics were conducted in R (R Core Team, 2023). Statistical models were designed, evaluated and visualized using LIMO EEG in EEGLAB and MATLAB functions (Delorme & Makeig, 2004; C. R. Pernet et al., 2011).

After preprocessing, ERP single trials data were processed in LIMO EEG. For a given subject, the first-level of the GLM has the general form$\mathbf{Y} = \mathbf{XB} + \epsilon$ where Ydenotes the single-trial ERP data in the form $channels * timeframes$,$\mathbf{X}$ is a design matrix coding for the paradigm-specific stimulus types, $\mathbf{B}$ are the beta coefficients to be estimated and ϵ is the residual term representing what is left when the effects of the beta coefficients are accounted for. Model parameter estimation was with weighted least squares (WLS), a robust extension to ordinary least-squares (OLS) which uses principal component projection to weigh down outlier trials (C. Pernet et al., 2022). In example, for response-locked ERPs in Flanker paradigm, the model was$ERP_{Flanker} = X_{correct}B_{correct} + X_{error}B_{error} + residual$where $X_{correct}B_{correct}$ models the correct trials,$X_{error}B_{error}$ the error trials and $residual$ is the residual signal. Accordingly, $ERN = X_{error}B_{error} + residual$. For ERPs which are evaluated as difference waves, e.g., the MMN, linear combinations of beta coefficients were evaluated. In all ten first-level GLM models were evaluated, each containing one or more classic ERP components (see Table 1).

At the second level, the group level, for each model, group effects were examined channel by channel with 1-way ANOVAs testing for differences at each time frame. These 64 N-ways ANOVAs produced an uncorrected statistical parametric map (SPM) of size $channels * timeframes$, e.g., 64*151 = 9664 F -values for ERPs in the AO paradigm. The$F$ -statistic was calculated with a generalization of Welch'sF -test, a modified F -test for distributions with unequal variances, applied to the 20% winsorized group mean (Delacre et al., 2019; C. R. Pernet et al., 2011; Wilcox & Rousselet, 2018). Multiple comparisons correction was conducted using a well-established and validated non-parametric bootstrap spatiotemporal clustering method at an alpha level of 5% (Maris & Oostenveld, 2007; C. R. Pernet, 2015). This produced a SPM of corrected p -values denoting significant clusters representing group effects across channels and time frames.

For the N-ways ANOVAs revealing such significant clusters, follow-up two-sample t-tests where conducted between each group. The$t$ -statistic was calculated with Yuen's t -test, an extension of Welch's t -test for winsorized means of distributions with unequal variances (C. R. Pernet et al., 2011; Wilcox & Rousselet, 2018; Yuen, 1974). Corrections for multiple comparisons was likewise with spatiotemporal clustering at an alpha level of 5%, yielding a SPM of corrected p -values.

Prior to the N-ways ANOVA, an N-ways ANCOVA as implemented in LIMO EEG was used to test for the effects of age, sex and medication status on group differences. For medication status, due to the variability in dosage and type of medication, two dummy variables encoded no prescription (0, 0), one prescription (1, 0) and more than one prescription (1, 1). Medication prescribed "as needed" was not considered since patients were instructed to avoid intake from the afternoon before the day of recording.

LIMO analysis windows (see Table 1) were chosen as to include the baseline period and to exclude activity immediately preceding or overlapping with the next epoch. In line with recent research, these wide analysis windows did not noticeably decrease power (Fields & Kuperberg, 2020). For example, group differences in ERN were significant at LIMO analysis windows between 0 and 200 ms as well as between -200 and 500 ms. Note that the Attended oddball paradigm had an ISI of 650 to 850 ms. Due to a slower P3b wave than expected when designing the experiment, the LIMO analysis window was extend to 700 ms in order

to capture the whole wave. As a result, the last 50 ms of some trials contained activity from the following epoch. Similarly, there is necessarily an overlap between the stimulus-locked Flanker P3b and the following response after approximately 300 to 500 ms. Therefore, the stimulus-locked Flanker P3b is included in the LIMO analysis window but results are not elaborated upon.

ERP grand averages are displayed for each stimulus type and group as the the 20% trimmed mean of subject-level weighted single-trial ERP data. Trimmed mean represents a robust central tendency estimate of the mean of the raw single-trial data and corresponds to a traditional grand average ERP waveform (C. R. Pernet et al., 2011; Wilcox & Rousselet, 2018). Instead of traditional frequentist confidence intervals (CI), which only gives the long-term probability of the true mean, LIMO EEG by default displays the 95% Bayesian Highest Density Interval (HDI), which is the 95% probability of the observed 20% trimmed mean (Morey et al., 2016; C. R. Pernet et al., 2011).

For models with significant clusters in the N-ways ANOVA, we display the group-wise mean beta coefficient time course with 95% CI at the channel where the corresponding ERP is traditionally evaluated, e.g., response-locked Flanker trials at FCz. Significant results from follow-up two sample t-tests are shown as as $channel * timeframe$t-statistic heat maps. Here, red denotes a positive difference (positive t-values) between the two compared groups, with the first denoted group as reference, and blue a negative difference. Shaded areas in heat maps indicate traditional evaluation windows of ERPs components elicited by the stimulus type. For simplicity, we denote maximum or minimum t-values as $max\,(|t|)$.

Results

Demographics, behavioral measures and self-report questionnaires

Table 2 shows demographics and behavioral measures.

>> Table 2 here <<

There was no difference between HC and the Patient group as a whole in sex and age, nor in number of correct and error trials in the Flanker paradigm. The patient group had significantly longer reaction times (RT) in both correct and error trials in the Flanker paradigm and to target tones in the AO paradigm. Within the patient group, the Distress group differed significantly in age and sex and had significantly longer RT in correct but not in error trials in the Flanker paradigm compared to the Fear group.

Table 3 shows medication status and results from self-report questionnaires for the Distress and Fear patient groups.

>> Table 3 here <<

The majority (86.1%) of the patients received psychiatric medication. Comorbidity, defined as more than one diagnosis, was 84% within Distress and 68% within Fear with a mean across both subfactors of 77%. Subfactor comorbidity, defined as the secondary diagnosis belonging to the other subfactor, was lower for both groups with Distress at 36% and Fear at 20% with a mean of 28%.

ERP grand average waveforms

Figures 1 to 4 shows the 20% trimmed mean of mean weighted subject-level single-trial ERP data across groups and stimulus types for all paradigms. Shaded areas indicate the 95% HDI. Established ERP components are marked on each plot and appear in much agreement with the literature. The ERPs from the AO paradigm have reached maximum peak well before 650 ms for all groups.

Response-locked Flanker

>> Figure 1 here <<

Stimulus-locked Flanker

>> Figure 2 here <<

Attended oddball

>> Figure 3 here <<

Unattended oddball

>> Figure 4 here <<

ERP group comparisons

Effects of age, sex and medication

N-ways ANCOVA with age, sex and medication status (none, one prescription and more than one prescription) as covariates showed no effects of sex and medication status in any of the models. For several of the ERPs, there was an effect of age. In agreement with the literature, age had an effect on the various P300 sub-components, e.g., P3a and P3b from the AO and stimulus-locked Flanker paradigms (Walhovd et al., 2008). Crucially, the N-ways ANCOVA found the same significant clusters spanning the same regions and time frames as clusters from the N-ways ANOVA. The exception to this was the cluster corresponding to P2 elicited to standard stimuli in the AO paradigm. Here, N-ways ANOVA found a cluster which the N-ways ANCOVA did not. On the other hand, there were no significant clusters in the P2 region for covariates age, sex or medication status either. We attribute this discrepancy to the reduced sensitivity of the ANCOVA compared to the ANOVA. Taken together, the results from the N-ways ANCOVA show that the significant clusters found in the N-ways ANOVA, on which we base the follow-up two-sample t-tests, are also significant when accounting for age, sex and medication status.

Response-locked Flanker

For correct trials (Figure 5, left panel), N-ways ANOVA revealed a significant effect of group at a cluster between -44 and 96 ms over central, frontal and parietal regions (maximum $F = 16.20$ at CP3 at 64 ms; corrected $p \leq 0.001$).

>> Figure 5 here <<

Follow-up two-sample t-tests (Figure 6, top panel) revealed a significant difference between HC and Distress at a cluster between -56 and 116 ms over central, frontal and parietal regions (maximum $t= 5.75$ at CP3 at 64 ms; corrected $p \leq 0.001$). The cluster corresponded well in location and time to the CRN, indicating a reduced CRN for Distress compared to HC. Note that $X_{correct}B_{correct}$ is positive-going because $CRN = X_{correct}B_{correct} + residual$ is less negative-going than $ERN = X_{error}B_{error} + residual$. Therefore, the observed difference in $X_{correct}B_{correct}$ amounts to a reduced, i.e., less negative, CRN, as is also seen on the grand average plot (Figure 1, left panel). Note also that the effect was weak at channel FCz and stronger at more central and posterior channels Cz and CPz. There was no difference between Fear and either Distress or HC.

>> Figure 6 here <<

For error trials (Figure 5, middle panel), N-ways ANOVA revealed a significant effect of group at a cluster between 120 and 260 ms over central, frontal and parietal regions (maximum $F = 12.89$ at P2 at 180 ms, corrected $p \leq 0.021$). Two-sample t-tests between HC and Distress (Figure 6, bottom panel) revealed two

significant clusters (cluster 1: from -16 to 56 ms over frontal, central and parietal regions, maximum t = -5.04 at Fz at 8 ms; cluster 2: from 120 to 296 ms over mainly central and parietal regions, maximum$t = 4.90$ at P2 at 180 ms; corrected $p \leq 0.021$). The first cluster corresponded well in location and time to ERN and indicated a reduced (less negative) ERN for Distress compared to HC. The second cluster corresponded well to Pe and indicated a reduced (less positive) Pe for Distress compared to HC. Note that N-ways ANOVA revealed only one significant cluster, corresponding to Pe, whereas follow-up two-sample t-tests found clusters corresponding to both ERN and Pe. Also note that the effect was stronger at FCz compared to more central and posterior channels Cz and CPz. Again, there was no difference between Fear and either Distress or HC.

In summary, results for the response-locked Flanker paradigm revealed a reduced CRN, ERN and Pe for Distress compared to HC.

Stimulus-locked Flanker

For congruent stimuli with correct response (Figure 7, left panel), N-ways ANOVA revealed a significant cluster between 256 and 472 ms over frontal, central and parietal regions (maximum $F = 18.95$ at CP2 at 360 ms, corrected $p \leq 0.001$).

>> Figure 7 here <<

Follow-up two-sample t-tests between HC and Distress (Figure 8, top panel) revealed a significant cluster from 248 to 480 ms over frontal, central and parietal regions (maximum $t = 5.67$ at FT7 at 332 ms,$p \leq 0.001$).

>> Figure 8 here <<

This cluster corresponded to P3b which cannot be reliably analyzed due to overlap with the following response. There was no difference for either congruent stimulus N2 or P3b between Fear and either Distress or HC.

For incongruent stimuli with correct response (Figure 7, middle panel), N-ways ANOVA revealed a significant cluster between 404 and 620 ms over frontal, central and parietal channels (maximum $F = 14.25$ at CP2 at 428 ms, corrected $p \leq 0.004$). Follow-up two-sample t-tests between HC and Distress (Figure 8, bottom panel) revealed two significant clusters, the first over mainly central and parietal regions and the second over mainly frontal and central regions (cluster 1: from 396 to 624 ms, maximum $t = 5.06$ at P2 at 444 ms; cluster 2: from 220 to 336 ms, maximum $t = 5.44$ at CP6 at 276 ms; corrected$p \leq 0.036$). Similar to the betas for congruent stimuli, the first cluster corresponded to P3b and was therefore not further analyzed. The second cluster corresponded to N2 and indicated a reduced (less negative) N2 for Distress compared to HC. There was no difference for either incongruent N2 or P3b between Fear and either Distress or HC.

In summary, results for the stimulus-locked Flanker paradigm revealed a reduced N2 to congruent stimuli preceding correct responses for Distress compared to HC.

Attended oddball

For standard stimulus (Figure 9, left panel), N-ways ANOVA revealed a significant effect of group at three clusters (cluster 1: from 116 to 220 ms over mainly frontal and central regions, maximum $F = 16.16$ at FC3 at 140 ms; cluster 2: from 276 to 348 ms, over mainly central and parietal regions, maximum $F = 18.08$ at C4 at 296 ms; cluster 3: from 488 to 700 ms over mainly parietal regions, maximum $F = 22.23$ at CP4 at 612 ms; corrected $p \leq 0.045$).

>> Figure 9 here <<

Follow-up two-sample t-tests revealed that these clusters were generated by differences between Distress and both HC and Fear (Figure 10, top panel).

>> Figure 10 here <<

Distress and Fear differed at a cluster from 272 to 396 ms over frontal, central and parietal regions with maximum $t = 5.89$ at C4 at 296 ms (corrected $p \leq 0.035$). This cluster (Figure 10, top panel, left) corresponded well to a somewhat late N2, which on the grand average (Figure 3, middle panel) at FCz can be seen peaking around 290 ms, and indicated a reduced (less negative) N2 to standard stimuli for Distress compared to Fear. HC and Distress differed at two clusters (cluster 1: from 472 to 700 ms over mainly parietal regions, maximum$t = 6.27$ at CP4 at 612 ms; Cluster 2: from 112 to 224 ms over frontal, central and parietal regions, maximum $t = 5.19$ at FC3 at 140 ms; corrected $p \leq 0.028$). The first cluster did not clearly correspond to any known ERP. Two possibilities are the late-positive potential (LPP) cut short by the following stimulus and polarity-reversed in parietal regions where HC and Distress differed (Figure 11) (Luck & Kappenman, 2011). Another possibility is the N400, which is typically elicited in language-related paradigms (Kutas & Federmeier, 2011).

>> Figure 11 here <<

The second cluster corresponded well to P2 and indicated a reduced (less positive) P2 for Distress compared to HC, which is also apparent on the grand average plot (Figure 3, middle panel).

For target stimulus (Figure 9, middle panel), N-ways ANOVA revealed a significant effect of group at a cluster from 316 to 696 ms over frontal, central and parietal regions (maximum $F = 20.93$ at P3 at 404 ms; $p \leq 0.001$). Follow-up two-sample t-tests revealed differences between Distress and both HC and Fear (Figure 10, bottom panel). Distress differed from HC at a cluster from 324 to 700 ms over frontal, central and parietal regions (maximum $F = 5.47$ at P3 at 404 ms; $p \leq 0.001$) and from Fear at two clusters, together covering the same time window and regions (cluster 1: from 448 to 688 ms, maximum $F = 5.33$ at F7 at 516 ms; cluster 2: from 296 to 440 ms, maximum $F = 5.32671$ at PO7 at 408 ms; $p \leq 0.019$). The clusters corresponded well to the P3b and indicated a reduced (less-positive) P3b for Distress compared to both HC and Fear, with no difference between the latter two.

In summary, results from the AO paradigm revealed a reduced P3b for Distress compared to both HC and Fear. In addition, Distress had a reduced N2 compared to Fear and a reduced P2 compared to HC. Lastly, Distress differed from HC also at an unidentified late, mainly parietal cluster, possibly the LPP or N400. Note that there were no significant clusters for ERPs to the distractor stimulus eliciting P3a.

Unattended oddball

For the Duration difference wave, N-ways ANOVA revealed a significant cluster from 128 to 192 ms over frontal, central and parietal regions (maximum $F = 13.89$ at P6 at 176 ms; $p \leq 0.002$). There were no significant group effects for the Combined and Frequency difference waves.

>> Figure 12 here <<

Follow-up two sample t-tests revealed differences between Distress and both HC and Fear (Figure 13).

>> Figure 13 here <<

Distress differed from Fear at two clusters (cluster 1: from 232 to 308 ms over frontal, central and parietal regions, maximum $t = 4.88$ at F8 at 276 ms; cluster 2: from 100 to 184 ms over frontal, central and parietal regions, maximum t = -5.17 at P6 at 176 ms; $p \leq 0.027$). The first cluster corresponded well to dP3a, indicating a reduced dP3a for Distress compared to Fear, while the second cluster corresponded well to dMMN, indicating a reduced dMMN for Distress compared to Fear, as is also evident on the grand average plot (Figure 4, middle panel). Distress also differed from HC at a barely significant cluster from 128 to 192 ms over parietal and occipital regions (maximum$t = 4.02$ at Iz at 184 ms; $p \leq 0.047$). This cluster corresponded to dP3a and indicated a reduced P3a for Distress compared to HC, but only at parietal and occipital regions.

In summary, results from the UO paradigm revealed reduced dMMN and dP3a for Distress compared to Fear. In addition, a barely significant cluster indicated a reduced dP3a for Distress compared to HC, but

only at parietal and occipital regions.

Discussion

In this study we investigated whether robust single-trial ERPs as recorded on the EEG can differentiate between groups based on the HiTOP classification of the emotional disorders into subfactors Distress and Fear. These two groups were also compared against a group of healthy comparison subjects (HC) matched with the whole patient group in age and sex. Addressing issues of traditional ERP preprocessing and analysis methods, we applied robust single-trial ERP analysis as implemented in the EEGLAB toolbox LIMO EEG.

We found several statistically significant differences between the three groups. For all ERPs differentiating between the two HiTOP subfactors, Distress had a weaker response in terms of a reduced absolute amplitude compared to Fear. For other ERPs, Distress differed only from HC, again with a reduced absolute amplitude, and in no cases did we find a difference between Fear and HC.

Contrary to our expectations, we found no difference between Distress and Fear in any of the response-locked Flanker ERPs. Hyperactive performance monitoring as indexed by enhanced ERN, and to a lesser extent, CRN, is consistently found in OCD and have been associated with broad symptoms in the Internalizing spectrum (Macedo et al., 2021; Pasion & Barbosa, 2019). Recent evidence from a large sample covering the full spectrum of OCD characteristics found more specific associations between ERN and a composite dimension of anxiety and neuroticism and between CRN and compulsivity (Riesel et al., 2022). In the HiTOP, the placement of OCD in Fear is debated due to heterogeneous symptoms cross-loading on both the Fear subfactor and on the Thought disorder spectrum (Faure & Forbes, 2021). Our results support this in indicating that enhanced response-locked Flanker ERPs are not markers of the whole Internalizing spectrum or of the more narrow Fear subfactor, but more likely of symptoms and traits at lower levels of the HiTOP hierarchy (Härpfer et al., 2022). However, the ERN has been shown to be sensitive to manipulations of experimental factors such as the amount, timing and type of feedback, the ratio between congruent and incongruent stimuli and the addition of response-dependent reward or punishment (Gloe & Louis, 2021; Larson et al., 2014; Nuñez-Estupiñan et al., 2022). Given this, it is possible that increased ERN can be observed uniformly in Fear at exposure leading to increased levels of anxiety, but not in the comparably relaxed and safe laboratory setting. As it stands, the wide 95% HDI on the grand average plot (Figure 1, right panel) suggests that Fear is not a sufficiently homogeneous group in terms of the ERN to render it different from either Distress or HC.

All of ERN, CRN and Pe were reduced in Distress compared to HC. The disorders in Distress included in this study (depression and GAD) are often comorbid, share a significant symptom overlap and have been suggested to be variations of the same etiology (Crocq, 2017; Hettema, 2008). Both enhanced and reduced ERN have been reported in depression and GAD (Cho et al., 2022; Dell'Acqua et al., 2023; Ladouceur et al., 2012; Weinberg et al., 2010; Z. Xiao et al., 2011). A recent meta-analysis using a p-curve method to detect publication bias found a significant albeit weak effect of reduced ERN in depression (Clayson et al., 2020). One study investigating GAD with and without comorbid depression in females found increased ERN compared to healthy comparison subjects only in GAD without depression (Weinberg et al., 2012). However, these results are not as robust as in OCD and are obscured by methodological issues associated with traditional ERP baseline subtraction methods. The ERN is preceded by and overlaps with the P3b, which is consistently reduced in depression (Francis et al., 2021). With traditional baseline correction methods, the subtraction of a smaller P3b in the baseline period will produce an artificially larger ERN (less is subtracted)(Klawohn, Santopetro, et al., 2020). This is not the case for the regression-based baseline correction method implemented in RELAX and utilized in this study. We see that the stimulus-locked Flanker P3b to correct trials is significantly reduced for Distress compared to HC, and so is the following response-locked CRN. Our results are therefore in line with recent evidence suggesting that reduced ERN is associated with specific cognitive anxiety - perhaps well-captured in our Distress sample - as opposed to physiological anxiety which is associated with an increased ERN (Macedo et al., 2021).

Finally, we see that the error positivity, Pe, indexing error awareness, is reduced for Distress compared to HC. Less studied than the ERN, the anatomical and functional correlates of the Pe are still being investigated, with some evidence indicating that Pe is a part of the P300 complex (Dali et al., 2023; Ridderinkhof et al., 2009). Although we are not aware of consistent results, evidence suggest an association between reduced Pe and symptoms of depression (Schroder et al., 2013). As such, our results support a reduced Pe in the Distress disorders depression and GAD.

We examined two versions of the N2, the exact biobehavioural correlates of which are not established and likely paradigm-dependent (Folstein & Van Petten, 2008; Larson et al., 2014). The N2 elicited in the stimulus-locked Flanker paradigm (FN2) is often described as an index of conflict monitoring and is larger for incongruent compared to congruent stimuli (Kałamała et al., 2018; Riesel et al., 2017). We note that not all studies using equiprobable congruent and incongruent stimuli, as we do in this study, find the Flanker N2 (Kałamała et al., 2018). FN2 is consistently enhanced in OCD (Riesel et al., 2017). Results on FN2 in depression are inconsistent and seem to be paradigm-dependent (Alderman et al., 2015; Luck & Kappenman, 2011). In GAD, somewhat more consistent results indicate a reduced FN2 compared to healthy comparison subjects (Larson et al., 2013; Nawani et al., 2018; Yu et al., 2018). We find that FN2 elicited to incongruent stimuli is reduced in Distress compared to HC. These results are in line with results from the response-locked Flanker ERPs in suggesting that Distress has globally reduced conflict-related ERPs compared to healthy comparison subjects. Given that ERN and FN2 are both generated in the dorsal anterior cingulate cortex, our results support that abnormalities in this region is a hallmark of Distress disorders (Hochman et al., 2014; J. Xiao et al., 2023).

The N2 from the Attended oddball paradigm (AN2) is elicited to auditory stimuli (standard, target and distractor stimuli) and is described as an index of stimulus identification and distinction (Patel & Azzam, 2005). Results on AN2 in Distress and Fear disorders are inconsistent and depends on the paradigm (Iwanami et al., 1997; Luck & Kappenman, 2011; Perera et al., 2019). Interestingly, we find that Distress and Fear differ in AN2 elicited to standard stimuli, and neither differ from HC. From the grand average plot (Figure 3, middle panel) it can be speculated that with more power Fear would differ also from HC. If so, increased AN2 would be able to differentiate Fear from Distress and HC.

The auditory P2 has mainly been studied in the context of the central N1-P2 auditory response to estimate auditory threshold or hearing loss (Crowley & Colrain, 2004; Lightfoot, 2016). Evidence indicate that the loudness dependency of the P2 predicts SSRI-treatment outcome in depression (Gallinat et al., 2000). However, we are not aware of abnormalities in P2 in either Distress or Fear disorders. We find that the auditory P2 elicited to standard stimuli is reduced for Distress compared to HC. On the grand average (Figure 3, middle panel), it would appear that P2 is significantly increased for Fear compared to Distress. However, as for the ERN, the wide 95% HDI suggests that Fear contains heterogeneous disorders in terms of P2.

The components in the P300 ERP complex elicited to attended stimuli in the AO paradigm are some of the most studied ERPs in neuroscience (Polich, 2007, 2020). We examined two components of the P300, the P3a elicited to distractor stimuli, indexing involuntary shift of attention, and the P3b elicited to target stimuli, indexing context updating and memory processing (Luck & Kappenman, 2011). We found no difference between groups in the P3a, a finding which is contrary to somewhat consistent results of reduced P3a in depression (Luck & Kappenman, 2011). It is possible that more power would have allowed us to distinguish between Distress and both Fear and HC, which appear very similar. Again, the wide 95% HDI on the grand average plot (Figure 3, left panel) indicates that Distress is not a homogeneous construct in terms of P3a.

The P3b elicited to target stimuli was significantly reduced for Distress compared to both Fear and HC. While we are not aware of studies examining P3b in GAD, our results are in line with somewhat consistent findings of reduced P3b in depression, especially when comorbid anxiety is accounted for (Bruder et al., 2002; Nan et al., 2018). For the disorders in Fear, results indicate that P3b is enhanced in OCD but not in panic disorder (Gohle et al., 2008; Howe et al., 2014). Our results indicate that Distress and Fear as transdiagnostic constructs are homogeneous enough as to render them different from each other in terms of

the P3b. As such, our findings corroborate a recent meta-analysis finding increased P3b in Fear compared to Distress11Note that in Botelho et al. (2023), OCD is placed in the Distress subfactor (Botelho et al., 2023). However, we find no difference between Fear and HC. The similar 95% HDIs suggest that more power would not allow us to differentiate between these two groups. In fact, the 95% HDI for Fear overlaps HC in both directions, indicating that our Fear sample contains disorders associated with both an increased and a decreased P3b compared to HC.

The mismatch negativity (MMN), an index of detection of change, and the difference wave P3a (dP3a), reflecting change of attention, are consistently reduced in chronic schizophrenia (Näätänen & Kähkönen, 2009). In depression, results on MMN are inconsistent and depends on comorbidity with anxiety and paradigm design (Bissonnette et al., 2020). We are not aware of consistent results for any of the other disorders included in this study (Luck & Kappenman, 2011). Interestingly, both duration MMN and dP3a are reduced in Distress compared to Fear. We can speculate that with more power Distress would also differ from HC, but HC and Fear appear similar on the grand average plot (Figure 4, middle panel).

From the above it can be argued that some of our results would have evaded studies based on traditional ERP analysis methods. For example, for the response-locked Flanker ERPs, we find that differences between groups are focused on different regions. Whereas the ERN differs at channels around FCz, differences in CRN are located more posteriously at Cz and CPz. A traditional ERP study with an *á priori* -defined region of interest of 0 to 100 ms at either FCz or Cz would have risked obtaining null results or a reduction in statistical power if analyzing several regions. Crucially, we can conclude that all but one of the significant clusters found by this method correspond to well-established ERP components. As such, the method translates directly to a richer version of traditional ERP analysis methods, the only caveat being that the direction of effects between beta coefficients must be interpreted alongside the residual term and grand average plot.

The HiTOP is based on factor analysis of self-report questionnaires and symptom checklists and not on biological measures such as ERPs. The subfactor level of the hierarchy describes dimensions, e.g. depression and panic disorder, which roughly correspond to diagnoses in the categorical taxonomies. ERPs have been shown to index specific brain functioning and have been associated with categorical diagnoses as well as specific symptoms and traits. Our results therefore, in the most general sense, validate the HiTOP classification of the emotional disorders into subfactors Distress and Fear in showing that two groups based hereof differ in specific measures of brain functioning. However, this validation is indirect for several reasons. First, allocation of patients into subfactor groups was based on primary diagnosis, a measure not derived from the HiTOP but from ICD-10, a categorical taxonomy. Indeed, diagnoses do not exist in the HiTOP and their placement in subfactors is for guiding purposes only. A more valid approach would have been to allocate patients according to results from self-report questionnaires directly assessing the Distress and Fear subfactors. However, such measures are not yet available. Comorbidity in our sample was consistent with established rates of comorbidity within the emotional disorders. Comorbidity between subfactors was markedly lower, indicating that the HiTOP captures some comorbidity by grouping similar disorders together. However, in several cases, the choice of a primary diagnosis was between diagnoses in opposing subfactors, e.g., a patient with chronic OCD (Fear) presenting with symptoms of depression (Distress). Therefore, it is likely that both groups contain patients who would have been allocated to the other subfactor had the 'ground truth' psychopathology been known. Second, both subfactors cover several diagnoses, Distress containing depression and GAD, Fear agoraphobia, OCD, PD and SP. Our results indicate that groups based on these subfactors are coherent enough to be distinguished by ERPs. However, even though our complete sample is rather large (N = 88), it does not permit comparisons between diagnoses within subfactors, or analysis of the contributions of each diagnosis to group differences. Therefore, we can not rule out that a few rather than all disorders are driving the observed differences between subfactors.

In addition to these limitations, Distress and Fear differed significantly in age and sex in that Fear consisted mainly of younger females. However, the N-ways ANCOVA suggested that the observed group effects were not due to differences in age or sex. Consequently, we could not corroborate that sex has an influence on the ERN and its associations with symptoms (Fischer et al., 2016; Imburgio et al., 2020; Sandre et al., 2020).

Future studies employing Distress and Fear groups matched in age and sex could elucidate this question.

In conclusion, the HiTOP subfactors Distress and Fear can be characterized by a set of abnormalities in brain functioning as indexed by well-established ERPs analyzed with robust single-trial ERP methods. Specifically, Distress has uniformly reduced ERP absolute amplitudes compared to Fear (AN2, MMN, dP3a and P3b) and HC (CRN, ERN, Pe, FN2, P2 and P3b). Contrary to our hypothesis, Fear did not show increased amplitudes compared to either Distress or HC. Future studies could utilize HiTOP-validated psychopathology measures rather than ICD-10 primary diagnosis to better characterize subfactor groups.

Tables

Trials	
ISI	
Baseline	
Epoch range	
LIMO analysis window	
LIMO model	
Traditional ERP	

Abbreviations: Congr., Congruent; Incongr., Incongruent; r.l., response-locked; s.l., stimulus-locked; w.c.r., with correct resp
Note: all measures in milliseconds (ms); * indicates a difference wave in that the beta coefficient to standard stimulus is sub

Table 1: Paradigm and ERP overview

	HC
Female/Male (%)	**25 (67.6)**
	12 (32.4)
Age	**38.5 (13.2)**
Flanker	
Correct trials	**428.1 (38.9)**
Error trials	**55.1 (24.3)**
RT correct trials	**410.1 (51.6)**
RT error trials	**318.1 (40.0)**
Attended oddball	
RT deviant tone	**404.8 (47.8)**
Abbreviations: RT, reaction time in miliseconds (ms)	**Abbreviations: RT, reaction time in miliseconds (m**

Table 2: Demographics and Behavioral Measures

HiTOP	Comorbidity
	(%)
Distress	36.00
Fear	20.00
Total/mean	28.00
Abbreviations: SRQ, self-report questionnaire	Abbreviations: SRQ, self-report questionnaire
Note: Comorbidity defined as comorbidity across HiTOP subfactors	Note: Comorbidity defined as comorbidity across HiT

Table 3: HiTOP factor breakdown, medication status and self-report questionnaires

Figures

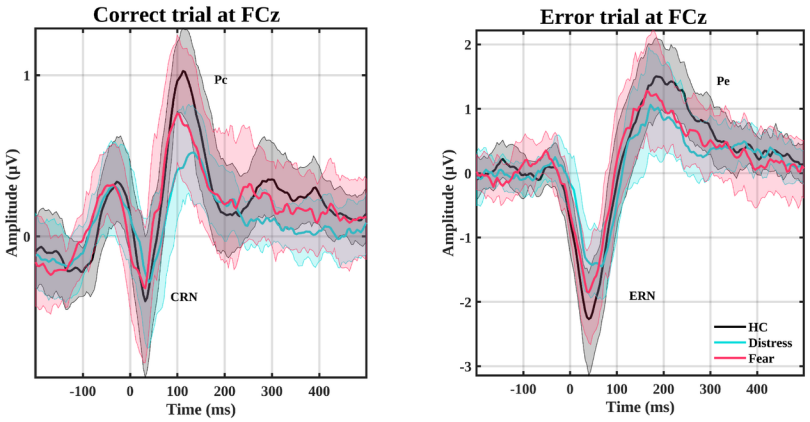


Figure 1: Response-locked Flanker grand average ERP waveforms at traditional evaluation channel FCz computed as the 20% trimmed mean of the mean subject-level single-trial ERP data. Shaded areas denote the 95% Bayesian Highest Density Interval (HDI). Typical ERP components are marked.

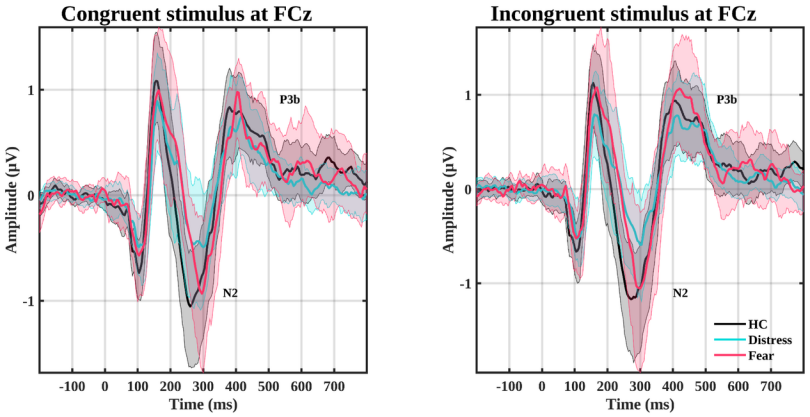


Figure 2: Stimulus-locked Flanker grand average ERP waveforms to stimuli preceding correct responses at traditional evaluation channel FCz computed as the 20% trimmed mean of the mean subject-level single-trial ERP data. Shaded areas denote the 95% Bayesian Highest Density Interval. Typical ERP components are marked.

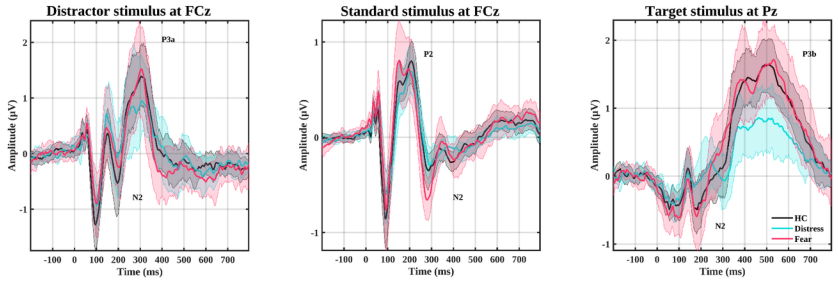


Figure 3: Attended oddball grand average ERP waveforms at traditional evaluation channels Pz and FCz computed as the 20% trimmed mean of the mean subject-level single-trial ERP data. Shaded areas denote the 95% Bayesian Highest Density Interval. Typical ERP components are marked.

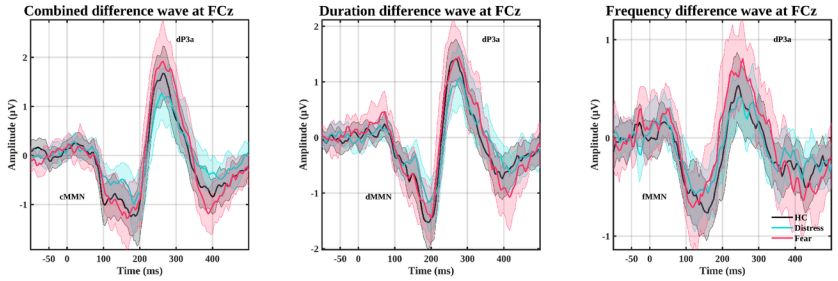


Figure 4: Unattended oddball grand average ERP waveforms at traditional evaluation channel FCz computed as the 20% trimmed mean of the mean subject-level single-trial ERP data. Shaded areas denote the 95% Bayesian Highest Density Interval. Typical ERP components are marked.

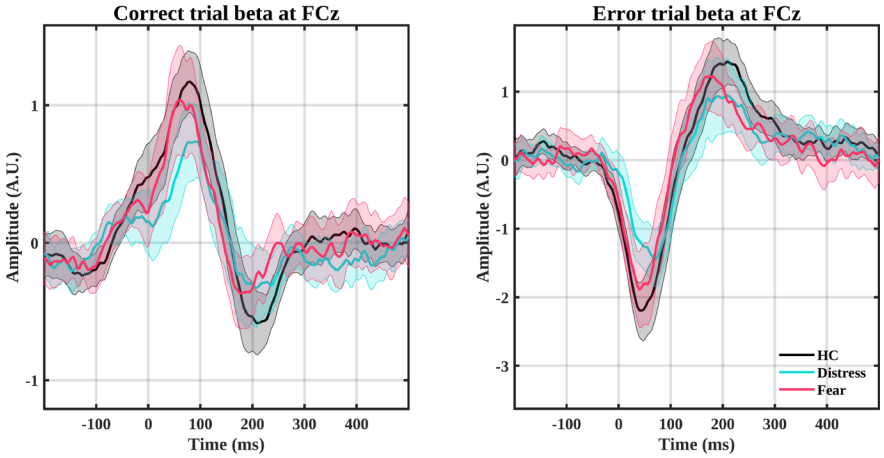
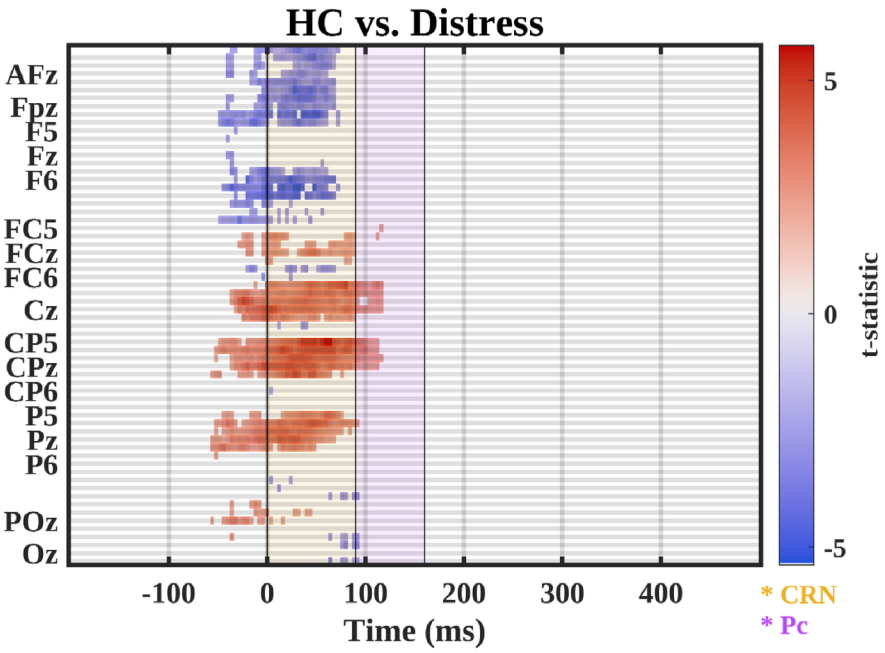


Figure 5: Time-courses for response-locked Flanker betas for which N-ways ANOVA revealed significant

group differences (left, correct trial; right, error trial). Shaded regions denote 95% Confidence Intervals. A.U., Arbitrary Units.

Correct trial



Error trial

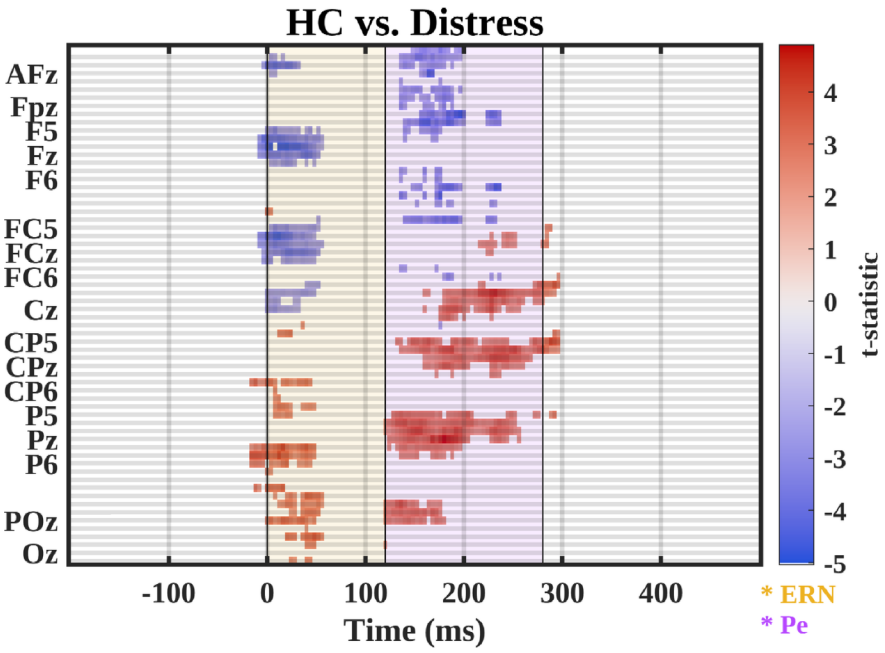


Figure 6: Results from follow-up two-sample t-tests for the response-locked Flanker betas (top, correct trial; bottom, error trial) shown as heat maps indicating significant regions (red, positive t-values; blue, negative t-values). Plot title indicate compared groups with the first denoted group as reference. Shaded regions indicate traditional evaluation windows. Results corrected with spatiotemporal clustering at an alpha level of 5%.

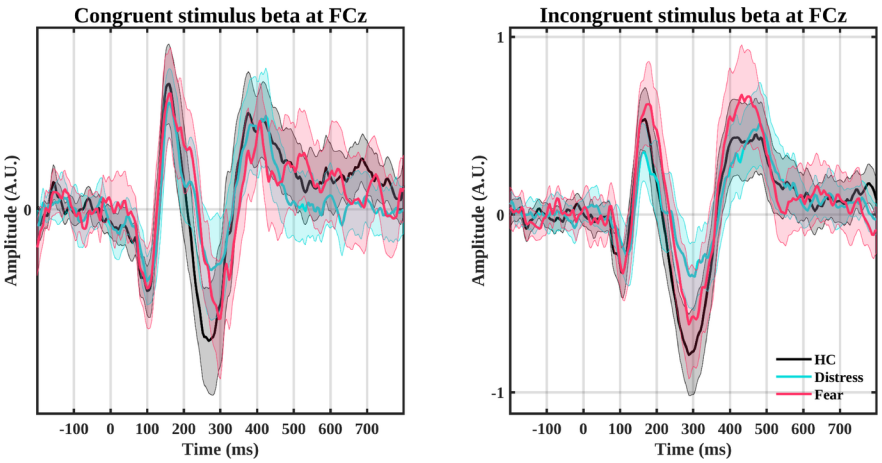
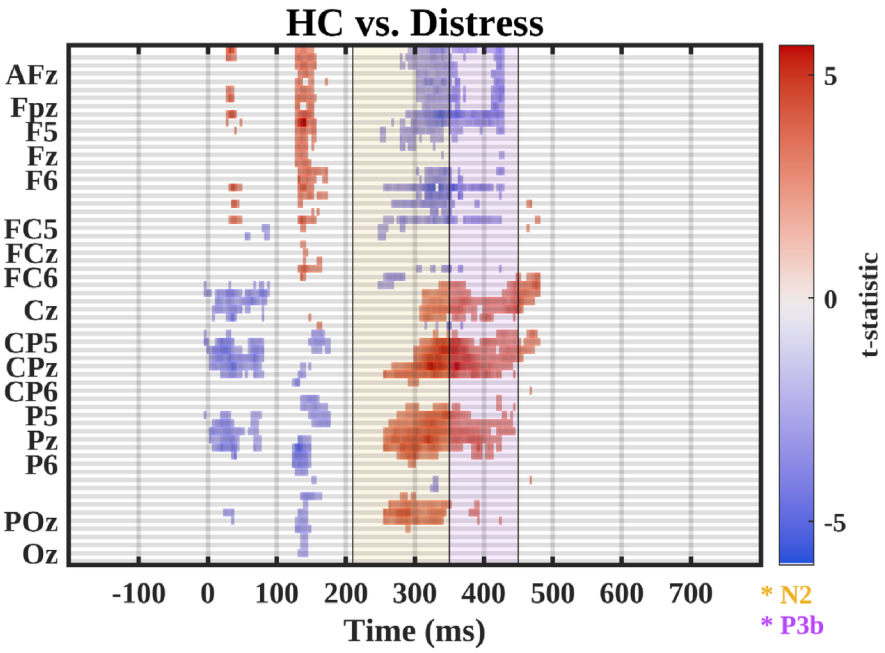


Figure 7: Time-courses for stimulus-locked Flanker betas for which N-ways ANOVA revealed significant group differences (left, congruent stimulus with correct response; right, incongruent stimulus with correct response). Shaded regions denote 95% Confidence Intervals. A.U., Arbitrary Units.

Congruent stimulus



Incongruent stimulus

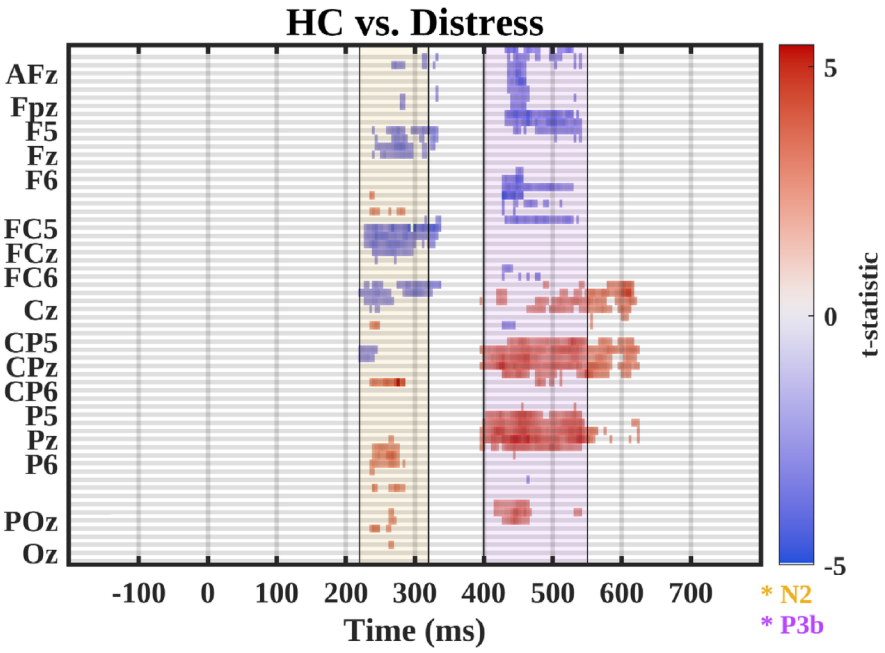


Figure 8: Results from follow-up two-sample t-tests for the stimulus-locked Flanker betas (top, congruent stimulus with correct response; bottom, incongruent stimulus with correct response) shown as heat maps indicating significant regions (red, positive t-values; blue, negative t-values). Plot title indicate compared groups with the first denoted group as reference. Shaded regions indicate traditional evaluation windows. Results corrected with spatiotemporal clustering at an alpha level of 5%.

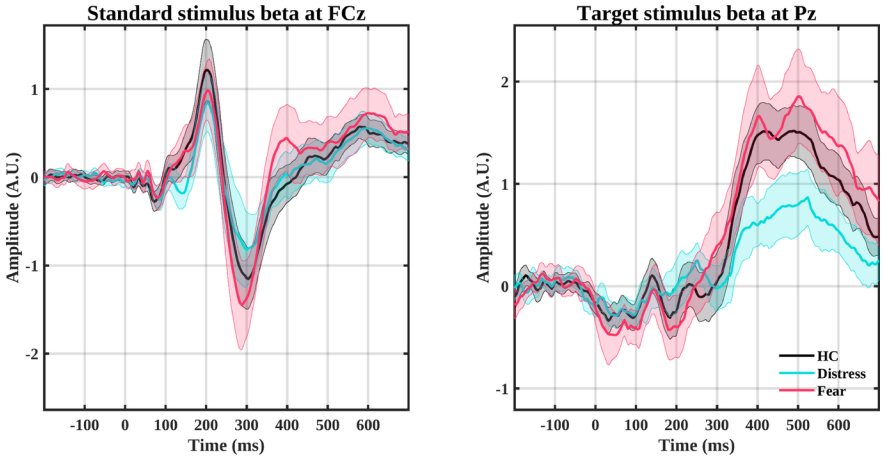


Figure 9: Time-courses for Attended oddball betas for which N-ways ANOVA revealed significant group differences (left, standard stimulus; right, target stimulus). Shaded regions denote 95% Confidence Intervals. A.U., Arbitrary Units.

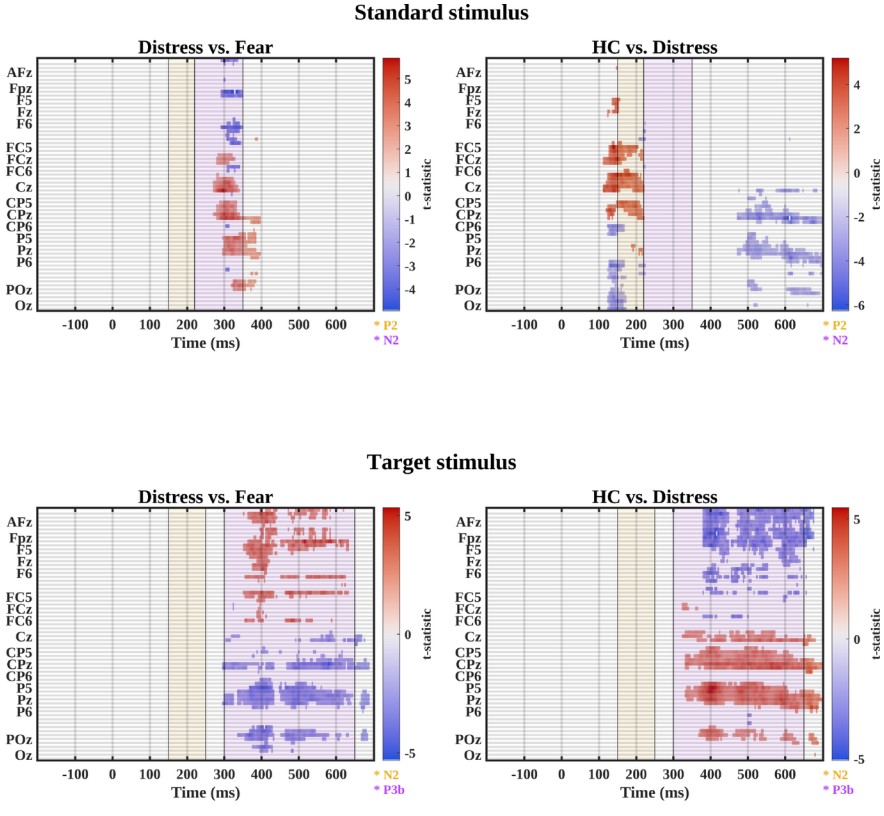


Figure 10: Results from follow-up two-sample t-tests for the Attended oddball betas (top, standard stimulus; bottom, target stimulus) shown as heat maps indicating significant regions (red, positive t-values; blue, negative t-values). Plot title indicate compared groups with the first denoted group as reference. Shaded regions indicate traditional evaluation windows. Results corrected with spatiotemporal clustering at an alpha level of 5%.

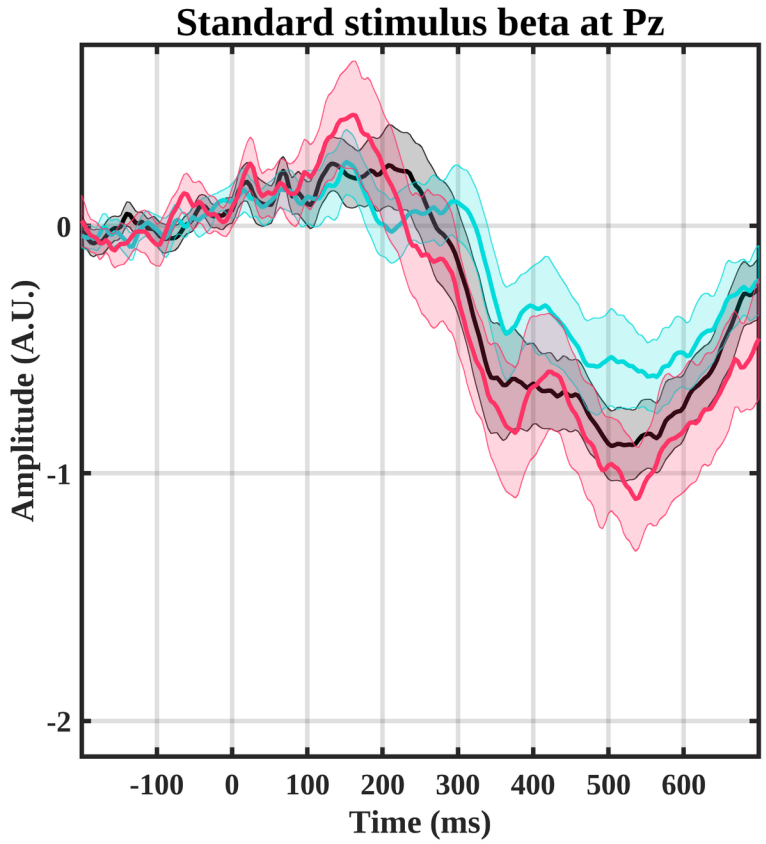


Figure 11: Standard stimulus beta coefficient time course at Pz showing a reduced (less-negative) unidentified late ERP for Distress (blue) compared to HC (black). The observed difference between Distress and Fear (red) was not significant. A.U., Arbitrary Units.

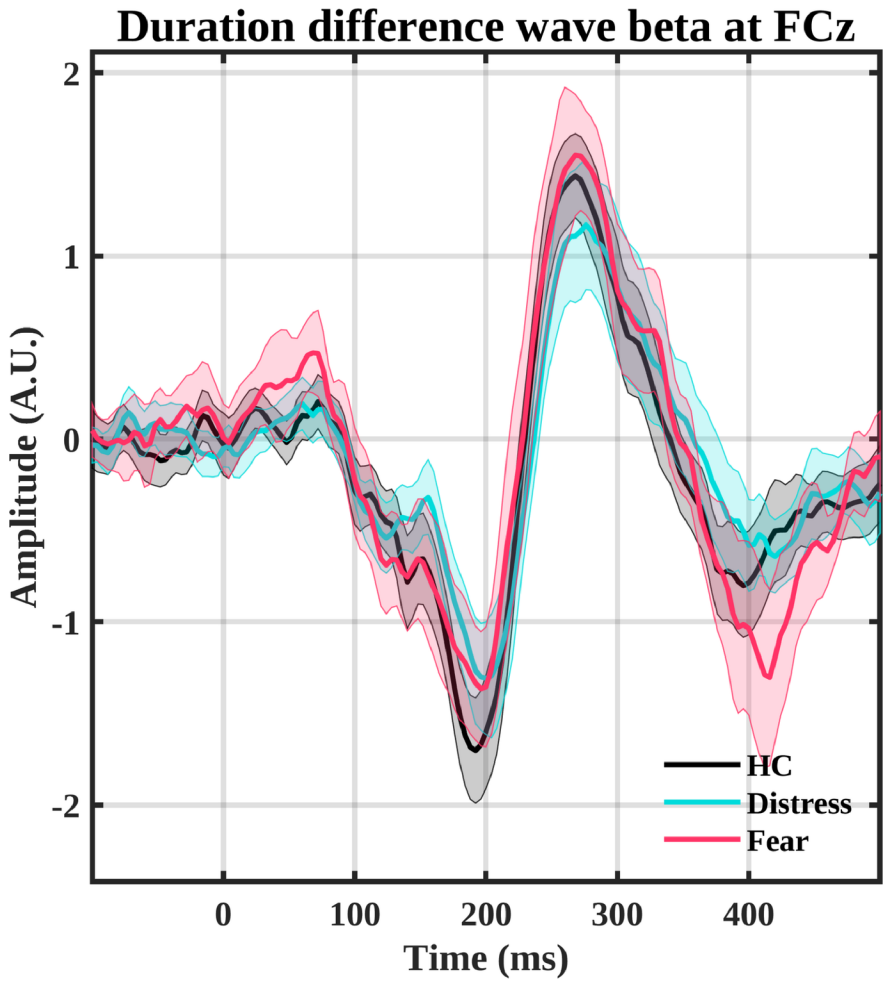


Figure 12: Time-course for the Duration difference wave beta from the Unattended oddball paradigm for which N-ways ANOVA revealed significant group differences. Shaded regions denote 95% Confidence Intervals. A.U., Arbitrary Units.

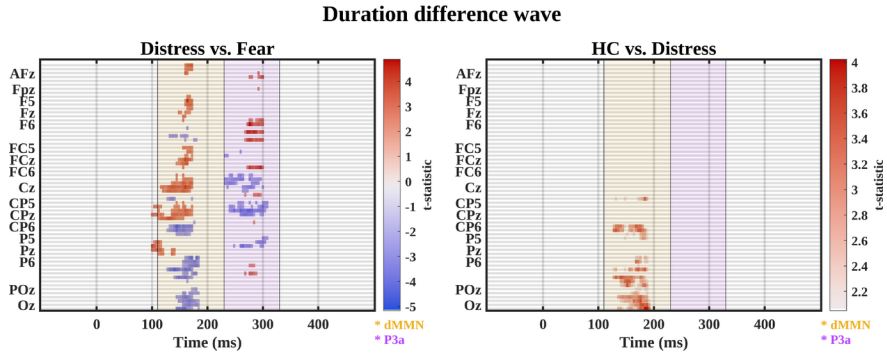


Figure 13: Results from follow-up two-sample t-tests for the Duration difference wave beta from the Unattended oddball paradigm shown as heat maps indicating significant regions (red, positive t-values; blue, negative t-values). Plot title indicate compared groups with the first denoted group as reference. Shaded re-

gions indicate traditional evaluation windows. Results corrected with spatiotemporal clustering at an alpha level of 5%.

Bibliography

Alday, P. M. (2019). How much baseline correction do we need in ERP research? Extended GLM model can replace baseline correction while lifting its limits. *Psychophysiology* , *56* (12), e13451. https://doi.org/10.1111/psyp.13451

Alday, P. M., & van Paridon, J. (2021). *Away from arbitrary thresholds: Using robust statistics to improve artifact rejection in ERP* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/wqrb5

Alderman, B. L., Olson, R. L., Bates, M. E., Selby, E. A., Buckman, J. F., Brush, C. J., Panza, E. A., Kranzler, A., Eddie, D., & Shors, T. J. (2015). Rumination in major depressive disorder is associated with impaired neural activation during conflict monitoring. *Frontiers in Human Neuroscience* , *9* .

Bailey, N. W., Biabani, M., Hill, A. T., Miljevic, A., Rogasch, N. C., McQueen, B., Murphy, O. W., & Fitzgerald, P. B. (2022).*Introducing RELAX (the Reduction of Electroencephalographic Artifacts): A fully automated pre-processing pipeline for cleaning EEG data - Part 1: Algorithm and Application to Oscillations* (p. 2022.03.08.483548). bioRxiv. https://doi.org/10.1101/2022.03.08.483548

Bailey, N. W., Hill, A. T., Biabani, M., Murphy, O. W., Rogasch, N. C., McQueen, B., Miljevic, A., & Fitzgerald, P. B. (2023). RELAX part 2: A fully automated EEG data cleaning algorithm that is applicable to Event-Related-Potentials. *Clinical Neurophysiology* , S1388245723000287. https://doi.org/10.1016/j.clinph.2023.01.018

Bejerot, S., Edman, G., Anckarsäter, H., Berglund, G., Gillberg, C., Hofvander, B., Humble, M. B., Mörtberg, E., Rastam, M., Stahlberg, O., & Frisén, L. (2014). The Brief Obsessive-Compulsive Scale (BOCS): A self-report scale for OCD and obsessive-compulsive related disorders.*Nordic Journal of Psychiatry* , *68* (8), 549–559. https://doi.org/10.3109/08039488.2014.884631

Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K.-M., & Robbins, K. A. (2015). The PREP pipeline: Standardized preprocessing for large-scale EEG analysis. *Frontiers in Neuroinformatics* , *9* .

BioSemi, Amsterdam. (2021). *BioSemi ActiveTwo Mark 2* .

Bissonnette, J. N., Francis, A. M., Hull, K. M., Leckey, J., Pimer, L., Berrigan, L. I., & Fisher, D. J. (2020). MMN-Indexed Auditory Change Detection in Major Depressive Disorder. *Clinical EEG and Neuroscience* , *51* (6), 365–372. https://doi.org/10.1177/1550059420914200

Botelho, C., Pasion, R., Prata, C., & Barbosa, F. (2023). Neuronal underpinnings of the attentional bias toward threat in the anxiety spectrum: Meta-analytical data on P3 and LPP event-related potentials.*Biological Psychology* , *176* , 108475. https://doi.org/10.1016/j.biopsycho.2022.108475

Bruder, G., Kayser, J., Tenke, C., Leite, P., Schneier, F., Stewart, J., & Quitkin, F. (2002). Cognitive ERPs in Depressive and Anxiety Disorders during Tonal and Phonetic Oddball Tasks. *Clinical EEG (Electroencephalography)* , *33* , 119–124. https://doi.org/10.1177/155005940203300308

C and H Distributors Inc., Milwaukee, WI, USA. (2021). *3M E-A-RTONE 3A and 5A Audiometric Insert Earphones* .

Carragher, N., Ruggero, C., Zimmerman, M., & Baillie, A. J. (2021).*A detailed hierarchical model of psychopathology: From individual symptoms up to the general factor of psychopathology* . 1–59. https://doi.org/10.31219/osf.io/3dp6f

Castellanos, N. P., & Makarov, V. A. (2006). Recovering EEG brain signals: Artifact suppression with wavelet enhanced independent component analysis. *Journal of Neuroscience Methods* ,*158* (2), 300–312. https://doi.org/10.1016/j.jneumeth.2006.05.033

Chambless, D. L., Caputo, G. C., Jasin, S. E., Gracely, E. J., & Williams, C. (1985). The Mobility Inventory for Agoraphobia.*Behaviour Research and Therapy* , *23* (1), 35–44. https://doi.org/10.1016/0005-7967(85)90140-8

Cho, H., Myruski, S., Denefrio, S., Mennin, D. S., & Dennis-Tiwary, T. A. (2022). Associations between GAD symptom severity and error monitoring depend on neural quenching variability. *Motivation and Emotion* , *46* (2), 254–263. https://doi.org/10.1007/s11031-021-09923-0

Clark, L. A., Cuthbert, B., Lewis-Fernández, R., Narrow, W. E., & Reed, G. M. (2017). Three Approaches to Understanding and Classifying Mental Disorder: ICD-11, DSM-5, and the National Institute of Mental Health's Research Domain Criteria (RDoC). *Psychological Science in the Public Interest* , *18* (2), 72–145. https://doi.org/10.1177/1529100617727266

Clayson, P. E. (2020). Moderators of the internal consistency of error-related negativity scores: A meta-analysis of internal consistency estimates. *Psychophysiology* , *57* (8), 1–27. https://doi.org/10.1111/psyp.13583

Clayson, P. E., Baldwin, S. A., & Larson, M. J. (2013). How does noise affect amplitude and latency measurement of event-related potentials (ERPs)? A methodological critique and simulation study.*Psychophysiology* , *50* (2), 174–186. https://doi.org/10.1111/PSYP.12001

Clayson, P. E., Brush, C. J., & Hajcak, G. (2021). Data quality and reliability metrics for event-related potentials (ERPs): The utility of subject-level reliability. *International Journal of Psychophysiology* , *165* , 121–136. https://doi.org/10.1016/J.IJPSYCHO.2021.04.004

Clayson, P. E., Carbine, K. A., Baldwin, S. A., & Larson, M. J. (2019). Methodological reporting behavior, sample sizes, and statistical power in studies of event-related potentials: Barriers to reproducibility and replicability. *Psychophysiology* , *56* (11), 1–17. https://doi.org/10.1111/psyp.13437

Clayson, P. E., Carbine, K. A., & Larson, M. J. (2020). A registered report of error-related negativity and reward positivity as biomarkers of depression: P-Curving the evidence. *International Journal of Psychophysiology* , *150* , 50–72. https://doi.org/10.1016/j.ijpsycho.2020.01.005

Conway, C. C., Forbes, M. K., Forbush, K. T., Fried, E. I., Hallquist, M. N., Kotov, R., Mullins-Sweatt, S. N., Shackman, A. J., Skodol, A. E., South, S. C., Sunderland, M., Waszczuk, M. A., Zald, D. H., Afzali, M. H., Bornovalova, M. A., Carragher, N., Docherty, A. R., Jonas, K. G., Krueger, R. F., ... Eaton, N. R. (2019). A Hierarchical Taxonomy of Psychopathology Can Transform Mental Health Research. *Perspectives on Psychological Science* , *14* (3), 419–436. https://doi.org/10.1177/1745691618810696

Conway, C. C., Forbes, M. K., & South, S. C. (2022). A Hierarchical Taxonomy of Psychopathology (HiTOP) Primer for Mental Health Researchers. *Clinical Psychological Science* , *10* (2), 236–258. https://doi.org/10.1177/21677026211017834

Crocq, M.-A. (2017). The history of generalized anxiety disorder as a diagnostic category. *Dialogues in Clinical Neuroscience* ,*19* (2), 107–116. https://doi.org/10.31887/DCNS.2017.19.2/macrocq

Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: Age, sleep and modality.*Clinical Neurophysiology* , *115* (4), 732–744. https://doi.org/10.1016/j.clinph.2003.11.021

Dali, G., Brosnan, M., Tiego, J., Johnson, B. P., Fornito, A., Bellgrove, M. A., & Hester, R. (2023). Examining the neural correlates of error awareness in a large fMRI study. *Cerebral Cortex* ,*33* (2), 458–468. https://doi.org/10.1093/cercor/bhac077

Delacre, M., Leys, C., Mora, Y. L., & Lakens, D. (2019). *Taking Parametric Assumptions Seriously: Arguments for the Use of Welch's F-test instead of the Classical F-test in One-Way ANOVA* . *32* (1), 13. https://doi.org/10.5334/irsp.198

Dell'Acqua, C., Hajcak, G., Amir, N., Santopetro, N. J., Brush, C. J., & Meyer, A. (2023). Error-related brain activity in pediatric major depressive disorder: An ERP and time-frequency investigation.*International Journal of Psychophysiology* , *184* , 100–109. https://doi.org/10.1016/j.ijpsycho.2023.01.005

Delorme, A., & Makeig, S. (2004). EEGLAB: An open sorce toolbox for analysis of single-trial EEG dynamics including independent component anlaysis. *Journal of Neuroscience Methods* . https://doi.org/10.1016/j.jneumeth.2003.10.009

DeYoung, C., Blain, S., Latzman, R., Grazioplene, R., Haltigan, J., Kotov, R., Michelini, G., Venables, N., Goghari, V., Kallen, A., Martin, E., Palumbo, I., Patrick, C., Perkins, E., Shackman, A., Snyder, M. E., & Tobin, K. (2023). *The Hierarchical Taxonomy of Psychopathology (HiTOP) and the search for neurobiological substrates of mental illness: A review and roadmap for future researc* . https://doi.org/10.31234/osf.io/yatw7

Donaldson, K. R., Novak, K. D., Foti, D., Marder, M., Perlman, G., Kotov, R., & Mohanty, A. (2020). Associations of mismatch negativity with psychotic symptoms and functioning transdiagnostically across psychotic disorders. *Journal of Abnormal Psychology* , *129* , 570–580. https://doi.org/10.1037/abn0000506

DSM-IV-TR. (2000). *Diagnostic and statistical manual of mental disorders* . American Psychiatric Association.

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

Faure, K., & Forbes, M. K. (2021). Clarifying the Placement of Obsessive-Compulsive Disorder in the Empirical Structure of Psychopathology. *Journal of Psychopathology and Behavioral Assessment* . https://doi.org/10.1007/s10862-021-09868-1

Feuerriegel, D., & Bode, S. (2022). Bring a map when exploring the ERP data processing multiverse: A commentary on Clayson et al. 2021.*NeuroImage* , *259* , 119443. https://doi.org/10.1016/j.neuroimage.2022.119443

Fields, E. C., & Kuperberg, G. R. (2020). Having your cake and eating it too: Flexibility and power with mass univariate statistics for ERP data. *Psychophysiology* , *57* (2), e13468. https://doi.org/10.1111/psyp.13468

Fischer, A. G., Danielmeier, C., Villringer, A., Klein, T. A., & Ullsperger, M. (2016). Gender Influences on Brain Responses to Errors and Post-Error Adjustments. *Scientific Reports* , *6* (1), 24435. https://doi.org/10.1038/srep24435

Folstein, J. R., & Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: A review.*Psychophysiology* , *45* (1), 152–170. https://doi.org/10.1111/j.1469-8986.2007.00602.x

Forbes, M. K. (2020). *A broad and detailed hierarchical model of psychopathology: From individual symptoms up to the general factor of psychopathology* .

Francis, A. M., Bissonnette, J. N., Hull, K. M., Leckey, J., Pimer, L., Berrigan, L. I., & Fisher, D. J. (2021). Alterations of novelty processing in major depressive disorder. *Journal of Affective Disorders Reports* , *4* , 100083. https://doi.org/10.1016/j.jadr.2021.100083

Gallinat, J., Bottlender, R., Juckel, G., Munke-Puchner, A., Stotz, G., Kuss, H.-J., Mavrogiorgou, P., & Hegerl, U. (2000). The loudness dependency of the auditory evoked N1/P2-component as a predictor of the acute SSRI response in depression. *Psychopharmacology* ,*148* (4), 404–411. https://doi.org/10.1007/s002130050070

Gloe, L. M., & Louis, C. C. (2021). The Error-Related Negativity (ERN) in Anxiety and Obsessive-Compulsive Disorder (OCD): A Call for Further Investigation of Task Parameters in the Flanker Task. *Frontiers in Human Neuroscience* , *15* .

Gohle, D., Juckel, G., Mavrogiorgou, P., Pogarell, O., Mulert, C., Rujescu, D., Giegling, I., Zaudig, M., & Hegerl, U. (2008). Electrophysiological evidence for cortical abnormalities in obsessive-compulsive disorder - A replication study using auditory event-related P300 subcomponents. *Journal of Psychiatric Research* , *42* (4), 297–303. https://doi.org/10.1016/j.jpsychires.2007.01.003

González-Robles, A., Díaz-García, A., Miguel, C., García-Palacios, A., & Botella, C. (2018). Comorbidity and diagnosis distribution in transdiagnostic treatments for emotional disorders: A systematic review of randomized controlled trials. *PLOS ONE* , *13* (11), e0207396. https://doi.org/10.1371/journal.pone.0207396

Granros, M. P. (2021). *Neural Reactivity to Affective Stimuli and Internalizing Symptom Dimensions in a Transdiagnostic Sample* [Thesis, University of Illinois at Chicago]. https://doi.org/10.25417/uic.17025860.v1

Härpfer, K., Carsten, H. P., Löwisch, K., Westermann, N., & Riesel, A. (2022). Disentangling the effects of trait and state worry on error-related brain activity: Results from a randomized controlled trial using worry manipulations. *Psychophysiology* , *n/a* (n/a), e14055. https://doi.org/10.1111/psyp.14055

Harris, P. A., Taylor, R., Minor, B. L., Elliott, V., Fernandez, M., O'Neal, L., McLeod, L., Delacqua, G., Delacqua, F., Kirby, J., & Duda, S. N. (2019). The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics* ,*95* , 103208. https://doi.org/10.1016/j.jbi.2019.103208

Hettema, J. M. (2008). The nosologic relationship between generalized anxiety disorder and major depression. *Depression and Anxiety* ,*25* (4), 300–316. https://doi.org/10.1002/da.20491

Hochman, E. Y., Vaidya, A. R., & Fellows, L. K. (2014). Evidence for a Role for the Dorsal Anterior Cingulate Cortex in Disengaging from an Incorrect Action. *PLoS ONE* , *9* (6), e101126. https://doi.org/10.1371/journal.pone.0101126

Houck, P. R., Spiegel, D. A., Shear, M. K., & Rucci, P. (2002). Reliability of the self-report version of the panic disorder severity scale. *Depression and Anxiety* , *15* (4), 183–185. https://doi.org/10.1002/da.10049

Howe, A. S., Pinto, A., & De Luca, V. (2014). Meta-analysis of P300 waveform in panic disorder. *Experimental Brain Research* ,*232* (10), 3221–3232. https://doi.org/10.1007/s00221-014-3999-5

Imburgio, M. J., Banica, I., Hill, K. E., Weinberg, A., Foti, D., & MacNamara, A. (2020). Establishing norms for error-related brain activity during the arrow Flanker task among young adults.*NeuroImage* , *213* (March), 116694. https://doi.org/10.1016/j.neuroimage.2020.116694

Iwanami, A., Isono, H., Okajima, Y., & Kamijima, K. (1997). Auditory event-related potentials in panic disorder. *European Archives of Psychiatry and Clinical Neuroscience* , *247* (2), 107–111. https://doi.org/10.1007/BF02900202

Kałamała, P., Szewczyk, J., Senderecka, M., & Wodniecka, Z. (2018). Flanker task with equiprobable congruent and incongruent conditions does not elicit the conflict N2. *Psychophysiology* , *55* (2), e12980. https://doi.org/10.1111/psyp.12980

Kiebel, S. J., & Friston, K. J. (2004). Statistical parametric mapping for event-related potentials (II): A hierarchical temporal model.*NeuroImage* , *22* (2), 503–520. https://doi.org/10.1016/j.neuroimage.2004.02.013

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., Santopetro, N. J., Meyer, A., & Hajcak, G. (2020). Reduced P300 in depression: Evidence from a flanker task and impact on ERN, CRN, and Pe. *Psychophysiology* , *57* (4), e13520. https://doi.org/10.1111/PSYP.13520

Klug, M., & Kloosterman, N. A. (2021). *Zapline-plus: A Zapline extension for automatic and adaptive removal of frequency-specific noise artifacts in M/EEG* . 20.

Kotov, R., Cicero, D. C., Conway, C. C., DeYoung, C. G., Dombrovski, A., Eaton, N. R., First, M. B., Forbes, M. K., Hyman, S. E., Jonas, K. G., Krueger, R. F., Latzman, R. D., Li, J. J., Nelson, B. D., Regier, D. A., Rodriguez-Seijas, C., Ruggero, C. J., Simms, L. J., Skodol, A. E., ... Wright, A. G. C. (2022). The Hierarchical Taxonomy of Psychopathology (HiTOP) in psychiatric practice and research.*Psychological Medicine* , *52* (9), 1666–1678. https://doi.org/10.1017/S0033291722001301

Kotov, R., Jonas, K. G., Carpenter, W. T., Dretsch, M. N., Eaton, N. R., Forbes, M. K., Forbush, K. T., Hobbs, K., Reininghaus, U., Slade, T., South, S. C., Sunderland, M., Waszczuk, M. A., Widiger, T. A., Wright, A. G. C., Zald, D. H., Krueger, R. F., & Watson, D. (2020). Validity and utility of Hierarchical Taxonomy of Psychopathology (HiTOP): I. Psychosis superspectrum. *World Psychiatry* , *19* (2), 151–172. https://doi.org/10.1002/wps.20730

Kotov, R., Krueger, R. F., Watson, D., Cicero, D. C., Conway, C. C., DeYoung, C. G., Eaton, N. R., Forbes, M. K., Hallquist, M. N., Latzman, R. D., Mullins-Sweatt, S. N., Ruggero, C. J., Simms, L. J., Waldman, I. D., Waszczuk, M. A., & Wright, A. G. C. (2021). The Hierarchical Taxonomy of Psychopathology (HiTOP): A Quantitative Nosology Based on Consensus of Evidence. *Annual Review of Clinical Psychology* ,*17* (1), 83–108. https://doi.org/10.1146/annurev-clinpsy-081219-093304

Kotov, R., Waszczuk, M. A., Krueger, R. F., Forbes, M. K., Watson, D., Clark, L. A., Achenbach, T. M., Althoff, R. R., Ivanova, M. Y., Michael Bagby, R., Brown, T. A., Carpenter, W. T., Caspi, A., Moffitt, T. E., Eaton, N. R., Forbush, K. T., Goldberg, D., Hasin, D., Hyman, S. E., ... Zimmerman, M. (2017). The hierarchical taxonomy of psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology* , *126* (4), 454–477. https://doi.org/10.1037/abn0000258

Kroenke, K., & Spitzer, R. L. (2002). The PHQ-9: A new depression diagnostic and severity measure. *Psychiatric Annals* ,*32* (9), 509–515. https://doi.org/10.3928/0048-5713-20020901-06

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual Review of Psychology* , *62* , 621–647. https://doi.org/10.1146/annurev.psych.093008.131123

Ladouceur, C. D., Slifka, J. S., Dahl, R. E., Birmaher, B., Axelson, D. A., & Ryan, N. D. (2012). Altered error-related brain activity in youth with major depression. *Developmental Cognitive Neuroscience* ,*2* (3), 351–362. https://doi.org/10.1016/j.dcn.2012.01.005

Lahey, B. B., Moore, T. M., Kaczkurkin, A. N., & Zald, D. H. (2021). Hierarchical models of psychopathology: Empirical support, implications, and remaining issues. *World Psychiatry* , *20* (1), 57–63. https://doi.org/10.1002/wps.20824

Lamers, F., Oppen, P. van, Comijs, H. C., Smit, J. H., Spinhoven, P., Balkom, A. J. L. M. van, Nolen, W. A., Zitman, F. G., Beekman, A. T. F., & Penninx, B. W. J. H. (2011). Comorbidity Patterns of Anxiety and Depressive Disorders in a Large Cohort Study: The Netherlands Study of Depression and Anxiety (NESDA). *The Journal of Clinical Psychiatry* , *72* (3), 3397. https://doi.org/10.4088/JCP.10m06176blu

Larson, M. J., Clawson, A., Clayson, P. E., & Baldwin, S. A. (2013). Cognitive conflict adaptation in generalized anxiety disorder.*Biological Psychology* , *94* (2), 408–418. https://doi.org/10.1016/j.biopsycho.2013.08.006

Larson, M. J., Clayson, P. E., & Clawson, A. (2014). Making sense of all the conflict: A theoretical review and critique of conflict-related ERPs. *International Journal of Psychophysiology* , *93* (3), 283–297. https://doi.org/10.1016/J.IJPSYCHO.2014.06.007

Levin-Aspenson, H. F., Watson, D., Clark, L. A., & Zimmerman, M. (2021). What Is the General Factor of Psychopathology? Consistency of the p Factor Across Samples. *Assessment* , *28* (4), 1035–1049.

https://doi.org/10.1177/1073191120954921

Lightfoot, G. (2016). Summary of the N1-P2 Cortical Auditory Evoked Potential to Estimate the Auditory Threshold in Adults. *Seminars in Hearing* , *37* (1), 1–8. https://doi.org/10.1055/s-0035-1570334

Luck, S. J. (2014). *An introduction to the event-related potential technique* .

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

Luck, S. J., & Kappenman, E. S. (2011). *The Oxford handbook of event-related potential components* . Oxford university press.

Macedo, I., Pasion, R., Barbosa, F., & Ferreira-Santos, F. (2021). A dimensional approach to the neuronal correlates of anxiety, depression, and perfectionism: A transdiagnostic dissociation of error-related brain activity. *Behavioural Brain Research* , *408* , 113271. https://doi.org/10.1016/j.bbr.2021.113271

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods* ,*164* (1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Mathworks. (2022). *MATLAB version 9.3.0.713579 (R2022a)* . The Mathworks, Inc.

Michelini, G., Palumbo, I. M., DeYoung, C. G., Latzman, R. D., & Kotov, R. (2021). Linking RDoC and HiTOP: A new interface for advancing psychiatric nosology and neuroscience. *Clinical Psychology Review* , *86* . https://doi.org/10.1016/j.cpr.2021.102025

Morey, R. D., Hoekstra, R., Rouder, J. N., Lee, M. D., & Wagenmakers, E.-J. (2016). The fallacy of placing confidence in confidence intervals.*Psychonomic Bulletin & Review* , *23* , 103–123. https://doi.org/10.3758/s13423-015-0947-8

Näätänen, R., & Kähkönen, S. (2009). Central auditory dysfunction in schizophrenia as revealed by the mismatch negativity (MMN) and its magnetic equivalent MMNm: A review. *International Journal of Neuropsychopharmacology* , *12* , 125–135. https://doi.org/10.1017/S1461145708009322

Nan, C., Wang, G., Wang, H., Wang, X., Liu, Z., Xiao, L., Bai, H., & Wu, S. (2018). The P300 component decreases in a bimodal oddball task in individuals with depression: An event-related potentials study. *Clinical Neurophysiology* , *129* (12), 2525–2533. https://doi.org/10.1016/j.clinph.2018.09.012

Nawani, H., Narayanaswamy, J. C., Basavaraju, S., Bose, A., Mahavir Agarwal, S., Venkatasubramanian, G., & Janardhan Reddy, Y. C. (2018). Enhanced error related negativity amplitude in medication-naïve, comorbidity-free obsessive compulsive disorder. *Psychiatry Research* , *262* (August 2016), 373–377. https://doi.org/10.1016/j.psychres.2017.09.010

Neurobehavioral Systems, B., Inc. (2021). *Presentation , Version 23.0* .

Nuñez-Estupiñan, X., Berticelli, L. Z., de Almeida, R. M. M., & Gauer, G. (2022). Aversiveness of errors and the error-related negativity (ERN): A systematic review on the affective states' manipulations findings. *Cognitive, Affective, & Behavioral Neuroscience* ,*22* (4), 754–776. https://doi.org/10.3758/s13415-022-01002-2

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

Patel, S. H., & Azzam, P. N. (2005). Characterization of N200 and P300: Selected Studies of the Event-Related Potential.*International Journal of Medical Sciences* , *2* (4), 147–154.

Perera, M. P. N., Bailey, N. W., Herring, S. E., & Fitzgerald, P. B. (2019). *Electrophysiology of obsessive compulsive disorder: A systematic review of the electroencephalographic literature* . Pergamon. https://doi.org/10.1016/j.janxdis.2018.11.001

Perkins, E. R., Joyner, K. J., Patrick, C. J., Bartholow, B. D., Latzman, R. D., DeYoung, C. G., Kotov, R., Reininghaus, U., Cooper, S. E., Afzali, M. H., Docherty, A. R., Dretsch, M. N., Eaton, N. R., Goghari, V. M., Haltigan, J. D., Krueger, R. F., Martin, E. A., Michelini, G., Ruocco, A. C., ... Zald, D. H. (2020). Neurobiology and the Hierarchical Taxonomy of Psychopathology: Progress toward ontogenetically informed and clinically useful nosology. *Dialogues in Clinical Neuroscience* , *22* (1), 51–63. https://doi.org/10.31887/DCNS.2020.22.1/eperkins

Perkins, E. R., Latzman, R. D., & Patrick, C. J. (2019). Interfacing neural constructs with the Hierarchical Taxonomy of Psychopathology : "Why" and "how." *Personality and Mental Health* , *6* , pmh.1460. https://doi.org/10.1002/pmh.1460

Pernet, C. R. (2015). *Cluster-based computational methods for mass univariate analyses of event-related brain potentials/fields: A simulation study | Elsevier Enhanced Reader* . https://doi.org/10.1016/j.jneumeth.2014.08.003

Pernet, C. R., Appelhoff, S., Gorgolewski, K. J., Flandin, G., Phillips, C., Delorme, A., & Oostenveld, R. (2019). EEG-BIDS, an extension to the brain imaging data structure for electroencephalography. *Scientific Data* , *6* (1), 6–10. https://doi.org/10.1038/s41597-019-0104-8

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., Martinez-Cancino, R., Truong, D., Makeig, S., & Delorme, A. (2021). From BIDS-Formatted EEG Data to Sensor-Space Group Results: A Fully Reproducible Workflow With EEGLAB and LIMO EEG. *Frontiers in Neuroscience* , *14* (January), 1–7. https://doi.org/10.3389/fnins.2020.610388

Pernet, C., Mas, I. S., Rousselet, G., Martinez, R., Wilcox, R., & Delorme, A. (2022). Electroencephalography robust statistical linear modelling using a single weight per trial. *Aperture Neuro* ,*2022* (7), 51. https://doi.org/10.52294/2e69f7cc-f061-40ad-ad77-017110464dfd

Pianowski, G., Carvalho, L. de F., & Miguel, F. K. (2019). Investigating the spectra constellations of the hierarchical taxonomy of psychopathology (HITOP) model for personality disorders based on empirical data from a community sample. *Revista Brasileira de Psiquiatria* , *41* (2), 148–152. https://doi.org/10.1590/1516-4446-2018-0015

Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage* , *198* , 181–197. https://doi.org/10.1016/j.neuroimage.2019.05.026

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology* , *118* (10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

Polich, J. (2020). 50+ years of P300: Where are we now? *Psychophysiology* , *57* (7). https://doi.org/10.1111/psyp.13616

R Core Team. (2023). *R: A language and environment for statistical computing* [Manual]. R Foundation for Statistical Computing.

Reinholt, N., Hvenegaard, M., Christensen, A. B., Eskildsen, A., Hjorthøj, C., Poulsen, S., Arendt, M. B., Rosenberg, N. K., Gryesten, J. R., Aharoni, R. N., Alrø, A. J., Christensen, C. W., & Arnfred, S. M. (2021). Transdiagnostic versus Diagnosis-Specific Group Cognitive Behavioral Therapy for Anxiety Disorders and Depression: A Randomized Controlled Trial. *Psychotherapy and Psychosomatics* , *91* (1), 36–49. https://doi.org/10.1159/000516380

Ridderinkhof, K. R., Ramautar, J. R., & Wijnen, J. G. (2009). To PE or not to PE: A P3-like ERP component reflecting the processing of response errors. *Psychophysiology* , *46* (3), 531–538. https://doi.org/10.1111/j.1469-8986.2009.00790.x

Riesel, A., Härpfer, K., Thoma, L., Kathmann, N., & Klawohn, J. (2022). Associations of neural error-processing with symptoms and traits in a dimensional sample recruited across the obsessivecompulsive spectrum.*Psychophysiology* , *60* (2), e14164. https://doi.org/10.1111/psyp.14164

Riesel, A., Klawohn, J., Kathmann, N., & Endrass, T. (2017). Conflict monitoring and adaptation as reflected by N2 amplitude in obsessivecompulsive disorder. *Psychological Medicine* ,*47* (8), 1379–1388. https://doi.org/10.1017/S0033291716003597

Ringwald, W. R., Forbes, M. K., & Wright, A. G. C. (2021). Meta-analysis of structural evidence for the Hierarchical Taxonomy of Psychopathology (HiTOP) model. *Psychological Medicine* . https://doi.org/10.1017/S0033291721001902

Robbins, K. A., Touryan, J., Mullen, T., Kothe, C., & Bigdely-Shamlo, N. (2020). How Sensitive are EEG Results to Preprocessing Methods: A Benchmarking Study. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* , *28* (5), 1–1. https://doi.org/10.1109/tnsre.2020.2980223

Rytwinski, N. K., Fresco, D. M., Heimberg, R. G., Coles, M. E., Liebowitz, M. R., Cissell, S., Stein, M. B., & Hofmann, S. G. (2009). Screening for social anxiety disorder with the self-report version of the liebowitz social anxiety scale. *Depression and Anxiety* ,*26* (1), 34–38. https://doi.org/10.1002/da.20503

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

Schroder, H. S., Moran, T. P., Infantolino, Z. P., & Moser, J. S. (2013). The relationship between depressive symptoms and error monitoring during response switching. *Cognitive, Affective, & Behavioral Neuroscience* , *13* (4), 790–802. https://doi.org/10.3758/s13415-013-0184-4

Seow, T. X. F., Benoit, E., Dempsey, C., Jennings, M., Maxwell, A., McDonough, M., & Gillan, C. M. (2020). A dimensional investigation of error-related negativity (ERN) and self-reported psychiatric symptoms.*International Journal of Psychophysiology* , *158* (October), 340–348. https://doi.org/10.1016/j.ijpsycho.2020.09.019

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., & Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): The Development and Validation of a Structured Diagnostic Psychiatric Interview for DSM-IV and ICD-10. *The Journal of Clinical Psychiatry* , *59* (suppl 20), 11980.

Smith, S. M., & Nichols, T. E. (2009). Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage* , *44* (1), 83–98. https://doi.org/10.1016/j.neuroimage.2008.03.061

Šoškić, A., Styles, S., Kappenman, E., & Kovic, V. (2022). *Garden of forking paths in ERP research effects of varying pre-processing and analysis steps in an N400 experimen* . https://doi.org/10.31234/osf.io/8rjah

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7.*Archives of Internal Medicine* , *166* (10). https://doi.org/10.1001/archinte.166.10.1092

Tanner, D., Norton, J. J. S., Morgan-Short, K., & Luck, S. J. (2016). On high-pass filter artifacts (they're real) and baseline correction (it's a good idea) in ERP/ERMF analysis. *Journal of Neuroscience Methods* , *266* , 166–170. https://doi.org/10.1016/j.jneumeth.2016.01.002

Walhovd, K. B., Rosquist, H., & Fjell, A. M. (2008). P300 amplitude age reductions are not caused by latency jitter. *Psychophysiology* ,*45* (4), 545–553. https://doi.org/10.1111/j.1469-8986.2008.00661.x

Waszczuk, M. A., Eaton, N. R., Krueger, R. F., Shackman, A. J., Waldman, I. D., Zald, D. H., Lahey, B. B., Patrick, C. J., Conway, C. C., Ormel, J., Hyman, S. E., Fried, E. I., Forbes, M. K., Docherty, A. R., Althoff, R. R., Bach, B., Chmielewski, M., DeYoung, C. G., Forbush, K. T., ... Kotov, R. (2020). Redefining phenotypes to advance psychiatric genetics: Implications from hierarchical taxonomy of psychopathology.*Journal of Abnormal Psychology* , *129* (2), 143–161. https://doi.org/10.1037/abn0000486

Watson, D., Levin-Aspenson, H. F., Waszczuk, M. A., Conway, C. C., Dalgleish, T., Dretsch, M. N., Eaton, N. R., Forbes, M. K., Forbush, K. T., Hobbs, K. A., Michelini, G., Nelson, B. D., Sellbom, M., Slade, T., South, S. C., Sunderland, M., Waldman, I., Witthöft, M., Wright, A. G. C., ... Zinbarg, R. E. (2022). Validity and utility of Hierarchical Taxonomy of Psychopathology (HiTOP): III. Emotional dysfunction superspectrum. *World Psychiatry : Official Journal of the World Psychiatric Association (WPA)* , *21* (1), 26–54. https://doi.org/10.1002/WPS.20943

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., 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

WHO. (2004). *ICD-10, International statistical classification of diseases and related health problems* . World Health Organization.

Wilcox, R. R., & Rousselet, G. A. (2018). A Guide to Robust Statistical Methods in Neuroscience. *Current Protocols in Neuroscience* ,*82* (1), 8.42.1–8.42.30. https://doi.org/10.1002/cpns.41

Winkler, I., Debener, S., Muller, K. R., & Tangermann, M. (2015). On the influence of high-pass filtering on ICA-based artifact reduction in EEG-ERP. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS* ,*2015-Novem* , 4101–4105. https://doi.org/10.1109/EMBC.2015.7319296

Xiao, J., Provenza, N. R., Asfouri, J., Myers, J., Mathura, R. K., Metzger, B., Adkinson, J. A., Allawala, A. B., Pirtle, V., Oswalt, D., Shofty, B., Robinson, M. E., Mathew, S. J., Goodman, W. K., Pouratian, N., Schrater, P. R., Patel, A. B., Tolias, A. S., Bijanki, K. R., ... Sheth, S. A. (2023). Decoding Depression Severity from Intracranial Neural Activity. *Biological Psychiatry* , *0* (0). https://doi.org/10.1016/j.biopsych.2023.01.020

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

Yu, Y., Jiang, C., Xu, H., Yang, Q., Li, J., Xu, Y., Xiang, W., Peng, L., Liu, B., Lv, F., & Li, M. (2018). Impaired Cognitive Control of Emotional Conflict in Trait Anxiety: A Preliminary Study Based on Clinical and Non-Clinical Individuals. *Frontiers in Psychiatry* ,*9* .

Yuen, K. K. (1974). The Two-Sample Trimmed t for Unequal Population Variances. *Biometrika* , *61* (1), 165–170. https://doi.org/10.2307/2334299