# Transdiagnostic psychopathology in the light of robust single-trial event-related potentials

Martin Randau[1], Bo Bach[1], Nina Reinholt[1], Cyril Pernet[1], Bob Oranje[1], Belinda Rasmussen[1], and Sidse Arnfred[2]

[1]Copenhagen University Hospital
[2]University of Copenhagen

September 26, 2023

**Abstract**

Recent evidence indicate that event-related potentials (ERPs) as measured on the electroencephalogram (EEG) are more closely related to transdiagnostic, dimensional measures of psychopathology (TDP) than to diagnostic categories. Given this, a comprehensive examination of correlations between well-studied ERPs and measures of TDP is called for. In this study, we recruited 50 patients with emotional disorders undergoing 14 weeks of transdiagnostic group psychotherapy as well as 37 healthy comparison subjects (HC) matched in age and sex. HCs were assessed once and patients three times throughout treatment (N = 172 datasets) with a battery of well-studied ERPs and psychopathology measures consistent with the TDP framework The Hierarchical Taxonomy of Psychopathology (HiTOP). ERPs were quantified using robust single-trial analysis (RSTA) methods and TDP correlations with linear regression models as implemented in the EEGLAB toolbox LIMO EEG. We found correlations at several levels of the HiTOP hierarchy. Among these, a reduced P3b as well as a reduced error-related negativity correlated strongly with worse symptomatology across the Internalizing spectrum. Conversely, increases in the correct-related negativity correlated with symptoms loading unto the Distress subfactor in the HiTOP. Increases in mismatch negativity were primarily related to maladaptive personality traits at the lowest levels of the HiTOP hierarchy. Our study highlights the advantages of RSTA methods and of using validated TDP constructs within a consistent framework such as the HiTOP. Future studies could utilize machine learning methods to predict TDP from a set of ERP features at the subject level.

[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 <camar@regionsjaelland.dk>

## Introduction

The electroencephalogram (EEG) shows promising clinical utility in being non-invasive, easy to record and cost-effective compared to other neuroimaging methods (Hajcak et al., 2019). Event-related potentials (ERPs) on the EEG are indexes of task-related brain activity time-locked to stimuli or other events and potential biomarkers of mental disorders (Hajcak et al., 2019). Several ERPs have been associated with diagnostic categories and severity of diagnosis-specific symptoms (Luck & Kappenman, 2011). However, for

the emotional disorders, one of the main causes of suffering worldwide, ERP findings remain inconsistent (González-Robles et al., 2018; Watson et al., 2022; Widiger & Oltmanns, 2017). While differences between ERP studies in paradigm design and preprocessing and analysis methods limit comparison of some results, discrepancies also stem from issues inherent in categorical taxonomies and which are especially abundant in the emotional disorders (Michelini et al., 2021). Recent evidence indicate that at least some ERPs are more closely related to transdiagnostic measures of psychopathology than to diagnostic categories (Donaldson et al., 2020; Macedo et al., 2021; Pasion & Barbosa, 2019; Riesel et al., 2022). Given this, a comprehensive investigation of the associations between classic ERPs, of which some were discovered in the 1960's, and transdiagnostic measures of psychopathology is called for (Latzman & DeYoung, 2020; Polich, 2020). In this study, we aim to do so by assessing a mixed sample for correlations between ERPs and transdiagnostic measures of psychopathology while accounting for the effects of medication and psychotherapy treatment.

Biomarkers in psychiatry could greatly improve clinical practice in providing objective measures of psychopathology and treatment outcome (Singh & Rose, 2009). The discovery of such markers requires a sound psychopathology framework from which to derive biobehavioral targets to examine (Latzman & DeYoung, 2020; Michelini et al., 2021). ERP studies have traditionally evaluated differences in ERP measures such as peak or average amplitude or latency between diagnostic groups based on categorical taxonomies ICD and DSM (DSM-IV-TR., 2000; WHO, 2004). These taxonomies posit that mental disorders are discrete entities with specific symptoms and clear-cut boundaries between healthy and ill and between diagnoses (Clark et al., 2017). However, clinical reality shows that comorbidity among the emotional disorders is in the range of 40 to 80% and that symptom profiles of patients with the same diagnosis varies greatly (González-Robles et al., 2018). This suggests that categorical taxonomies do not capture the true nature of psychopathology (Clark et al., 2017; González-Robles et al., 2018). There being no straightforward way to account for comorbidity in case-control designs, most ERP studies merely report concurrent diagnoses and rely on the primary diagnosis as a sufficient classification of the sample (Petrolini & Vicente, 2022; Zald & Lahey, 2017). To see why this can be problematic, consider the error-related negativity (ERN) which is robustly enhanced (increased amplitude) in some anxiety disorders such as obsessive-compulsive disorder (OCD) compared to healthy comparison subjects (Macedo et al., 2021). Somewhat inconsistent results indicate that ERN is attenuated (decreased amplitude) in depression (Klawohn et al., 2020). It is clear to see how a study examining ERN in depression while not accounting for anxiety-related comorbidity might end up with null results. Heterogeneity within disorders and arbitrary boundaries between healthy and ill pose similar loss-of-information problems in studies based on categorical taxonomies (Michelini et al., 2021). In fact, with the notable exception of ERN in OCD, decades of research has revealed no robust deviations in ERPs or other EEG measures in any of the emotional disorders as defined in the categorical taxonomies, e.g., depression or major depressive disorder (MDD) (de Aguiar Neto & Rosa, 2019), generalized anxiety disorder (GAD) (Maron & Nutt, 2022), panic disorder (PD) (Howe et al., 2014) and social anxiety disorder (SAD) (Al-Ezzi et al., 2020).

Alternative frameworks of psychopathology transcends arbitrary diagnostic boundaries in considering transdiagnostic symptoms which are shared among disorders as the basic building blocks of mental disorders (Clark et al., 2017). The Hierarchical Taxonomy of Psychopathology (HiTOP) is an empirical and data-driven attempt to describe the full range of psychopathology (Kotov et al., 2017). In the HiTOP, transdiagnostic symptoms and maladaptive traits at the lowest level of the hierarchy are clustered based on shared features into subfactors roughly corresponding to categorical diagnoses, which in turn are joined into higher-level spectra such as the Internalizing and Thought disorder spectra. Accordingly, the emotional disorders share core symptoms and traits but are further up in the hierarchy allocated to the *Fear*and *Distress* subfactors, the latter containing depression and GAD. The HiTOP comes with several advantages for neuroimaging research (Conway et al., 2022; Corr & Mobbs, 2023; Kotov et al., 2022; Latzman & DeYoung, 2020; Michelini et al., 2021; Perkins et al., 2019). By design, the HiTOP deals with comorbidity, symptom heterogeneity within disorders and arbitrary boundaries between healthy and ill. In contrast to categorical taxonomies, the HiTOP encourages studies of mixed samples characterized at different levels of the hierarchy capturing the full range of psychopathology (Conway et al., 2022; Latzman & DeYoung, 2020). In other words, subjects included in a study based on the HiTOP do not need to fulfill some diagnostic criteria or score above

some threshold, but are fully characterized in terms of homogeneous dimensional constructs. A given ERP measure can thereby be investigated in terms of being a marker of a transdiagnostic symptom or trait, of a subfactor or of a whole spectrum. Relying on the HiTOP is also advantageous when selecting biobehavioral targets with which associations to ERPs are sought. For transdiagnostic measures, rather than using sub scales of rating scales developed in categorical setting, measures consistent with the HiTOP would directly place results in the context of a comprehensive empirical model of mental disorders (Perkins et al., 2019).

Ample evidence support that biological measures align more closely to transdiagnostic constructs than to diagnostic categories (Kotov et al., 2020; Waszczuk et al., 2020; Watson et al., 2022). The Research Domain Criteria (RDoC) was launched to encourage research into such biobehavioral constructs cutting across diagnostic boundaries (Cuthbert & Insel, 2010). In line with this, several ERPs have recently been recast as markers of transdiagnostic psychopathology. ERN and its counter-part, the correct-related negativity (CRN), previously solely associated with OCD, are now conceived as markers of specific transdiagnostic measures of anxiety and negative affect in the Internalizing spectrum (Macedo et al., 2021; Pasion & Barbosa, 2019; Riesel et al., 2022). Mismatch negativity (MMN) and some other ERP components in auditory oddball paradigms do not seem to be uniquely related to a chronic diagnosis of schizophrenia but to symptoms shared by a range of psychotic disorders (Donaldson et al., 2020; Parker et al., 2021). These developments being very recent, only the ERN, CRN, and to a lesser extent the late-positive potential, have thus far been cast in the light of transdiagnostic psychopathology in the emotional disorders (Granros, 2021). A comprehensive investigation of the associations between other classic ERPs and transdiagnostic markers of psychopathology in the emotional disorders is lacking. Conversely, further validation of the HiTOP with biological measures is called for (Perkins et al., 2019).

The aim of the present study was to examine the associations between a set of transdiagnostic measures of psychopathology and a range of ERPs elicited by thee classic paradigms (the Eriksen Flanker, the auditory Attended Oddball and the auditory Unattended Oddball) (Luck & Kappenman, 2011). Measures of transdiagnostic psychopathology were assessed with validated self-report measures covering symptoms and traits consistent with the HiTOP Internalizing spectra. We included 50 patients with emotional disorders undergoing 14 weeks of UP transdiagnostic group cognitive behavioral psychotherapy and 37 healthy comparison subjects (HC) matched in age and sex (Barlow et al., 2017; Reinholt et al., 2021). Patients were assessed with EEG and self-report questionnaires three times: before, 10 weeks into, and within one week after treatment. The majority of HCs were assessed once but some a second time after at least two months in order to account for normal variation in the models.

To evaluate the associations between ERPs and measures of transdiagnostic psychopathology, we conducted robust mass univariate linear regression based on single-trial ERP analysis as implemented in the EEGLAB toolbox LIMO EEG (Delorme & Makeig, 2004; C. R. Pernet et al., 2011). LIMO EEG is based on statistical parametric mapping (SPM), as in the analysis of fMRI data, and provides a complete workflow from preprocessed EEG data to the evaluation of single-trial subject-level ERPs at group level with a range of robust statistical measures (Kiebel & Friston, 2004; C. R. Pernet et al., 2021). Employing a hierarchical generalized linear model (GLM) approach, the method makes redundant several choices required in traditional ERP methods known to inflate false positives and influence group level statistics (Feuerriegel & Bode, 2022; Luck & Gaspelin, 2017). Instead of requiring the *á priori* selection of channel and time window regions of interest, as well as methods for peak or average amplitude extraction, LIMO EEG models the subject-level single-trial GLM across all channels and time points concurrently. False positives are controlled through bootstrap methods and threshold-free cluster enhancement (TFCE) (Maris & Oostenveld, 2007; Mensen & Khatami, 2013; C. R. Pernet, 2015). Consequently, the investigate scope is vastly expanded without loss of statistical power and can reveal effects at other channels and time periods than what is traditionally investigated (Fields & Kuperberg, 2020). Recognizing current issues in the preprocessing of ERP data, we relied on a novel cleaning pipeline based on an empirical evaluation of other well-established pipelines, the Reduction of Electroencephalographic Artifacts (RELAX) Bailey et al. (2022); Bailey et al. (2023). Given evidence that robust single-trial methods allows for less aggressive cleaning of ERP data, thereby preserving more brain activity, we applied a less strict than default cleaning of artifacts and noise (Alday & van Paridon, 2021;

Delorme, 2022).

Establishing associations between measures of transdiagnostic psychopathology and ERPs, many of which are related to specific neural functioning, would be an important step toward biomarkers in psychiatry and would increase our understanding of the neural basis of mental disorders (Hajcak et al., 2019; Lavoie et al., 2019).

Given the exploratory nature of the study, we refrain from making specific hypotheses. However, as found in two recent studies, we expected the ERN to be related to one or more measures in the Internalizing spectrum (Macedo et al., 2021; Riesel et al., 2022).

# Methods

## Ethics

The study was approved by the Danish National Committee on Health Research Ethics (VEK journal ID 74188). Data management and privacy policies was approved by the Danish Data Protection Agency (journal ID REG-131-2020). Informed consent was provided by all participants. An honorary fee of 500 Danish Crowns per EEG session was provided to all participants.

## Participants

Patients (N = 50) of both sexes aged 18 to 59 with a primary 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 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 medical history, and normal mental capabilities as estimated by having completed Danish primary school.

Healthy comparison subjects (HC, N = 37), matched with the patient group 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.

## Clinical 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 by a psychiatry trainee (MR) (Sheehan, 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 screened for the absence of symptoms fulfilling criteria for any psychiatric diagnosis.

## Psychometrics

The battery of psychopathology measures consisted of several validated self-report questionnaires. While rating-scales derived directly from the HiTOP are under development and require local translation and validation, the items selected in this study were deemed to be sufficiently consistent with the HiTOP (Wendt et al., 2021).

The Multidimensional Emotional Disorder Inventory (MEDI, 49 items ranged 0 to 8) assesses nine empirically-supported transdiagnostic symptom dimensions within the Internalizing spectrum: autonomic arousal, avoidance, depression, intrusive cognitions, neurotic temperament, positive temperament, somatic anxiety, social anxiety and traumatic re-experiencing (Rosellini & Brown, 2019). Note that when calculating the MEDI total score, Positive temperament is subtracted rather than added.

The Modified Personality Inventory for DSM-5 and ICD-11 – Brief Form Plus (PID5BF+M, here referred to as PID36, 36 items ranged 0 to 3) assesses six personality trait domains in the Internalizing and Thought disorder spectra: anankastia, antagonism, detachment, disinhibition, negative affectivity and psychoticism (Bach et al., 2020). Note that this version is developed in accordance with the coming ICD-11, which is moving toward a dimensional understanding of the personality disorders (Bach & Mulder, 2022).

In addition, two shorter self-report questionnaires were administered to assess the severity of the transdiagnostic dimensions personality pathology and psychological distress, respectively: the Level of Personality Functioning Scale-Brief Form (LPFS, 12 items ranged 1 to 4) (Hutsebaut et al., 2016) and the K10 distress scale (K10, 10 items ranged 1 to 5) (Kessler et al., 2003).

For both HC and patients, this battery was administered in conjunction with the respective EEG recordings. Participants were instructed to answer +/- 1 week from the EEG recording. All psychopathology measures were obtained using the online survey and database management web application REDCap licensed to Region Zealand, Denmark (Harris et al., 2019).

## 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. The first session, baseline, lasted approximately three hours including information, electrode and cap application, EEG recording (˜1 hour), M.I.N.I diagnostic interview and breaks. Sessions two and three, for patients after 10 and 14 weeks, respectively, lasted at most two hours and consisted only of the EEG recording. In order to account for normal variation in the statistical models, some HCs were invited to a second recording after at least 8 weeks. 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 extended 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 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 (N = 1500 in 5 blocks) delivered monoaurally 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 resarch, 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 (N = 480 in 10 blocks) could be either congruent (<<<<< or >>>>>) or incongruent (<<><< or >><>>) and were presented for 200 ms. Trials were 50% congruent and 50% incongruent 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 (N = 1800 in 6 blocks) were delivered monoaurally 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 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. Across all sessions, on average 59.7 channels remained for the AO, 59.8 for the Flanker and 60.2 for the UO paradigm. There was no difference between groups in number of removed channels at baseline.

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.

Across all sessions, on average 1070 epochs of all trial types remained for the AO, 417 for the Flanker paradigm and 972 for the UO paradigm. At baseline, in the AO paradigm, there was a significant difference between groups in number of remaining epochs (HC: 1093, Patients: 1056; $t(85) = 3.32$ , $p = 0.001$). This difference was driven by nearly non-significant differences in the number of remaining Standard stimuli epochs ($t(85) = 1.90$, $p = 0.061$) and Target stimuli epochs ($t(85) = 1.77$, $p = 0.08$). This small difference was deemed 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). ERP 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-trial data were processed in LIMO EEG. For a given subject, session and channel, this first-level of the GLM has the general form where denotes the single-trial ERP data in the form , is a design matrix coding for the paradigm-specific stimulus types, are the first-level beta coefficients to be estimated and is the residual term representing what is left when the effects of the beta coefficients are accounted for.

The term , in LIMO EEG referred to as the adjusted mean, warrants special attention, as effects on the other beta parameters are modulations around this constant term. For example, the response-locked Flanker model is , where is the beta coefficient corresponding to correct responses and corresponds to error responses. Accordingly, . Given the near-identical triphasic waveform of the CRN and the ERN, if is more negative-going than , necessarily lies in-between. Therefore, will be positive-going even though the CRN is a negative-going wave. As such, if a result indicates that is modulated negatively by a psychopathology measure, the

interpretation is that a greater, or more negative, CRN correlates with higher scores.

First-level 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 all ten ERP beta models were evaluated, each containing one or more classic ERP components (see *Table 1* for stimulus types and associated ERP models).

At the second level, for each of these ERP beta models, we applied mass univariate robust linear regression as implemented in LIMO EEG. Age, gender, group, medication status, session number and psychopathology measure corresponding to each session were explanatory variables. The general form of the model was:

where are the second-level beta coefficients to be estimated, is the explanatory variables data matrix and is the first-level ERP beta model defined above.

In this linear regression model, gender was coded as female = 1, male = -1. Group was coded as HC = 1, Patient = -1. Due to the large variety in dosage and type, medication status was coded with two dummy variables denoting no prescription (-1, -1), one prescription (1, -1) 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. Session was coded with three columns indicating with 1 or 0 whether the particular entry of the data matrix belonged to baseline, week 10 or week 14. Psychopathology measure was likewise coded in three columns indicating scores for the associated session. Accordingly, had nine columns and as many rows as there were data sets (N = 172). For example, a given row corresponding to data set 125 had the form denoting the subject's age = 25, gender = -1 (male), group = -1 (Patient), medication = [1 1] (more than one prescription), Session = [... 0 1 0 ...] (week 10) and psychopathology measure = [... 0 39 0] (week 10).

Maximum likelihood estimates of were computed at each time frame, channel by channel, using iterative re-weighted least squares (IRLS). IRLS is a robust extension to OLS adding weights to outlier subjects, and has been shown to increase sensitivity in the analysis of neuroimaging data (Wager et al., 2005). had the form representing the effects of each of the explanatory variables (plus a constant term) on the ERP model at each data point.

Next, a linear combination of these second-level beta coefficients was used to test for significant effects of the psychopathology measures (Kiebel & Friston, 2004). Specifically, we defined a reduced model by applying the contrast and tested, channel by channel, at each time frame the null hypothesis where is the transpose of the contrast vector. In other words, we tested the null hypothesis of no effect on of the psychopathology measures while accounting for the other explanatory variables. Note that this contrast model did not assess whether psychotherapy treatment changes ERP features or modulates the association with psychopathology measures, or whether associations are present only at a given session, e.g. at baseline. On the upside, the model allowed us to state that detected associations were present across groups and sessions irrespective of effects of psychotherapy.

The associated one-sided t-test was:

where is the variance of the full model and the weights estimated with IRLS applied to .

The result of these many one-sided t-tests was an uncorrected statistical parametric map (SPM) of size , e.g., $t$ -values for ERP models in the AO paradigm. Correction for multiple comparisons (MC) was conducted using threshold-free cluster enhancement (TFCE) as implemented in LIMO EEG using 1000 bootstrap iterations (Mensen & Khatami, 2013; C. R. Pernet, 2015; Smith & Nichols, 2009). TFCE builds on traditional bootstrap or permutation-based cluster MC correction methods commonly used in neuroimaging research, e.g. spatiotemporal clustering (Maris & Oostenveld, 2007; Sassenhagen & Draschkow, 2019). However, instead of pre-specifying a cluster-forming threshold and assigning to a cluster all connected data points whose corresponding $t$ -value is above this threshold, the method considers clusters formed at all possible thresholds. The more clusters a given data points belongs to within the range of thresholds, the higher is the assigned TFCE score. As a result, whereas in traditional spatiotemporal clustering methods, the threshold

would influence what type of clusters are detected, in TFCE narrow clusters with high t-values are equalized with broad clusters with lower t-values (Smith & Nichols, 2009). At each data point, , the TFCE score is given by:

where and are the minimum and maximum t-values in the data, respectively, is the cluster extent, is the cluster height and and are scaling constants, which in LIMO EEG are fixed to 0.5 and 2, respectively (C. R. Pernet, 2015). To arrive at the final corrected SPM of significant $t$ -values, the method proceeds with estimating the empirical TFCE distribution through bootstrapping. Importantly, sampling is with replacement of all datasets belonging to a subject. Then, the maximum TFCE values from each bootstrap are sorted and the value at is the estimated TFCE threshold, where is a pre-determined significance threshold and are the number of bootstrap iterations. Data points whose TFCE score exceeds this threshold are deemed significant at the level and the corresponding $t$ -values are included in the SPM. Note that a trade-off for the increased cluster-detection capabilities of TFCE is that one cannot state which of the included data points make a cluster significant (Smith & Nichols, 2009).

For our main analysis, we tested the associations between the 10 ERP beta models defined in *Table 1* and the four transdiagnostic psychopathology measures (K10, LPFS, MEDI and PID36) at an level of or 0.1% In case of significant results for MEDI and PID36 total scores, we also show results for the respective sub scales. These results are presented in full in Supplementary Materials and commented upon in Discussion.

Results are presented as heat maps indicating in red or blue with varying intensity positive or negative $t$ -values, respectively. Clusters of these $t$ -values denote spatiotemporal regions where the effects of psychopathology measures on the ERP model were significant. As such, interpretation of results is in terms of direction of effects at relevant regions of interest. To this end we also display shaded regions indicating traditional ERP analysis time windows.

ERP grand averages are displayed for each stimulus type and group (HC and Patient) 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).

Finally, we also show results from statistical analyses of demographic and psychopathology measures. To test the internal consistency of MEDI and PID36 we estimated McDonald's Omega using functions from the R package *semTools* applied to the baseline dataset (N = 87) (Jorgensen et al., 2022). McDonald's Omega has been suggested to be a more reliable estimate than the commonly used Cronbach's Alpha (Flora, 2020). To test for group differences in demographics and psychopathology measures we applied Welch's two-sample t-tests for continuous variables and Fisher's exact test for the categorical variable Gender. To test for change of psychopathology measures across sessions, from baseline to week 10 and 14, respectively, we applied mixed linear regression models using the R package *lme4* , e.g., (Bates et al., 2015). Confidence intervals and p-values were computed with a Wald t-distribution approximation (Luke, 2017).

# Results

A total of 172 datasets entered the analysis. These were distributed accordingly: 37 HC and 50 patients had baseline data; 15 HC and 34 patients (68%) had data from session 2; 36 patients (72%) had data from session 3. All cases of dropout were related to issues with treatment, e.g., the patient was, contrary to initial evaluation, deemed not suitable for psychotherapy or was rejected due to too low attendance. Note that for HC, session 2 was not after 10 weeks but after $> 8$ weeks, as HC did not participate in treatment.

## Demographics and behavioral measures

All participants were right-handed save for 5 in the HC group and 1 in the patient group and all participants reported normal hearing. Diagnoses were distributed accordingly: agoraphobia: 4; depression: 20; GAD: 5; OCD: 12; PD: 8; and SAD: 1. Seven patients (14%) received no medication, 34 patients (68%) received one type of medication and the remaining nine patients (18%) received more than one medication. In total 86% of patients received at least one psychiatric medication, of which all received at least one type of selective serotonin re-uptake inhibitors (SSRI). No patients were treated with anti-psychotic medication.

Table 2 shows demographics and behavioral measures at baseline for the two groups. There was no difference between groups in sex or 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 stimuli in the AO paradigm. The mean number of Error trials for each group was high, 53.6 for HC and 47.4 for patients, on average well above the recommended minimum of 17 trials for reliable (traditional) estimation of the ERN (Clayson, 2020). However, four patients were below 17 remaining trials (5, 8, 13 and 16, respectively). Note that in WLS, parameter estimation is on the total number of trials, which in the Flanker paradigm was well above 250 trials for all subjects. Upon visual inspection of the grand average and beta coefficient plots for these subjects, the data and parameter estimation were deemed to be of sufficient quality, in all cases showing the characteristic response-locked Flanker triphasic waveform. Nevertheless, although no empirical value exists, it must be noted that the more available trials, the more precise the estimation of variables (C. Pernet et al., 2022).

>> Table 2 here <<

## Psychometrics

Table 3 shows McDonald's Omega as a measure of internal consistency as well as results from group comparisons for all self-report questionnaires at baseline.

>> Table 3 here <<

All psychopathology measures showed good internal consistency (Omega $> 0.7$). The patient grouped scored significantly higher than HC on all items with two exceptions: Positive temperament in MEDI, where the patient group scored significantly lower than HC, and the Antagonism personality trait in PID36 where there was no significant difference between the groups.

Table 4 shows change in psychopathology measures across sessions assessed using mixed linear models with subject as random factor and baseline (Session 1) as reference, e.g., (Bates et al., 2015).

>> Table 4 here <<

For MEDI, treatment significantly reduced total score and almost all of the of sub scale scores. For some dimensions, e.g., Neurotic Temperament, a reduction from baseline was only significant at week 14 after the end of treatment. For others, e.g., Intrusive Cognitions, significant and lasting reduction could be measured already at week 10. Interestingly, Positive Temperament improved significantly at 10 weeks, but effects diminished to not significant compared to baseline at week 14. Traumatic Re-experiencing was the only symptom dimension in MEDI not significantly changed by treatment.

Somewhat surprisingly, treatment did not lead to an overall improvement in PID36 total score. Among the PID36 sub scales, reduction was only in Detachment.

Finally, both K10 and LPFS showed significant improvement due to treatment already at 10 weeks.

## ERP grand average waveforms

Figures 1 to 4 show the group-wise 20% trimmed mean of mean weighted subject-level single-trial ERP data across stimulus types for all paradigms at baseline. Shaded areas indicate the 95% HDI. Well-known ERP components are marked on each plot and appear in agreement with the literature. For the UO ERPs, it is interesting to note that cMMN, the MMN in the Combined difference wave, is a mixture of fMMN and dMMN in that both the earlier-detected frequency change as well as the later-detected duration change are captured in the waveform. As such, the cMMN has two peaks.

### Attended oddball

>> Figure 1 here <<

### Response-locked Flanker

>> Figure 2 here <<

### Stimulus-locked Flanker

>> Figure 3 here <<

### Unattended oddball

>> Figure 4 here <<

## ERP beta coefficients

Figure 5 shows beta time courses at FCz corresponding to stimuli in each paradigm.

>> Figure 5 here <<

, in LIMO EEG called the adjusted mean, is in itself not a measure of brain activity but depends on the other beta coefficients which are modulations around this constant term. Note that the beta coefficient to Target stimuli eliciting the P3b is more suitably plotted at Pz where it reaches maximum (see Figure 7 below). Note also that the Correct trial beta coefficient is positive-going even though the CRN is a negative-going wave. Finally, note that plots for the UO paradigm are difference contrasts between deviant and standard beta coefficients.

## ERP correlations with psychometrics

Due to the many correlation analyses (10 ERP models * 4 psychopathology measures), we present only those for which TFCE detected large significant regions at the corrected level of 0.1%. Full results at an uncorrected TFCE of 5% are available in Supplementary materials. Due to TFCE often detecting many and overlapping significant regions and since we cannot know which of the voxels in each cluster are significant, we use visual inspection to describe significant regions in terms of start and end times and general cortical regions. Throughout the analysis we interpret effects on ERPs in terms of changes in amplitudes even though we cannot rule out that some effects are due to differences in peak latency rather than peak amplitude. That being said, correlations between higher scores and changes in amplitude within a given time interval is valid irrespective of whether effects are due to increased peak latency or reduced peak amplitude *within the significant region* . In addition, most significant regions span at least 50 ms, making it unlikely that the observed effects are due to increases in peak latency.

**K10 Distress Scale (K10)**

Figure 5 shows results for K10 assessing general psychological distress. We found several correlations between K10 and ERPs from the AO and response-locked Flanker paradigms. There were no significant correlations between K10 and ERP models from the stimulus-locked Flanker or UO paradigms at the main analysis level.

>> Figure 6 here <<

For Target stimuli in the AO paradigm, we found a large significant region from 324 ms to 696 ms covering both frontal, central and parietal channels. Figure 7 shows the Target stimulus beta along with the adjusted mean at Pz.

>> Figure 7 <<

This region clearly corresponded to the P3b and the negative$t$ -values at central-parietal and parietal channels, where is more positive-going than , indicated that a reduced P3b is associated with higher scores in K10.

For Standard stimuli in the AO paradigm, we found a significant region from 304 to 372 ms centered at CPz, Pz and POz. This region immediately succeeds the N2, which is commonly analyzed at more frontal regions. As such, the significant region does not correspond to a known ERP. To aid in interpretation, Figure 8 shows the Standard stimulus grand average and beta time course at Pz.

>> Figure 8 here <<

It can be seen that the significant region corresponds to a negative peak starting just after 300 ms on the grand average plot (Figure 8, left). Because is more negative-going than in this region (Figure 8, right), the positive $t$ -values indicate that decreased or less negative-going amplitudes at this region correlates with higher scores in K10.

For Standard stimuli, we also found a large region, not corresponding to any known ERPs either, starting at 396 ms and ending at 588 ms covering both frontal, central and parietal channels. Figure 8 shows a negative-going slow wave at Pz and the positive $t$ -values indicate that less negative-going amplitudes across this wave correlates with higher scores in K10.

For Correct trials, we observed a significant region from -16 to 64 ms at central electrodes, centered at FCz, Cz and CPz. The time range and involved regions clearly corresponded to the CRN. The negative$t$- values indicated that higher scores in K10 correlated with a increased, or more negative-going, . As can be deduced from Figure 5 (top right) and because , a more negative translates to an increased, or more negative-going, CRN. In other words, an *increased* CRN correlated with higher scores in K10.

For Error trials, we observed a significant region from -20 to 56 ms, more frontal than the corresponding regions for the CRN, centered at Fz, FCz and Cz. This region clearly corresponded to the ERN and the positive$t$ -values indicated that higher scores in K10 are associated with a *reduced* (less negative-going) ERN.

Finally, also for Error trials, we observed a large significant region from 284 ms to 468 ms, mainly at central-parietal and parietal regions, and with opposite effects at frontal regions. However, no obvious peak or wave is shown in neither the grand average plot (Figure 2, right) nor in the beta coefficient time course plot for Error trials and adjusted mean (Figure 5, top right). We interpreted this region as a continuation of the Pe, and the negative $t$ -values indicated that a reduced late part of the Pe correlates with higher scores in K10.

**Level of Personality Functioning Scale (LPFS)**

Figure 8 shows results for LPFS assessing severity of personality pathology. We found correlations between LPFS and ERPs from the AO, response-locked Flanker and UO paradigms. There were no significant correlations with ERPs from the stimulus-locked Flanker at the main analysis level.

>> Figure 9 here <<

For ERPs from the AO paradigm, we found similar correlations as for K10. For Target stimuli, we found a large region from 312 ms to 688 ms covering frontal, central and parietal channels, clearly corresponding to the P3b. As for K10, reduced (less positive-going) P3b at parietal channels correlated with increased scores in LPFS. For Standard stimuli, we found that reduced amplitudes in the same wave covering a large region - from 312 to 680 ms, at both frontal, central and parietal channels - correlated with increased scores in LPFS. So did the negative peak just after 300 ms described for K10 above.

For the response-locked Flanker ERPs, as for K10, an increased CRN and a reduced ERN correlated with increased scores in LPFS. However, effects for the CRN were more frontally distributed this time, with centers at Fz, FCz and Cz.

We also noted a reversal of effects, with positive $t$ -values correlating with increased scores in parietal regions. However, we also noted that the effect was absent at parietal channels along the midline, e.g., CPz and Pz. A plot of the Correct trial beta at P6, where the effect was strong, did not reveal a CRN-like waveform. As such, results were not straightforward to interpret. It is possible that these regional effects reflected activity from the brain processes generating the CRN, which were then picked up by lateral electrodes through volume conduction. A similar but weaker effect was observed for the ERN.

For the Combined deviant difference wave from the UO paradigm, we found a significant region from 156 to 200 ms centered around Fz. This region corresponded to the cMMN and the negative $t$ -values indicated than an *increased* or more negative-going cMMN correlated with higher scores in LPFS. In addition, we found a significant region from 212 ms to 256 ms, centered around Fz and FCz. The positive $t$ -values indicated that an *increased* combined deviant dP3a11Note that the d in dP3a denotes the difference wave ERP component corresponding to P3a. dP3a elicited to one of the three deviant types is spelled out as such, e.g., frequency deviant dP3a. correlated with increased LPFS scores.

**Multidimensional Emotional Disorder Inventory (MEDI)**

Figure 9 shows results for MEDI assessing symptom dimensions within the Internalizing spectrum.

>> Figure 9 here <<

For MEDI total scores, in a similar fashion to results for K10 and LPFS, we found that an increased CRN, a reduced ERN and a reduced P3b correlated with higher scores in MEDI. Again we found a significant region corresponding to the above-described late wave in response to Standard stimuli in the AO paradigm, but the effect was weaker. In addition, for Error trials we found a significant region from 184 ms to 388 ms centered at Cz and CPz. This region corresponded to Pe and because is above at this time interval, the negative $t$ -values indicated that a *reduced* Pe correlated with higher scores in MEDI.

Because we found significant correlations for MEDI total score, we also analysed the MEDI sub scales at the less conservative level of 5%, each assessing a specific symptom dimension within the Internalizing spectrum. These results are available in Supplementary materials.

Not surprisingly, at this less conservative level, all of the regional effects for MEDI total score described above were present and more pronounced in covering more channels and longer time frames. In addition, higher scores correlated with a reduced stimulus-locked Flanker P3b and a reduced duration deviant dP3a from the UO paradigm.

*Post-hoc* , then, a reduced P3b elicited to Target stimuli in the AO paradigm correlated with *all* of the MEDI sub scales. Similarly, a reduced ERN correlated with higher scores in all MEDI sub scales except Traumatic Re-experiencing, albeit weakly for Somatic Anxiety and Avoidance. An increased CRN correlated with higher scores in Depression and less strongly with higher scores in Avoidance, Social Anxiety and Somatic Anxiety.

As perhaps could be expected, a reversal of effects was seen for Positive temperament, which is the only dimension where healthy comparison subjects scored higher than patients (Table 3). Here we observed that,

e.g., a reduced CRN and, more weakly, an increased ERN correlated with higher scores.

For Error trials in the response-locked Flanker paradigm, in addition to the results already described above, we found that a *reduced* Pe correlated, more or less strongly, with higher scores in the sub scales Autonomic Arousal, Avoidance, Depression (the late part of Pe), Intrusive Cognitions, Somatic Anxiety, Social Anxiety as well as Traumatic Re-experiencing (weak and the late part of Pe). In addition, an *increased* Pc correlated with increased scores in Social Anxiety.

*Post-hoc* , we also found significant effects for the ERPs in the stimulus-locked Flanker paradigm, e.g., N2 and P3b, which did not survive the conservative level of 0.1%. However, because the P3b in the stimulus-locked Flanker paradigm falls immediately before button press, it cannot be reliably analysed due to some trials overlapping. Therefore, here we consider only results for the N2. Interestingly, for MEDI sub scales Intrusive Cognitions and Traumatic Re-experiencing a*reduced* N2 correlated with higher scores, whereas for Neurotic Temperament, an *increased* N2 correlated with higher scores.

*Post-hoc* , we also found that a *reduced* Novelty P300 at CPz and Pz correlated with increased scores in MEDI Depression.

For ERPs from the UO paradigm, the strongest result was a correlation between MEDI Avoidance and a *reduced* (less positive) duration deviant dP3a and, more weakly, a reduced combined deviant dP3a. A reduced dP3a also correlated with increased scores in Intrusive Cognitions (duration deviant) and with increased scores in Traumatic Re-experiencing (both combined and duration deviants). For the MMN, correlations were quite specific in that only Social Anxiety correlated with an *increased* (more negative) cMMN and more weakly with an increased fMMN.

Finally, for the N1-P2-N2 complex, which is elicited to both Standard, Distractor and Target stimuli in varying degrees and with different overlap and proximity of peaks, results are a bit more difficult to analyze. Given that little is known about the properties of this wave complex in terms of, e.g., polarity-reversal, we resort to reporting that, *post-hoc* , we found correlations for each of these peaks with several of the MEDI sub scales (Winkler et al., 2013). The interested reader is referred to Supplementary materials. Plots of the beta coefficients at relevant channels can be supplied at request.

### Modified Personality Inventory for DSM-5 and ICD-11 (PID36)

Figure 10 shows results for PID36 indexing maladaptive personality trait dimensions.

>> Figure 10 here <<

Results for PID36 largely mimicked those already described above. We found that a reduced P3b and ERN and an increased CRN correlated with higher scores. For Standard stimuli in the AO paradigm, we found the same regional effects corresponding to reductions in the negative peak just after 300 ms and the following slow negative wave in parietal regions.

As for MEDI, since we found significant correlations for the PID36 total score, *post-hoc* we also analysed each PID36 sub scale at the uncorrected 5% level. Again, the correlations found for PID36 total score were stronger, involving more channels and longer time ranges. A reduced P3b correlated strongly with increased scores in all sub scales except for Psychoticism. A corresponding pattern was observed for the late parietal wave elicited to Standard stimuli in the AO paradigm. Of the response-locked Flanker ERPs, an increased CRN and a decreased ERN correlated with higher scores in the same three sub scales: Anankastia, Detachment and Negative Affect. In addition, an increased CRN correlated with increased scores in Psychoticism, but only at Cz and CPz.

For Negative Affect, a reduced Pc in frontal-central regions correlated with higher scores while a reduced Pe in the same regions correlated with higher scores in Anankastia.

Also for Negative Affect, we found correlations between a *reduced*Novelty P300 in frontal-central regions and higher scores. We noted that, along with the results for MEDI Depression described above, this was the only

significant correlation for Novelty P300 in our entire analysis.

We also observed some interesting correlations for the UO paradigm. For Anankastia, increases in dMMN and the following duration deviant dP3a, both centered around Fz and FCz, correlated with higher scores, albeit weakly. For Antagonism, reductions in cMMN (although only at CPz and Pz) and dMMN and the following duration deviant dP3a correlated with higher scores. For Detachment, increased cMMN at Fz correlated with higher scores. For Disinhibition, increases in all three MMN measures as well as decreases in all three corresponding dP3a measures correlated with higher scores. Finally, of opposite direction, reductions in cMMN and fMMN correlated with higher scores in Negative Affect.

Other than these go-to results, as in the *post-hoc* analysis of the MEDI sub scales described above, several correlations involving the N1-P2-N2 complex were observed. Again we invite the reader to study the results in the Supplementary materials.

# Discussion

In this study, we utilized linear regression models based on robust single-trial ERP analysis as implemented in the EEGLAB toolbox LIMO EEG to test for correlations between a range of well-studied ERPs and symptom and trait measures of psychopathology compatible with the transdiagnostic framework the HiTOP. Recruiting a mixed sample of 50 patients with emotional disorders undergoing 14 weeks of transdiagnostic psychotherapy as well as 37 healthy comparison subjects matched in age and sex, we assessed longitudinal correlations across the full psychopathology spectrum. In the following, we pragmatically follow a top-down discourse in that we first treat results covering the HiTOP spectrum and subfactor levels. After this, we look at ERPs which only correlated with a single or a few sub scales and therefore are of relevance to the lowest symptom and maladaptive trait level.

The most consistent result in this study was that a reduced P3b elicited to Target stimuli in the AO paradigm correlated with worse symptomatology. In fact, significant correlations between higher scores and reduced P3b was found in *all* psychopathology measures with two exceptions: MEDI Positive temperament - where, as expected, we found the opposite pattern - and PID36 Psychoticism. These findings were corroborated by results from the Correct response stimulus-locked Flanker ERPs elicited to Congruent stimuli. Although not significant at the corrected level, at 5% a reduced stimulus-locked Flanker P3b correlated with higher scores in several of the psychopathology measures, including MEDI total score11Note that because the P3b in the Flanker paradigm falls immediately before button press, it cannot be as reliably analysed as the P3b from the AO paradigm due to some trials overlapping.. The P3b is thought to be an index of cognitive processes such as context updating and memory processing (Luck & Kappenman, 2011; Polich, 2007). While P3b is consistently reduced in chronic schizophrenia and less consistently reduced in depression, we are not aware of studies relating P3b to specific symptoms in any of the emotional disorders included in this study (Klawohn et al., 2020; Onitsuka et al., 2013). In schizophrenia, some studies have reported associations between a reduced P3b and increased symptoms of cognitive deficits (Giordano et al., 2021; Kruiper et al., 2019). Our results are in line with these findings and support that cognitive impairment is a mainstay also of the emotional disorders. In the HiTOP, our results suggest that a reduced P3b is a marker not of subfactor or lower symptom and trait levels but of the Internalizing spectrum as a whole. Given the above as well as evidence of reductions also in Externalizing disorders, it can be speculated that a reduced P3b is a marker not only of the Internalizing, but also of the Externalizing and Thought disorder spectra in the HiTOP (Pasion & Barbosa, 2019; Patrick et al., 2006). Indeed, cognitive impairment is a symptom of most, if not all, psychiatric disorders (Etkin et al., 2013). As such, a reduced P3b could be a marker of the general p-factor sometimes included at the top of the HiTOP hierarchy (Levin-Aspenson et al., 2021). However, there is some evidence of an increased P3b in OCD (Gohle et al., 2008; Mavrogiorgou et al., 2002). This highlights the unique features of OCD in a transdiagnostic framework (see below) (Faure & Forbes, 2021). Finally, it should be noted that cognitive impairment is not only a core symptom of psychiatric disorders, but also a known side effect to treatment with psychotropic medication (Cowen & Sherwood, 2013; Paterniti

et al., 1999). Even though we accounted for medication status in our analysis, 86% of the patient population was treated with at least one psychiatric medication. Given this high proportion, we can't be certain that cognitive impairment, as indexed by a reduced P3b, were due to disorder or psychotropic medication. However, a separate analysis with a model excluding medication status did not yield stronger P3b regional effects, suggesting that medication status did not contribute to the results. Indeed, evidence indicate that the P3b is not altered by SSRI treatment, which was the most prevalent psychotropic medication in our sample (d'Ardhuy et al., 1999; Oranje et al., 2008-06-31; however, see Wienberg et al., 2010).

After the P3b, the second most consistent findings involved the response-locked Flanker ERPs, indexing conflict or performance monitoring (Larson et al., 2014). Starting with the ERN, elicited to errors, we had hypothesized than an increased ERN would correlate with higher scores, especially in the psychopathology measures indexing different forms of anxiety and symptoms related to OCD. This was based on findings of correlations between increases in the ERN and transdiagnostic dimensions related to OCD and anxiety, especially Fear-based anxiety, within the Internalizing spectrum (Gorka et al., 2017; Pasion & Barbosa, 2019; Riesel et al., 2022). Contrary to this hypothesis, however, we found consistent and strong correlations between a *reduced* ERN and the majority of symptom dimensions in MEDI, as well as the Anankastia, Detachment and Negative Affect maladaptive traits in PID36. This apparent discrepancy with the literature invites to several interpretations. First, a few studies have indeed found correlations between a reduced ERN and Internalizing psychopathology. Tanovic et al. (2017) found associations between a reduced ERN and symptoms of ruminations, but only when controlling for effects of anxiety. Olvet et al. (2010) found associations between depression severity and both an increased CRN and a reduced dERN, the difference wave between ERN and CRN. Our results therefore corroborate these results in finding correlations with increased scores in MEDI Intrusive Cognitions and MEDI Depression, the latter for which we also found correlations with an increased CRN as in Olvet et al. (2010) (see below). Second, owing to the nature of linear regression models as utilized in this study, regional effects can either indicate a reduced or an increased ERN, but not both simultaneously. Consequently, if the 'ground truth' is that both types of ERN deviations correlate with higher scores, the dominant feature 'wins', or, alternatively, the effects cancel out and the correlation is not significant. Perhaps this is why Seow et al. (2020) failed to find associations between the ERN and transdiagnostic measures in a community sample. In Riesel et al. (2022), correlations between increases in the ERN and an anxious-misery dimension was in a mixed sample across the OCD spectrum. In that study, the patient group had a significantly increased ERN compared to the healthy comparison group. For our present sample, we have recently shown that a sub sample based on the HiTOP subfactor Distress - containing the ICD-10 diagnoses of Depression and GAD - had a *reduced* ERN at baseline compared to healthy comparison subjects (Randau et al., 2023). In that study, the ERN of the Fear subfactor - containing agoraphobia, OCD, PD and SAD - was not significantly different to either the Distress or HC group. As such, our sample characteristics differs from Riesel et al. (2022) in that the group contributing the most to the psychopathology variance is not defined by an increased ERN, but rather by a decreased one. Third, it is established that the ERN is sensitive to not only manipulations of experimental factors, but also to preprocessing and analysis methods (Clayson, 2020; Clayson et al., 2021; Feuerriegel & Bode, 2022). In all paradigm-related aspects, our study closely followed established conventions in the literature. However, we cannot rule out that a desensitizing effect across sessions took place for the patient group, essentially decreasing the perceived threat of errors and thereby the ERN. That being said, patients showed a reduced ERN compared to HC already at baseline and the ERN has been shown to be stable across sessions (Olvet & Hajcak, 2009a; Randau et al., 2023). Neither can we rule out a fatigue effect from the rather long paradigm, even though such an effect has not been demonstrated (Olvet & Hajcak, 2009b). In addition, our paradigm was not longer in duration than the one used in Riesel et al. (2022) and divided into 10 rather than 6 blocks. Taken together, we do not find it likely that our divergent effects were paradigm- or study design-related. While we believe that the robust single-trial analysis method utilized in this study constitutes an improvement, we did not conduct a formal analysis comparing our methods against traditional methods. However, we can report that applying traditional baseline subtraction methods at intervals commonly reported in the literature (-500 to -300 and -200 to 0 ms pre-stimulus, respectively) did not noticeably alter the group-wise grand average ERN waveforms in terms of maximum peak amplitude

or latency. In Gorka et al. (2017) on a transdiagnostic sample, OCD was an exclusion criteria and yet higher scores in a derived Fear dimension - but not in a Distress dimension - correlated with increases in a residual ERN-measure, 22Defined as the ERN activity when the CRN is 'regressed out (Gorka et al., 2017).. In our study, even though the ERN beta coefficient represents residual activity when effects of the CRN beta coefficient (and the adjusted mean) are accounted for, we also informally tested an explicit ERN-CRN difference contrast. However, results for this dERN was in the same direction as the ERN. Therefore, we must conclude that when it comes to the ERN, 'all roads lead to Rome' in the sense that both decreases and increases can be observed in clinical populations and that both types of deviations are associated with worse symptomatology. In this regard, it must be noted that past studies have utilized rating scales based on a categorical understanding of psychopathology and converted these into transdiagnostic dimensions through factor analysis conducted on the study sample or from weights derived from previous studies. Needless to say, into what latent dimension a given rating scale is allocated will affect the direction of correlations, if any. While neither MEDI nor PID36 are directly derived from the HiTOP, both measures are validated in large populations and index distinct transdiagnostic dimensions consistent with the HiTOP framework. In terms of the HiTOP, then, as we saw for the P3b, a reduced ERN seems to be associated with worse symptomatology across the whole Internalizing spectrum. In addition to ruminations and depressive symptoms, reductions in the ERN have been associated with symptoms and traits belonging to the *Externalizing* spectrum (Hall et al., 2007; Lutz et al., 2021; Pasion & Barbosa, 2019). However, we found no significant correlations between the ERN and the two Externalizing traits in PID36 (Disinhibition and Antagonism). Then again, our sample did not include Externalizing disorders. Indeed, scores for Antagonism were comparably low and did not differ between the two groups. Finally, in the HiTOP, the placement of OCD within the Internalizing spectrum is not fully established, with results indicating that symptoms cross-load on the Fear subfactor within the Internalizing spectrum but also on the Thought-disorder spectrum (Faure & Forbes, 2021). As such, we can raise the possibility that a *reduced* ERN is specific to the Internalizing and the Externalizing spectra, as conceptualized in the HiTOP, while an *increased* ERN is a specific marker of some other construct encompassing both anxiety symptoms contained in the Fear subfactor as well as symptoms related to the Thought disorder spectrum. The uniqueness of OCD in terms of ERP abnormalities is also supported by associations with an increased P3b discussed above (Gohle et al., 2008). We can conclude that more studies are needed to understand the associations between the ERN and psychopathology, especially studies looking to discern divergent effects in different patient populations and using validated transdiagnostic measures.

Results for the CRN, elicited to correct responses, were somewhat more specific and in the expected direction. Firstly, we can corroborate results from Riesel et al. (2022) in finding correlations between an increased CRN and PID36 Anankastia. PID36 Anankastia must be considered to capture much of the same psychopathology as the dimensions Compulsiveness and Personal standards examined in Riesel et al. (2022). In addition, we can corroborate results from Olvet et al. (2010) in finding correlations between an increased CRN and depressive symptoms as indexed by MEDI Depression. Indeed, we find that increased CRN correlates strongly with transdiagnostic dimensions which can be considered to load unto the HiTOP Distress subfactor (MEDI Depression and PID36 Negative Affect, but also PID36 Detachment33We note that Detachment is in itself a spectrum in the HiTOP.), whereas it correlates more weakly or not at all with Fear subfactor dimensions, e.g., MEDI Autonomic Arousal, Neurotic Temperament, Social Anxiety44Here, correlations were present but considerably weaker than for MEDI Depression and PID36 Negative Affect., Somatic Anxiety and PID36 Avoidance. As such, we find some evidence of an increased CRN as a marker of symptoms and traits in the HiTOP Distress subfactor. At first sight, this contradicts the results from our above-mentioned study, where the Distress subfactor - containing patients with a primary ICD-10 diagnosis of depression and GAD - had a significantly*reduced* CRN compared to healthy comparison subjects (Randau et al., 2023). However, group-level differences between HiTOP subfactors based on primary ICD-10 diagnosis do not necessarily translate to correlations with transdiagnostic symptomatology. While the above-mentioned Distress dimensions must be considered to be a core part of both ICD-10 Depression and GAD, both HCs as well as patients with disorders which would be allocated to the Fear subfactor contributed to the correlations. Therefore, it is not a contradiction to state that categorical diagnoses associated with Distress show a reduced CRN compared to healthy comparison subjects, but when considering the full spectrum of psychopathology, an *increased*

CRN correlates with dimensions which primarily load unto the Distress subfactor. It is likely that other factors which are common to Distress disorders and not captured by our psychopathology measures, such as cognitive impairment, influence to reduce the CRN. Some evidence of a negative (reduced) effect of cognitive impairment on the response-locked Flanker ERPs exist (Eppinger et al., 2008; Simó et al., 2018; Swainston et al., 2021).

ERN and CRN are followed by the Pe and Pc, which are believed to index the conscious awareness of correct and error responses, respectively (Overbeek et al., 2005; Wessel, 2012). In the main analysis, only correlations between a reduced Pe and MEDI total survived the corrected level of 0.1%. *Post-hoc* , a reduced Pe correlated strongly with increased scores in MEDI Avoidance and Social Anxiety and more weakly with Autonomic Arousal, Intrusive Cognitions and Somatic Anxiety. Even though PID36 total scores did not survive the main analysis, a look at the sub scales reveals significant but considerably weaker correlations with Anankastia, Disinhibition, Negative Affect and Psychoticism. Given this, and the absence of strong correlations with Distress subfactor dimensions and traits, it is possible that a reduced Pe is a marker of the HiTOP Fear subfactor rather than of the whole Internalizing spectrum. Again, OCD would not fit well into Fear in having an increased Pe compared to healthy comparison subjects (Bellato et al., 2021). As for Pc, correlations with MEDI and PID36 total scores were not significant in the main analysis. *Post-hoc* , we find quite specifically that a reduced Pe correlates only with increased scores in PID36 Negative Affect. This speaks for the specificity of the Correct response-locked Flanker ERPs in that both CRN and Pc correlate with Distress symptoms, the latter perhaps at the lowest symptom and trait level.

Having discussed the major results spanning HiTOP spectra and subfactor levels, we now turn to more specific sub scale results representing effects at the symptom and maladaptive trait level. It should be noted that the following results were not tested at the main analysis corrected level of 0.1% but at the more conventional 5%. Speaking for their specificity, neither of the MMN measures (with the exception of cMMN with LPFS) nor the Novelty P300 elicited to Distractor stimuli in the AO paradigm or the stimulus-locked Flanker N2 yielded significant, strong correlations with any of the four major psychopathology measures (K10, LPFS, MEDI and PID36) at either level.

The Novelty P300 is rightfully distinguished from the more typical P3a elicited to deviant stimuli in the Unattended oddball paradigm (Polich, 2007). Perhaps due to some confusion in the literature, we are not aware of studies examining correlations between the Novelty P300 and measures of psychopathology. We find, quite specifically, that a reduced Novelty P300 at Fz, FCz and Cz correlates with increased scores in PID36 Negative Affect. For MEDI Depression, a reduced Novelty P300 correlates with higher scores, but weaker and more posterior at CPz and Pz. We note that this effect of central versus posterior effects might to some extent be due to correlations with latency in addition to amplitude. Nevertheless, PID36 Negative Affect and MEDI Depression can be considered to index roughly similar symptom and trait level dimensions. As such, given the absence of other correlations, we can postulate that a reduced Novelty P300 is a marker of a negative affect dimension at the lowest level of the HiTOP hierarchy, or alternatively, of the Distress subfactor.

For LPFS at the main analysis level, increases in the latter part of the cMMN as well as in the following combined deviant dP3a correlated with higher scores. This curiously indicates an association between MMN/dP3a and measures of personality functioning or maladaptive traits. MMN is an index of pre-attentive auditory processing and is reduced in chronic schizophrenia. However, studies examining the role of MMN in personality disorders and associated symptoms are, to our knowledge, rare or inconclusive. Increases in MMN have been associated with schizotypal and antisocial personality disorders as well as with treatment-resistant depression when controlling for comorbid borderline personality disorder (He et al., 2010; Liu et al., 2007). Given this potential connection between MMN and the personality disorders, it is interesting to have a look at the PID36 sub scales. Here, increased MMN correlated with Detachment, Disinhibition and more weakly with Psychoticism and Anankastia. Decreased MMN correlated with Antagonism and Negative Affect, the latter consistent with reduced MMN in depression (Tseng et al., 2021). These results stand in an interesting contrast to results for the MEDI sub scales, where the only significant correlation was between increased

MMN and Social Anxiety. Taken together, we find compelling evidence of MMN being a quite specific marker of maladaptive personality traits loading unto the Internalizing (Detachment, Negative Affect), Externalizing (Antagonism, Disinhibition) and Thought-disorder (Psychoticism) spectra in the HiTOP. Such a specificity was not seen in the following dP3a, which, even though an increased dP3a correlated with LPFS, was equally related to both MEDI and PID36 sub scales. Curiously, dP3a was most strongly associated with with MEDI Avoidance and PID36 Antagonism where a decreased dP3a correlated with higher scores. However, we believe it is beyond the scope of this paper to discuss similarities between these two measures.

The last ERPs we consider in this discussion are the stimulus-locked Flanker N2s elicited to correct response to congruent and incongruent stimuli, respectively. Like the CRN and the ERN, the Flanker N2 is an index of conflict monitoring and cognitive control (Larson et al., 2014). First, we note that only the Flanker N2 elicited to congruent stimuli yielded strong correlations. Second, while both the CRN and the ERN correlated with maladaptive traits as indexed by LPFS and PID36, as well as with symptom dimensions as indexed by K10 and MEDI total score, the Flanker N2 correlated only with a few sub scales in MEDI. Specifically, a reduced Flanker N2 correlated with higher scores in Intrusive Cognitions and Traumatic Re-experiencing. Even though the dimensions in MEDI are distinct and validated, these two sub scales must be considered to capture closely related psychopathology. As such, we can postulate that a reduced Flanker N2 to congruent stimuli is a marker of a single or a few specific symptom dimensions at the lowest level of the HiTOP hierarchy. However, it is unclear to us to what extent Intrusive Cognitions and Traumatic Re-experiencing loads unto the Internalizing and Thought-disorder spectra (Kotov et al., 2020). We also saw that an increased Flanker N2 to congruent stimuli correlated with increased scores in MEDI Neurotic Temperament, a core part of the Internalizing spectrum (Watson et al., 2022).

Our study has several limitations. First, our setup did not allow us to infer to what extent treatment and group influenced the correlations. As such, we cannot rule out that our results are driven by correlations which are the strongest in, e.g., patients at baseline. Second, while we believe that we controlled for false positives with the conservative level in the main analysis, we did not define how large a significant region shall be for it to be considered a true correlation. Add to this that we found several significant effects at regions not corresponding to traditional ERP evaluation windows and channels. Rather than considering only the strongest correlations clearly corresponding to traditional ERPs, we opted to describe all significant regions above some arbitrary visual threshold and to quantify these correlations in terms of strong or weak. It is likely that with more data, some of these regions would either become larger or vanish.

To our knowledge, this is the first comprehensive examination of the associations between ERPs and transdiagnostic psychopathology. The ERPs included in the study are easily measured in a clinical setting and index pre-attentive auditory processing, cognition and performance monitoring. Some ERPs, e.g., the MMN, appear to be exclusively related to maladaptive personality traits at the lowest level of the hierarchy, whereas others, e.g., the P3b, cut across and are related to entire spectra or even the general p-factor. The ERN remains elusive in that we found solid evidence of a reduced ERN correlating with higher scores at the spectrum level. Conversely, increases in the CRN correlated with worse symptomatology at the subfactor level, results which are in line with the literature. In showing that abnormalities in such basic brain processes are associated with transdiagnostic symptoms and traits at several levels of the HiTOP hierarchy we have taken yet another small step toward biomarkers in psychiatry. We have also shown the advantages of utilizing a consistent framework such as the HiTOP, which allowed us to pinpoint associations between ERPs and diagnostic measures to specific levels in the hierarchy. While we did not directly compare our results against traditional ERP methods, we can state that robust single-trial ERP analysis as implemented in LIMO EEG is an excellent tool for a pragmatic analysis of ERP features across channels and time frames. In future steps, after replication, machine learning and related advanced method would be obvious candidates in translating ERP features into transdiagnostic symptom profiles at the subject level (Nielsen et al., 2020).

# 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 miliseconds (ms). Also note that the d in dP3a denotes a difference wave.

Table 1: Paradigm and ERP overview

| | **HC** |
|---|---|
| **Female/Male (%)** | **25 (67.6)** |
| | **12 (32.4)** |
| **Age** | **38.5 (13.2)** |
| **Attended oddball** | |
| **RT Target stimuli** | **403.9 (47.3)** |
| **Flanker** | |
| **Correct trials** | **429.3 (38.1)** |
| **Error trials** | **53.6 (24.3)** |
| **RT Correct trials** | **406.3 (52.8)** |
| **RT Error trials** | **315.6 (40.2)** |
| **All measures except Sex reported as mean (standard deviation)** | **All measures except Sex reported as mea** |
| **Abbreviations: RT, reaction time in miliseconds (ms)** | **Abbreviations: RT, reaction time in milis** |

Table 2: Demographics and behavioural measures at baseline

| | **Omega** | **HC** | **Patient** | p |
|---|---|---|---|---|
| **K10** | 0.94 | 14.5 (4.4) | 28.3 (8.2) | <0.001 |
| **LPFS** | 0.90 | 17.6 (5.6) | 27.0 (6.7) | <0.001 |
| **MEDI** | 0.92 | 19.8 (45.9) | 145.8 (51.3) | <0.001 |
| Autonomic Arousal | 0.80 | 4.6 (5.8) | 18.2 (8.3) | <0.001 |
| Avoidance | 0.92 | 10.1 (9.7) | 32.1 (12.8) | <0.001 |
| Depression | 0.88 | 3.9 (5.9) | 20.5 (10.3) | <0.001 |
| Intrusive Cognitions | 0.88 | 3.8 (4.9) | 20.6 (10.2) | <0.001 |

| | **Omega** | **HC** | **Patient** | p |
|---|---|---|---|---|
| Neurotic Temperament | 0.83 | 8.4 (7.0) | 24.5 (8.0) | <0.001 |
| Positive Temperament | 0.88 | 27.4 (7.3) | 19.9 (7.5) | <0.001 |
| Social Anxiety | 0.93 | 8.6 (9.7) | 19.4 (11.2) | <0.001 |
| Somatic Anxiety | 0.87 | 4.2 (6.2) | 15.5 (9.9) | <0.001 |
| Traumatic Re-experiencing | 0.90 | 3.7 (5.4) | 14.9 (9.1) | <0.001 |
| **PID36** | 0.77 | 17.6 (13.0) | 33.5 (13.0) | <0.001 |
| Anankastia | 0.88 | 4.9 (4.3) | 7.9 (4.5) | 0.002 |
| Antagonism | 0.77 | 2.5 (2.9) | 2.3 (2.4) | 0.779 |
| Detachment | 0.78 | 3.1 (3.3) | 5.4 (3.8) | 0.003 |
| Disinhibition | 0.73 | 2.4 (2.5) | 5.1 (3.2) | <0.001 |
| Negative Affect | 0.86 | 2.8 (3.1) | 8.9 (4.2) | <0.001 |
| Psychoticism | 0.73 | 1.9 (2.0) | 3.8 (3.3) | 0.001 |

Table 3: Self-report questionnaire internal consistency and group comparison at baseline

| | **Week 10** | **Week 10** | **Week 14** |
|---|---|---|---|
| | **Beta (95% CI)** | **Beta (95% CI)** | **Beta (95%** |
| **K10** | -4.90*** | (-7.01, -2.78) | -4.95*** |
| **LPFS** | -2.07* | (-4.02, -0.12) | -2.84** |
| **MEDI** | -28.53*** | (-42.85, -14.21) | -33.97*** |
| Autonomic Arousal | -2.03 | (-4.26, 0.20) | -2.21* |
| Avoidance | -5.15** | (-8.62, -1.68) | -6.54*** |
| Depression | -5.68*** | (-8.26, -3.10) | -7.18*** |
| Intrusive Cognitions | -4.58** | (-7.46, -1.70) | -5.21*** |
| Neurotic Temperament | -2.06 | (-4.17, 0.04) | -3.78*** |
| Positive Temperament | 2.64* | (0.47, 4.80) | 0.37 |
| Social Anxiety | -1.81 | (-4.14, 0.52) | -3.55** |
| Somatic Anxiety | -2.99** | (-5.03, -0.96) | -3.07** |
| Traumatic Re-experiencing | -1.56 | (-4.26, 1.13) | -2.02 |
| **PID36** | -1.07 | (-4.26, 2.13) | -1.68 |
| Anankastia | 0.01 | (-0.88, 0.89) | -0.30 |
| Antagonism | -3 | (-0.62, 0.61) | 0.01 |
| Detachment | -0.87 | (-1.78, 0.05) | -1.36** |
| Disinhibition | 0.64 | (-0.15, 1.44) | 0.36 |
| Negative Affect | -0.57 | (-1.65, 0.51) | -0.49 |
| Psychoticism | -0.21 | (-0.97, 0.54) | 0.13 |
| p < .05; ** p < .01; *** p < .001 | * p < .05; ** p < .01; *** p < .001 | * p < .05; ** p < .01; *** p < .001 | * p < .05; |

Table 4: Regression coefficients for psychopathology measures at weeks 10 and 14

# Figures

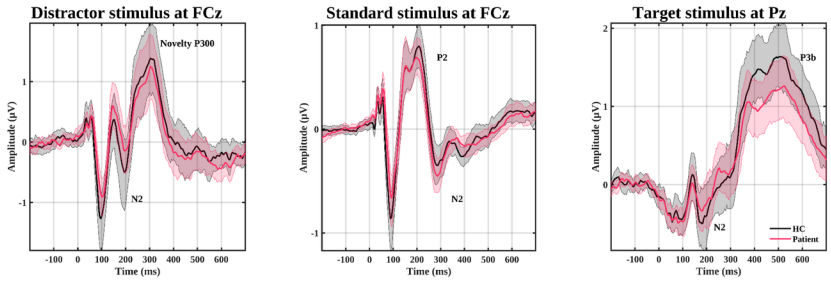



Figure 1: Attended oddball grand average ERP waveforms at traditional evaluation channels FCz and Pz computed as the 20% trimmed mean of the mean subject-level single-trial ERP data. Shaded areas denote the 95% Bayesian Highest Density Interval (HDI). Typical ERP components are marked.

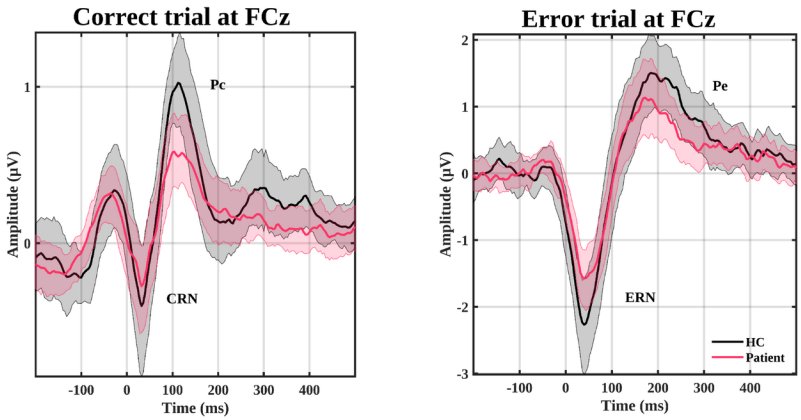



Figure 2: Response-locked Flanker grand average ERP waveforms at traditional evaluation channel FCz computed as the 20% trimmed mean of the mean subject-level single-trial ERP data. Shaded areas denote the 95% Bayesian Highest Density Interval (HDI). Typical ERP components are marked.

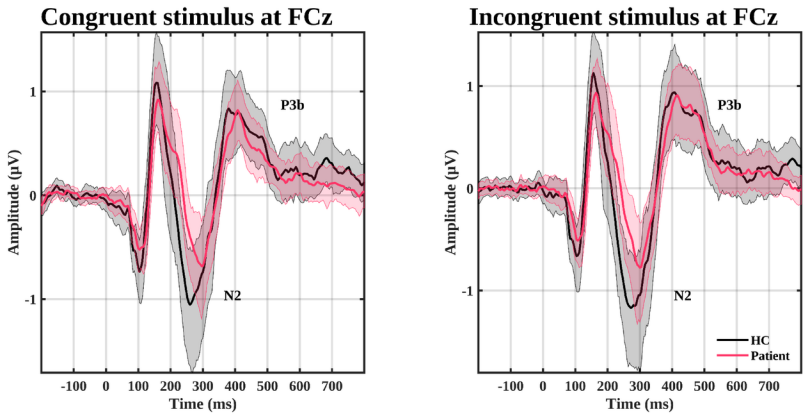


Figure 3: Stimulus-locked Flanker grand average ERP waveforms at traditional evaluation channel FCz computed as the 20% trimmed mean of the mean subject-level single-trial ERP data. Shaded areas denote the 95% Bayesian Highest Density Interval (HDI). Typical ERP components are marked.

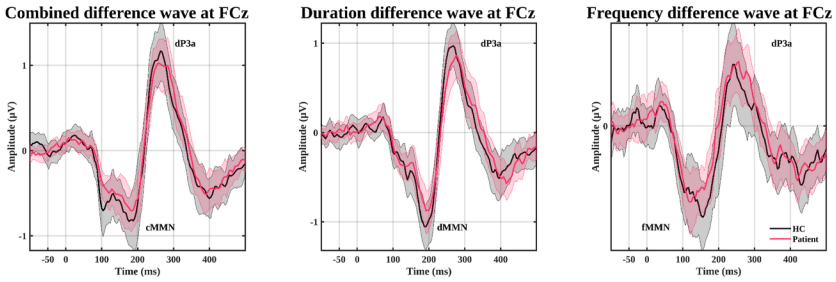



Figure 4: Unattended oddball grand average ERP difference waveforms at traditional evaluation channel FCz computed as the difference between the 20% trimmed mean of the mean subject-level single-trial ERP to standard and deviant stimuli, respectively. Shaded areas denote the 95% Bayesian Highest Density Interval (HDI). Typical ERP components are marked. Note that the d in dP3a denotes a difference wave ERP component. As such, each type of MMN is followed by a corresponding dP3a.

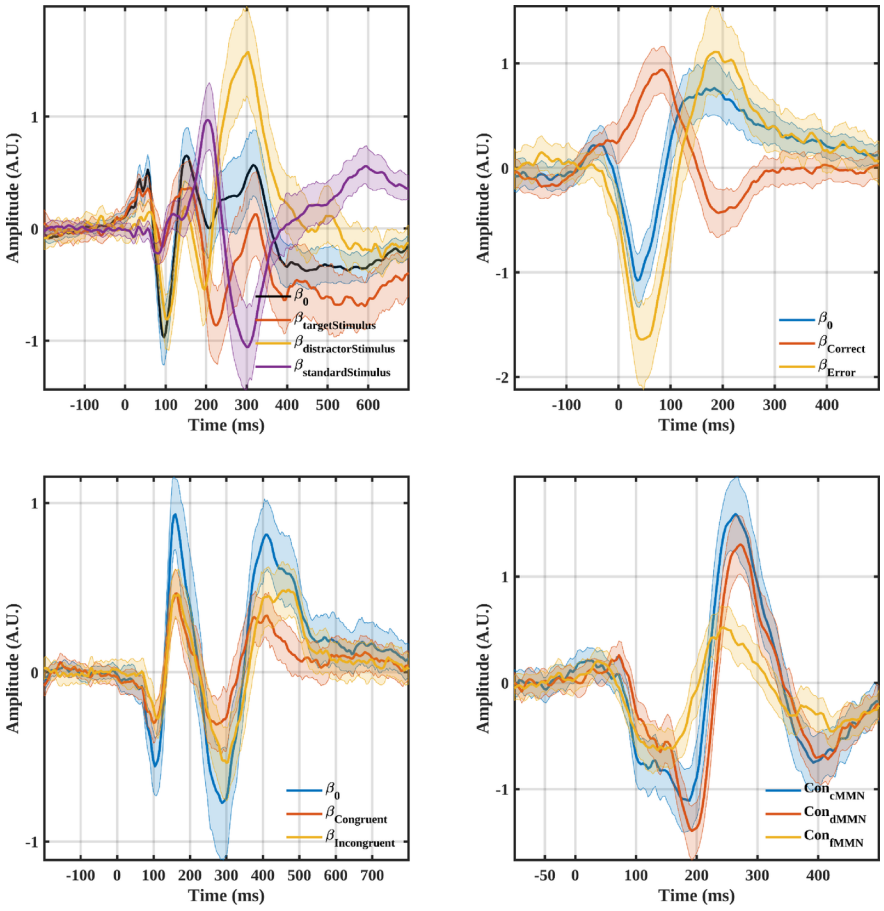



Figure 5: Beta coefficient time courses at FCz corresponding to stimuli in each paradigm for both groups

concatenated. Top left: Attended oddball; top right: Response-locked Flanker; bottom left: Stimulus-locked Flanker; bottom right: Unattended oddball. Shaded areas denote the 95% Bayesian Highest Density Interval (HDI). Note that the cMMN, dMMN and fMMN beta coefficients are difference contrasts between deviant and standard stimuli beta coefficients. A.U., Arbitrary units.

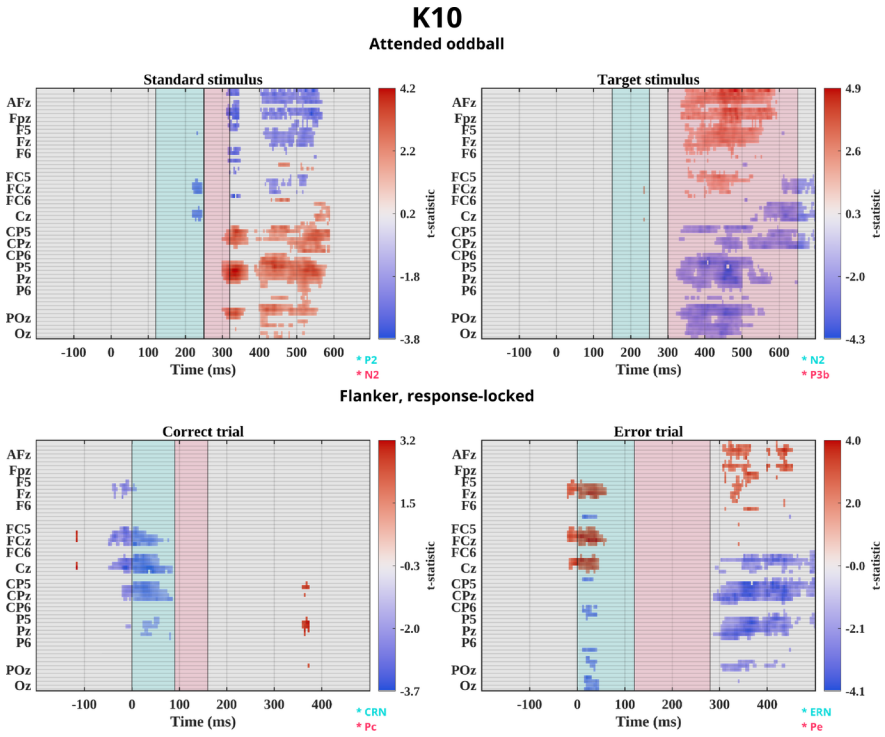



Figure 6: Results for K10 showing correlations with ERP components as heat maps indicating significant regions after correction for multiple comparisons using threshold-free cluster-enhancement at an alpha level og 0.1%. Subplot title indicates LIMO model corresponding to conditions in the respective paradigms. Shaded regions indicate traditional ERP component evaluation windows.

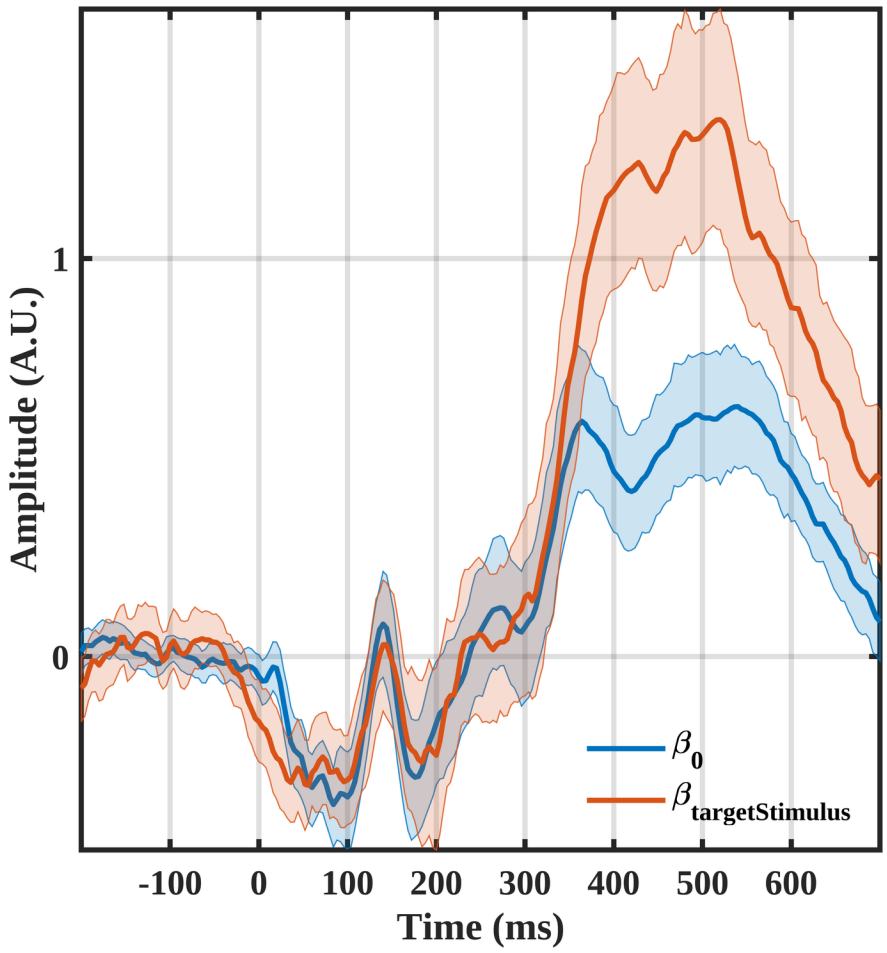



Figure 7: Time courses for Target stimulus and adjusted mean betas at Pz where correlation analysis with K10 revealed a significant region. Shaded areas denote the 95% Bayesian Highest Density Interval (HDI). Note that groups are concatenated in the beta time course plot. A.U., Arbitrary units.

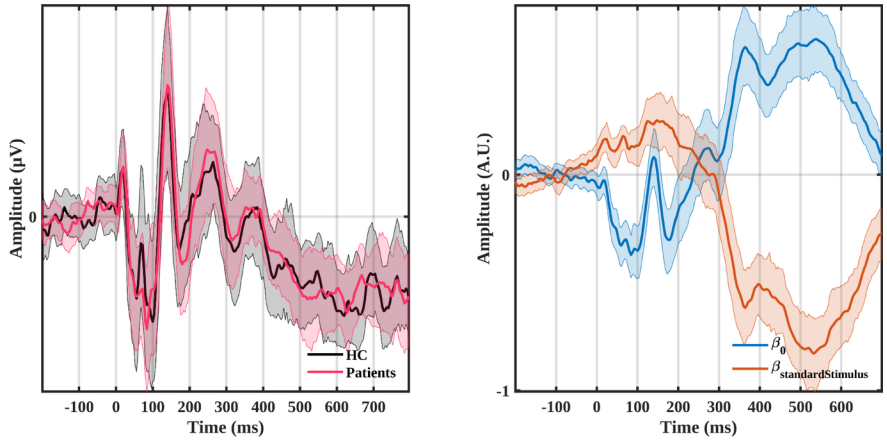



Figure 8: Group-wise standard stimulus grand average (left) and time courses for Standard stimulus and

adjusted mean betas (right) at Pz where correlation analysis with K10 revealed a significant region. Shaded areas denote the 95% Bayesian Highest Density Interval (HDI). Note that groups are concatenated in the beta time course plot. A.U., Arbitrary units.

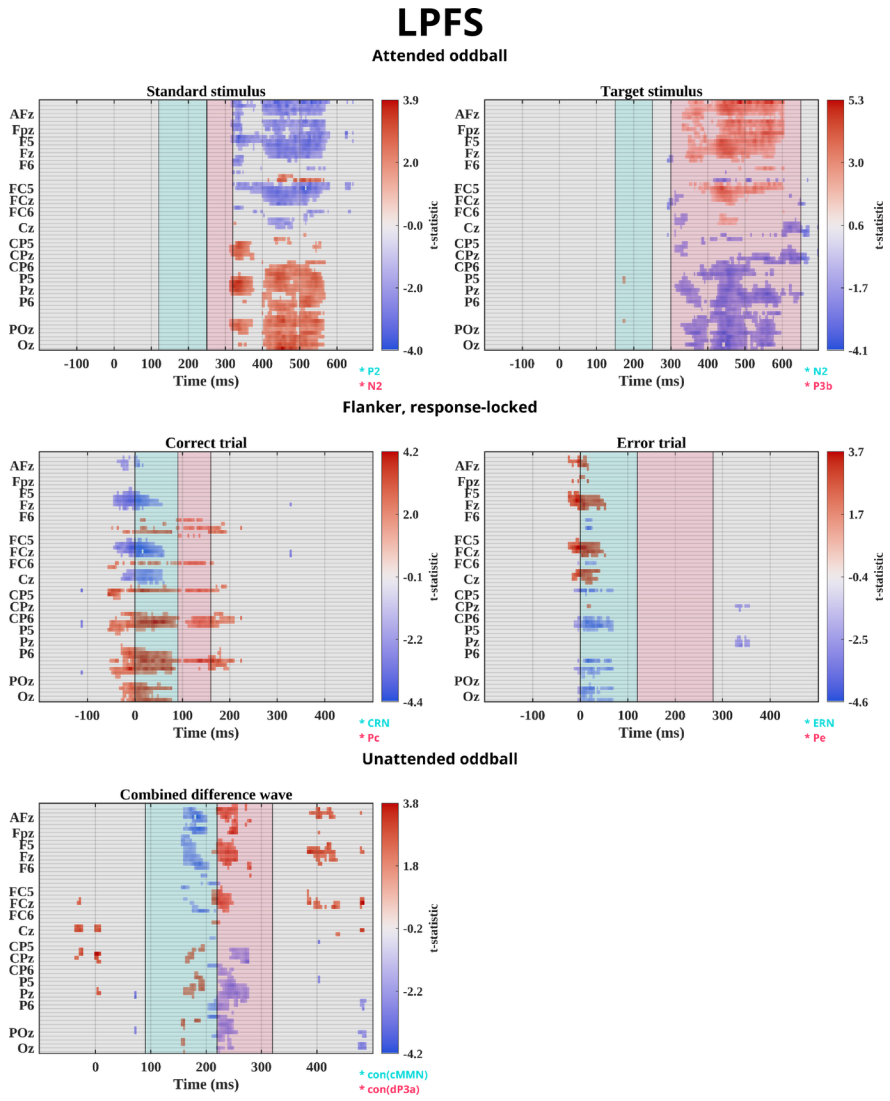



Figure 9: Results for LPFS showing correlations with ERP components as heat maps indicating significant regions after correction for multiple comparisons using threshold-free cluster-enhancement at an alpha level og 0.1%. Subplot title indicates LIMO model corresponding to conditions in the respective paradigms. Shaded regions indicate traditional ERP component evaluation windows.

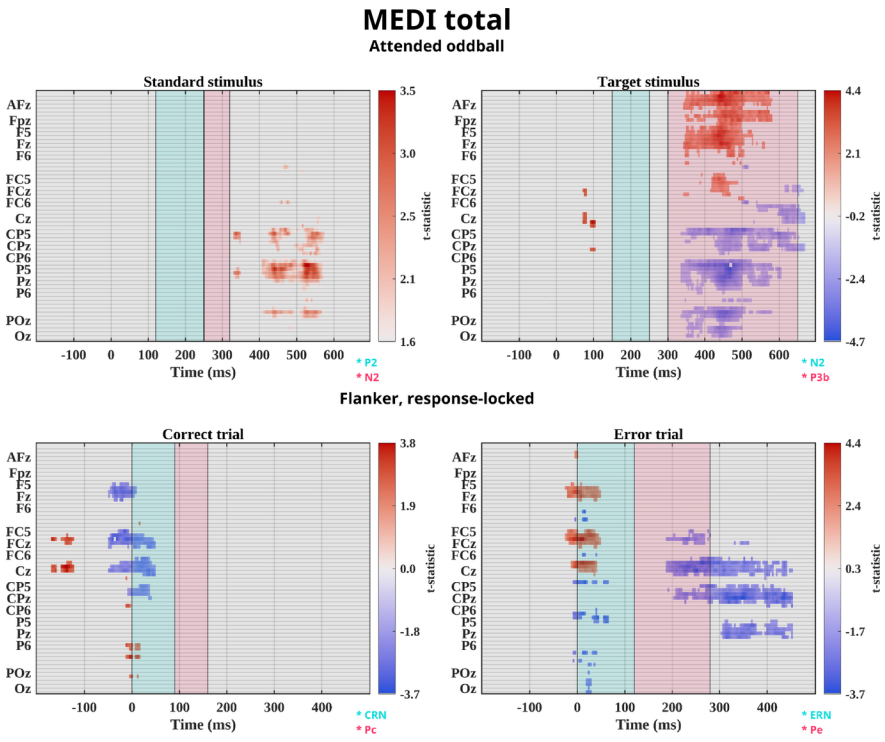



Figure 10: Results for MEDI total score showing correlations with ERP components as heat maps indicating significant regions after correction for multiple comparisons using threshold-free cluster-enhancement at an alpha level og 0.1%. Subplot title indicates LIMO model corresponding to conditions in the respective paradigms. Shaded regions indicate traditional ERP component evaluation windows.

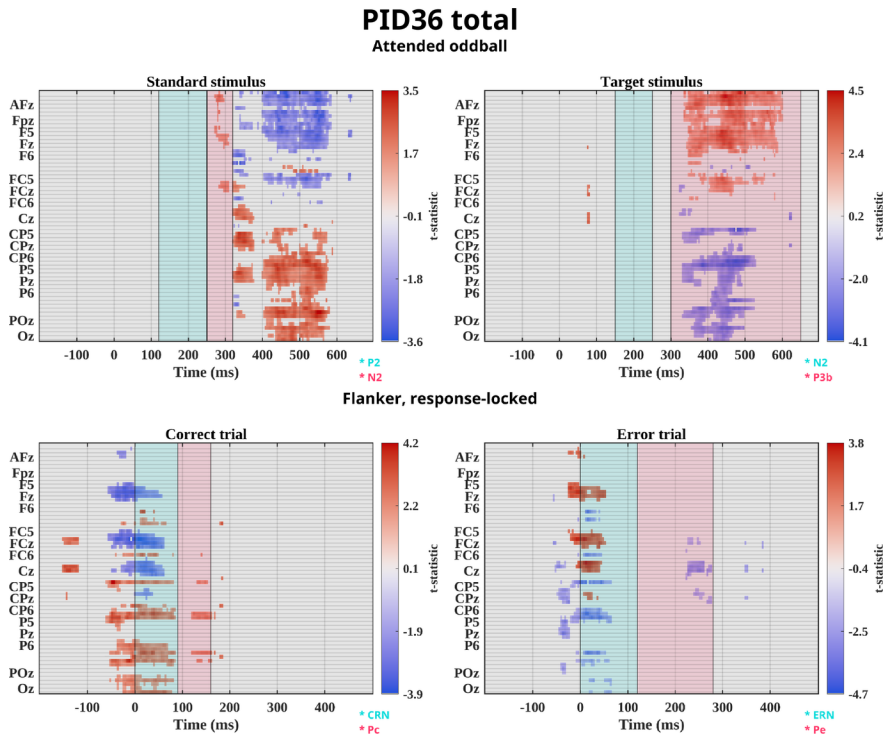



Figure 11: Results for PID36 total score showing correlations with ERP components as heat maps indicating significant regions after correction for multiple comparisons using threshold-free cluster-enhancement at an alpha level og 0.1%. Subplot title indicates LIMO model corresponding to conditions in the respective paradigms. Shaded regions indicate traditional ERP component evaluation windows.

# Supplementary materials

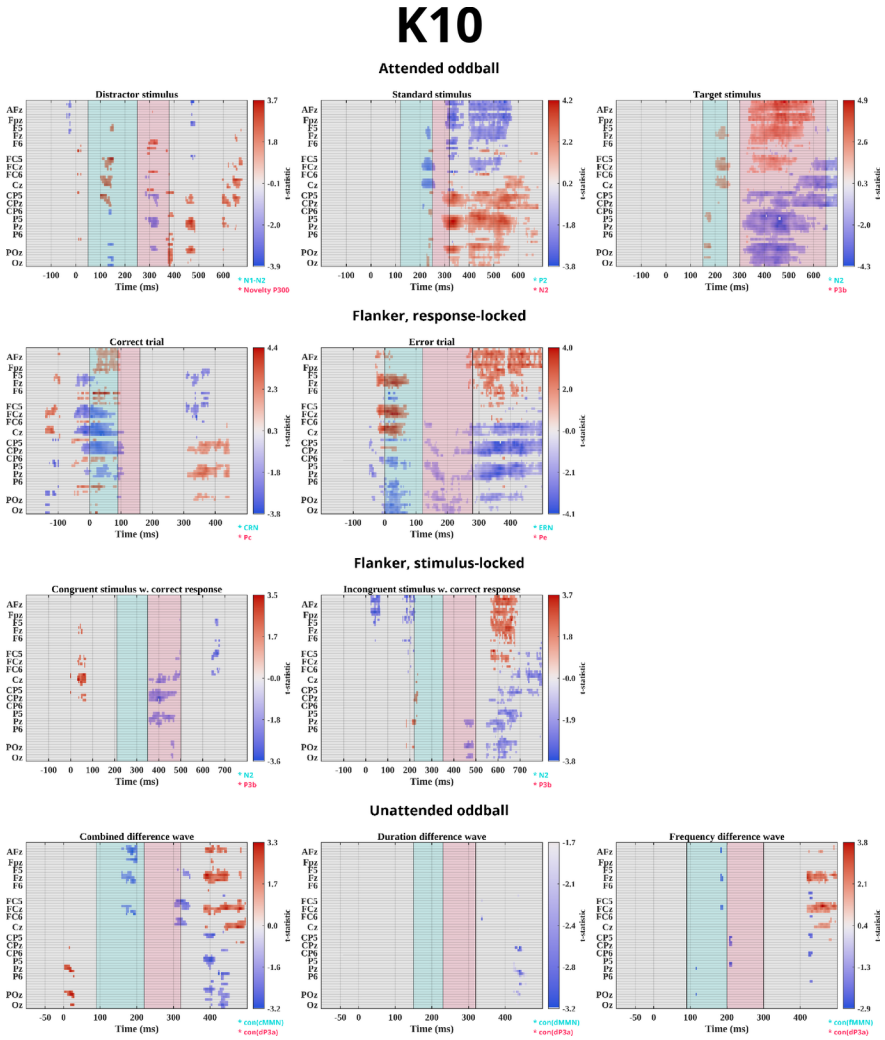



Figure S1: Post-hoc results for K10 showing correlations with ERP components as heat maps indicating significant regions after correction for multiple comparisons using threshold-free cluster-enhancement at an alpha level og 5%. Subplot title indicates LIMO model corresponding to conditions in the respective paradigms. Shaded regions indicate traditional ERP component evaluation windows.

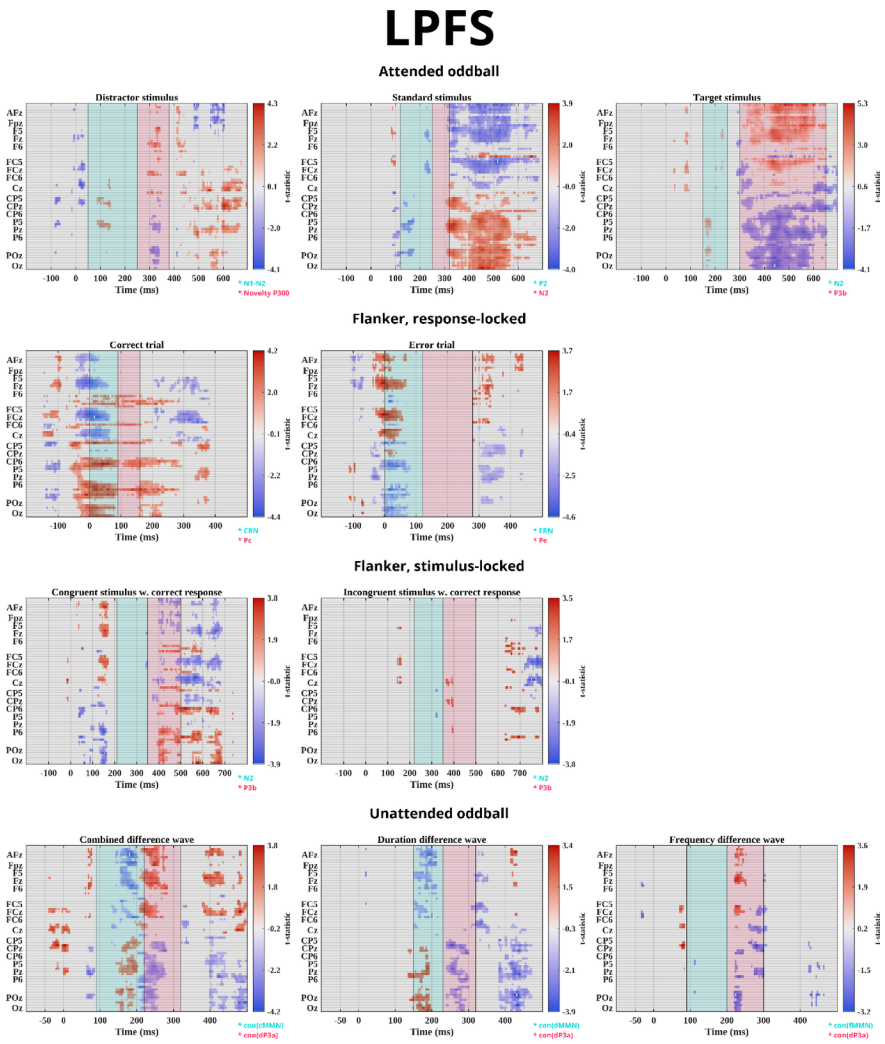
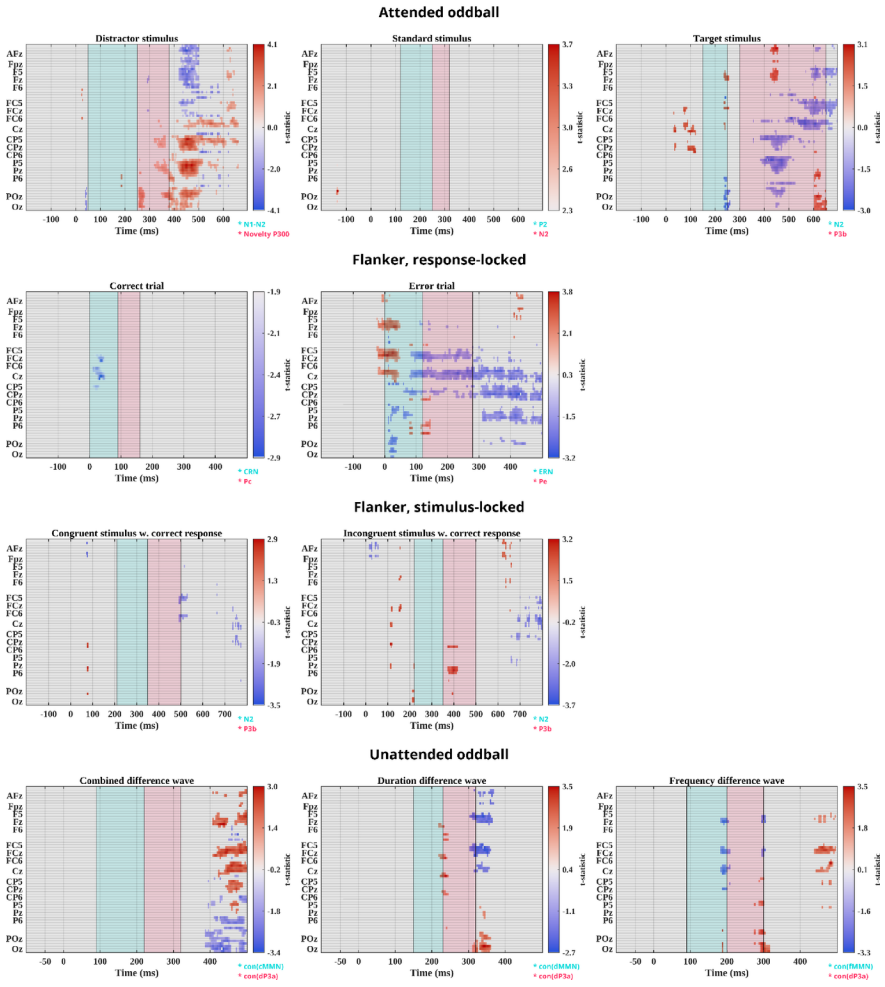



Figure S2: Post-hoc results for LPFS showing correlations with ERP components as heat maps indicating significant regions after correction for multiple comparisons using threshold-free cluster-enhancement at an alpha level og 5%. Subplot title indicates LIMO model corresponding to conditions in the respective paradigms. Shaded regions indicate traditional ERP component evaluation windows.

# MEDI Autonomic Arousal



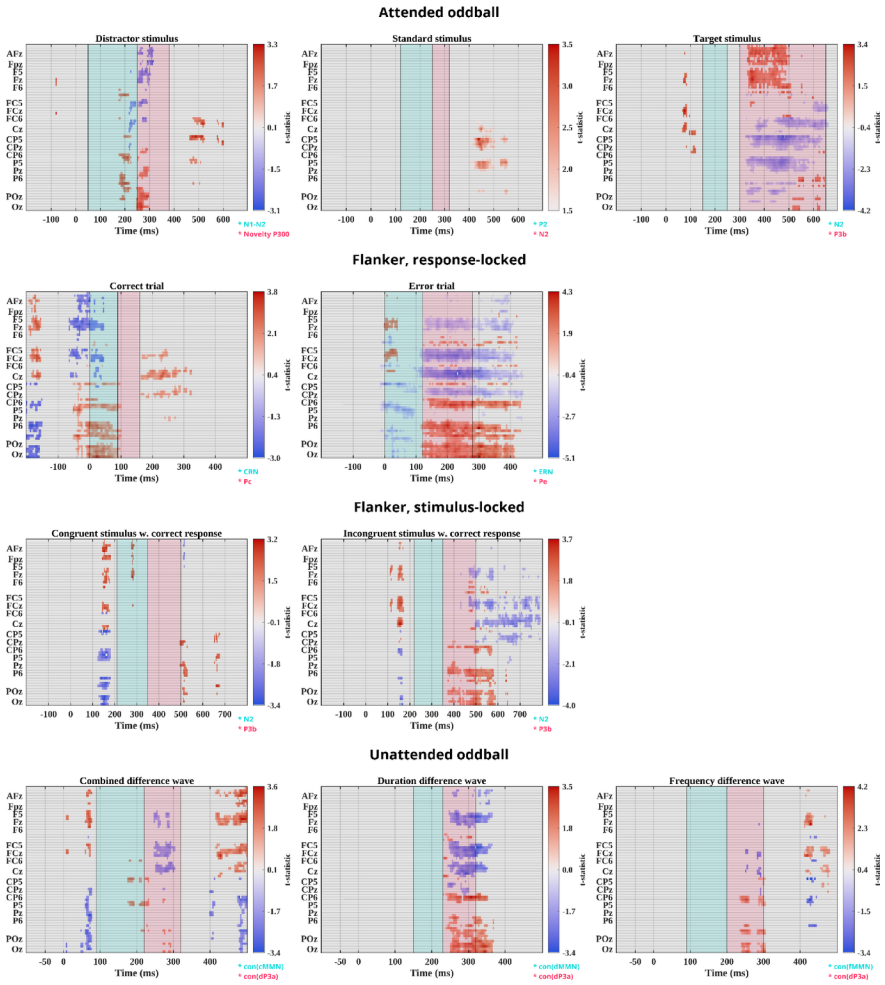



Figure S3: Post-hoc results for MEDI Autonomic Arousal showing correlations with ERP components as heat maps indicating significant regions after correction for multiple comparisons using threshold-free cluster-enhancement at an alpha level og 5%. Subplot title indicates LIMO model corresponding to conditions in the respective paradigms. Shaded regions indicate traditional ERP component evaluation windows.

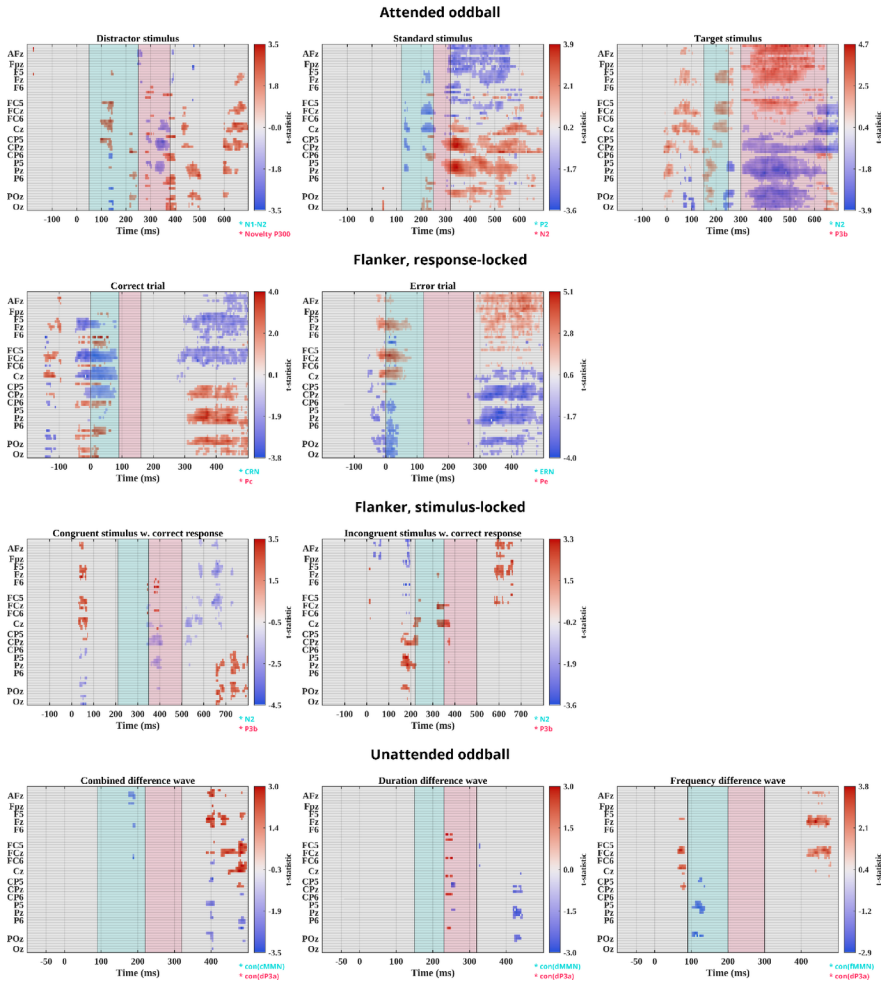



Figure S4: Post-hoc results for MEDI Avoidance showing correlations with ERP components as heat maps indicating significant regions after correction for multiple comparisons using threshold-free cluster-enhancement at an alpha level og 5%. Subplot title indicates LIMO model corresponding to conditions in the respective paradigms. Shaded regions indicate traditional ERP component evaluation windows.

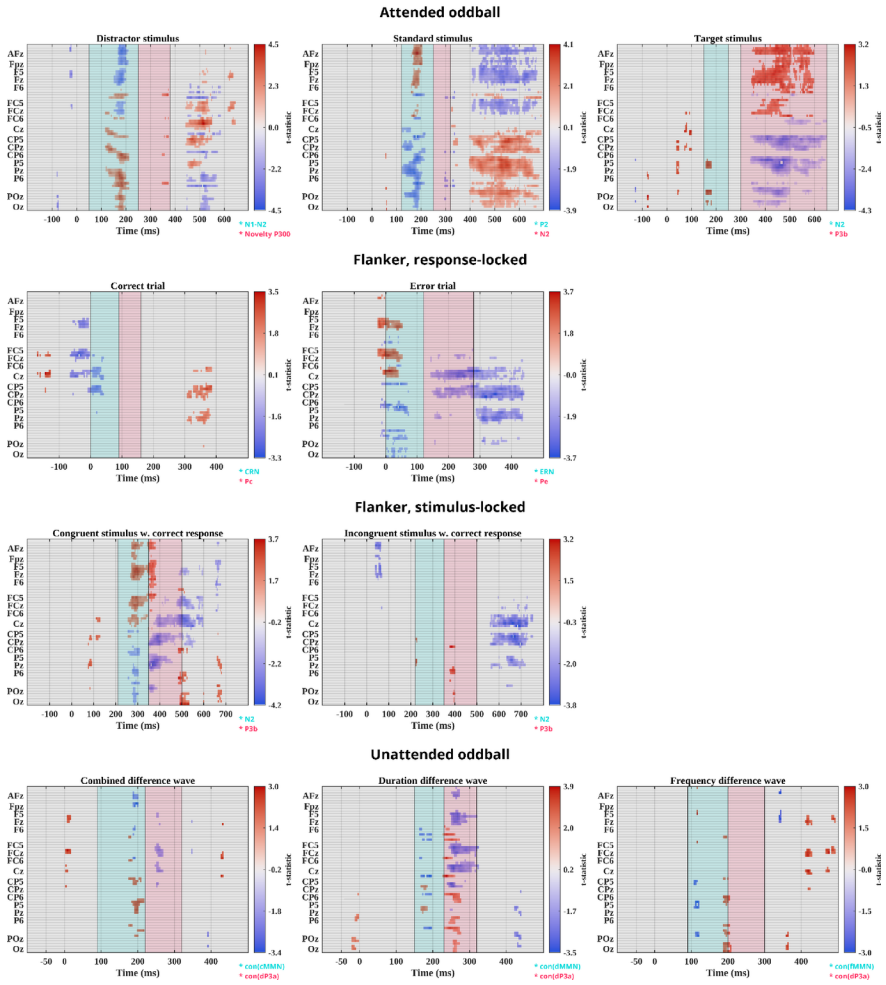



Figure S5: Post-hoc results for MEDI Depression showing correlations with ERP components as heat maps indicating significant regions after correction for multiple comparisons using threshold-free cluster-enhancement at an alpha level og 5%. Subplot title indicates LIMO model corresponding to conditions in the respective paradigms. Shaded regions indicate traditional ERP component evaluation windows.

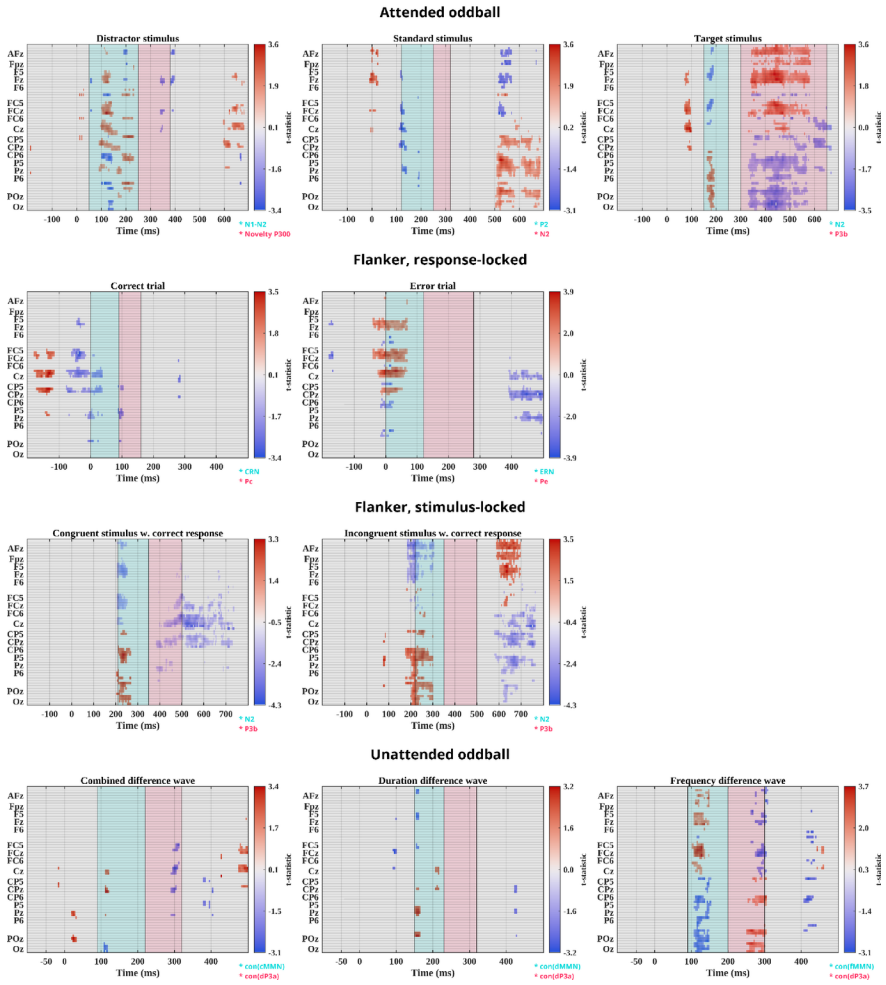



Figure S6: Post-hoc results for MEDI Intrusive Cognitions showing correlations with ERP components as heat maps indicating significant regions after correction for multiple comparisons using threshold-free cluster-enhancement at an alpha level og 5%. Subplot title indicates LIMO model corresponding to conditions in the respective paradigms. Shaded regions indicate traditional ERP component evaluation windows.

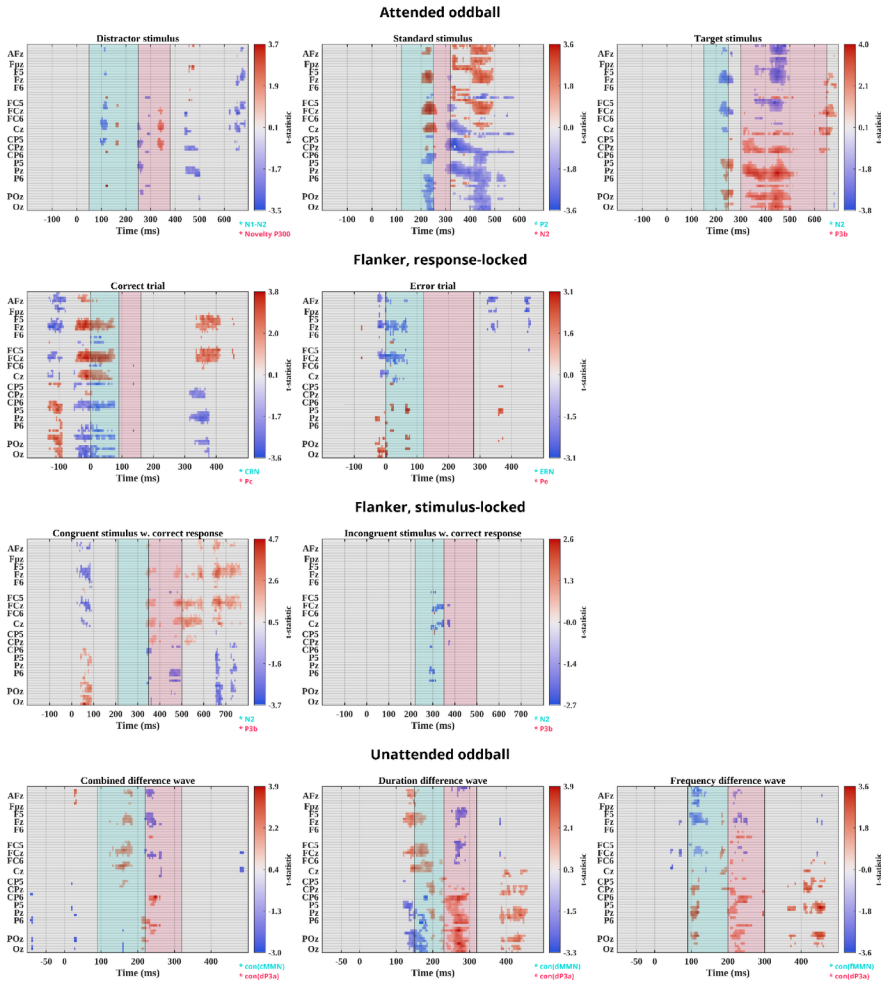



Figure S7: Post-hoc results for MEDI Neurotic Temperament showing correlations with ERP components as heat maps indicating significant regions after correction for multiple comparisons using threshold-free cluster-enhancement at an alpha level og 5%. Subplot title indicates LIMO model corresponding to conditions in the respective paradigms. Shaded regions indicate traditional ERP component evaluation windows.

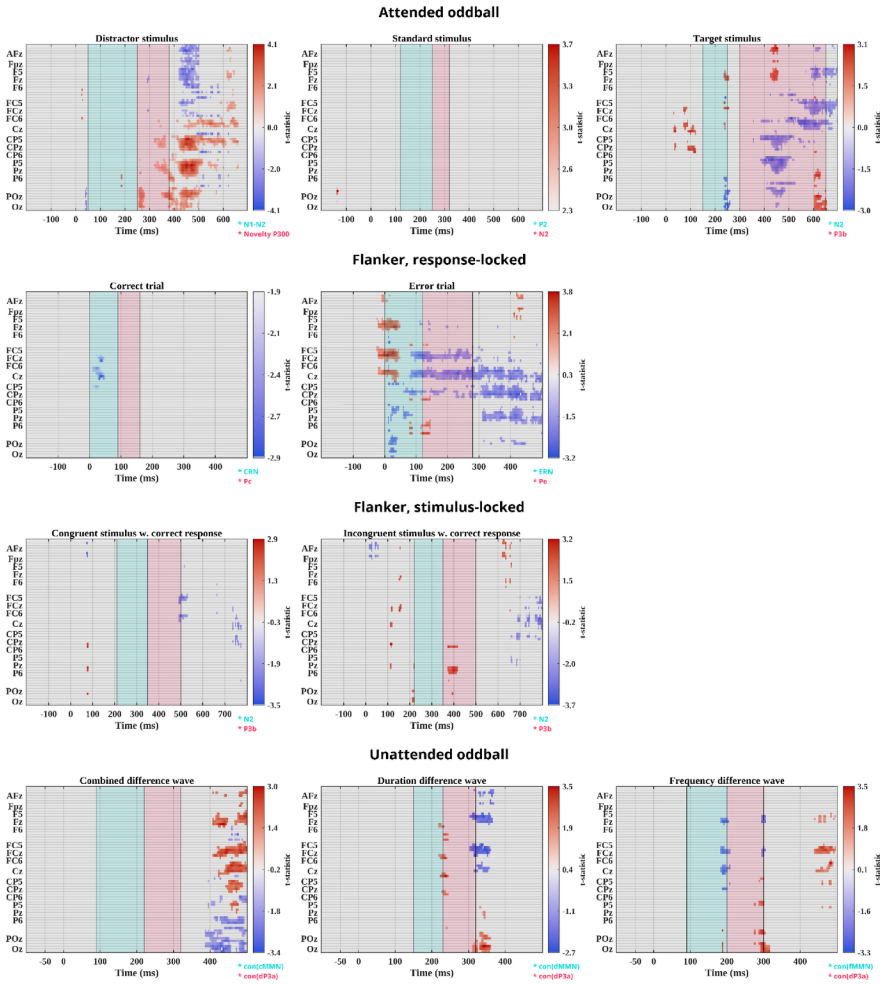



Figure S8: Post-hoc results for MEDI Positive Temperament showing correlations with ERP components as heat maps indicating significant regions after correction for multiple comparisons using threshold-free cluster-enhancement at an alpha level og 5%. Subplot title indicates LIMO model corresponding to conditions in the respective paradigms. Shaded regions indicate traditional ERP component evaluation windows.

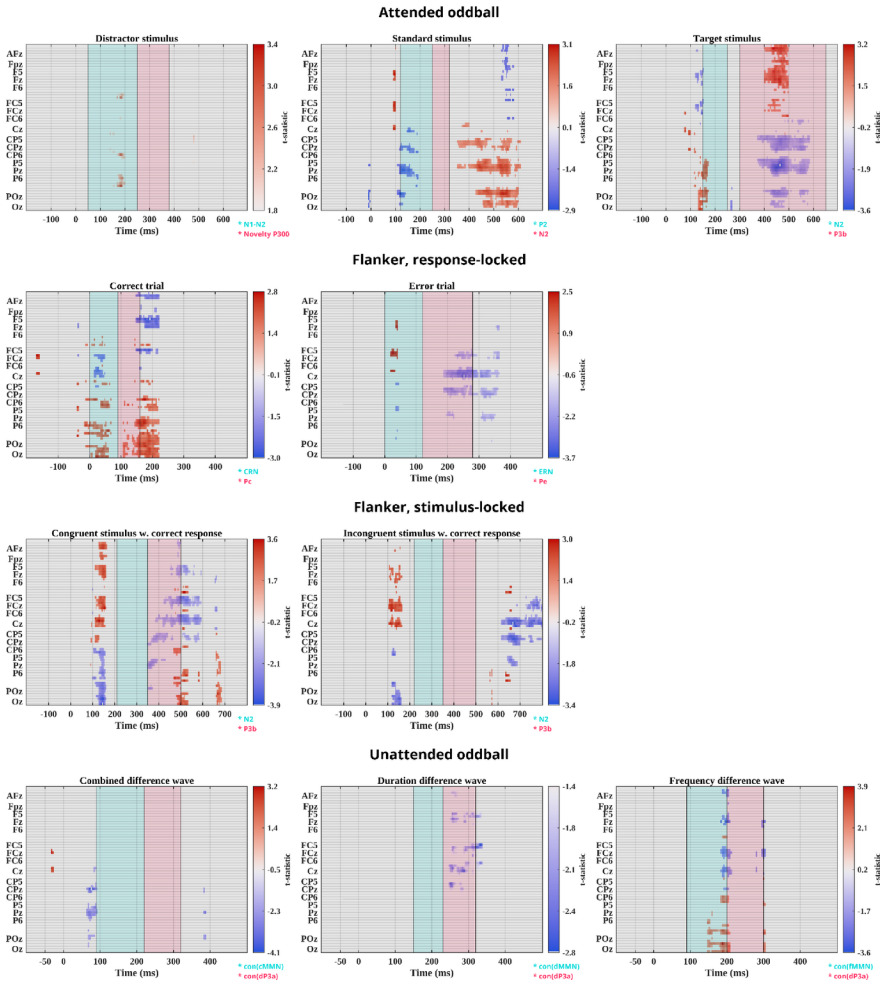



Figure S9: Post-hoc results for MEDI Social Anxiety showing correlations with ERP components as heat maps indicating significant regions after correction for multiple comparisons using threshold-free cluster-enhancement at an alpha level og 5%. Subplot title indicates LIMO model corresponding to conditions in the respective paradigms. Shaded regions indicate traditional ERP component evaluation windows.

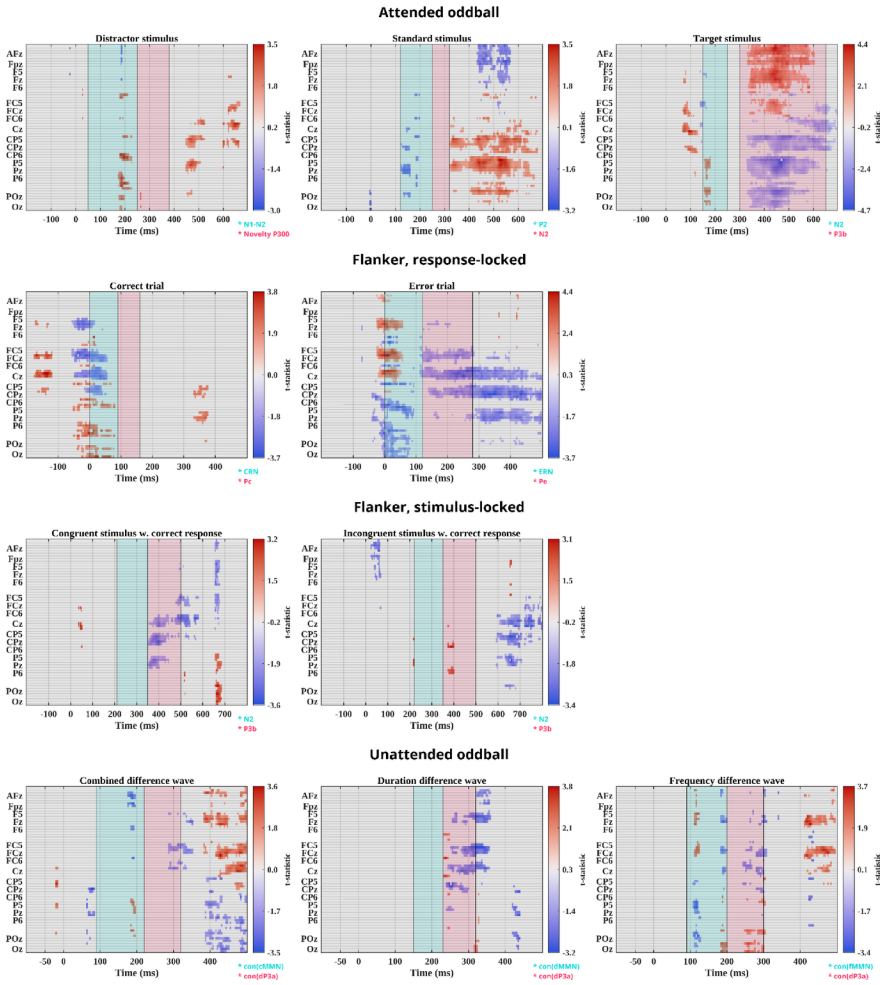



Figure S10: Post-hoc results for MEDI Somatic Anxiety showing correlations with ERP components as heat maps indicating significant regions after correction for multiple comparisons using threshold-free cluster-enhancement at an alpha level og 5%. Subplot title indicates LIMO model corresponding to conditions in the respective paradigms. Shaded regions indicate traditional ERP component evaluation windows.

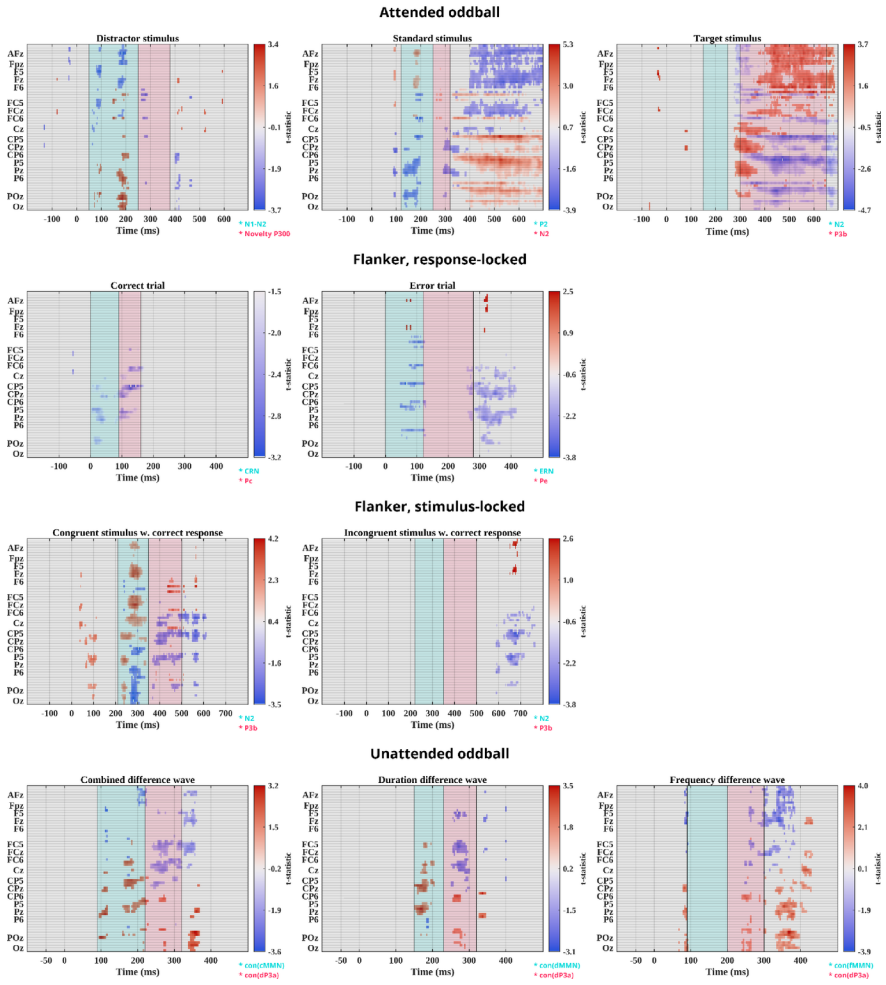



Figure S11: Post-hoc results for MEDI total showing correlations with ERP components as heat maps indicating significant regions after correction for multiple comparisons using threshold-free cluster-enhancement at an alpha level og 5%. Subplot title indicates LIMO model corresponding to conditions in the respective paradigms. Shaded regions indicate traditional ERP component evaluation windows.

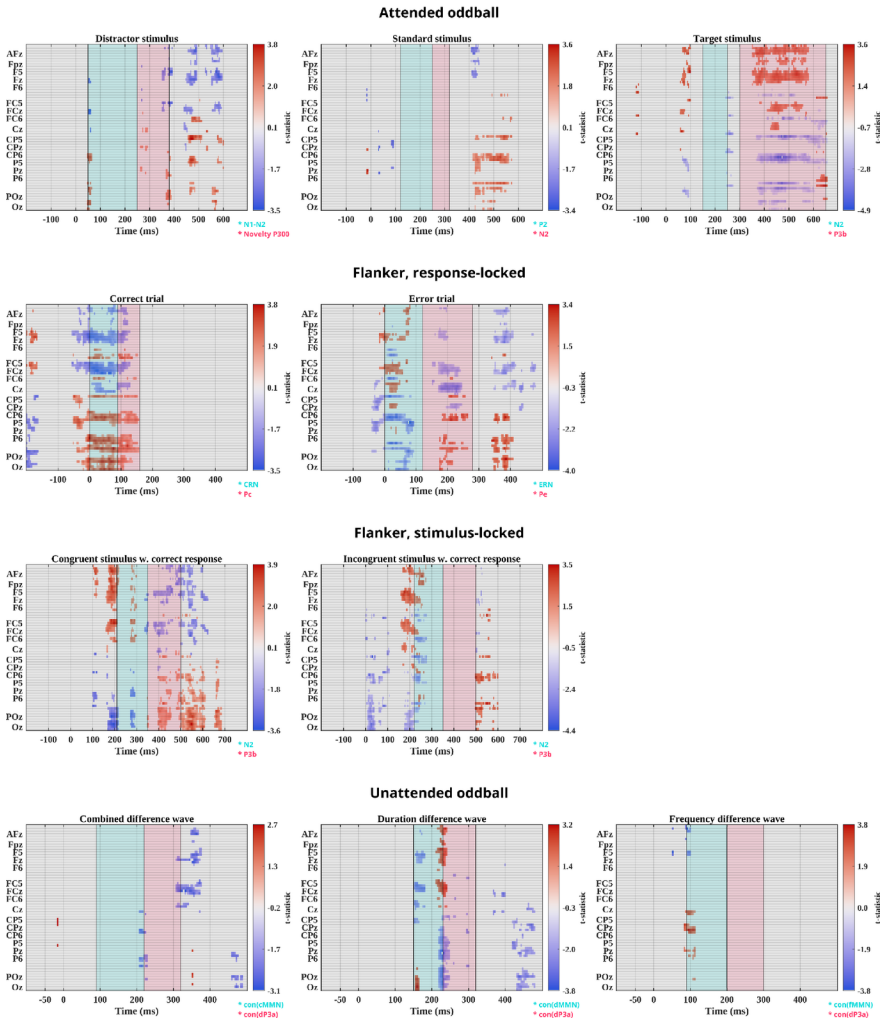



Figure S12: Post-hoc results for MEDI Tramatic Re-experiencing showing correlations with ERP components as heat maps indicating significant regions after correction for multiple comparisons using threshold-free cluster-enhancement at an alpha level og 5%. Subplot title indicates LIMO model corresponding to conditions in the respective paradigms. Shaded regions indicate traditional ERP component evaluation windows.

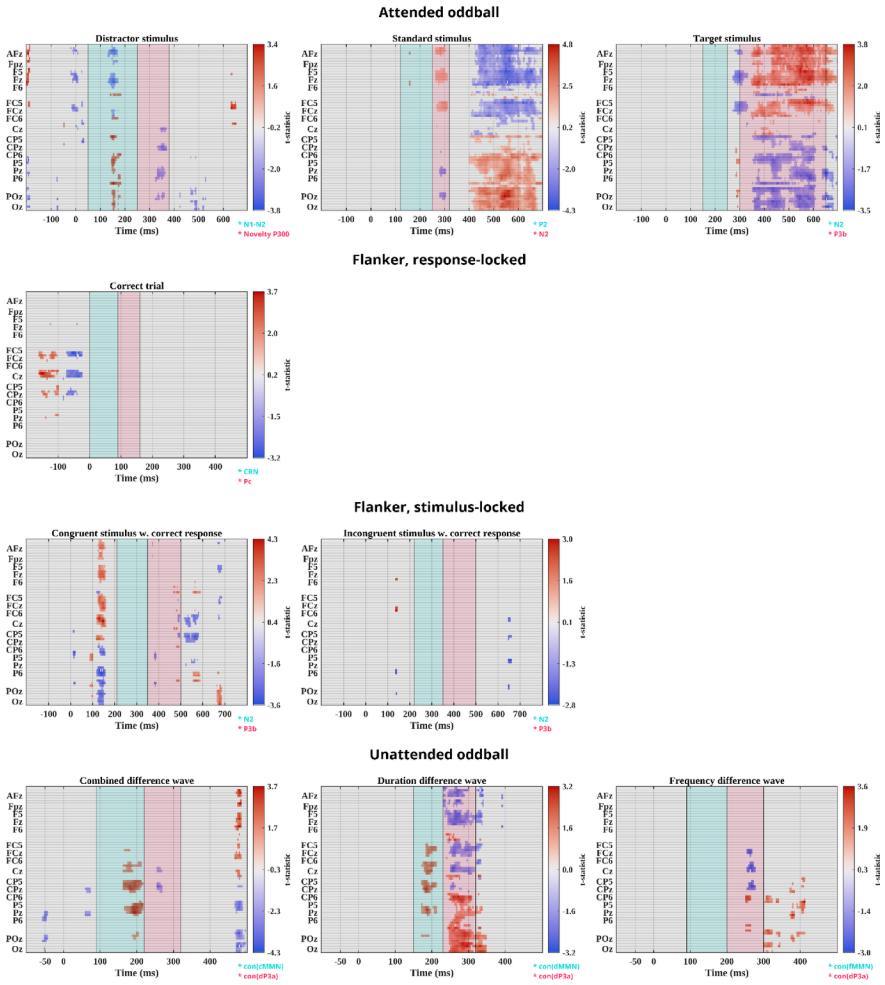



Figure S13: Post-hoc results for PID36 Anankastia showing correlations with ERP components as heat maps indicating significant regions after correction for multiple comparisons using threshold-free cluster-enhancement at an alpha level og 5%. Subplot title indicates LIMO model corresponding to conditions in the respective paradigms. Shaded regions indicate traditional ERP component evaluation windows.

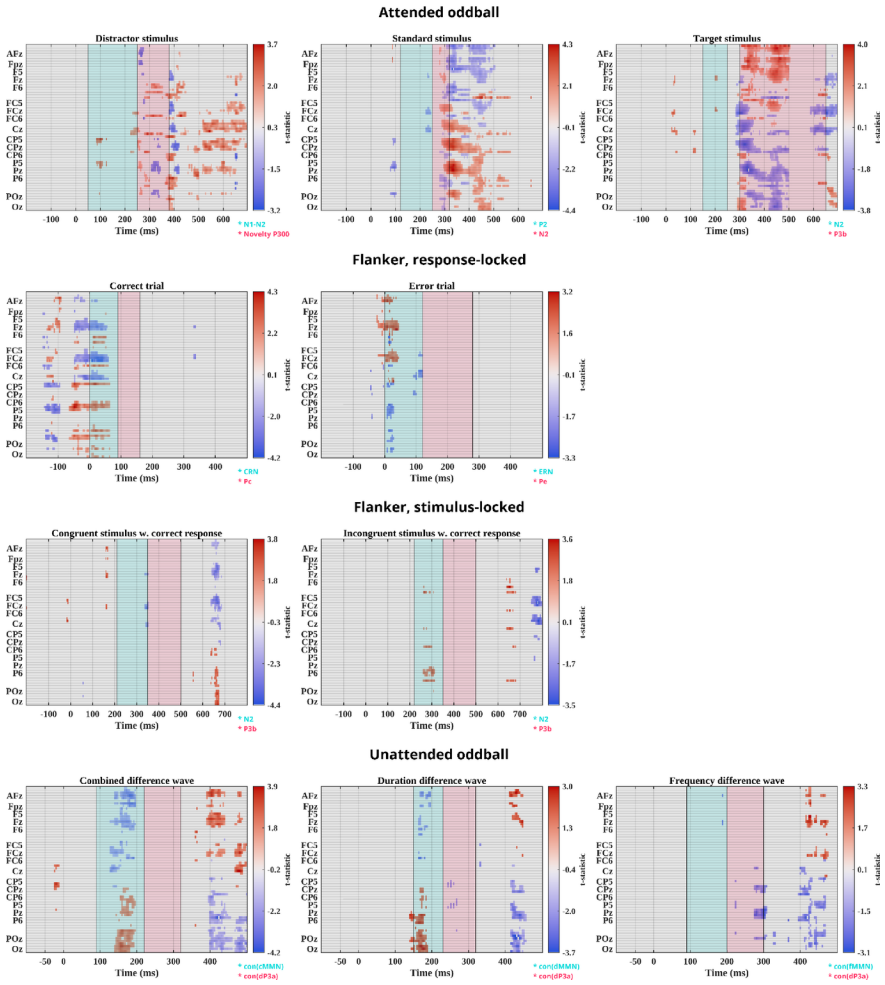



Figure S14: Post-hoc results for PID36 Antagonism showing correlations with ERP components as heat maps indicating significant regions after correction for multiple comparisons using threshold-free cluster-enhancement at an alpha level og 5%. Subplot title indicates LIMO model corresponding to conditions in the respective paradigms. Shaded regions indicate traditional ERP component evaluation windows.

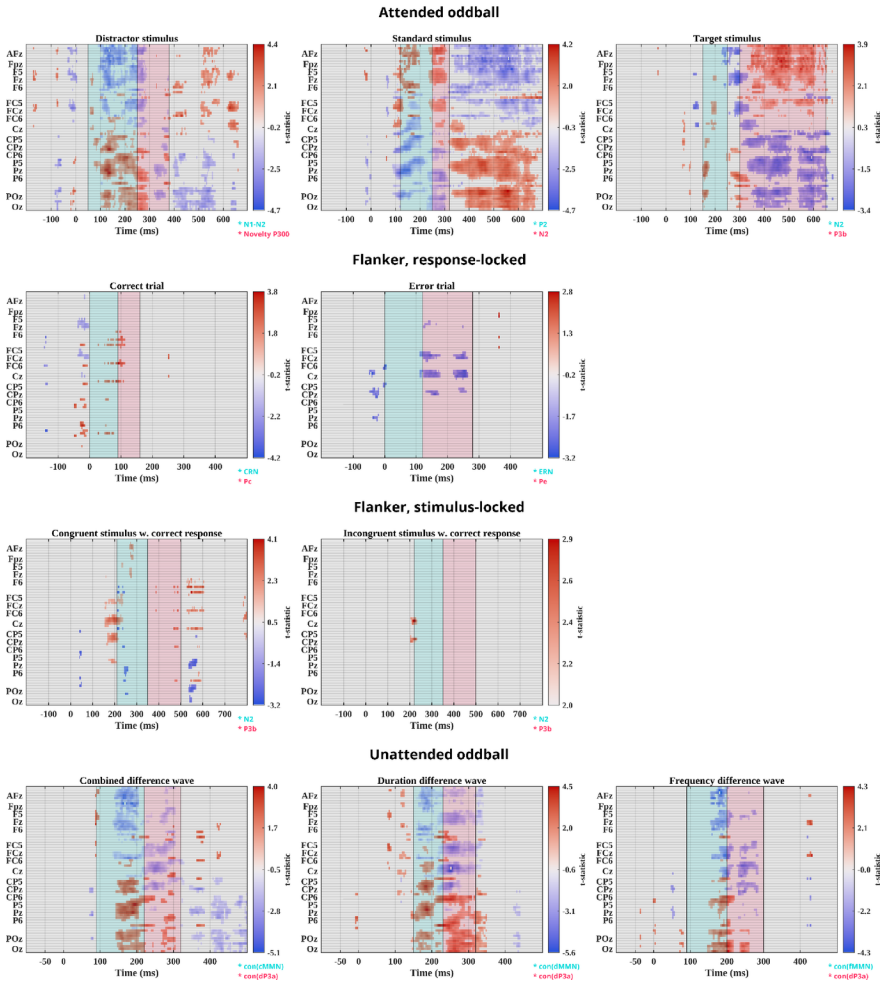



Figure S15: Post-hoc results for PID36 Detachment showing correlations with ERP components as heat maps indicating significant regions after correction for multiple comparisons using threshold-free cluster-enhancement at an alpha level og 5%. Subplot title indicates LIMO model corresponding to conditions in the respective paradigms. Shaded regions indicate traditional ERP component evaluation windows.

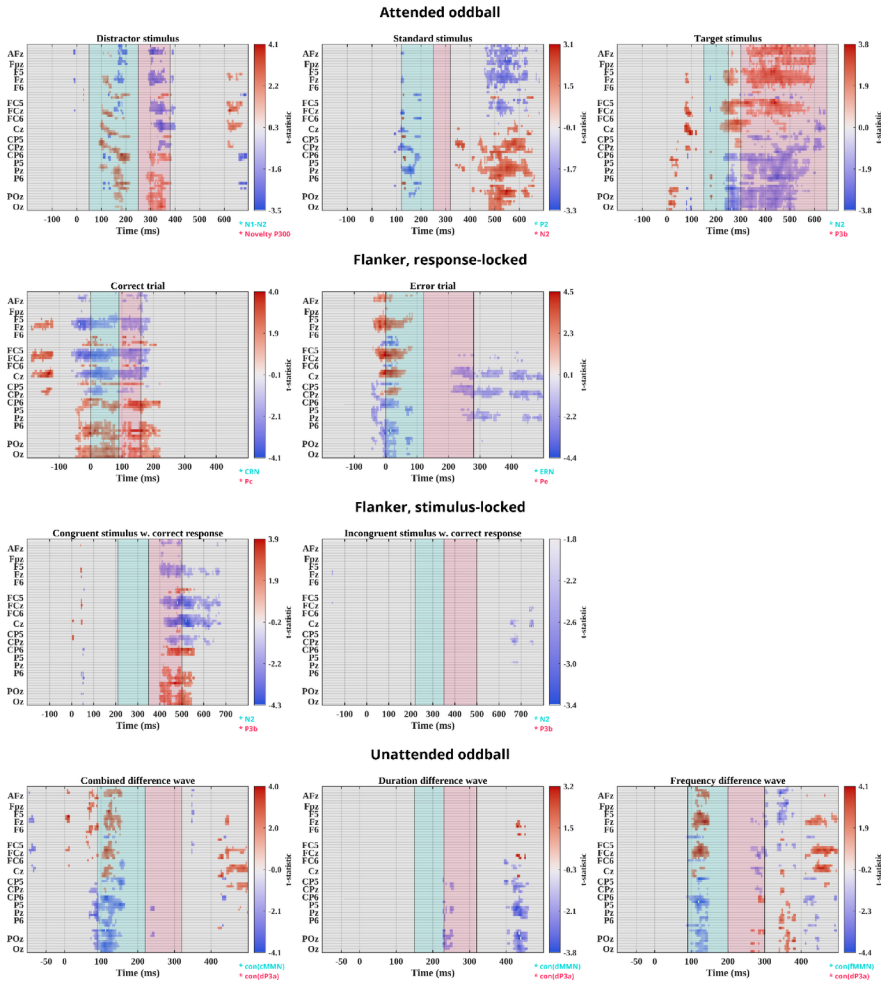



Figure S16: Post-hoc results for PID36 Disinhibition showing correlations with ERP components as heat maps indicating significant regions after correction for multiple comparisons using threshold-free cluster-enhancement at an alpha level og 5%. Subplot title indicates LIMO model corresponding to conditions in the respective paradigms. Shaded regions indicate traditional ERP component evaluation windows.

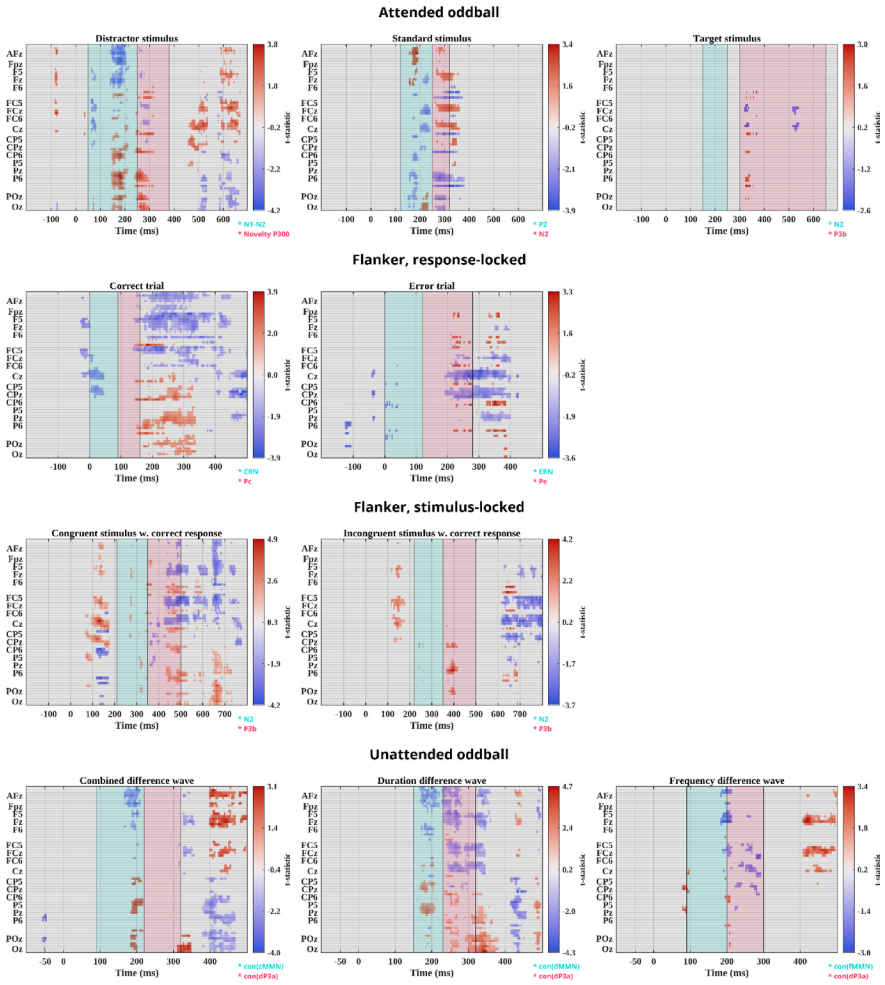



Figure S17: Post-hoc results for PID36 Negative Affect showing correlations with ERP components as heat maps indicating significant regions after correction for multiple comparisons using threshold-free cluster-enhancement at an alpha level og 5%. Subplot title indicates LIMO model corresponding to conditions in the respective paradigms. Shaded regions indicate traditional ERP component evaluation windows.

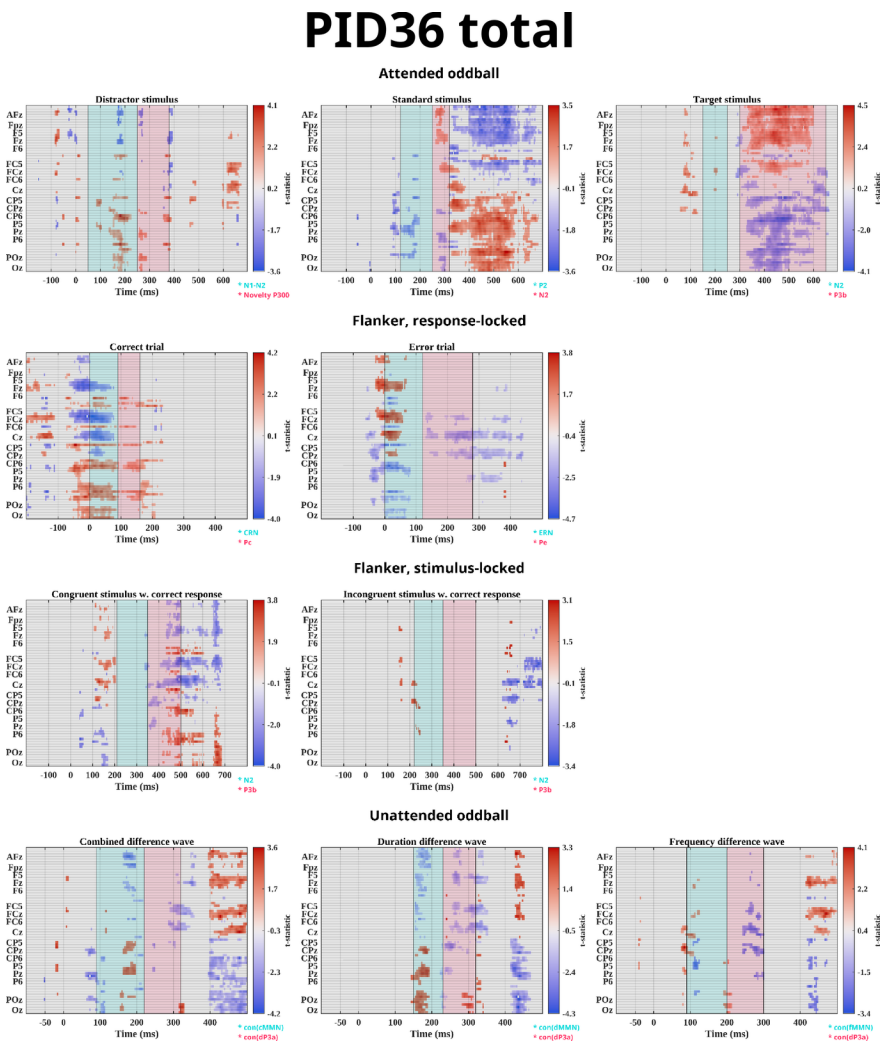



Figure S18: Post-hoc results for PID36 Psychoticism showing correlations with ERP components as heat maps indicating significant regions after correction for multiple comparisons using threshold-free cluster-enhancement at an alpha level og 5%. Subplot title indicates LIMO model corresponding to conditions in the respective paradigms. Shaded regions indicate traditional ERP component evaluation windows.

Figure S19: Post-hoc results for PID36 total showing correlations with ERP components as heat maps indicating significant regions after correction for multiple comparisons using threshold-free cluster-enhancement at an alpha level og 5%. Subplot title indicates LIMO model corresponding to conditions in the respective paradigms. Shaded regions indicate traditional ERP component evaluation windows.

# 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

Al-Ezzi, A., Kamel, N., Faye, I., & Gunaseli, E. (2020). Review of EEG, ERP, and Brain Connectivity Estimators as Predictive Biomarkers of Social Anxiety Disorder.*Frontiers in Psychology* , *11* (May). https://doi.org/10.3389/fpsyg.2020.00730

Bach, B., Kerber, A., Aluja, A., Bastiaens, T., Keeley, J. W., Claes, L., Fossati, A., Gutierrez, F., Oliveira, S. E. S., Pires, R., Riegel, K. D., Rolland, J.-P., Roskam, I., Sellbom, M., Somma, A., Spanemberg, L., Strus, W., Thimm, J. C., Wright, A. G. C., & Zimmermann, J. (2020). International Assessment of DSM-5 and ICD-11 Personality Disorder Traits: Toward a Common Nosology in DSM-5.1. *Psychopathology* , 1–10. https://doi.org/10.1159/000507589

Bach, B., & Mulder, R. (2022). Empirical foundation of the ICD-11 classification of personality disorders. In *Personality disorders and pathology: Integrating clinical assessment and practice in the DSM-5 and ICD-11 era* (pp. 27–52). American Psychological Association. https://doi.org/10.1037/0000310-003

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

Barlow, D. H., Farchione, T. J., Sauer-Zavala, S., Latin, H. M., Ellard, K. K., Bullis, J. R., Bentley, K. H., Boettcher, H. T., & Cassiello-Robbins, C. (2017). *Unified protocol for transdiagnostic treatment of emotional disorders: Therapist guide* . Oxford University Press.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using **Lme4** . *Journal of Statistical Software* , *67* (1). https://doi.org/10.18637/jss.v067.i01

Bellato, A., Norman, L., Idrees, I., Ogawa, C. Y., Waitt, A., Zuccolo, P. F., Tye, C., Radua, J., Groom, M. J., & Shephard, E. (2021). A systematic review and meta-analysis of altered electrophysiological markers of performance monitoring in Obsessive-Compulsive Disorder (OCD), Gilles de la Tourette Syndrome (GTS), Attention-Deficit/Hyperactivity disorder (ADHD) and Autism. *Neuroscience & Biobehavioral Reviews* ,*131* , 964–987. https://doi.org/10.1016/J.NEUBIOREV.2021.10.018

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* .

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

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

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., Rocha, H. A., & Larson, M. J. (2021). The data-processing multiverse of event-related potentials (ERPs): A roadmap for the optimization and standardization of ERP processing and reduction pipelines. *NeuroImage* , *245* , 118712. https://doi.org/10.1016/j.neuroimage.2021.118712

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

Corr, P., & Mobbs, D. (2023). Editorial: An emerging field with bright prospects. *Personality Neuroscience* , *6* , e1. https://doi.org/10.1017/pen.2022.6

Cowen, P., & Sherwood, A. C. (2013). The role of serotonin in cognitive function: Evidence from recent studies and implications for understanding depression. *Journal of Psychopharmacology* , *27* (7), 575–583. https://doi.org/10.1177/0269881113482531

Cuthbert, B. N., & Insel, T. R. (2010). Toward new approaches to psychotic disorders: The NIMH research domain criteria project. *Schizophrenia Bulletin* , *36* (6), 1061–1062. https://doi.org/10.1093/schbul/sbq108

d'Ardhuy, X. L., Boeijinga, P. H., Renault, B., Luthringer, R., Rinaudo, G., Soufflet, L., Toussaint, M., & Macher, J.-P. (1999). Effects of Serotonin-Selective and Classical Antidepressants on the Auditory P300 Cognitive Potential. *Neuropsychobiology* , *40* (4), 207–213. https://doi.org/10.1159/000026621

de Aguiar Neto, F. S., & Rosa, J. L. G. (2019). Depression biomarkers using non-invasive EEG: A review. *Neuroscience & Biobehavioral Reviews* , *105* , 83–93. https://doi.org/10.1016/j.neubiorev.2019.07.021

Delorme, A. (2022). *EEG is better left alone* [Preprint]. Neuroscience. https://doi.org/10.1101/2022.12.03.518987

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

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.

Eppinger, B., Kray, J., Mock, B., & Mecklinger, A. (2008). Better or worse than expected? Aging, learning, and the ERN. *Neuropsychologia* ,*46* (2), 521–539. https://doi.org/10.1016/j.neuropsychologia.2007.09.001

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

Etkin, A., Gyurak, A., & O'Hara, R. (2013). A neurobiological approach to the cognitive deficits of psychiatric disorders. *Dialogues in Clinical Neuroscience* , *15* (4), 419–429. https://doi.org/10.31887/DCNS.2013.15.4/aetkin

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

Flora, D. B. (2020). Your Coefficient Alpha Is Probably Wrong, but Which Coefficient Omega Is Right? A Tutorial on Using R to Obtain Better Reliability Estimates. *Advances in Methods and Practices in Psychological Science* , *3* (4), 484–501. https://doi.org/10.1177/2515245920951747

Giordano, G. M., Perrottelli, A., Mucci, A., Di Lorenzo, G., Altamura, M., Bellomo, A., Brugnoli, R., Corrivetti, G., Girardi, P., Monteleone, P., Niolu, C., Galderisi, S., & Maj, M. (2021). Investigating the Relationships of P3b with Negative Symptoms and Neurocognition in Subjects with Chronic Schizophrenia. *Brain Sciences* ,*11* (12), 1632. https://doi.org/10.3390/brainsci11121632

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

Gorka, S. M., Burkhouse, K. L., Afshar, K., & Phan, K. L. (2017). Error-related brain activity and internalizing disorder symptom dimensions in depression and anxiety. *Depression and Anxiety* ,*34* (11), 985–995. https://doi.org/10.1002/da.22648

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

Hajcak, G., Klawohn, J., & Meyer, A. (2019). The Utility of Event-Related Potentials in Clinical Psychology. *Annual Review of Clinical Psychology* , *15* (1), 71–95. https://doi.org/10.1146/annurev-clinpsy-050718-095457

Hall, J. R., Bernat, E. M., & Patrick, C. J. (2007). Externalizing Psychopathology and the Error-Related Negativity. *Psychological Science* , *18* (4), 326–333. https://doi.org/10.1111/j.1467-9280.2007.01899.x

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

He, W., Chai, H., Zheng, L., Yu, W., Chen, W., Li, J., Chen, W., & Wang, W. (2010). Mismatch negativity in treatment-resistant depression and borderline personality disorder. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* , *34* (2), 366–371. https://doi.org/10.1016/j.pnpbp.2009.12.021

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

Hutsebaut, J., Feenstra, D. J., & Kamphuis, J. H. (2016). Development and Preliminary Psychometric Evaluation of a Brief Self-Report Questionnaire for the Assessment of the DSM-5 Level of Personality Functioning Scale: The LPFS Brief Form (LPFS-BF). *Personality Disorders: Theory, Research, and Treatment* , *7* (2), 192–197. https://doi.org/10.1037/per0000159

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., Rosseel, Y., Miller, P., Quick, C., Garnier-Villarreal, M., Selig, J., Boulton, A., Preacher, K., Coffman, D., Rhemtulla, M., Robitzsch, A., Enders, C., Arslan, R., Clinton, B., Panko, P., Merkle, E., Chesnut, S., ... Johnson, A. R. (2022). *semTools: Useful Tools for Structural Equation Modeling* .

Kessler, R. C., Barker, P. R., Colpe, L. J., Epstein, J. F., Gfroerer, J. C., Hiripi, E., Howes, M. J., Normand, S. L. T., Manderscheid, R. W., Walters, E. E., & Zaslavsky, A. M. (2003). Screening for serious mental illness in the general population. *Archives of General Psychiatry* , *60* (2), 184–189. https://doi.org/10.1001/archpsyc.60.2.184

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., 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., 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

Kruiper, C., Fagerlund, B., Nielsen, M. Ø., Düring, S., Jensen, M. H., Ebdrup, B. H., Glenthøj, B. Y., & Oranje, B. (2019). Associations between P3a and P3b amplitudes and cognition in antipsychotic-naïve first-episode schizophrenia patients. *Psychological Medicine* , *49* (5), 868–875. https://doi.org/10.1017/S0033291718001575

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

Latzman, R. D., & DeYoung, C. G. (2020). Using empirically-derived dimensional phenotypes to accelerate clinical neuroscience: The Hierarchical Taxonomy of Psychopathology (HiTOP) framework.*Neuropsychopharmacology* , *45* (7), 1083–1085. https://doi.org/10.1038/s41386-020-0639-6

Lavoie, S., Polari, A. R., Goldstone, S., Nelson, B., & McGorry, P. D. (2019). Staging model in psychiatry: Review of the evolution of electroencephalography abnormalities in major psychiatric disorders.*Early Intervention in Psychiatry* , *December 2018* , 1–10. https://doi.org/10.1111/eip.12792

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

Liu, Y., Shen, X., Zhu, Y., Xu, Y., Cai, W., Shen, M., Yu, R., & Wang, W. (2007). Mismatch negativity in paranoid, schizotypal, and antisocial personality disorders. *Neurophysiologie Clinique/Clinical Neurophysiology* , *37* (2), 89–96. https://doi.org/10.1016/j.neucli.2007.03.001

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.

Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods* , *49* (4), 1494–1502. https://doi.org/10.3758/s13428-016-0809-y

Lutz, M. C., Kok, R., Verveer, I., Malbec, M., Koot, S., Van Lier, P. A. C., & Franken, I. H. A. (2021). Diminished error-related negativity and error positivity in children and adults with externalizing problems and disorders: A meta-analysis on error processing. *Journal of Psychiatry and Neuroscience* , *46* (6), E615–E627. https://doi.org/10.1503/jpn.200031

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

Maron, E., & Nutt, D. (2022). Biological markers of generalized anxiety disorder. *Dialogues in Clinical Neuroscience* , *19* (2), 147–158. https://doi.org/10.31887/DCNS.2017.19.2/dnutt

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

Mavrogiorgou, P., Juckel, G., Frodl, T., Gallinat, J., Hauke, W., Zaudig, M., Dammann, G., Möller, H. J., & Hegerl, U. (2002). P300 subcomponents in obsessive-compulsive disorder. *Journal of Psychiatric Research* ,*36* (6), 399–406. https://doi.org/10.1016/S0022-3956(02)00055-9

Mensen, A., & Khatami, R. (2013). Advanced EEG analysis using threshold-free cluster-enhancement and non-parametric statistics. *NeuroImage* ,*67* , 111–118. https://doi.org/10.1016/j.neuroimage.2012.10.027

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

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

Nielsen, A. N., Barch, D. M., Petersen, S. E., Schlaggar, B. L., & Greene, D. J. (2020). Machine Learning With Neuroimaging: Evaluating Its Applications in Psychiatry. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* , *5* (8), 791–798. https://doi.org/10.1016/j.bpsc.2019.11.007

Olvet, D. M., & Hajcak, G. (2009a). Reliability of error-related brain activity. *Brain Research* , *1284* , 89–99. https://doi.org/10.1016/j.brainres.2009.05.079

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

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

Onitsuka, T., Oribe, N., Nakamura, I., & Kanba, S. (2013). Review of neurophysiological findings in patients with schizophrenia. *Psychiatry and Clinical Neurosciences* , *67* (7), 461–470. https://doi.org/10.1111/pcn.12090

Oranje, B., Jensen, K., Wienberg, M., & Glenthøj, B. Y. (2008-06-31). Divergent effects of increased serotonergic activity on psychophysiological parameters of human attention. *The International Journal of*

*Neuropsychopharmacology / Official Scientific Journal of the Collegium Internationale Neuropsychopharmacologicum (CINP)* , *11* (4), 453–463. https://doi.org/10.1017/S1461145707008176

Overbeek, T. J. M., Nieuwenhuis, S., & Ridderinkhof, K. R. (2005). Dissociable Components of Error Processing. *Journal of Psychophysiology* ,*19* (4), 319–329. https://doi.org/10.1027/0269-8803.19.4.319

Parker, D. A., Trotti, R. L., McDowell, J. E., Keedy, S. K., Hill, S. K., Gershon, E. S., Ivleva, E. I., Pearlson, G. D., Keshavan, M. S., Tamminga, C. A., & Clementz, B. A. (2021). Auditory Oddball Responses Across the Schizophrenia-Bipolar Spectrum and Their Relationship to Cognitive and Clinical Features. *American Journal of Psychiatry* ,*178* (10), 952–964. https://doi.org/10.1176/appi.ajp.2021.20071043

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

Paterniti, S., Dufouil, C., Bisserbe, J.-C., & Alpérovitch, A. (1999). Anxiety, depression, psychotropic drug use and cognitive impairment.*Psychological Medicine* , *29* (2), 421–428. https://doi.org/10.1017/S0033291798008010

Patrick, C. J., Bernat, E. M., Malone, S. M., Iacono, W. G., Krueger, R. F., & McGue, M. (2006). P300 amplitude as an indicator of externalizing in adolescent males. *Psychophysiology* , *43* (1), 84–92. https://doi.org/10.1111/j.1469-8986.2006.00376.x

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

Petrolini, V., & Vicente, A. (2022). The challenges raised by comorbidity in psychiatric research: The case of autism. *Philosophical Psychology* , *35* (8), 1234–1263. https://doi.org/10.1080/09515089.2022.2052829

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.

Randau, M., Reinholt, N., Pernet, C., Oranje, B., Rasmussen, B., & Arnfred, S. (2023). *Robust single-trial event-related potentials differentiate between Distress and Fear disorder* . https://doi.org/10.22541/au.168998357.78763078/v1

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

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

Rosellini, A. J., & Brown, T. A. (2019). The Multidimensional Emotional Disorder Inventory (MEDI): Assessing transdiagnostic dimensions to validate a profile approach to emotional disorder classification.*Psychological Assessment* , *31* (1), 59–72. https://doi.org/10.1037/pas0000649

Sassenhagen, J., & Draschkow, D. (2019). Cluster-based permutation tests of MEG/EEG data do not establish significance of effect latency or location.*Psychophysiology* , *56* (6), e13335. https://doi.org/10.1111/psyp.13335

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. (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. In *J Clin Psychiatry* (suppl 20; Vol. 59, pp. 0–0). Physicians Postgraduate Press, Inc.

Simó, M., Gurtubay-Antolin, A., Vaquero, L., Bruna, J., & Rodríguez-Fornells, A. (2018). Performance monitoring in lung cancer patients pre- and post-chemotherapy using fine-grained electrophysiological measures.*NeuroImage: Clinical* , *18* , 86–96. https://doi.org/10.1016/j.nicl.2017.12.032

Singh, I., & Rose, N. (2009). Biomarkers in psychiatry. *Nature* ,*460* (7252), 202–207. https://doi.org/10.1038/460202a

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

Swainston, J., Louis, C., Moser, J., & Derakshan, N. (2021). Neurocognitive efficiency in breast cancer survivorship: A performance monitoring ERP study. *International Journal of Psychophysiology* , *168* , 9–20. https://doi.org/10.1016/j.ijpsycho.2021.06.013

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

Tanovic, E., Hajcak, G., & Sanislow, C. A. (2017). Rumination is associated with diminished performance monitoring. *Emotion* , *17* (6), 953–964. https://doi.org/10.1037/emo0000290

Tseng, Y.-J., Nouchi, R., & Cheng, C.-H. (2021). Mismatch negativity in patients with major depressive disorder: A meta-analysis. *Clinical Neurophysiology* , *132* (10), 2654–2665. htt-

ps://doi.org/10.1016/j.clinph.2021.06.019

Wager, T. D., Keller, M. C., Lacey, S. C., & Jonides, J. (2005). Increased sensitivity in neuroimaging analyses using robust regression.*NeuroImage* , *26* (1), 99–113. https://doi.org/10.1016/j.neuroimage.2005.01.011

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

Wendt, L. P., Jankowsky, K., Zimmermann, J., Schroeders, U., Nolte, T., Fonagy, P., Montague, P. R., & Olaru, G. (2021). Established self-report questionnaires of psychopathology can be efficiently summarized with HiTOP. *PsyArXiv* , *January* .

Wessel, J. R. (2012). Error awareness and the error-related negativity: Evaluating the first decade of evidence. *Frontiers in Human Neuroscience* , *6* .

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

Widiger, T. A., & Oltmanns, J. R. (2017). Neuroticism is a fundamental domain of personality with enormous public health implications. *World Psychiatry* , *16* (2), 144–145. https://doi.org/10.1002/wps.20411

Wienberg, M., Glenthoj, B., Jensen, K., & Oranje, B. (2010). A single high dose of escitalopram increases mismatch negativity without affecting processing negativity or P300 amplitude in healthy volunteers. *Journal of Psychopharmacology* , *24* (8), 1183–1192. https://doi.org/10.1177/0269881109102606

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

Winkler, I., Denham, S., & Escera, C. (2013). Auditory Event-related Potentials. In D. Jaeger & R. Jung (Eds.), *Encyclopedia of Computational Neuroscience* (pp. 1–29). Springer. https://doi.org/10.1007/978-1-4614-7320-6_99-1

Zald, D. H., & Lahey, B. B. (2017). Implications of the Hierarchical Structure of Psychopathology for Psychiatric Neuroimaging.*Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* ,*2* (4), 310–317. https://doi.org/10.1016/j.bpsc.2017.02.003