The role of performance monitoring in psychopathology

PhD Dissertation

Erasmus University Rotterdam

Miranda C. Lutz

July 2023

Table of Contents

Chapter 1

General Introduction

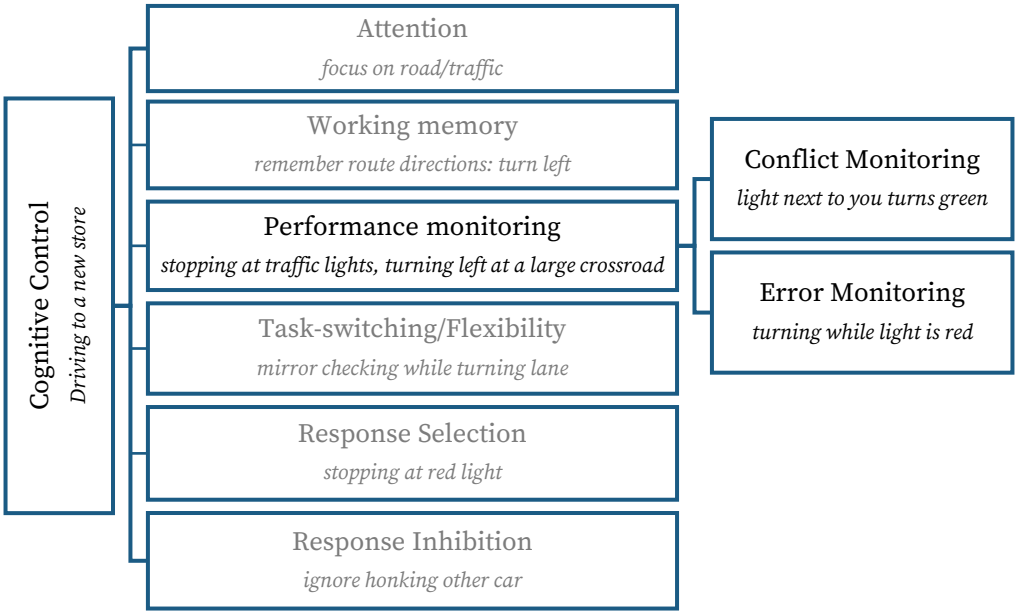
How Performance Monitoring Can Help Us Understand Complex Disorders

Cognitive control is an umbrella term for the cognitive processes that allow for the intrinsic ability to change behavior easily, willingly, and adaptively according to self-defined goals, while being subjected to the ever-changing demands of the environment (Badre, 2011; Shackman et al., 2011). Cognitive control processes are for example, selective attention, task-switching, working memory, response selection, response inhibition, and performance monitoring (Luna et al., 2015). Cognitive control processes are reactive and proactive, and contribute to maintaining goal-oriented actions, suppressing irrelevant or competing information, and evaluating current actions. In this thesis, I study one cognitive control component: performance monitoring. There are several running definitions of performance monitoring, for a detailed overview, see Ullsperger et al. (2014). Here, performance monitoring is considered the continuous supervision of one's activities and the initiation of behavioral adaptations to ensure goal-directed behavior. There are two important and partly intertwined processes of performance monitoring: conflict and error monitoring, depicted in Figure 1. The two concepts are the central topics studied in this dissertation.

Performance monitoring is essential for learning, behavior, and emotion regulation. Identifying how performance monitoring facilitates observable behavior will help us understand goal-directed behavior and learning. In individuals with psychopathology, deviant information processing is a mechanism thought to be underlying symptomatology and maladaptive behavior. Studying performance monitoring, along with other indicators of cognitive control, is, therefore, considered important in understanding how symptoms or maladaptive behavior typically observed in psychopathology develops, persists, or worsens. This is why the current dissertation demonstrates the role of performance monitoring in psychopathology.

Figure 1.

Visual presentation of performance monitoring in cognitive control processes and a practical example of how they steer behavior.



In this dissertation, I refer to psychopathology when considering mental disorders that are clinically assessed. Otherwise, I considered the subclinical level of symptoms as psychological problems. Understanding why individuals develop psychopathology is an increasingly important endeavor. This is because approximately one in every eight individuals in the world is affected by psychopathology (Institute of Health Metrics and Evaluation. Global Health Data Exchange; GHDx, 2022), causing severe problems for the individual, their direct surroundings and society. The traditional categorical classification systems, such as the DSM and ICD, do not aid in understanding and treating psychopathology (Kotov et al., 2017), because of symptom overlap (Forbes et al., 2023), outdated standards (in the case of depression: Fried, 2017), limited reliability of observed phenomena, and large heterogeneity of diagnosis within categories. This has led researchers and practitioners to shift from a categorical classification to a more dimensional approach to psychopathology (Conway et al., 2019). The dimensional approach proposes common hierarchical clusters of comprehensive and corresponding disorders, where most disorders load onto an internalizing and externalizing spectrum (Hierarchical Taxonomy of Psychopathology, HiTOP; Krueger et al., 2018). The problem behavior observed in internalizing disorders is characterized by emotional, harmful, fearful, depressive, and somatic symptoms typically directed toward the individuals themselves. Disorders like major depression, anxiety, and fear disorders are part of this dimension. Externalizing disorders (Krueger et al., 2007; Krueger et al., 2009) involve maladaptive, impulsive, and disruptive behavior directed outwards. Addiction, conduct, and attention deficit hyperactivity disorders are part of the externalizing spectrum. There is a

notion that the deviant behavior observed in patients with internalizing or externalizing disorders can be driven by variations in performance monitoring ability (amongst the other cognitive control processes, see RDoC: National Institute of Mental Health, NIMH, 2008; Cuthbert, 2014; HiTOP). Understanding when and how performance monitoring is a common transdiagnostic marker across these disorders will help researchers and clinicians improve the nosology of psychopathology. Also, it deepens our understanding of comorbidity (Krueger & Markon, 2006) and heterogeneity, which ultimately helps us to better support patients and develop targeted (pharmacological) treatment interventions. With the research presented in this dissertation, I contribute to our understanding of the role of performance monitoring in the etiology of psychopathology.

Studying Performance Monitoring

Performance monitoring is an automatic, elementary, and unconscious process, which makes it challenging to capture its components. Yet, we can study observable behavior and brain activity as a proxy of these cognitive processes. Cognitive scientists have designed speeded tasks that stimulate components of performance monitoring by simplifying and mimicking real-life actions that require performance monitoring. In these computerized tasks, the individual needs to pay close attention to the stimuli, process conflicting information, inhibit predominant responses and avoid errors. The performance monitoring tasks can be easily modified to allow for different experimental manipulations (e.g., create conditions that evoke certain emotions) or can be adjusted to accommodate sample characteristics (e.g., age). Examples of these performance monitoring tasks are the flanker task (Eriksen, & Eriksen, 1974), go-nogo task (e.g. Falkenstein et al., 1999), Stroop task (based on the Stroop color-word test; Jensen, & Rohwer, 1966), stop-signal task (Logan et al., 1984; Verbruggen & Logan, 2008), error awareness task (e.g., Orr & Hester, 2012) and the (AX-)continuous performance task (CPT; Conners, 1985; Gonthier et al., 2016; Rosvold et al., 1956). These performance monitoring tasks can be used to gain insight into information processing by measuring behavioral indicators and/or by measuring brain activity. First, the actual task behavior can be observed, e.g., the number of mistakes one makes or the slowing of response after a mistake is made. Second, performance monitoring is examined on a neurophysiological level with instruments like electroencephalogram (EEG, for the study of event-related potentials, ERP's) and functional magnetic resonance imaging (fMRI; identifying the activity of brain regions). This dissertation combines different methodologies to investigate performance monitoring on both behavioral and neurophysiological levels to increase our understanding of its role in psychopathology.

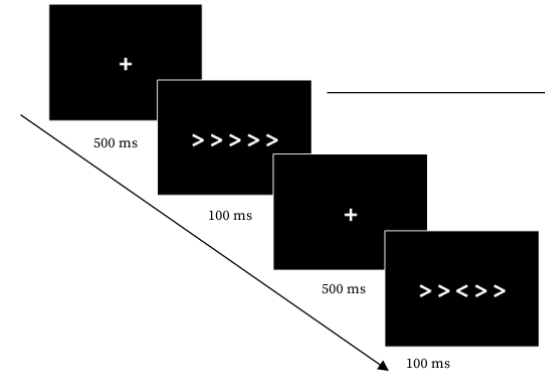
Behavioral Indices of Performance Monitoring

The performance task that is predominantly applied in this dissertation is the flanker task. The flanker task aims to stimulate two components of performance monitoring: conflict monitoring or adaptation and error monitoring or processing (explained in the next

paragraphs). In this task, shown in Figure 2, participants are instructed to look at a string of symbols (such as arrows or letters) that point in a certain direction (in the case of arrows) or that are identical (in the case of symbols). In a typical arrow flanker task, there are two trial types: 1) all the symbols point in the same direction (compatible or congruent) or 2) the middle symbol points in another direction than the surrounding symbols (incompatible or incongruent, thus conflicting information). The participant is instructed to look at the middle symbol and press a button from the corresponding side (e.g., when an arrow is pointing to the left, the participant presses the left button). These symbols are presented at a fast rate to trigger error-making. In this way, the processing of conflicting information (incongruent trials) and errors can be studied.

Figure 2.

A presentation of how the flanker task allows for the measurement of performance monitoring.



Flanker task	Performance monitoring				
	Conflict monitoring			Error monitoring	
	Trial type	Response options	Behavioral	Behavioral	Event-related potential
	Congruent	Correct		Post-correct accuracy	Correct-related negativity
		Error		Post-error slowing Post-error accuracy	Error-related negativity/Error positivity
	Incongruent	Correct		Post-correct accuracy	Correct-related negativity
		Error	Congruency effect: Longer response time & more errors made	Post-error slowing Post-error accuracy	Error-related negativity/Error positivity

Behavioral indices of performance monitoring are response time (the time it takes the participant to press a button), accuracy (correct or incorrect), and calculations derived from these variables. For instance, to study conflict monitoring, we need to know the correct or incorrect responses on congruent and incongruent trials. Incongruent trials contain interfering or conflicting information and trigger what is called the congruency effect. The congruency effect requires more cognitive control, evidenced by an increased chance of error-making and longer response times. There are several theories that explain performance monitoring on a behavioral level (see Ullsperger et al., 2014), such as the *reinforcement learning theory* (Holryod & Coles, 2002), *adaptive orienting theory of error processing* (Wessel, 2018), and the *conflict monitoring theory* (Botvinick et al., 2001). Most relevant in the current dissertation is the conflict monitoring theory (Botvinick et al., 2001), which states that the congruency effect stimulated in these tasks involves 1) the effective detection of conflicting information and 2) behavioral adaptation after response to the conflict. When planning to carry out a particular action, the conflict monitoring system focuses on the identification of current behavior and encourages attention processes to adjust consecutive behavior to execute the 'correct' action.

Another frequently used behavioral index of error processing is post-error slowing (PES; Danielmeier & Ullsperger, 2011; Rabbitt, 1966). This observable phenomenon, which occurs after an individual makes an error, is the slowing of response time reflecting the processing of an error behavioral adjustment. Another measure that can be derived from the behavioral indices of task performance is the speed-accuracy trade-off. Speed-accuracy trade-off (SAT) describes the 'strategy' of the participant: it is the inverse relationship between the speed of the response and the response accuracy (Heitz, 2014). When an individual focuses on avoiding mistakes, the response time increases. When an individual favors speed, the response time decreases at the expense of more error-making.

Although the investigation of behavioral indices of performance monitoring tasks is usually a secondary aim in neurocognitive studies, task behavior indices are reported to investigate patterns of performance monitoring in individuals with psychopathology. For instance, the aggregation of behavioral indices in flanker and go-nogo tasks show deficits in performance monitoring in patients with juvenile and adult ADHD in the study of Geburek et al. (2013). Differences in PES between healthy controls and patients with an externalizing disorder have been reported for ADHD (Balogh & Czobor, 2016) and substance use disorder (SUD: Sullivan et al., 2019). In both meta-analyses with ADHD and SUD samples, the patient groups had reduced accuracy, slower response times, and diminished PES when compared to controls. In samples with internalizing disorders, behavioral indices of performance monitoring are less studied, and existing studies show mixed results. In two studies (Rueppel et al., 2022; Van der Borght et al., 2016), task behavior and PES appeared to be affected in patients with anxiety disorders or individuals with trait anxiety. More specifically, patients with anxiety or individuals with high anxiety symptoms had lower accuracy, slower RTs, and larger PES compared to controls or individuals with low anxiety. Yet others do not report task performance differences between individuals with and without high anxiety traits (Aarts & Pourtois, 2010; Cavanagh et al., 2017; Hseih et al., 2021). Last, speed-accuracy trade-off has not been widely studied in relation to psychopathology. Initial studies on this behavioral indicator of performance monitoring show that healthy young adolescents value speed as much as accuracy (Ladouceur et al, 2007) yet children with ADHD prefer speed over accuracy (Mulder et al., 2010), reflecting the impulsivity trait of the disorder.

To better understand how behavioral indices of performance monitoring play a role in the development of psychopathology, we can explore the development of performance monitoring in children. To date, very few studies exist on this development on a behavioral level and, in particular, its role in psychopathology. The studies of Davies et al. (2004) and Gavin et al. (2019) are two cross-sectional investigations covering a large time span in childhood and adolescence. The two studies indicate that children's performance monitoring enhances on a behavioral level: both errors and response times decrease. In addition, several studies indicate that emerging internalizing and externalizing problems in children are related to behavioral indices of performance monitoring (e.g., Meyer et al., 2012; Woltering et al., 2011). Yet a combination of the examination of performance monitoring development and its

link to the development of psychological problems in childhood is lacking. Chapter 4 describes an exploration of the trajectory of flanker performance in a sample of children of mainstream elementary school children and how this trajectory links with psychological problem development.

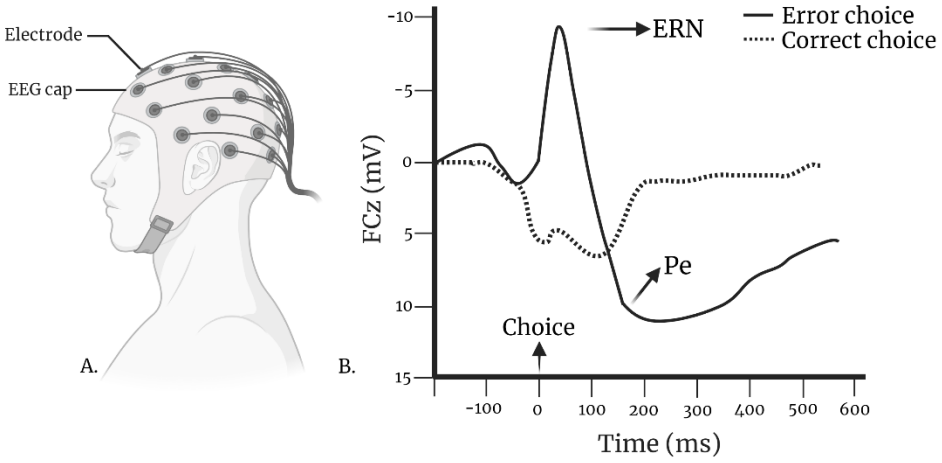
The behavioral responses evoked by performance monitoring tasks are driven by processes in the brain. Here, the neurophysiological correlates of error processing that are relevant to the current dissertation are discussed below. Other relevant physiological measures like heart rate and skin conductance or other performance- and attention-related measures (stimulus-locked N200, N450 and P300, neural oscillations, correct-related negativity, feedback-related negativity, etc.) are outside of the scope of this dissertation.

Event-Related Potentials of Performance Monitoring

A well-known method that detects the electrical activity of the brain through a cap of electrodes is the electroencephalogram (EEG, figure 3). EEG is a non-invasive instrument known for its high temporal resolution and the study of event-related potentials (ERP's) in psychology. ERPs are derivatives or averages of the brain waves that are elicited through stimuli or responses (time-locked ERP's) during the tasks. There are several performance monitoring, or more specifically, error processing ERP's relevant for the two review articles in the current dissertation (Chapter 2 and 3): the response-locked error-related negativity (ERN) and error positivity (Pe), shown in Figure 3.

Figure 3.

An illustration of error-related brain activity measured by EEG. A). EEG cap with electrodes; B). An averaged ERN and Pe waveform at the frontocentral (FCz) for error and correct trials.



Error-related Negativity. Brain activity in response to an error can be observed via EEG. A sharp negative deflection (shown in Figure 3), appearing about 50 to 100 ms after the error (Coles et al., 2001; Falkenstein et al., 2000), is called error-related negativity (ERN;

Gehring et al., 1993; 2018) or Error Negativity (Ne; Fankenstein et al., 1991). The ERN is usually the largest (more negative) in the frontal and central electrodes and originates from the anterior cingulate cortex (ACC) in the brain (Dehaene et al., 1994; Ridderinkhof et al., 2004; Van Veen & Carter, 2002b). The ERN indicates the initial and trait-like reaction of the brain toward the error and marks the commencement of error processing in the brain (Holroyd & Coles, 2002; Olvet & Hajcak, 2009). The ERN appears irrespective of the participant being conscious of the error (Hester et al., 2005; Nieuwenhuis et al., 2001) and has proven to be a stable, robust, and valid measure of error processing (Larson et al., 2010; Riesel et al., 2013; Rietdijk et al., 2014). There are several running theories and models (see Lo, 2018; Olvet & Hajcak, 2008) that describe the functional significance of the ERN, such as *reinforcement learning* (Holroyd & Coles, 2002), *mismatch theory* (Bernstein et al. 1995; Falkenstein et al., 1991) and *motivational significance theory* (Olvet & Hajcak, 2008), *conflict monitoring theory* (Botvinick et al., 2001; Larson et al., 2014), and *threat sensitivity* (Weinberg et al., 2016). Finally, there is evidence that conflict and error processing closely relate to other mechanisms, such as personality, affective functions or other traits like anxiety (Hajack, 2012; Segalowitz & Dywan, 2009). For example, Dignath et al. (2020) suggest that conflict and errors trigger a negative reaction, which drives the 'control adaptation' (increased attention to change behavioral performance). The different theories on the ERN collectively explain how the ERN is generated by a network of brain regions and clarify how error processing relates to the dopamine system, and the motivational, threatening, and affective aspects that ultimately drive learning and behavioral adjustment.

Error positivity. There is another yet independent error-related ERP: the error positivity, abbreviated as Pe (Figure 3). Pe is a larger, slower positive wave that appears 200 to 600 ms after the error, largest in the central-parietal electrodes (Arbel & Dochin, 2009; Falkenstein et al, 1991). The Pe relies on a network containing the ACC, anterior insular, and medial frontal cortex in the brain (Overbeek et al., 2005; Ullsperger, et al., 2010; Van Veen & Carter, 2002a; Vocat et al., 2008). The Pe can be interpreted as the conscious awareness of the error (Ficarella et al., 2019; Nieuwenhuis et al., 2001, Ullsperger et al., 2010) and the processing thereof (Overbeek et al., 2005). Currently, there is no consensus on the running hypotheses of the functional significance of Pe (proposed by Falkenstein, 2004; discussed by Overbeek et al., 2005). Namely, the Pe could indicate the emotional appraisal of the error and its consequences according to the *affective processing hypothesis*; the Pe could be reflective of the change in behavior after error-making according to the *behavior-adaption hypothesis*; and according to the *error awareness hypothesis*, the Pe could mark the conscious recognition of the error.

The ERN has previously been studied as a candidate transdiagnostic marker for several psychopathologies (Riesel et al., 2019), specifically in internalizing disorders like anxiety (e.g., Riesel, 2019) and obsessive-compulsive disorders (Gilian et al., 2017). Patients with internalizing disorders show an enhanced reaction to error-making, evidenced by a greater ERN, compared to healthy controls. For externalizing disorders, there were inconsistent findings. There is a growing interest in the role of Pe in psychopathology

(Donoghue & Voytek, 2021) as evidenced by an increase in reviews on error positivity, such as the study by Boen et al. (2021), even though Pe is usually not reported as a main error-related ERP. There are several reasons for this, such as a) the Pe has similarities with another ERP: the P300 (see for discussion Overbeek et al., 2005), b) it remains unclear what the Pe really means. There is no consensus yet on whether and to what extent internalizing disorders are related to Pe. In externalizing disorders, there is an indication that it is reduced, for example, in ADHD (Kaiser et al., 2020). By compiling studies on error-related ERP's in individuals with externalizing problems, I contribute to the literature on the role of error processing in externalizing disorders (Chapter 2). In both Chapter 2 and 3, I evaluate the bias in reporting, address the inconsistent findings, and provide future directions for research. In Chapter 2, I conducted meta-analyses on ERN and Pe studies with samples of patients with externalizing disorders. In Chapter 3, I performed a narrative review to evaluate whether the error-related ERP's can serve as biomarkers for externalizing psychopathology.

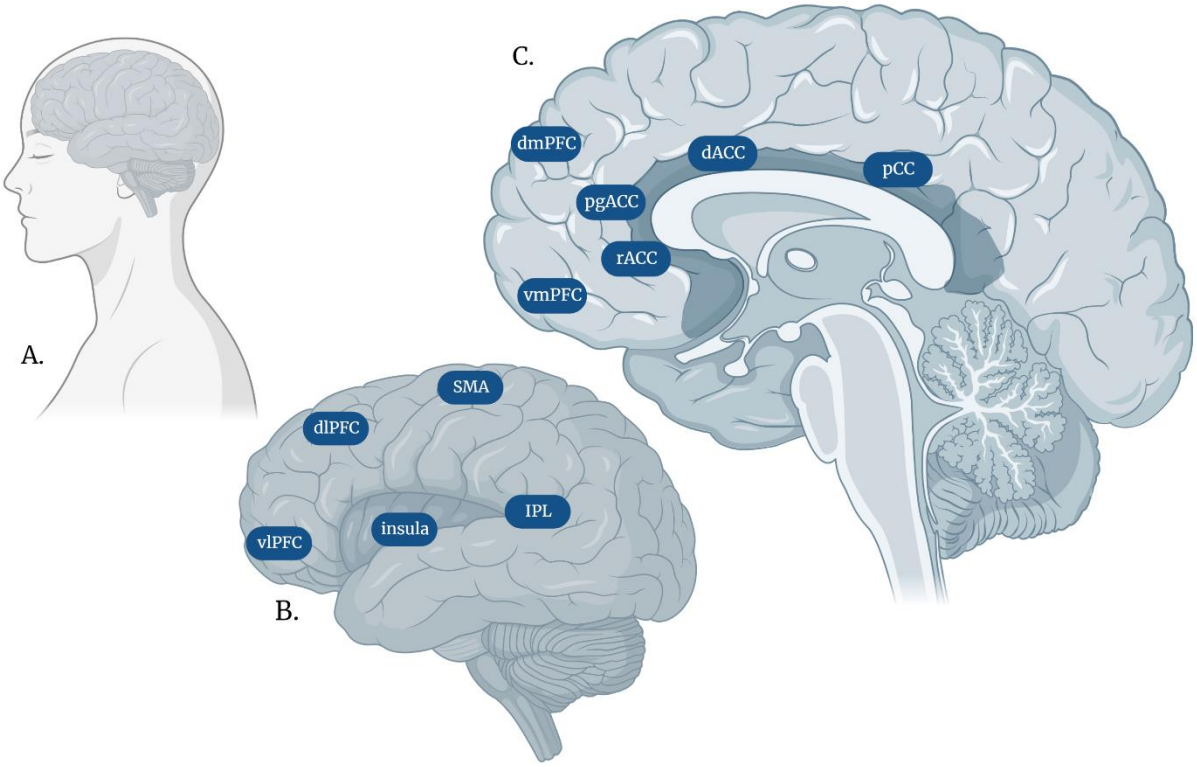
Brain Regions Involved in Performance Monitoring

Magnetic Resonance Imaging (MRI) is an imaging technique that makes scans of the anatomy, tissue type, and physiological processes of the body. MRI relies on strong magnetic fields and radio waves which allows for high special resolution imagery. Functional MRI (fMRI) detects the blood-oxygen-level-dependent (BLOD) signal, which is the oxygenated blood and blood flow and reflects the neural activity. fMRI has been useful in the identification of activation patterns of brain regions where information, such as performance monitoring, is being processed.

Figure 5 shows the brain with the anterior cingulate cortex as the main center for performance monitoring (Holroyd & Yeung, 2012; Kerns et al., 2004; Shenhav et al., 2013; van Veen & Carter, 2002b). Several studies found specific regions of the ACC to be involved during error processing, such as the dorsal region (Gilbertson et al., 2021), anterior mid-cingulate cortex (Wessel et al., 2012), ventral and rostral part of the ACC, and the posterior medial prefrontal cortex (Dehaene et al., 1994). Several other brain regions appear to be involved during performance monitoring, such as the supplementary motor area (Taylor et al., 2007), dorsal lateral prefrontal and superior frontal cortex (Pourtois et al., 2010; Ridderinkhof et al., 2004; particularly for behavioral adjustment in Kerns et al., 2004), the ventrolateral prefrontal cortex, the insula, and the inferior parietal lobule (Ham et al., 2013; Klein et al., 2013; Taylor et al., 2007). Initial connectivity analyses (type of analysis to test the network of brain regions working together) reveal that change in ERN amplitude (incorrect vs correct) is related to the connectivity between the dorsal ACC and supplementary motor area (Gilbertson et al., 2021). For an elaborate discussion of the performance network, see Ullsperger et al. (2014). Abnormal functioning of the ACC region is indicative of disorders such as addiction (Luijten et al., 2014; McTeague et al., 2017), ADHD, depression and obsessive-compulsive disorder (Holroyd & Umemoto, 2016).

Figure 5.

Brain regions involved in performance monitoring. A. Brain in the skull; B. lateral view of the brain; C. Sagittal view of the brain. Abbreviations are vlPFC: ventrolateral prefrontal cortex, dlPFC: dorsolateral prefrontal cortex, SMA: supplementary motor area, IPL: inferior parietal lobule, vmPFC: ventromedial prefrontal cortex, dmPFC: dorsomedial prefrontal cortex, rACC: rostral anterior cingulate cortex, pgACC: pregenual anterior cingulate cortex, dACC: dorsal anterior cingulate cortex, pCC: posterior cingulate cortex.



Performance Monitoring Across the Lifespan

To understand when performance monitoring could play a role in the development of disorders, it is crucial to investigate it across different age samples and explore potential changes in performance monitoring over time. Performance monitoring can be detected as early as 4 years old (Brooker et al., 2011; Morales et al., 2021). During childhood and adolescence, the ERN develops substantially (Boen et al., 2021; Lo et al., 2018). As children age, the ERN becomes larger (more negative), but the Pe appears to be stable across childhood (Boen et al., 2021). The development of the error-related ERPs is driven by neural changes (e.g., synaptic pruning and neuroplasticity) in brain regions such as the medial frontal cortex, the dorsal (caudal) ACC and posterior cingulate cortex (Tamnes et al., 2013), and the maturation of the network of performance monitoring. There is some evidence that the

changes in functional maturation during childhood facilitate the improvement in performance monitoring (see for more details Tamnes et al., 2013).

In healthy adult samples, the ERN appears to be stable, evidenced by the lack of correlation between the ERN and age (e.g., Fischer et al., 2016; Imburgio et al., 2020; Larson et al., 2016; Niessen et al., 2017). Age-related changes in Pe are studied to a lesser extent. Initial studies are inconsistent and show that the Pe magnitude might be consistent over time (Larson et al., 2016) or related to age (when age increased, Pe decreased: Imburgio et al., 2020). The adult age-related changes in the ERN and Pe are also driven by neural changes in the brain, for instance, dopamine levels (Segalowitz & Dywan, 2009).

In addition to age-related changes in error-related brain activity, we can examine whether these neurophysiological markers can be seen as predictors of psychopathology across development (Loo et al., 2016). This can be investigated by comparing the ERN in healthy children with the ERN in children at risk for a disorder. For instance, the ERN is decreased in children who are at risk for depression (Meyer et al., 2018) and addiction (Cádenas et al., 2023; Euser et al., 2013), whereas the ERN is increased in children at risk for anxiety disorders (Meyer, 2017) or OCD (Riesel et al., 2019). This has made the ERN a promising transdiagnostic endophenotype, at least for internalizing disorders, yet it remains unclear whether this is also the case for externalizing disorders. This is why I address, in Chapter 3, whether error processing correlates can be considered a biomarker in externalizing psychopathology.

Key Variables for the Performance Monitoring and Psychopathology Link.

There are several variables that are related to performance monitoring. Performance monitoring, together with variables like defense reactivity (Weinberg et al., 2012) and other measures of cognitive control (working memory: Meyer & Hajcak, 2019), may predict psychopathology. Furthermore, there are relevant moderators like gender (e.g., Hill et al., 2018) or personality (e.g. conscientiousness; Pailing & Segalowitz, 2004), that can modulate the role of performance monitoring in psychopathology (in particular, the ERN). The temperamental trait behavioral inhibition (BI, not to be confused with the cognitive measure, for which the same term can be used) reflects shy and withdrawn behavior in unfamiliar situations and is a strong predictor of the development of anxiety problems and disorders (Sandstorm et al., 2020). Since performance monitoring is also affected by anxiety problems, it is possible that infant BI affects social anxiety and performance monitoring later in life. Therefore, Chapter 5 will address the influence of infant BI on social anxiety and error processing by the brain later in adulthood.

Another manner of testing the effect of variables on performance monitoring is by adjusting the cognitive task through manipulations of the addition of stimuli. Task adaptations can trigger different task performances and neurophysiological responses, but they also provide insight into possible mechanisms between the two. For instance, the presence of

feedback on performance may influence task behavior (Grützmann et al., 2014). In the meta-analysis in Chapter 2, I test several moderators that can systematically influence error processing in externalizing samples that have applied EEG. In Chapter 5, a social flanker is adopted to test the effect of a social context on performance monitoring. By stimulating a social context, it is possible to investigate the 'pressure' to perform well by inducing peer observation. Social manipulation mimics real-life social situations in which an individual should avoid making mistakes to evade social evaluation. For individuals with social anxiety, the fear of social evaluation is one of the fundamental underpinnings of the disorder. The social flanker task has only been adopted in a handful of studies, yet never in an adult sample during MRI. That is why in Chapter 5, I investigate the social error processing by the brain in adults.

Current Dissertation

The role of performance monitoring in the development of psychopathology has been studied extensively. The research in the current dissertation builds upon this line of research, using different sets of existing data and different methods (such as experimental and meta-analyses) and instruments (EEG and MRI). First of all, the growing amount of research on the role of error processing in externalizing psychopathology encourages the need to summarize the literature. Also, it allows the study of the extent to which we can consider error-related ERPs as biomarkers for the externalizing spectrum. This question was central to this dissertation: **What is the role of error processing in externalizing psychopathology?** In **Chapter 2**, we compiled EEG experiments that investigated two error-related ERP's, ERN and Pe in externalizing samples. The objective of this study was to systematically investigate whether the ERN and Pe were different in patients with externalizing disorders or individuals with externalizing problem behavior versus healthy controls. We tested several sample and task-related variables to explain heterogeneity in the data and to inform future experimental designs. We expected that the ERN and Pe were diminished, reflecting the reduced ability to monitor and respond to errors specifically for individuals with externalizing psychopathology. In **Chapter 3**, we reflected on the current state of literature to answer a central question: Can error-related ERPs be considered biomarkers in externalizing psychopathology? In this review, we integrated research to evaluate whether ERN and Pe can serve as potential biomarkers for externalizing disorders. In **Chapters 4 and 5**, two longitudinal samples were used: the "Happy Child, Happy Adolescents?" study and a prospective longitudinal study on the influence of infant temperament on socioemotional development (sample recruited in Washington D.C., United States). Abundant performance monitoring research has relied on cross-sectional and experimental paradigms. Longitudinal study designs allow for the examination of emerging psychopathology and the long-term effects of important predictors of behavior. The following two research questions are answered in the remainder of this dissertation: **What is the development of flanker performance in children, and how is it related to behavioral and emotional problems?** In **Chapter 4**, we study a sample of mainstream elementary school children that have performed the same flanker task across 5 years. We explore the

developmental trajectory of flanker performance and its association with the development of teacher-reported behavioral and emotional problems. The study described in **Chapter 5** aims to answer the third research question: **Is there an association between infant temperament, current social anxiety and social performance monitoring during adulthood?** In this longitudinal prospective study, we investigate the effect of infant behavior inhibition on current social anxiety problems and the performance of a social flanker task performed in the MRI 30 years later. Finally, **Chapter 6** summarizes and discusses the main findings of this dissertation, as well as the implications for future research and clinical practice.

Chapter 2

Diminished Error-Related Negativity and Error Positivity in Children and Adults with Externalizing Problems and Disorders: A Meta-Analysis on Error Processing

This chapter has been published as:

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

Abstract

Background: Deficits in error processing are reflected in an inability of people with externalizing problems to adjust their problem behaviour. The present study contains 2 meta-analyses, testing whether error processing — indexed by the event-related potentials error-related negativity (ERN) and error positivity (Pe) — is reduced in children and adults with externalizing problems and disorders compared to healthy controls. **Methods**: We conducted a systematic search in PubMed (1980 to December 2018), PsycInfo (1980 to December 2018) and Scopus (1970 to December 2018), identifying 328 studies. We included studies that measured error processing using the Eriksen flanker task, the go/no-go task or the stop-signal task in healthy controls and in adults or children with clearly described externalizing behavioural problems (e.g., aggression) or a clinical diagnosis on the externalizing spectrum (e.g., addiction). **Results**: Random-effect models (ERN: 23 studies, 1739 participants; Pe: 27 studies, 1456 participants) revealed a reduced ERN amplitude (Hedges' $g = 0.44$, 95% confidence interval [CI] 0.29 to 0.58) and a reduced Pe amplitude (Hedges' $g = -0.27$, 95% CI −0.44 to −0.09) during error processing in people with externalizing problems or disorders compared to healthy controls. Type of diagnosis, age and the presence of performance feedback or comorbidity did not moderate the results. The employed cognitive task was a moderator for Pe but not for ERN. The go/no-go task generated a greater amplitude difference in Pe than the Eriksen flanker task. Small-sample assessment revealed evidence of publication bias for both event-related potentials. However, a *p* curve analysis for ERN showed that evidential value was present; for Pe, the *p* curve analysis was inconclusive. **Limitations**: The moderators did not explain the potential heterogeneity in most of the analysis, suggesting that other disorder- and patient-related factors affect error processing. **Conclusion**: Our findings indicate the presence of compromised error processing in externalizing psychopathology, suggesting diminished activation of the pre-frontal cortex during performance monitoring.

Diminished Error-Related Negativity and Error Positivity in Children and Adults with Externalizing Problems and Disorders: A Meta-Analysis on Error Processing

Externalizing problem behaviour has been associated with problems in cognitive control (Morgan & Lilienfeld, 2000), of which error processing is an important component (Luna et al., 2015; Ridderinkhof et al., 2004). People with externalizing problems or disorders are characterized by disruptive and problematic behaviour that is directed outward to the environment and are further referred to as externalizing samples. Specific diagnoses and behaviours that belong to the category of externalizing samples include attention-deficit/hyperactivity disorder (ADHD), oppositional defiant disorder, psychopathy, conduct disorder, aggression, antisocial personality disorder, substance use disorder (SUD) and delinquency (Kruger & South, 2009). Error-processing, which refers to the ability to detect errors and evaluate performance, allows for the adaptation of behaviour to correctly react to stimuli from the environment (Holroyd & Coles, 2002). Deficits in error processing can be reflected in a failure to adjust behaviour, which is indicative of externalizing psychopathology. Deficits in error processing can be detected by electroencephalography (EEG) and are reflected in a diminished amplitude of the event-related potentials (ERP) error-related negativity (ERN) and error positivity (Pe). Several studies have investigated differences in the ERN (e.g., Lo 2018, in children and adolescents with externalizing problem behaviour) and Pe (e.g., Luijten et al., 2014 in substance use disorder), but a systematic review is lacking in children and adults that combines externalizing samples and includes comparison with healthy controls. This study investigates the ERN and Pe across different externalizing samples to determine error processing deficits in children and adults with externalizing problems or disorders compared to healthy controls.

Error-Related Negativity

The ERN (Gehring et al., 1993; 2018) or negativity error (Ne, Falkenstein et al., 1991) is a negative deflection that occurs approximately 50 ms to 100 ms after commission of an error (Coles et al., 2001; Falkenstein et al., 2000). This ERP waveform peaks at the frontocentral electrodes, reflecting the neuronal activity of the anterior cingulate cortex (Dehaene et al., 2010) during error processing (Ridderinkhof et al., 2004; Holroyd & Coles, 2002; Van Veen & Carter, 2002a). The ERN is a robust and reliable (Riesel et al., 2013; Rietdijk et al., 2014) neurobiological marker that reflects the brain's initial reaction to an error and the start of error processing, whether or not the person is conscious of the error Hester et al, 2005; Nieuwenhuis et al., 2001). Several theories outline the functional significance of the ERN (for an overview, see Loo et al., 2016; Olvet & Hajcak, 2008), including mismatch theory, motivational significance theory, reinforcement and learning-based theory, and conflict monitoring theory. These theories describe different processes of error and conflict detection, as well as the role of the dopaminergic system of the brain and the anterior cingulate cortex. Previous work has suggested that the ERN can serve as a candidate endophenotype for psychopathology, especially for internalizing disorders. Meta-analyses of internalizing samples show that the

ERN appears to be increased in patients with anxiety (Moser et al.,2016), obsessive-compulsive disorder (Riesel, 2019) and depression disorders (Moran et al., 2017) compared to healthy controls. A recent meta-analysis concluded that the ERN can serve as a transdiagnostic marker for internal and externalizing disorders (Pasion & Barbosa, 2019). The current study is an extension of this meta-analysis, including child samples as well as adult samples, and also including psychopathy samples. In addition, we have used effect size calculation rather than effect size estimation, and included the late error processing component Pe as well as the ERN.

Deviating activation patterns in the brain with respect to cognitive control have been found in externalizing behaviour (Olvet & Hajcak, 2008; Rudo-Hutt, 2015). Some studies have reported decreased ERN amplitude in ADHD (e.g., Wiersema et al., 2009) and addiction (e.g., Zhou et al., 2013), but other reports have found no differences in the ERN compared to healthy controls (e.g., in addiction: Franken et al., 2017; ADHD: Van de Voorde et al., 2010). Some studies have even reported increased ERN amplitude in ADHD (Wiersema et al., 2005) and addiction (Schellekens et al., 2010) compared to controls. The presence of comorbid internalizing problems has been suggested as a possible explanation for these mixed results, as illustrated by the study of Schellekens et al. (2010) in a sample of patients with alcohol dependence. Although medication (e.g., Groom et al., 2013), age (Lo, 2018) and the experimental paradigm (Riesel et al., 2013) have been studied as moderating factors for ERN amplitude, it remains unclear whether these variables influence ERN results across externalizing samples. Furthermore, several studies have reported that performance feedback during tasks can influence error processing (Grützmann et a., 2014; Hajcak et al., 2005; Riesel et al., 2012). When participants receive feedback on performance, they become cautious of their response accuracy, inducing greater reactions to errors. By conducting a meta-analysis, we were able to integrate inconsistent findings to shed light on the role of the ERN in externalizing behaviours. Moreover, by explicitly testing medication use, age, comorbidity, experimental paradigm and performance feedback as moderators, we were able to investigate whether or how they account for variability in ERN studies.

Error Positivity

Another ERP component relevant for performance monitoring is Pe amplitude. The Pe is a slow, positive deflection, peaking at approximately 200 ms to 600 ms after an error; it is measured across the centroparietal area (Arbel & Donchin, 2009; Falkenstein et al., 2000; Overbeek et al., 2005) The Pe is said to reflect the conscious awareness of errors and error processing (Overbeek et al., 2005). It is an independent ERP component, despite the fact that it follows directly after the ERN and shows similarities with the P300 component (for example, the latency window). For further reading on the similarities and differences between these components, see Arbel and Donchin (2009), Davies et al. (2001), Overbeek et al. (2005) and Ridderinkhof et al. (2009). The functional significance of the Pe has been described in several hypotheses (Falkenstein et al, 2000; Overbeek et al., 2005) including the affective processing hypothesis (in which the Pe reflects the emotional appraisal of the error), the behaviour-

adaptation hypothesis (in which the Pe indicates performance adjustment after error) and the error awareness hypothesis (in which the Pe reflects the conscious recognition of the error committed). Although empirical evidence is needed to support these hypotheses (specifically in terms of the neural generators of Pe), they suggest that diminished Pe could be related to deviant activity of the rostral part of the anterior cingulate cortex (Van Veen & Carter, 2002a; Hermann et al., 2004).

Compared to controls, Pe reductions in error processing for people with externalizing problems or disorders have been observed more consistently than ERN reductions, although some discrepancies have been found across studies. For instance, some studies have found diminished Pe amplitudes in people with ADHD (e.g., Albrecht et al., 2008) and substance use (e.g., Franken et al., 2007), but these were not replicated in subsequent studies (e.g., addiction studies: Franken et al., 2017; Luijten et al., 2011). Moreover, other studies have found a reverse effect, indicating that increased Pe amplitudes are related to externalizing behaviour (e.g., in addiction: Rass et al., 2014 and ADHD: Wild-Wall et al., 2009). This is the first meta-analysis to summarize Pe findings in externalizing samples compared to healthy controls. We have examined the moderators suggested for the ERN, above, to try to better understand these discrepancies in study findings.

In the current study, we aimed to investigate whether the ERN and Pe were different in children and adults with externalizing problems or disorders compared to healthy controls. We use meta-analysis and focused on the mean amplitude of the ERN at the midline frontocentral electrode (FCz; for subsequent analyses to investigate the effect of other midline electrode sites Fz and Cz, see supplementary materials section 5) and the Pe at the midline central electrode (Cz). We expected that both the ERN and Pe amplitudes would be reduced in the externalizing groups, indicating deficits in error processing. To explain the mixed results found in this field of research, we investigated potential heterogeneity by adding type of diagnosis, presence of comorbidity, experimental paradigm, age and medication use as moderators in both analyses. Except for comorbidity (comorbid internalizing or externalizing symptoms or a combination of both) and performance feedback, we expected that the moderators would not influence ERP amplitudes. We did not expect that effect size variability would be explained by the experimental paradigm, because these tasks often elicit highly correlated amplitudes and have high construct validity (Riesel et al., 2013; Segalowitz et al., 2010). In cases of comorbidity, we expected that differences in ERN and Pe between the clinical and control groups would be smaller for samples that had internalizing comorbid symptoms, and greater for samples with externalizing symptoms. We also expected that the presence of performance feedback would elicit a greater ERN and Pe than no performance feedback.

Method

We did not preregister this study, but to enhance reproducibility and accommodate the open science community, our data and code are available at the Open Science Framework

(https://osf.io/dkxtp/). We determined a search strategy and inclusion and exclusion criteria before our literature search. Secondary to the steps undertaken as described in this report, we reviewed the relevant literature (Siddaway et al., 2019), consulted experts and compiled study-related factors to ensure that we were informed about the state of art in this field. We intended to identify as many EEG studies that evaluated ERN and Pe magnitudes in case–control (externalizing samples vs. healthy volunteers) studies.

Search Strategy

We conducted the literature search using 3 databases: PsycInfo (1980 to December 2018), PubMed (1980 to December 2018) and Scopus (1970 to December 2018). Search terms included the following: inhibit*, cognitive or inhibitory control, error processing or monitoring, external* symptoms, disorders and problems, alcohol, cocaine, stimulants, heroin, smoking, cannabis, substance abuse, substance use-, dependence-, misuse, alcoholism, ADHD, ADD, antisocial personality disorder, oppositional defiant disorder, aggression, psychopathy, intermittent explosive disorder, conduct disorder, antisocial behaviour, behavioural problems or disorders, psychopathic traits and callous-unemotional traits. We cross referenced the above terms with the following: error-related negativity, error positivity, ERP, EEG, Eriksen flanker task, go/no-go task and stop-signal task. For complete search strategy queries by database, see supplementary materials section 1.

Eligibility Criteria

We assessed studies identified from the literature search, together with studies identified from other sources, according to the following inclusion criteria: studies were published in peer-reviewed journals in English and performed in human volunteers of any age; studies addressed error processing using the EEG components ERN (at Fz, FCz and Cz) or Pe (Cz), irrespective of the latency window; the ERN and Pe were measured during the Eriksen flanker task (Eriksen & Eriksen, 1974), the go/no- go task or the stop-signal task; studies included a healthy control group of participants with no clinical or neurologic diagnosis; and participants in the patient groups were recruited because they had a clinical diagnosis of an externalizing disorder (based on the Diagnostic and Statistical Manual of Mental Disorders, fifth edition [DSM-5] or the International Classification of Diseases, 10th revision [ICD-10], or earlier versions) or they showed subclinical levels of externalizing problems. Studies were excluded according to the following criteria: means and standard deviations of the ERN or Pe amplitude for both groups (derived from averaging ERN and Pe epochs where the peak was the maximum from error trials) were unavailable from the published report or after contact with the authors; studies used adjusted paradigms (e.g., lack of neutral stimuli presented) or stimuli that were not presented visually; and studies used the continuous performance task or error awareness task (despite including go/no-go elements).

Data Extraction

Records identified through the literature search were imported to Mendeley. In this program, we screened titles and abstracts using our inclusion criteria. To avoid unwanted exclusion, all articles in which the abstract did not contain full information were kept for further reading. Next, we downloaded and read the full text of articles that had passed the screening stage, and we reviewed their reference lists to identify additional studies for potential inclusion. The first author (ML) extracted relevant information from the included studies, retrieving sample characteristics such as sample size, age, sex ratio, disorder (ADHD, addiction or other externalizing disorders), condition (clinical v. subclinical), the presence of comorbidity (if known, coding for externalizing, internalizing or mixed problems), the use of medication and pretesting group differences. We assessed patients' diagnostic status (clinical or subclinical) by extracting the diagnostic tools used (DSM III or IV or ICD-10) and details about the informants (specialist, self-report, parent, teacher, or medical or legal reports). A study was considered clinical when the diagnosis was obtained by a trained psychologist or psychiatrist or when participants were recruited from inpatient treatment facilities. Offenders incarcerated for serious crimes were also considered to be clinical. We categorized the studies into 5 diagnosis groups: child and adult ADHD, clinical and subclinical addiction (adults only) and "other." Studies with samples of offenders, people with multi-problem behaviour and people with high scores on psychopathy or aggression measures were considered "other," leaning toward a sample with forensic characteristics. Participants were considered subclinical when no diagnosis was determined, but when diagnostic tools or self-reports indicated heightened levels of externalizing problems. To be included in the final analysis, the eligible study had to report a cut-off score or level for the diagnostic tool. People were considered healthy controls when no clinical or neurologic diagnosis or disabilities were reported. Comorbidity was coded as any co-occurring symptom or (sub)clinical level of other internalizing (e.g., anxiety) or externalizing (e.g., conduct) problems.

We also gathered information about the experiment, including the cognitive task used, whether the task was adjusted (e.g., instructions the participants received) and latency windows. For studies that used multiple experimental manipulations, we systematically selected the first or baseline time point, the neutral stimuli trials and, when multiple tasks were presented, the Eriksen flanker task. We requested the mean and standard deviation of ERN and Pe amplitudes of error trials by contacting authors when articles did not provide this information. We also requested unpublished data, but those requests did not lead to viable data for our analysis. Two authors (I.V. and M.M.) independently extracted information from the manuscripts to verify the work of the first author. For categorical variables, Cohen's κ was between 0.79 and 0.81, indicating strong level of agreement. For continuous variables, intraclass correlation was between 0.97 and 0.99, which was near-perfect agreement. We evaluated the selected studies primarily on their choice of sample and their experimental design. We (M.L., I.V., M.M. and I.F.) discussed whether the selected studies adhered to our inclusion and exclusion criteria and whether they were similar enough to be compiled.

Data Analysis and Small Sample Bias Assessment

For both the ERN and Pe meta-analyses, we assumed a random model because of variance in the estimates due to different clinical disorders and experimental tasks administered (Viechtbauer, 2005). We used restricted maximum likelihood estimation to estimate between-study variance (Novianti et al., 2014). As recommended by Veroniki and colleagues (2016), we ran analyses with the DerSimonian– Laird and Sidik and Jonkman estimators to determine sensitivity, but restricted maximum likelihood estimation resulted in a better model fit. We computed standardized mean difference (SMD; Hedges' *g*, Hedges, 1981) from the means and variances of the ERN and Pe amplitudes, including factor *J* to reduce overestimation of the bias induced by small sample sizes. For studies with multiple externalizing groups, we adjusted the weights appointed to effect sizes by splitting the N of the control group (Harrer et al., 2019). This was to avoid unit-of-analysis errors or doublecounting problems evoked by multiple testing of the control groups. For the ERN, a positive SMD indicated reduced amplitude for the externalizing group. For the Pe, a negative SMD indicated reduced amplitude for the externalizing group. Both SMDs indicated a diminished electrocortical reaction after the error. Effect sizes of 0.2 to 0.3 were considered small; effect sizes of approximately 0.5 were considered medium; and effect sizes of 0.8 and higher were considered large (Cohen, 1977). We investigated influential or outlier studies based on the recommendations of Viechtbauer and Cheung (2010). We evaluated the degree of heterogeneity using I^2, where a larger value indicated increasing variety in effect sizes Higgins & Thompson, 2002; Higgins et al., 2003). We performed subgroup or moderation analyses when heterogeneity Cochran's Q was significant. Moderation analyses for clinical disorder, comorbidity, medication use, experimental task and a meta-regression of age were determined a priori. Reviewers also suggested that we test the effect of performance feedback and electrode site (see supplementary materials, section 4) as additional moderators.

We examined small sample study bias by assessing asymmetry in funnel plots, applying Egger's test of the intercept (Egger et al., 1997) and Duval and Tweedie's trim-and-fill procedure (Duval & Tweedie, 2000). To detect whether small samples distorted the funnel plots created using SMD (Zwetsloot et al., 2017), we performed a sensitivity analysis using adjusted funnel plots with $1 \sqrt{n}$ on the y axis as a precision estimate, rather than the standard error. In the end, the adjusted funnel plots also detected asymmetry as a result of publication bias, not only small sample size. We also evaluated the robustness of the effects we found with the fail-safe N calculation using the Orwin approach (Orwin, 1983). However, these assessments have their limitations (Simonsohn et al., 2014a), so we also performed a *p* curve analysis (Simonsohn et al., 2014a; Simonsohn et al., 2014b, Simonsohn et al., 2015), see supplementary materials section 5) to inspect whether significant *p* values ($p < 0.0562$) provided proof of evidential value (for a full description and application of this assessment, see Harrer et al., 2019). As recommended by van Aert et al (2016), a *p* curve analysis is conducted only when I^2 is less than 50% and studies' effects are in one direction, to allow for robust conclusions. All analyses were performed in R (version 1.3.959), using meta (version 4.12.0; Schwarzer, 2007), metafor

(version 2.4.0; Viechtbauer, 2010) and dmetar (version 0.0.9000), guided by the instructions of Harrer et al (2019). All significance tests were conducted at a significance level of 5%.

Results

Figure 1

PRISMA flow diagram. ERN = error-related negativity; Pe = error positivity; PRISMA = Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

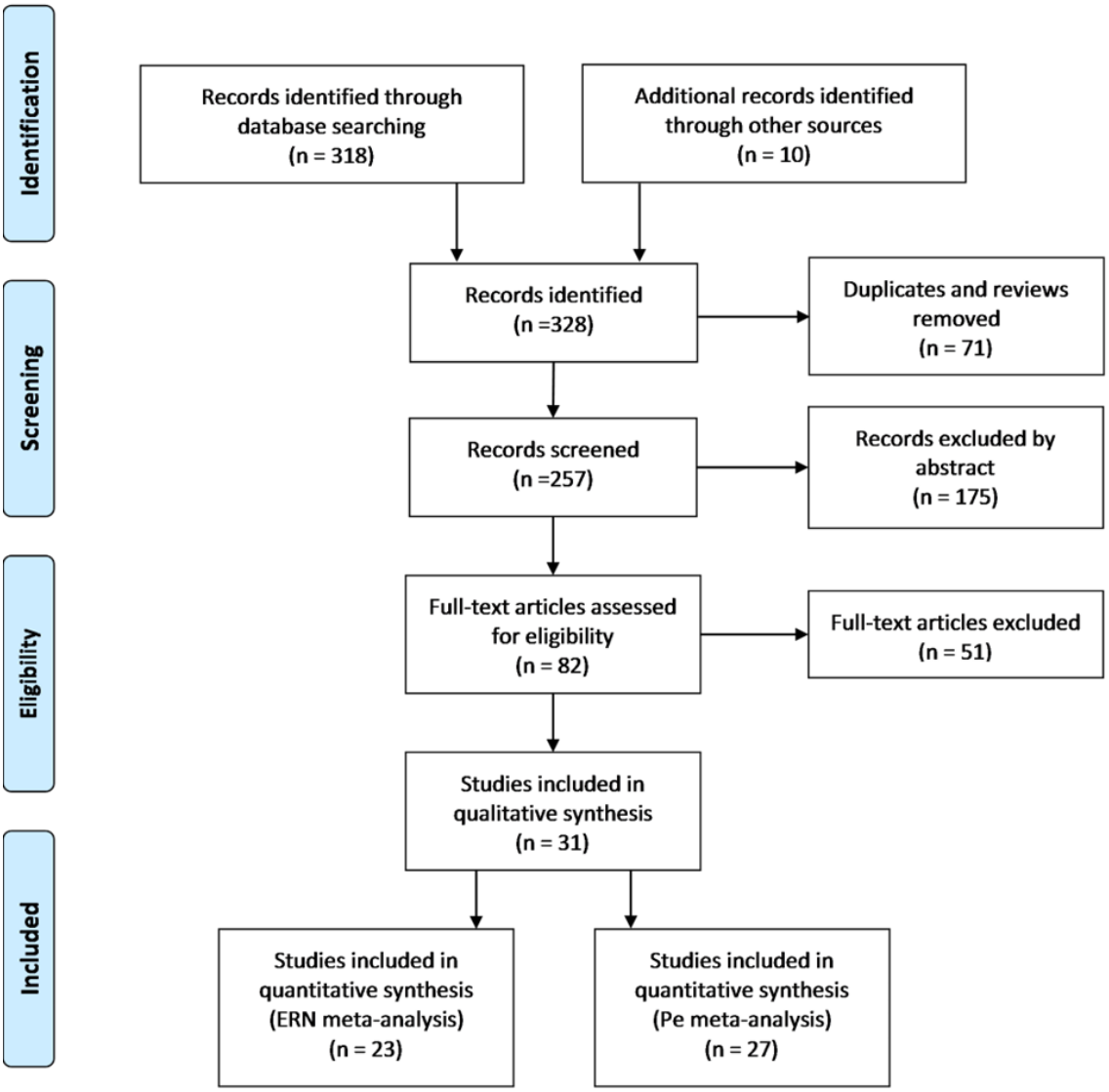


Table 1. Characteristics of included studies.

Study	ERP	Diagnosis	Ne	Nc	% Male	Me (SD)	Mc (SD)	Comorbidity*	Experimental Paradigm	Medication	Number of trials§	ERN latency window (ms)	Pe latency window (ms)
Albrecht et al. 2008	Both	Child ADHD	68	22	100	11.3 (1.6)	11.2 (1.7)	Yes, mix	Flanker	Yes, Off	400	0-150	200-500
Balogh et al. 2017	Both	Adult ADHD	26	14	78	26.7 (5.7)	31.5 (11.4)	Yes, int	Go-NoGo	Yes, Off	240	20-70	100-300
Brazil et al. 2009	Both	Other: violent offenders	16	18	100	39 (9.5)	37 (6.4)	No	Flanker	No	400	0-150	250-400
Chang et al. 2009	Both	Adult ADHD	36	32	50	23.7 (3.7)	23.7 (3.7)	Yes, mix	Flanker	Yes, Off	480	(-10)-180	120-400
Chen et al. 2013	ERN	Clinical Addiction	20	15	100	37.1 (9.5)	32.5 (10)	No	Flanker	Unknown	800	0-100	NA
Czobor et al. 2017	Pe	Adult ADHD	22	29	71	30.6 (9.7)	30.1 (9)	No	Go-NoGo	Yes, On	478	NA	200-400
Franken et al. 2007	Both	Clinical Addiction	14	13	78	38.1 (10.2)	32 (13.8)	No	Flanker	Unknown	400	25-75	200-400
Franken et al. 2010	Pe	Subclinical Addiction	23	28	48	21.7 (2.7)	21.3 (2.8)	No	Flanker	Unknown	400	NA	200-400
Franken et al. 2017	Both	Subclinical Addiction	48	49	49	23.4 (10)	11.9 (8.5)	No	Flanker	Unknown	400	25-75	200-400
Franken et al. 2018	Both	Subclinical Addiction	34	34	12	19.9 (1.7)	20.8 (3)	No	Flanker	No	400	NA	200-400
Groom et al. 2010	Both	Child ADHD	23	19	74	16.2 (0.3)	16.1 (2)	Yes, ext	Go-NoGo	Yes, Off	304	(-50)-100	100-350
Groom et al. 2013	ERN	Child ADHD	28	28	96	12.5 (1.8)	12.5 (1.8)	Yes, mix	Go-NoGo	Yes, On	40	(-10)-100	NA
Herrmann et al. 2010[a]	Pe	Adult ADHD	17	9	50	25.2 (4.4)	24.2 (3.1)	No	Flanker	Yes, Off	NA	NA	110 -450
Herrmann et al. 2010[b]	Pe	Adult ADHD	17	9	56	40.9 (6.8)	39.7 (6.6)	No	Flanker	Yes, Off	NA	NA	110 -450
Jonkman et al. 2007	Pe	Child ADHD	10	10	NA	9.5 (2.1)	10.76 (1.2)	No	Flanker	Yes, Off	480	NA	200-450
Littel et al. 2012	Both	Subclinical Addiction	25	27	63	20.5 (3)	21.42 (2.6)	No	Go-NoGo	No	636	0-75ms	200-400
Luijten et al. 2011	Both	Subclinical Addiction	13	14	70	20.7 (1.3)	21.4 (2.6)	No	Flanker	No	900	25-75	250-350
Maij et al. 2017[a]	Both	Clinical Addiction	35	39	76	21.7 (2.1)	22.1 (2.1)	Yes, ext	Flanker	No	400	25-75	150-250
Maij et al. 2017[b]	Both	Clinical Addiction	38	39	68	21.4 (2.5)	22.1 (2.1)	Yes, ext	Flanker	No	400	25-75	150-250

Table 1. (continued)

Study	ERP	Diagnosis	Ne	Nc	% Male	Me (SD)	Mc (SD)	Comorbidity*	Experimental Paradigm	Medication	Number of trials	ERN latency window (ms)	Pe latency window (ms)
Marhe et al. 2013	Both	Clinical Addiction	49	23	84	39.6 (8.4)	39.9 (9.4)	No	Flanker	Unknown	400	25-100	NA
Marquardt et al. 201[9]	Both	Adult ADHD	27	28	51	35.3 (8.8)	33.4 (7)	Yes, mix	Flanker	Yes, Off	520	20-60	180-220
McLoughlin et al. 2009[a,]	Both	Adult ADHD	21	20	100	32.5 (5.8)	30 (6.5)	No	Flanker	Yes, Off	400	0-150	200-500
McLoughlin et al. 2009[b,]	Both	Adult ADHD	20	20	100	45.9 (4.2)	30 (6.5)	No	Flanker	No	400	0-250	200-500
Michelini et al. 2016[a,]	Both	Adult ADHD	87	169	67	18.3 (3)	18.8 (2.2)	Yes	Flanker	Yes, Off	400	0-150	NA
Michelini et al. 2016 [b,]	Both	Adult ADHD	23	169	79	18.9 (3)	18.8 (2.2)	Yes	Flanker	Yes, Off	400	0-150	NA
Morie et al. 2014	Both	Clinical Addiction	23	27	72	44 (6.6)	41 (8.5)	Yes, ext	Go-NoGo	No	1260	30-70	100-300
Munro et al. 2007	Both	Other: violent offenders	15	15	100	45.9 (13.6)	46.6 (6.9)	No	Flanker	Yes, On	480	0-150	150-350
Rass et al. 2014[a,]	Both	Clinical Addiction	22	15	52	27.2 (5.3)	25.2 (4.3)	No	Flanker	No	400	(-50)-100	100-250
Rass et al. 2014[b,]	Both	Subclinical Addiction	31	15	43	23.9 (4.4)	25.2 (4.3)	No	Flanker	No	400	(-50)-100	100-250
Sokhadze et al. 2008	ERN	Clinical Addiction	19	15	56	42.1 (5.5)	37 (9.4)	Yes, int	Flanker	No	960	50-200	NA
Vilà-Balló et al. 2014	Both	Other: violent offenders	17	17	100	18.3 (0.3)	18.6 (0.3)	No	Flanker	No	1920	65-115	135-285
Wiersema et al. 2005	Pe	Child ADHD	22	15	65	10.3 (1.6)	10.2 (2)	Yes, ext	Go-NoGo	Yes, Off	NA	NA	200-500
Wiersema et al. 2009	Pe	Child ADHD	23	19	57	29.3 (11)	30.9 (11)	Yes, mix	Go- NoGo	Yes, Off	NA	NA	200-400
Wild-Wall et al. 2009	Both	Child ADHD	15	12	71	13.9 (1.6)	13.2 (1.5)	Yes, int	Flanker	Unknown	840	(-50)-200	200-250
Xue et al. 2017	Pe	Other: aggression	13	14	44	21.3 (0.9)	21.3 (1.3)	No	Go-NoGo	Unknown	220	NA	100-500
Zhang et al. 2009	Pe	Child ADHD	16	16	NA	7.5 (1.4)	7.6 (1.8)	No	Go-NoGo	Unknown	320	NA	200-400
Zijlmans et al. 2019	Both	Other: multi-problem	119	26	100	22.5 (2.4)	23.1 (2.6)	Yes, ext	Flanker	Unknown	400	25-100	250-400

Note. ERP = event related potential, ERN = Error-related negativity, Pe = Error positivity, ADHD = attention deficit hyperactivity disorder, Ne = sample size externalizing group, Nc = sample size control group, Me = mean age externalizing group in years, Mc = mean age control group in years, SD = standard deviation age in years, * = comorbid diagnosis or symptoms: ext = externalizing, int = internalizing, mix = both external and internalizing, [§]= number of trials of full paradigm, ms = milliseconds, NA = not applicable (not measured or unknown).

Selected Studies

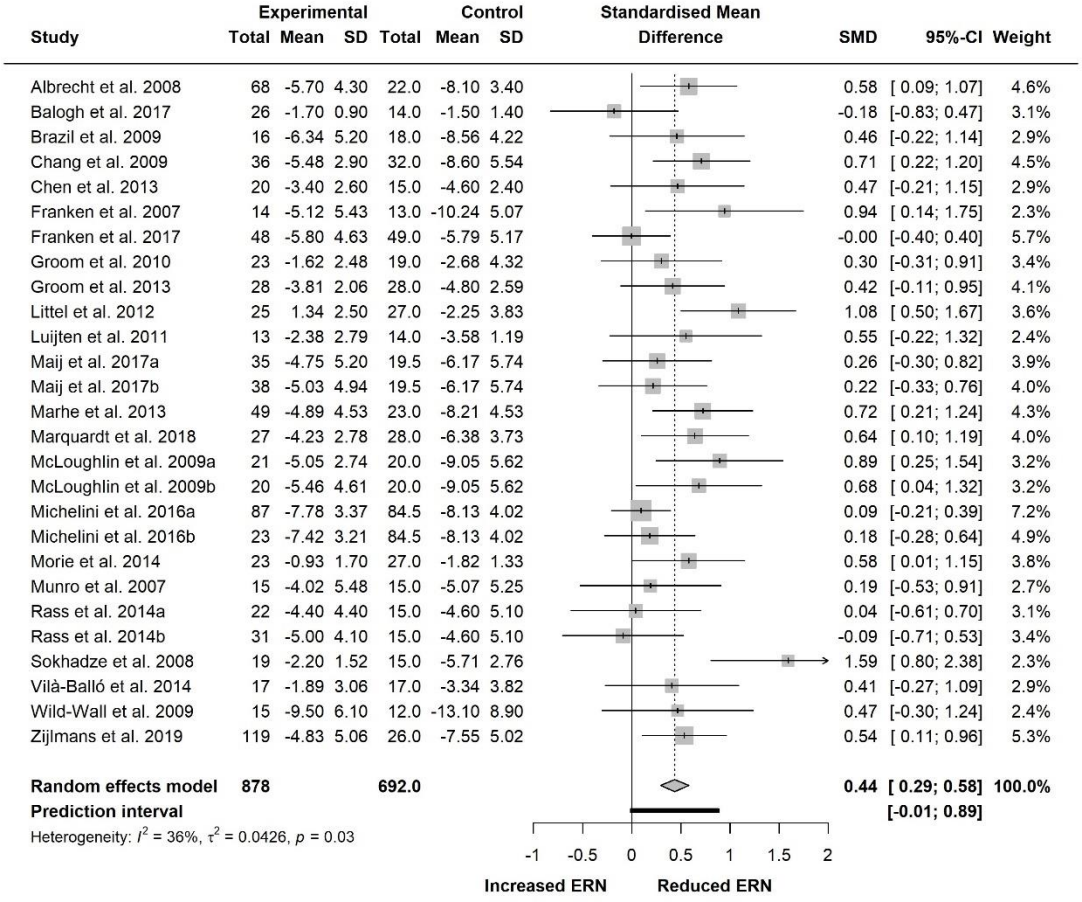
Figure 1 shows a flow chart of the literature search. Where applicable and possible, we adhered to PRISMA guidelines (supplementary materials, section 2). The search of databases and additional sources yielded a total of 328 records. After removing duplicates and reviews (n = 71), we screened the abstracts of 257 studies. We then assessed the full text of the 82 articles that met our inclusion criteria. We included 31 articles for qualitative analysis, of which 23 were ERN studies at the FCz electrode (27 effect sizes; n = 1739) and 27 were Pe studies at the Cz electrode (31 effect sizes; n = 1456). We found no studies that used the stop-signal task. Descriptive information for the included studies is shown in Table 1; further details of the included studies are shown in supplementary materials section 3.

ERN Summary Effect

The ERN meta-analysis included 23 studies and 1739 participants. We found a small to medium overall effect size (g = 0.44, 95% CI 0.29 to 0.58, p < 0.01). This indicated that in the patient group, the ERN had a decreased negative amplitude compared to healthy controls. Between-study variability was 36%, indicating a low to moderate amount of variability in effect sizes. The test for heterogeneity was significant (Q_{26} = 40.69, p = 0.03), which gave us cause to perform moderation analysis. We identified the studies of Sokhadze et al (2008; high effect size) and Michelini et al (2016; large sample size) as influential cases. However, we kept these studies in the overall analysis because they did not influence the overall model. A forest plot for the ERN is presented in Figure 2.

Figure 2.

Overall ERN meta-analysis, including a forest plot. CI = confidence interval; ERN = error-related negativity; SD = standard deviation; SMD = standardized mean difference.



ERN Subgroup Analyses

Moderation analyses revealed no significant difference in ERN amplitudes between diagnosis groups (Q_4 = 0.66, p = 0.96). Comorbidity did not significantly influence the ERN amplitudes (Q_3 = 5.11, p = 0.16), and the type of experimental paradigm was not a moderator (Q_1 = 0.01, p = 0.91). The presence of performance feedback also did not account for variability in ERN effect size (Q_1 = 0.08, p = 0.78). We had initially intended to test the effect of medication, but this variable was confounded in the sample of ADHD participants, making further investigation futile. For study details, see supplementary materials section 4 for medication, and section 5 for electrode site. Table 2 provides an overview of the moderation results for the categorical variables. A meta-regression with age as a predictor revealed that age was not associated with the effect sizes (F 1,25 = 2.30, p = 0.14).

Table 2.

Results for the multiple moderator analyses for ERN.

		ERN				
Moderator	Categories (k)	SMD	95% CI	*Q*	I^2	*p*
Clinical diagnosis group	Child ADHD (4)	0.46	[0.26; 0.65]	0.51	0%	0.92
	Adult ADHD (7)	0.40	[0.05; 0.75]	13.69	56%	**0.03**
	Clinical addiction (8)	0.56	[0.18; 0.94]	12.93	46%	0.07
	Subclinical addiction (4)	0.36	[-0.51; 1.24]	11.03	73%	**0.01**
	Other (4)	0.44	[0.22; 0.67]	0.67	0%	0.88
Comorbidity	Yes, mixed (5)	0.58	[0.43; 0.73]	0.76	0%	0.94
	Yes, internalizing (2)	-	-	-	-	-
	Yes, externalizing (5)	0.40	[0.19; 0.61]	1.54	0%	0.82
	No (15)	0.40	[0.19; 0.60]	23.80	41%	**0.04**
Experimental paradigm	Flanker (22)	0.43	[0.27; 0.59]	31.92	34.2%	**0.04**
	Go-NoGo (5)	0.46	[-0.10; 1.01]	8.52	53%	0.07
Performance feedback	Yes (17)	0.43	[0.27; 0.58]	16.97	5.7%	0.39
	No (10)	0.47	[0.13; 0.82]	23.61	61.9%	**<0.01**

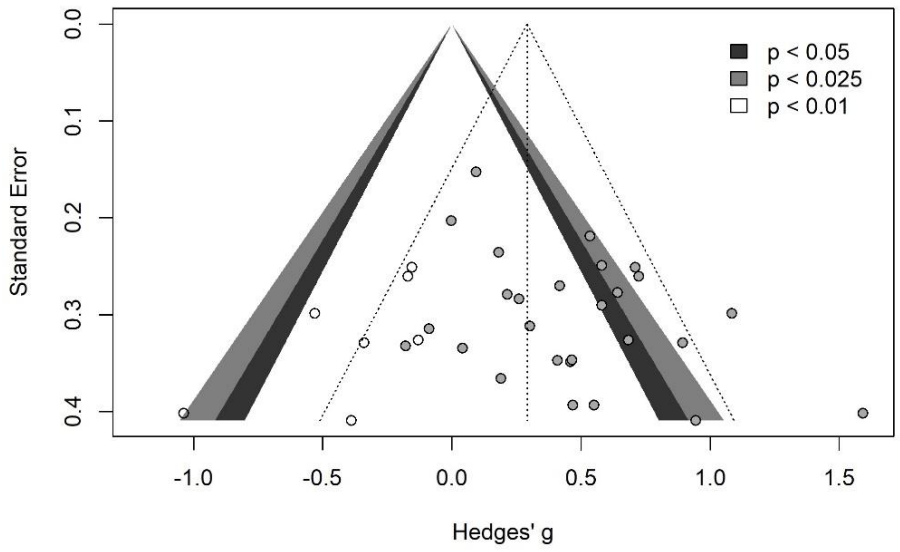
Note. Random models with moderators for ERN studies, with exception for internalizing comorbidity (due to fewer than three studies). *k* = number of studies included in the model; ADHD = attention deficit hyperactivity disorder; ERN = error-related negativity; SMD = Standardized mean difference; CI = confidence interval; *Q* = Cochran's test of heterogeneity; I^2 = measure of heterogeneity; *p* = significance of Cochran's *Q* statistic, bold if heterogeneity is significant.

ERN Small Sample Study Bias

To investigate publication bias, we visually inspected the funnel plots of the effect sizes. The Egger's intercept of the funnel plot was significant ($B = 2.09$, $p = 0.03$), indicating evidence of publication bias. The funnel plot in Figure 3 applies Duval and Tweedie's trim-and-fill procedure. Application of this procedure revealed that by filling 7 studies, the overall effect would be reduced to small ($g = 0.29$, 95% CI 0.12 to 0.45, $p = 0.001$), indicating the presence of bias. Despite this bias, 27 studies were needed to get to an unweighted average effect size of 0.24 using the fail-safe N test. However, a *p* curve analysis (k = 9) revealed the presence of right skewness of the significant *p* values, and of evidential value (half: $Z = -2.33$, $p = 0.009$; full: $Z = -1.43$, $p = 0.08$). The flatness test was not significant (half: Z = 2.73, $p = 0.99$; full Z = –0.32, $p = 0.37$). Although the analysis was underpowered (25%), typical in this field, it indicated that there was most likely no selective reporting of *p* values. The *p* curve estimate of the average "true" effect size was 0.32, which was similar to the trim-and-fill result of 0.29 and lower than the combined effect size of the overall analysis (0.44). For a full report, a *p* value distribution figure and results of the *p* curve analysis, see supplementary materials section 6.

Figure 3

Funnel plot including filled studies for error-related negativity

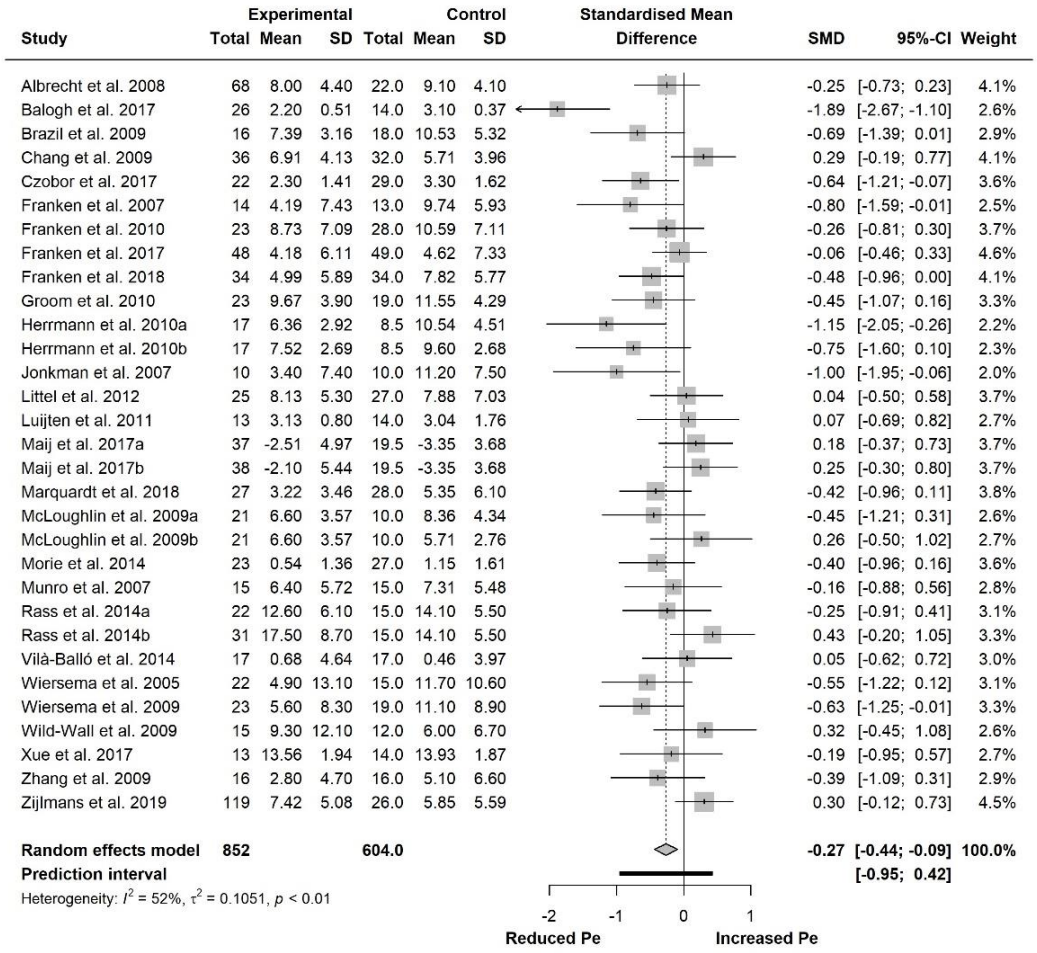


Pe Summary Effect

The Pe meta-analysis included 27 studies, incorporating 31 effect sizes and 1456 participants. We found a small to medium overall effect size (g = –0.27, 95% CI –0.44 to –0.09, p = 0.004), indicative of decreased amplitude of the Pe waveform for the externalizing group compared to controls. We observed a moderate degree of heterogeneity (I^2 = 52%, Q_{30} = 62.74, p = 0.004), which gave us cause for further exploration of effect size variability through subgroup analysis. A forest plot for the Pe is presented in Figure 4.

Figure 4

Overall Pe meta-analysis, including a forest plot. CI = confidence interval; Pe = error positivity; SD = standard deviation; SMD = standardized mean difference.



Pe Subgroup Analysis

Diagnosis was not a moderator for the Pe effect sizes (Q_4 = 5.17, p = 0.22), nor was comorbidity (Q_3 = 1.61, p = 0.66). The presence of feedback did not account for variability in Pe effect size (Q_1 = 2.58, p = 0.12). The experimental paradigm was a moderator in this meta-analysis. The go/no-go task generated a greater difference in Pe amplitudes (SMD = –0.54, k = 9) than the Eriksen flanker task (SMD = –0.15, k = 22; Q_1 = 4.17, p = 0.041). Similar to the ERN, medication use was confounded in the ADHD sample, making further moderation analysis ineffective (supplementary materials section 4). Age did not explain the variability in effect sizes for Pe (F 1,27 = 0.02, p = 0.88). Table 3 shows the results of the moderation analysis for the categorical variables.

Table 3.

Results for the multiple moderator analyses for Pe.

		Pe				
Moderator	Categories (k)	SMD	95% CI	*Q*	I^2	*p*
Clinical diagnosis group	Child ADHD (7)	-0.39	[-0.70; -0.08]	6.04	0%	0.42
	Adult ADHD (8)	-0.56	[-1.15; 0.04]	28.86	76%	**<0.01**
	Clinical addiction (5)	-0.15	[-0.66; 0.36]	6.81	41%	0.15
	Subclinical addiction (5)	-0.09	[-0.40; 0.23]	5.9	15%	0.30
	Other (6)	-0.08	[-0.55; 0.40]	6.14	35%	0.19
Comorbidity	Yes, mixed (4)	-0.23	[-0.85; 0.40]	6.59	54%	0.09
	Yes, internalizing (2)	-	-	-	-	-
	Yes, externalizing (19)	-0.07	[-0.48; 0.34]	9.65	48%	0.09
	No (6)	-0.28	[-0.47; -0.10]	23.84	24%	0.16
Experimental paradigm	Flanker (22)	-0.15	[-0.33; 0.04]	36.01	42%	**0.02**
	Go-NoGo (9)	-0.53	[-0.92; -0.15]	17.14	53%	**0.03**
Performance feedback	Yes (20)	-0.16	[-0.34; 0.02]	28.40	33.1%	0.08
	No (11)	-0.47	[-0.86; -0.08]	30.08	66.8%	**<0.01**

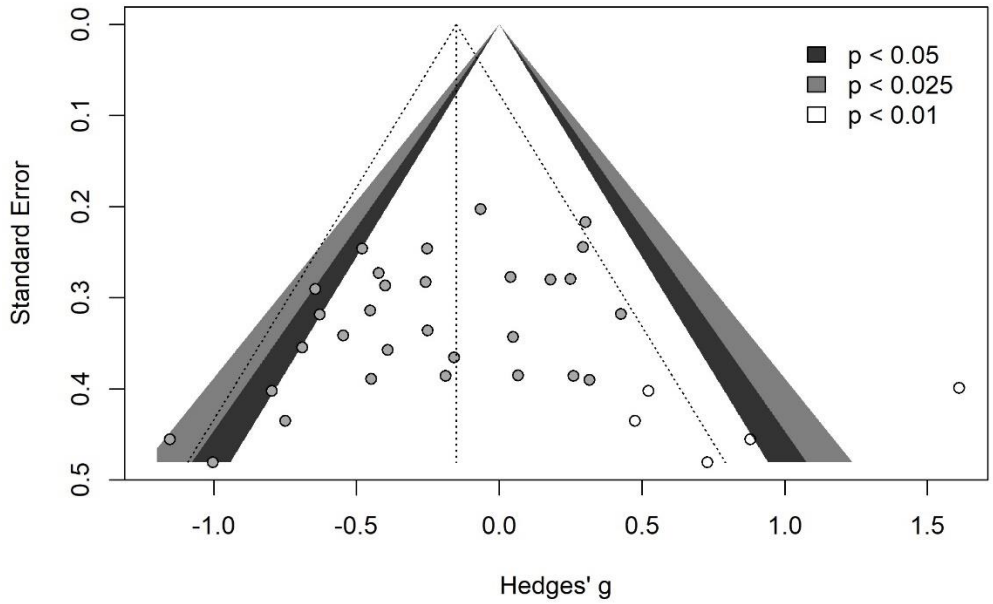
Note. Random models with moderators for Pe studies, with exception for internalizing comorbidity (due to fewer than three studies). *k* = number of studies included in the model; ADHD = attention deficit hyperactivity disorder; Pe = error positivity; SMD = Standardized mean difference; CI = confidence interval; *Q* = Cochran's test of heterogeneity; I^2 = measure of heterogeneity; *p* = significance of Cochran's *Q* statistic, bold if heterogeneity is significant.

Pe Small Sample Study Bias

The Egger's intercept (*B* = –2.45, *p* = 0.019) was significant, indicating asymmetry in the funnel plot for the Pe studies. The funnel plot in Figure 5 includes studies that needed to be added to make the plot symmetrical using Duval and Tweedie's trim-and-fill procedure. This procedure revealed that by filling 5 studies, the overall effect was reduced (*g* = –0.15, 95% CI –0.35 to 0.05), meaning that the Pe meta-analysis could be contaminated by publication bias. However, the fail-safe N assessment revealed that there needed to be 31 effect sizes to achieve the unweighted effect size of –0.16. We performed a preliminary *p* curve analysis, but it was inconclusive (see supplementary materials section 6 for explanation).

Figure 5

Funnel plot including filled studies for error positivity.



Discussion

This meta-analysis collated current EEG studies on error processing to test whether the ERN and Pe amplitude were different in people with externalizing problems or disorders compared to controls. As expected, we found diminished ERN and Pe amplitude for people with externalizing problems or disorders compared to controls. These findings confirmed compromised error processing in the externalizing spectrum, regardless of a specific diagnosis or problem behaviour. For both ERP components, we found a considerable degree of heterogeneity. The variation in results was not explained by comorbidity, the presence of performance feedback, age or type of clinical disorder. The experimental paradigm was a moderator for the Pe studies, but not for the ERN studies. Our results for the ERN were in line with a recent meta-analysis by Pasion and Barbosa (2019) and studies that described error processing deficits in separate externalizing disorders (such as Luijten et al., 2014 and Olvet & Hajcak, 2008 for substance use disorders; Shiels and Hawk, 2010 for ADHD; and Lo, 2018 for children with externalizing symptoms). This was the first meta-analysis to explore age effects in error processing and to confirm deficits in the late error processing component (Pe) for children and adults with externalizing problems and disorders.

Diminished ERN and Pe imply a deviant activation pattern of the dorsal anterior cingulate cortex Botvinick et al. (2004). More specifically, when we consider reinforcement and

learning-based theories of the function of the ERN, decreased ERN amplitude could be indicative of abnormal dopamine activity in the midbrain, affecting processes of error and conflict detection. Dysfunction in the dorsal anterior cingulate cortex is also indicative of deficits in inhibitory control and conflict monitoring (Luijten et al., 2014; Morie et al., 2015; Ridderinkhof et al., 2004). Problems in inhibitory control and conflict monitoring (among other cognitive impairments) have been related to symptomatology in externalizing disorders, such as craving in addiction (Czermainski et al., 2017). Based on the proposed hypotheses for Pe (Overbeek et al., 2005), affected Pe amplitudes reflect differences in conscious recognition of the error committed, differences in emotional appraisal of the error and its consequences, or distinct behavioural adjustment after errors. In turn, such deviation in the processing of errors could imply reduced insight in aberrant and unwanted behaviour for externalizing samples.

The data revealed a considerable amount of heterogeneity for both ERPs, but type of diagnosis, age and the presence of performance feedback or comorbidity did not moderate the results. Studies that controlled for medication use in their samples appeared mostly to be in ADHD, preventing us from examining medication use across other externalizing samples. Although medication use is evident in externalizing samples, many studies have not reported or controlled for medication. Future experiments and systematic reviews should consider the effect of medication on error processing components for different externalizing samples. The experimental task did moderate the association of externalizing problems with Pe and not ERN: we found a greater Pe amplitude difference between the patient group and the control group in the go/no-go task than in the Eriksen flanker task. Although both tasks are known to reliably elicit an electrophysiological reaction after an error (Riesel et al., 2013; Segalowitz et al., 2010), it is possible that the go/no-go task elicits a stronger reaction to an error than the Eriksen flanker task. As well, it is possible that the Eriksen flanker task allows the participant to be more unconscious of an error than the go/no-go task, because the Pe is said to reflect conscious awareness of the error (Overbeek et al., 2005). Finally, contrary to our expectations, comorbidity did not affect the ERN or Pe amplitude in this sample. Although we proposed to test whether internalizing comorbid problems influenced the ERPs, not enough studies were included to properly test this hypothesis and draw firm conclusions.

We performed small bias assessment to investigate the effect of the published data in this study. For the ERN, the result of the trim-and-fill procedure remained significant even after adding 5 studies. The estimated "true" effect size was included in the confidence interval of the overall model, along with evidence of the evidential value from the *p* curve analysis, indicating that the combined effect size for the ERN was robust. However, for the Pe, we found evidence of publication bias (although an inconclusive *p* curve analysis), because the trim-and-fill analysis reduced the effect size to non-significant. Although this nonsignificant effect could have been because of large heterogeneity, reflected in the broad confidence interval, we should be cautious in drawing firm conclusions about the Pe meta-analysis.

Limitations

We should acknowledge the limitations of this study and the methodological considerations of EEG research in general. The present meta-analysis included studies with ERN amplitudes on the FCz electrode and Pe amplitudes on the Cz electrode, generated by the Eriksen flanker and go/no- go tasks. These criteria allowed for solid results (supplementary analysis revealed no effect of electrode site), but other electrodes (see Arbel & Donchin, 2009 for a summary of reliable electrodes for error processing), neurophysiological measures (e.g., functional MRI) and experimental paradigms can be used to examine error processing. Future research should be directed at investigating whether similar error processing deficits are found at other electrode sites and using other tasks. In addition, investigating deficits in biobehavioural markers of performance monitoring, such as post-error slowing (indicative of response caution for maintaining accuracy) and cortisol involvement (Tops & Boksem, 2011) can shed light on responses and behaviour after errors, elucidating different behaviour patterns. Although we considered a substantial number of potential moderators in this study, other sample characteristics could have accounted for differences in the ERP findings. For example, the global assessment of functioning of patients with a specific psychiatric disorder could influence the magnitude of deficits in error processing. Future studies could examine whether the severity of symptoms within a disorder is related to the degree of diminished reaction (that is, correlational measures with ERP amplitudes). As well, individual differences such as personality traits have been known to influence the ERN and Pe (Lo, 2018; Overbeek et al., 2005) and we did not control for these factors in this study. Furthermore, we should be cautious of the results (particularly indicated by the publication bias assessment) because of variations in quality in the EEG experiments. Although the current study evaluated the included experiments, differences between experiments could have influenced our results. We could not assess some aspects of experimental design, such as the manner in which the ERN or Pe were quantified or the effects of task adjustments, because this information was not provided in the reports. To address the possible effect of experimental design differences on the associations between ERP or Pe with externalizing problems in the future, we encourage researchers to disclose the following information: the minimum number of trials or errors for a reliable ERP calculation (e.g., 816 or 696 trials for a reliable ERN); which trials were used (incongruent or error trials) to calculate the ERP; and a clear description of the task instructions and adjustments (e.g., error rate to ensure task difficulty, participant instructions or feedback to influence performance); and other potential confounding variables such as medication use and latency window.

Future work could investigate error processing in specific externalizing disorders that are underexplored in the current literature, such as antisocial personality disorder, specific addictions (e.g., Internet addiction disorder) and double diagnoses (e.g., addiction and a personality disorder). We recommend that future work examine the predictive value of the ERN and Pe using large-scale longitudinal designs to elucidate their role in the etiology of these disorders. We also encourage researchers to assess the feasibility of interventions aimed at

improving error processing. To improve error processing abilities in patients, next steps for future experiments could include examinations of the effectiveness of behavioural training, medication and neuro-modulation techniques.

Conclusion

Our meta-analysis showed that the neurophysiological correlates of error processing, ERN and Pe, were reduced in children and adults with externalizing problems or disorders. However, we found considerable heterogeneity that could not be explained by the moderators explored in this study; this warrants further exploration and limits strong conclusions. Future research can elucidate the role of individual differences, symptom severity and experimental characteristics in error processing deficits specific to externalizing disorders. With the knowledge that the EEG correlates of error processing are affected in people with internalizing problems and could serve as a possible marker for these disorders, we propose that reduced ERN and Pe could be considered markers for the externalizing spectrum of disorders.

Supplementary Materials

Contents

Section 1: Full search query for literature search strategy.

Table S1. Search query for literature search.

Database	Query	Number of studies
PsychInfo	1. (external* symptoms or external* disorders or external* problems or "substance use" or substance abuse or substance misuse or substance dependence or alcohol addiction or alcoholism or cocaine addiction or "cocaine use" or "stimulant use" or heroin addiction or "heroin use" or smoking or "cannabis use" or cannabis addiction or drug addiction or drug dependence or ADHD or attention deficit hyper activity disorder or "attention deficit-hyper activity disorder" or attention deficit disorder or antisocial personality disorder or antisocial behavior or oppositional defiant disorder or "oppositional-defiant disorder" or aggression or conduct disorder or psychopathic traits or psychopathy or "callous-unemotional traits" or behavioral disorders or behavioral problems).mp. 2. (Eriksen Flanker paradigm task or Eriksen Flanker task or modified Eriksen Flanker or "Eriksen Flanker task-modified" or Simon Eriksen Flanker task or Eriksen Flanker test or "Eriksen-Flanker task" or stop signal task or "stop-signal task" or "go-no go task").mp 3. (ERN or error related negativity or Ne or error negativity or Pe or error positivity or evoked potentials or ERP or event related potentials or EEG or electroencephalography).mp. 4. (inhibit* or error processing or error monitoring or cognitive control or inhibitory control, behavioral or response inhibition).mp. 5. 1 and 2 and 3 and 4	110
Scopus	((TITLE-ABS-KEY ("behavioral inhibit*" OR "response inhibit*" OR "error processing" OR "error monitoring")) AND (TITLE-ABS-KEY ("error related negativity" OR "error negativity" OR "error positivity" OR "evoked potentials" OR "event related potentials" OR electroencephalography)) AND (TITLE-ABS-KEY ("external* symptoms" OR "external* disorders" OR "external* problems" OR "substance use" OR "substance abuse" OR "substance misuse" OR "substance dependence" OR "alcohol addiction" OR alcoholism OR "cocaine addiction" OR "cocaine use" OR "stimulant use" OR "heroin addiction" OR "heroin use" OR smoking OR "cannabis use" OR "cannabis addiction" OR "drug addiction" OR "drug dependence" OR adhd OR "attention deficit hyper activity disorder" OR "attention deficit-hyper activity disorder" OR "attention deficit disorder" OR "antisocial personality disorder" OR "antisocial behavior" OR "oppositional defiant disorder" OR "oppositional-defiant disorder" OR aggression OR "conduct disorder" OR "psychopathic traits" OR "psychopathy" OR "callous-unemotional traits" OR "behavioral disorders" OR "behavioral problems"))) AND NOT ((((TITLE-ABS-KEY ("behavioral inhibit*" OR "response inhibit*" OR "error processing" OR "error monitoring")) AND (TITLE-ABS-KEY ("error related negativity" OR "error negativity" OR "error positivity" OR "evoked potentials" OR "event related potentials" OR electroencephalography))) AND (TITLE-ABS-KEY ((flanker AND task) OR (go/no-go AND task) OR (stop-signal AND task)))) AND (TITLE-ABS-KEY ("external* symptoms" OR "external* disorders" OR "external* problems" OR "substance use" OR "substance abuse" OR "substance misuse" OR "substance dependence" OR "alcohol addiction" OR alcoholism OR "cocaine addiction" OR "cocaine use" OR "stimulant use" OR "heroin addiction" OR "heroin use" OR smoking OR "cannabis use" OR "cannabis addiction" OR "drug addiction" OR "drug dependence" OR adhd OR "attention deficit hyper activity disorder" OR "attention deficit-hyper activity disorder" OR "attention deficit disorder" OR "antisocial personality disorder" OR "antisocial behavior" OR "oppositional defiant disorder" OR "oppositional-defiant disorder" OR aggression OR "conduct disorder" OR "psychopathic traits" OR "psychopathy" OR "callous-unemotional traits" OR "behavioral disorders" OR "behavioral problems")))	155
PubMed	((behavioral inhibit*[Text Word] OR behavioural inhibit*[Text Word] OR response inhibit*[Text Word] OR error processing[Text Word] OR error monitoring[Text Word])) AND ((error related negativity[Text Word] OR error negativity[Text Word] OR error positivity[Text Word] OR evoked potentials[Text Word] OR event related potentials[Text Word] OR electroencephalography[Text Word] OR EEG[Text Word])) AND (Eriksen Flanker paradigm task[Text Word] OR Eriksen Flanker task[Text Word] OR modified Eriksen Flanker[Text Word] OR Eriksen Flanker task-modified[Text Word] OR Simon Eriksen Flanker task[Text Word] OR Eriksen Flanker test[Text Word] OR Eriksen-Flanker task[Text Word] OR stop signal task[Text Word] OR stop-signal task[Text Word] OR go-no go task[Text Word]) AND (externalizing symptoms[Text Word] OR externalising symptoms[Text Word] OR external symptoms[Text Word] OR externalizing disorders[Text Word] OR externalising disorders[Text Word] OR external disorders[Text Word] OR externalizing problems[Text Word] OR externalising problems[Text Word] OR external problems[Text Word] OR substance use[Text Word] OR substance abuse[Text Word] OR	53

	substance misuse[Text Word] OR substance dependence[Text Word] OR alcohol addiction[Text Word] OR alcoholism[Text Word] OR cocaine addiction[Text Word] OR cocaine use[Text Word] OR stimulant use[Text Word] OR heroin addiction[Text Word] OR heroin use[Text Word] OR smoking[Text Word] OR cannabis use[Text Word] OR cannabis addiction[Text Word] OR drug addiction[Text Word] OR drug dependence[Text Word] OR adhd[Text Word] OR attention deficit hyper activity disorder[Text Word] OR attention deficit-hyper activity disorder[Text Word] OR attention deficit disorder[Text Word] OR antisocial personality disorder[Text Word] OR antisocial behavior[Text Word] OR antisocial behaviour[Text Word] OR oppositional defiant disorder[Text Word] OR oppositional-defiant disorder[Text Word] OR aggression[Text Word] OR conduct disorder[Text Word] OR psychopathic traits[Text Word] OR psychopathy[Text Word] OR callous-unemotional traits[Text Word] OR behavioral disorders[Text Word] OR behavioural disorders[Text Word] OR behavioral problems[Text Word] OR behavioural problems[Text Word])	

Note. This is the initial search query, therefore the number of studies identified include duplicates and review papers.

Section 2: PRISMA checklist for PRISMA guidelines.

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	4, 8
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	7,6
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	NA
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	9,10
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	9
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Supplementary material, section 1
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	8,9
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	10,11
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	12,13
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	NA
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	12,13
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	12,13

Section 2 continued

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	NA
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	13
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	Figure 1
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	Table 1
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	NA
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	Figure 2, 4
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	15-18
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	Figure 3, 5
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	16,18
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	20-22
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	23
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	25
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	1

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group*, T. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Annals of internal medicine, 151(4), 264-269. https://doi.org/10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.

Section 3: Experimental details included studies

Table S2. Experimental characteristics of included studies.

Study	ERP	Experimental Paradigm	Number of Trials[§]	Medication	ERN Latency window (ms)	Pe Latency window (ms)	Presence of Feedback
Albrecht et al. 2008[41]	Both	Flanker	400	Yes, Off	0-150	200-500	Yes
Balogh et al. 2017[66]	Both	Go-NoGo	240	Yes, Off	20-70	100-300	No
Brazil et al. 2009[67]	Both	Flanker	400	No	0-150	250-400	NA
Chang et al. 2009[68]	Both	Flanker	480	Yes, Off	(-10)-180	120-400	No
Chen et al. 2013[69]	ERN	Flanker	800	Unknown	0-100	NA	No
Czobor et al. 2017[70]	Pe	Go-NoGo	478	Yes, On	NA	200-400	No
Franken et al. 2007[42]	Both	Flanker	400	Unknown	25-75	200-400	Yes
Franken et al. 2010[62]	Pe	Flanker	400	Unknown	NA	200-400	Yes
Franken et al. 2017[28]	Both	Flanker	400	Unknown	25-75	200-400	Yes
Franken et al. 2018[72]	Both	Flanker	400	No	NA	200-400	Yes
Groom et al. 2010[73]	Both	Go-NoGo	304	Yes, Off	(-50)-100	100-350	No
Groom et al. 2013[32]	ERN	Go-NoGo	40	Yes, On	(-10)-100	NA	Yes
Hermann et al. 2010[a,74]	Pe	Flanker	NA	Yes, Off	NA	110-450	Yes
Hermann et al. 2010[b,74]	Pe	Flanker	NA	Yes, Off	NA	110-450	Yes
Jonkman et al 2007[75]	Pe	Flanker	480	Yes, Off	NA	200-450	Yes
Littel et al. 2012[76]	Both	Go-NoGo	636	No	0-75	200-400	No
Luijten et al. 2011[43]	Both	Flanker	900	No	25-75	250-350	Yes
Maij et al. 2017[a,77]	Both	Flanker	400	No	25-75	150-250	Yes
Maij et al. 2017[b,77]	Both	Flanker	400	No	25-75	150-250	Yes
Marhe et al. 2013[78]	Both	Flanker	400	Unknown	25-100	NA	Yes
Marquardt et al. 2018[79]	Both	Flanker	520	Yes, Off	20-60	180-220	Yes
McLoughlin et al. 2009[a,80]	Both	Flanker	400	Yes, Off	0-150	200-500	Yes
McLoughlin et al. 2009[b,80]	Both	Flanker	400	No	0-250	200-500	Yes
Michelini et al. 2016[a,81]	Both	Flanker	400	Yes, Off	0-150	NA	No
Michelini et al. 2016[b,81]	Both	Flanker	400	Yes, Off	0-150	NA	No
Morie et al. 2014[82]	Both	Go-NoGo	1260	No	30-70	100-300	Yes
Munro et al. 2007[83]	Both	Flanker	480	Yes, On	0-150	150-350	Yes
Rass et al. 2014[a,44]	Both	Flanker	400	No	(-50)-100	100-250	Yes
Rass et al. 2014 [b,44]	Both	Flanker	400	No	(-50)-100	100-250	Yes
Sokhadze et al. 2008[84]	ERN	Flanker	960	No	50-200	NA	No
Vilà-Balló et al. 2014[85]	Both	Flanker	1920	No	65-115	135-285	No
Wiersema et al. 2005[30]	Pe	Go-NoGo	NA	Yes, Off	NA	200-500	No
Wiersema et al. 2009[26]	Pe	Go-NoGo	NA	Yes, Off	NA	200-400	No
Wild-Wall et al. 2009[45]	Both	Flanker	840	Unknown	(-50)-200	200-250	Yes
Xue et al. 2017[86]	Pe	Go-NoGo	220	Unknown	NA	100-500	No
Zhang et al. 2009[87]	Pe	Go-NoGo	320	Unknown	NA	200-400	Yes
Zijlmans et al. 2019[88]	Both	Flanker	400	Unknown	25-100	250-400	Yes

Note. ERP = event related potential, ERN = Error-related negativity, Pe = Error positivity, §= number of trials of full paradigm, ms = milliseconds, NA = not applicable (not measured or unknown).

Section 4: Inspection possible medication effects

Table S3. Moderation analysis of medication

Medication as a Moderator	Categories (k)	SMD	95% CI	*Q*	I^2	*p*
ERN	Yes, temporarily off (8)	0.38	[0.10; 0.67]	12.62	44.5%	0.08
	Yes, on during experiment (2)	0.34	-	-	-	-
	No (11)	0.50	[0.20; 0.81]	18.82	46.9%	**0.04**
	Unknown (6)	0.47	[0.12; 0.81]	7.64	34.6	0.18
Pe	Yes, temporarily off (13)	-0.51	[-0.85; -0.18]	30.78	61.0%	**<0.01**
	Yes, on during experiment (2)	-0.04	-	-	-	-
	No (9)	-0.45	[-0.32; 0.23]	11.00	27.4%	0.20
	Unknown (7)	-0.09	[-0.42; 0.24]	8.51	29.5	0.20

Note. Random models with medication as a possible moderator. *k* = number of studies included in the model; ADHD = attention deficit hyperactivity disorder; ERN = error-related negativity; SMD = Standardized mean difference; CI = confidence interval; *Q* = Cochran's test of heterogeneity; I^2 = measure of heterogeneity; *p* = significance of Cochran's *Q* statistic, bold if heterogeneity is significant.

Conclusion:

Only two studies for each ERP (ERN: Groom et al., 2013 and Munro et al., 2007; Pe: Czobor et al., 2017 and Munro et al., 2007) reported that participants were on medication, so it is not possible to interpret these results. Also, this analysis should be interpreted with caution: the studies that reported yes on medication, either chronical use but temporarily off or on during experimentation, were predominantly children and adults with ADHD. We do not know what medication they were on. In addition, substantial studies did not report their medication check and were classified as unknown, but it is possible that this group included studies that had participants on medication during experimentation.

Moderation analysis for ERN revealed no effect of medication, Q (3)= 1.12, p = 0.77. For Pe, the moderation analysis did not reach significance level, Q (3)= 7.64, p = 0.0542. For both ERP's, there was still evidence of heterogeneity, suggesting that this moderator cannot account for the variability of effect sizes in the current sample.

We recommend that future experiments report on the role of medication in their results and urge future reviews in this field to investigate the effects of medication in this spectrum of disorders.

Section 5: Inspection possible electrode site effects

To investigate the impact of electrode site in this meta-analysis, we added the ERN means and standard deviations amplitudes of Fz and Cz, and we conducted a moderation analysis. Full model (k = 63, excluding 6 outliers) revealed a pooled effect size of g = 0.37 with 95% CI [.29, .46], p < 0.0001. The lower pooled effect size (from 0.44 to 0.37) already indicates that the effect is still present but that, as expected, the effects sizes on other electrodes are a bit weaker. Moderation analyses revealed that there was not a significant difference of electrode site on amplitudes, Q (2) = 1.88, p = 0.39. The pooled effect size of electrode Fz (k=16) is g = 0.31, 95% CI [.13, .49] and at Cz (k = 21), the pooled effect size was g = 0.33, 95% CI [.17, .49]. We can conclude that the strongest difference between the experimental and control group is visible at FCz, and to a lesser extend at Fz and Cz.

However, there is a possible dependency between the electrodes (finding an effect on one electrode increases the chance of finding another effect on surrounding electrodes) and some samples were tested multiple times in the model, possibly inflating results. We opted to inspect the effect of electrode site through a multi-level meta-analysis. This approach allows to control for dependency and provides an adequate display of the heterogeneity (which was considerate for ERN) at each level of the data.

In the multi-level model, electrode site (differences in effect sizes due to electrode site within persons) was level 1, between group differences (experimental vs. controls) at level 2 and between study differences at level 3. The overall multi-level model revealed a pooled effect size of g = 0.39 with 95% CI [.28, .50], p < 0.0001. This is similar to the full model mentioned above.

In the multi-level model, 58% of the variances in effect sizes can be attributed to level 1, 42% at level 2 and 0% at level 3. High variance at level 1 means that the effect sizes vary to a large degree, reflecting substantial magnitude differences at electrode site. The variance at level 2 (between group differences) is similar to the heterogeneity found in the ERN at FCz meta-analysis (36%). This degree of variance suggests that there are possible subgroups in the data that could explain the range of effect sizes. No variation at level 3 was observed, as the lower levels take up all the variation, suggesting that between study differences are less influential than the within study differences.

Data and code are available by accessing the project on the Open Science Framework: https://osf.io/dkxtp/.

Section 6: *P*-curve analysis.

For the disclosure table of included studies for the *p*-curve analyses and data script, please find the current project on first author's Open Science Framework profile.

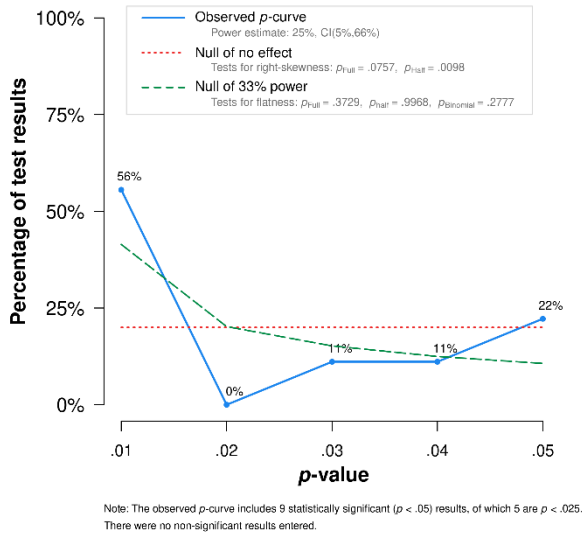


Figure S1 P-value distribution ERN

Interpretation

Note: the *p*-curve analysis is sensitive to the number of studies/sample size, the following interpretation should keep this in mind.

The ERN *p*-value analysis indicated the half *p*-curve significant, Z = -2.33, p = 0.0098 and the full *p*-curve to be not significant, Z = -1.43, p = 0.0757 (figure S5). However, both values are below 0.1 suggesting that the ERN studies contain the evidential value. Also, the evidential was not inadequate, as both the half and full *p*-curve of the 33% power test were not significant, Z_{full}=-0.32, p_{full}=.3729, Z_{half}=2.73, p_{half}=.996. The estimated power of the selected studies was 25%, which is considered low but expected as this is typical for EEG experimentation.

Taken together with the results of the other assessments mentioned in the manuscript, only this analysis can conclude that the studies included contain the evidential value and it is unlikely that selective reporting of significant *p*-values for ERN has occurred.

The Pe *p*-curve analysis indicated that both the half and full *p*-curve were not significant, Z=1, p=.8411 Z=-0.36, p=.3593, suggesting that the selected studies do not contain the evidential value (figure S6). Also, the evidential is inadequate or absent, as the 33% power test of both the full and half curve were below the p<0.1, Z_{full}=-2.16, p_{full}=.0153, Z_{half}=1.37, p_{half}=.9147. This can imply that replicating these results will be difficult. Similar to the ERN, the value of power was low, 5%.
It is worth mentioning the following; for the Pe, many studies did not qualify for the p-curve analysis, as they reported non-significant effect, but yet were published. Interpreting the following results together with the results mentioned in the manuscript suggests that we cannot conclude that selective reporting

of *p*-values has occurred, but that the *p*-values that are reported are random. For now, this means that there were simply a few studies that met the 'criteria' for this analysis, because of lack of hypothesis reporting or lack significant *p*-values.

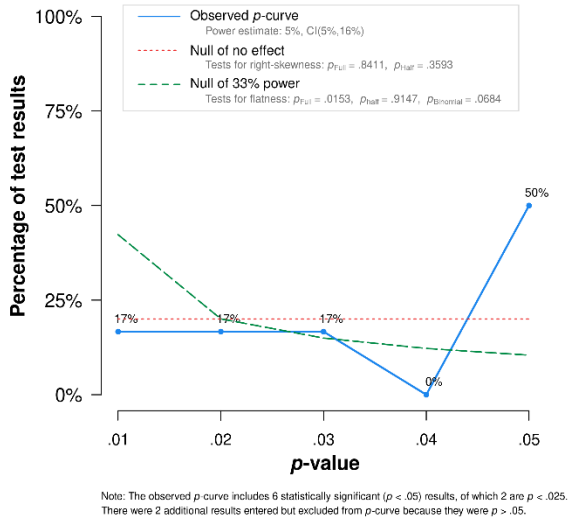


Figure S2 P-value distribution Pe.

Chapter 3

Event-Related Potential (ERP) Measures of Error Processing as Biomarkers of Externalizing Disorders: A Narrative Review

This chapter is published as:

Lutz, M. C., Kok, R., & Franken, I. H. A. (2021). Event-related potential (ERP) measures of error processing as biomarkers of externalizing disorders: A narrative review. *International Journal of Psychophysiology, 166*, 151-159. https://doi.org/10.1016/j.ijpsycho.202

Abstract

Previous studies have shown that electrophysiological measures of error processing are affected in patients at risk or diagnosed with internalizing disorders, hence, suggesting that error processing could be a suitable biomarker for internalizing disorders. In this narrative review, we will evaluate studies that address the role of event-related potential (ERP) measures of error-processing in externalizing disorders and discuss to what extend these can be considered a biomarker for externalizing disorders. Currently, there is evidence for the notion that electrophysiological indices of error processing such as the error-related negativity (ERN) and error positivity (Pe) are reduced in individuals with substance use disorders, attention-deficit/hyperactivity disorder, and in forensic populations. However, it remains unclear whether this is also the case for other understudied disorders such as behavioral addiction. Furthermore, to fully understand how these deficits affect day to day behavior, we encourage research to focus on testing current theories and hypotheses of ERN and Pe. In addition, we argue that within an externalizing disorder, individual differences in error processing deficits may be related to prognosis and gender of the patient, methodological issues and presence of comorbidity. Next, we review studies that have related treatment trajectories with ERP measures of error processing, and we discuss the prospect of improving error processing as a treatment option. We conclude that ERP measures of error processing are candidate biomarkers for externalizing disorders, albeit we strongly urge researchers to continue looking into the predictive value of these measures in the etiology and treatment outcome through multi-method and longitudinal designs.

Event-Related Potential (ERP) Measures of Error Processing as Biomarkers of Externalizing Disorders: A Narrative Review.

The observation that some persons make more repeating mistakes than others was already a subject of study millennia ago by Roman philosophers such as Seneca the younger ("To err is human, to repeat error is of the devil") and Cicero ("Anyone can err, but only the fool persists in his fault"). These observations already constituted the idea that the persistence of making errors might be related to abnormal behavior. In modern times, the question why some people do not learn from their mistakes is still relevant and unanswered. The increasing knowledge about cognitive neuroscience can help to answer this question and enables us to investigate error processing by making use of modern psychophysiological techniques such as electroencephalography (EEG) and magnetic resonance imaging (MRI).

In the current paper we assume that the repeated making of errors is a common hallmark for externalizing behaviors, which are characterized by a pattern of inability to inhibit unwanted behaviors and properly adapt to new situations. Disorders within the externalizing spectrum traditionally include diagnosis such as attention-deficit/hyperactivity disorder (ADHD), substance use disorder (SUD), oppositional defiant disorder (ODD), conduct disorder (CD), and anti-social personality disorder (Krueger and South, 2009). Patients with these diagnoses share a variety of facets (e.g. impulsivity, irresponsibility) that relate to common behavior such as aggression and substance use (Krueger et al., 2007). In turn, these shared characteristics explain the proneness to maladaptive behaviors, such as risk-taking, un-empathic and delinquent behavior. Although this pattern can be both voluntary as well as involuntary, it could fit a pattern of not learning from these 'mistakes', or their experiences with negative outcomes. The cognitive processing of errors, i.e., the ability to detect and respond to a committed error (consciously or not), is an important regulating component in adjusting behavior and could thus be crucial for learning processes. It has been proposed (Olvet and Hajcak, 2008) that neurophysiological error processing is dysfunctional (i.e., reduced) in externalizing disorders and that this dysfunction is an etiological marker for these disorders. In this narrative review, we will explore and evaluate the empirical evidence for the hypothesis that reduced error processing, reflected as smaller error-related electrophysiological brain responses, could be a common biomarker for externalizing behavior and disorders.

Externalizing problems can cause tremendous harm to people displaying these behaviors, as well as to their environment and to society at large. Societal costs due to for example substance abuse, have been estimated to be more than $700 billion annually in the United States alone (Volkow et al., 2016). Treatment of patients with externalizing disorders has proven to be challenging, and is characterized by negative prognosis, treatment drop-out, high relapse rates, and increased chance of incarceration and recidivism. The origins of these externalizing disorders are often already observed in childhood and adolescence and continue or even aggravate later in life. Therefore, early identification of individuals at risk for

developing an externalizing disorder and/or the identification of patients who are resistant to treatment might be helpful when tackling externalizing disorders. Investigating candidate biomarkers might be an important step in aiding to identify persons at risk or to further clarify pathophysiological mechanisms of these complex disorders.

There is a growing body of literature in search of candidate biomarkers for externalizing disorders and behaviors. There are several definitions for the term 'biological marker', but biomarkers generally refer to any objective characteristic that can be measured accurately, is reproducible and is sensitive and specific enough to be observed in a large heterogenous population from patients with a disorder as opposed to healthy individuals (Mehta et al., 2020). Neurobiological parameters, such as genes, hormones, skin-conductance, heart rate, and event-related potentials (ERP's), are examples of such biomarkers. Biomarker and endophenotypic research in the psychiatric and psychological field has greatly contributed to diagnostic and prognostic understanding of disorders and has identified provisional neuro-biological parameters that drive the etiology of disorders (Miller & Rockstroh, 2013). In turn, these markers can inform diagnostic tools, treatment options and future research. In the study of psychopathology, evidence suggests that ERP's are appropriate neurophysiological biomarkers (Hajcak et al., 2019). For example, dysfunction of the prefrontal cortex has been found in addiction (in magnetic resonance imaging studies; Goldstein and Volkow, 2011) and a diminished P300 in patients with substance abuse (Euser et al., 2012; Houston and Schlienz, 2018; Iacono and Malone, 2011) or with ADHD (Mehta et al., 2020). We can learn from the advances made in research on other neurophysiological biomarkers to pinpoint us what still needs to be done, such as investigating the role of genetic markers modulating error processing (e.g. Beste et al., 2010; Monoach & Agam, 2013) or investigating whether ERP's can be an index for behavioral change in treatment settings (Houston & Schlienz, 2018). In this review, we will focus on two response-locked EEG components as candidate electrophysiological biomarkers for error processing: the error-related negativity (ERN or Ne, Gehring et al., 1993, 2018; Falkenstein et al., 1991) and error positivity (Pe, Arbel and Donchin, 2009; Falkenstein et al., 1991; Overbeek et al., 2005). Here, we will not address behavioral performance nor other neurophysiological markers such as time-frequency theta or MRI measures and other behavioral related indices, such as reward processing. Where the ERN is an index for monitoring action outcome, the Pe can reflect error awareness and the more motivational significance of an error. Both ERP's serve as mechanistic markers for behavior adaptation. Research focusing on ERN and to a lesser extent on Pe, have revealed associations with psychological conditions. Previous reviews and meta-analyses (e. g., Gilian et al., 2017; Moser et al., 2016; Pasion and Barbosa, 2019; Riesel, 2019; Riesel et al., 2019), found that error processing can be affected in patients with internalizing conditions such as anxiety disorder and obsessive compulsive disorder, often evidenced by increased ERN potentials in patients compared to healthy controls. In the field of major depression disorder, inconsistencies of error processing effects still need to be resolved (Gilian et al., 2017). For instance, error related brain activity did not differ between healthy subjects and major depression patients in the

study of Schrijvers et al. (2009) or was blunted in the study of Weinberg et al. (2016), yet ERN was related to symptom severity in patients in treatment of their depression (Schrijvers et al., 2009). In other studies, patients with depression disorder were more sensitive to error making reflected by an enhanced ERN compared to controls (Chi & Deldin, 2007; Moran et al., 2017). Similarly, the ERN seems to be a suitable biomarker for anxiety in children and adolescents (Hanna et al., 2020; Meyer, 2017).

This narrative review will not provide a systematic overview of ERN and Pe studies in externalizing disorders. Instead, we take a step back from previous findings and discuss the current state of the art for the notion that reduced ERN and Pe are neurophysiological biomarkers for externalizing disorders by answering the following questions: 1) What is the evidence that the ERN/Pe is reduced in externalizing populations? 2) Can a reduced ERN/Pe be a predictor of externalizing disorders? 3) Are there individual differences in ERN/Pe amplitudes within externalizing disorder populations, and what do these individual differences indicate? 4) Can the ERN/Pe be indicative of treatment trajectories? 5) Can error processing be improved by targeted interventions?

1) What is the evidence that the ERN/Pe is reduced in externalizing populations?

To investigate biomarkers is to explore deficits or sensitivities in comparing patient samples and healthy individuals. There are several case-control studies showing that ERN and Pe are reduced among externalizing populations (e.g., Brazil et al., 2009; Marquardt et al., 2018; Morie et al., 2014). In these studies, the typical design is to contrast the amplitude of the ERP's, measured by a typical cognitive task such as the Eriksen Flanker or Go-noGo task, between the clinical and control groups. The robustness of the reduced ERN and Pe in externalizing populations has been very recently confirmed by two meta-analyses compiling case-control studies (Lutz et al., 2021; Pasion & Barbosa, 2019). The meta-analysis of Pasion and Barbosa (2019) found an overall effect size of g = -0.65 (based on 32 ERN studies, 44 effect sizes, n = 1921), where a negative g indicates a decreased amplitude for the clinical/subclinical group. Similarly, Lutz et al. (2021) found a small to medium effect for the ERN and Pe, observing a decreased amplitude for the clinical/subclinical group when compared to healthy controls. These studies provide the initial evidence for reduced ERN and Pe in specific subgroups within the externalizing spectrum, i.e., substance use disorder, ADHD and personality disorders, as well as in individuals with subclinical levels of externalizing behavioral problems (e.g., symptoms of aggression, psychopathy, and impulsivity, Hall et al., 2007; Zijlmans et al., 2019). Also, in the two meta-analyses, several important moderators, such as ERN peak scoring, type of diagnosis, disorder severity, task and presence of comorbidity and performance feedback and age, were tested to elucidate on the heterogeneity of the data. Although these meta-analyses differ in inclusion criteria and approach, the results of the meta-analyses highlight that performance monitoring processes are indeed compromised in individuals on the externalizing spectrum. This is in correspondence with results from disorder-specific systematic reviews on error processing, such as the meta-analysis of Geburek et al. (2013) and

Kaiser et al. (2020) in ADHD and the systematic review of Luijten et al. (2014) on substance use disorders. In sum, evidence is building to support the hypothesis that reduced error processing, indexed by the ERN and Pe, is present in externalizing disorders and could indeed be considered a suitable biomarker. When gathering the evidence for the ERN and Pe as a neuro-biomarker in the externalizing spectrum, we have to consider possible moderators, that are for instance population or research related. When looking at the patients with externalizing disorder, sex is a possible moderator. In healthy samples, performance monitoring at a behavioral and neurophysiological level is moderated by sex (Fischer et al., 2016; Hill et al., 2018; Larson et al., 2011; Li et al., 2009). That is, it appears that men show more error related activity than females (Fischer et al., 2016; Hill et al., 2018) and that women show different activation and deactivation patterns of the brain than men (Li et al., 2009). This could be due to sex related morphometric differences of the brain, e.g. a larger anterior cingulate cortex (ACC) in men (Ruigrok et al., 2014). It is known that disorders in the externalizing spectrum are more prevalent in males (Becker & Hu, 2008; Eaton et al., 2012; Krueger and South, 2009) which is reflected in the unbalanced sampling of the (sub)clinical participants in error processing studies. Including participants in studies of both sexes appears to be challenging, making it not yet possible to draw solid conclusions when systematically investigating error processing deficits between male and female patients, as Kaiser et al. (2020) rightfully discuss in their ERP meta-analysis in ADHD. Already, there are indications that error processing is affected differently for males than females in externalizing samples (for psychopathy see Efferson & Glenn, 2018; in internet gaming disorder; Dong et al., 2018; for food addiction see Hsu et al., 2017). For instance, when predicting cocaine relapse and early relapse time, a reduced activity in the dACC and thalamus was indicative for females, whereas the reduced activity in the dACC (dorsal anterior cingulate cortex) and left insula was indicative for males (Luo et al., 2013). More extensive research with this moderator is warranted to pinpoint which brain area and activity is affected for males or females in particular, and whether this is disorder specific or generalizable for the externalizing spectrum. Another possible moderator that should be taken into account when studying error processing pertains to methodological choices in ERP (pre)processing. Recently, Klawohn et al. (2020) investigated the effect of different quantification ERP methods, and although there are differences, most methods had acceptable to good internal consistencies. Indeed, there are several important moderators in error processing studies that affect the internal consistencies of studies (Sandre et al., 2020), carefully investigated in the meta-analysis of Clayson (2020). Choices relating to EEG referencing, the scoring procedure, electrode (cluster), (ocular) artifact rejection and number of trials for ERP calculation, can influence the internal consistency of the ERN (Clayson, 2020). An adequate solution for this moderation issue in EEG research is the disclosure of hypotheses, data collection, processing and analyses through pre-registration (Paul et al., 2021) and open science practices, which not only allows for reproducibility and transparency but provides a control mechanism for these possible moderators.

In order to further validate whether neurophysiological markers of error processing are suitable as biomarkers of externalizing disorders, we need to better understand how defiant error-related brain activity acts as an underlying mechanism in these disorders. The functional significance of error processing relies on several theories or hypotheses that might be complementary and exclusive at the same time. The mismatch, reinforcement and learning based, conflict monitoring and motivational significance theory, as outlined in Olvet and Hajcak (2008) and Weinberg et al. (2012) for ERN and the affective-processing, behavior-adaptation and error awareness hypotheses, as discussed in Overbeek et al. (2005) for Pe, could help us to understand the underlying mechanisms for the disruptive behavior of patients with externalizing disorders. Briefly, these hypotheses attempt to explain how errors are (not) processed and evaluated by the brain, how errors elicit learning behavior which in turn leads to adjustment in behavior. When this system does not adequately work, error processing is affected and leads to the inability to adjust disadvantageous behavior, which is often observed in individuals with externalizing disorders. However, it is unclear how these hypotheses explain day to day behavior that we see in patients with externalizing disorders or, how we could use these hypotheses to improve error processing deficits (e.g. training performance monitoring using feedback or error awareness training).

A continuation of exploring the neural network behind performance monitoring is encouraged (Wessel, 2012; Wessel et al., 2012). For instance, an insufficiently working salience network of the brain, indicated by hypoactivity of the insula or ACC (as described in Ham et al., 2013) can explain the performance deficits seen in externalizing disorders. Also, the interplay of ACC and other brain regions on functional and structural level contributing to regulating behavior, cannot be omitted. For instance, distinct activation patterns of the insula, rostralACC and the dorsolateral prefrontal cortex in patients with cocaine addiction and intermittent explosive disorder (Moeller et al., 2014a, 2014b) explained the behavior in the performance tasks.

So far, we interpreted reduced error related brain activity as an indicator of the inability to adjust behavior to avoid future errors, which in turn are related to symptoms (e.g. the continuation of substance abuse: Crane et al., 2018; Easdon et al., 2005; Franken et al., 2007; Hajcak, 2012; Luijten et al., 2011; Sokhadze et al., 2008). The late component of error processing, Pe, has only recently been subject of experimental studies (e.g., Di Gregorio et al., 2018), and therefore less is known about the functional significance of the Pe in externalizing disorders. In the study of Rosburg et al. (2018), it has been proposed that a reduced Pe could reflect the reduced awareness of the committed errors in child sexual offenders, which in turn could contribute to their delinquent behavior. A reduction in the recognition (that is the awareness) of, or the motivational significance of an error, might explain why individuals with externalizing disorders are less inclined to change their behavior because of that error. Clearly, more studies are needed to clarify the role of the reduced ERN and Pe in externalizing conditions. One interesting possibility, that has become available with mobile EEG and mobile

cognitive assessments, is the investigation of error-processing in 'daily life', particularly in relation to externalizing behaviors. With this method, we could gain knowledge about the significance of error processing deficits in daily cognitive processes.

Another outstanding question is whether the ERN and Pe are suitable biomarkers for the externalizing spectrum or whether it is specific to certain externalizing disorders or problem behaviors. There is a substantial number of studies providing evidence that error processing is affected in individuals with ADHD or in substance use disorder. However, other externalizing conditions have been understudied, such as anti-social personality disorder or psychopathy (Vallet et al., 2021). A few incidental case-control studies have indicated that error processing is reduced in behavioral addictions, in for example computer gaming addiction (for the ERN, Littel et al., 2012) or internet gaming addiction (Park et al., 2020; Zhou et al., 2013), and food addiction (Franken et al., 2018). Taken from these four studies, the participants with addictive behavior made more errors, were more impulsive, and showed decreased ERN, indicating to reduced reduce performance monitoring. Although these results are in accordance with results from substance addiction, we cannot yet draw firm conclusions based on four studies. Yet, we have reasons to believe that future studies will find error processing deficits in patients with behavioral addiction as brain studies examining the functional activity, structure, and connectivity already have shown that the ACC or orbitofrontal cortex and connectivity with the insula are affected (for instance in internet gaming disorder: Dong et al., 2015; Ko et al., 2014; Lee et al., 2018; Xing et al., 2014; Zhou et al., 2011). More research is needed on both ERN and Pe in diverse externalizing populations to explore whether the two ERP's are a general biomarker for externalizing disorders, or only related to specific externalizing behaviors. Another outstanding and relevant issue is the role of error processing in explaining comorbidity as externalizing and internalizing conditions can co-occur (Krueger & Markon, 2006). Since externalizing disorders are often characterized by a decreased ERN and internalizing disorders by an increased ERN, it is interesting to investigate how error processing plays a role in comorbid conditions (as for example investigated in the study of Schellekens et al., 2010 and Gorka et al., 2016). Although it has been proposed that error processing can be considered as a transdiagnostic marker which is also relevant for individuals presenting with comorbid disorders (Ladouceur, 2016; Pasion & Barbosa, 2019; Weinberg et al., 2015), the direction of the association (reduced vs. increased) is one important aspect that needs to be examined. When trying to understand the comorbidity issue, the p factor or a generalized psychopathology factor might offer insight (Caspi et al., 2013; Caspi & Moffitt, 2018). The p factor, incorporating the internalizing and externalizing and thought disorders/psychotic experiences allows for the co-occurrence of problems from all disorders, which is applicable in our discussion here. However, until now researchers have been devoted to test such nosology in large populations, and a few studies have validated this model with global executive functioning in children (Bloemen et al., 2018; Martel et al., 2017; Shiels et al., 2019). We can only speculate how error processing or cognitive control fits in the general factor. The error processing effects found in both internalizing and externalizing disorders

support the fundamental idea of the bifactor or hierarchical factor, proposed by Caspi and Moffitt (2018). It is therefore important that error processing components are incorporated in the Research Domain Criteria of the NIMH (as discussed in Weinberg et al., 2015). For now, experimental studies are needed to investigate the role of error processing in the etiology of either externalizing or internalizing disorders as well as comorbidity between these disorders by controlling for both symptoms in terms of onset, severity, and genetic predispositions.

2) Can a reduced ERN/Pe be a predictor of externalizing disorders?

In order to be able to determine whether reductions of the ERN/Pe can be considered an etiological biomarker for externalizing disorders, the role of error processing should be studied at an early age. When establishing deficits in adulthood, the notion that error processing was already affected in childhood, should be tested. Several cross-sectional or case-control reports have found reduced error processing in children with elevated subclinical levels of externalizing behavioral problems and children with clinical externalizing disorders (e.g., Burgio-Murphy et al., 2007; Kessel et al., 2016; Meyer & Klein, 2018; Moadab et al., 2010; Stieben et al., 2007) and ADHD (Groen et al., 2008; Senderecka et al., 2012). The next step is to test these associations in longitudinal designs in children and adolescents, as illustrated by the review study of Meyer (2017) in anxiety. These designs should keep in mind the normal trajectory of error processing indices (such as the increase of the ERN over time found in the reviews of Lo, 2018, and Tamnes et al., 2013) but are essential to determine whether error processing could be a predictor for externalizing problems.

Case-control studies do not provide information about the causal role of error processing deficits in externalizing disorders. It has been proposed that error processing deficits could indeed be one of the causal factors of externalizing disorders. To study this, the level of error processing in at-risk samples (that is children, family members, or adults that have an increased chance for developing externalizing problems due to their parental conditions or exposure) can indicate a possible causal effect. The study of Euser et al. (2013) is an example of such a study, where the hypothesis that error processing is an antecedent and reflects biological predisposition to the disorder (in this case SUD). High risk adolescents, who had a parent undergoing treatment for SUD showed smaller ERN amplitudes than normal risks (healthy controls). In this design, one can find evidence whether the deficits found in error processing contribute to the disorder that is diagnosed at a later stage in their lives. This idea has been studied previously with other candidate neurophysiological markers such as the reduced P3 (Euser et al., 2012; Iacono et al., 2002) within the externalizing spectrum. One the other hand, it is also conceivable that error processing deficits could be a consequence of psychopathology. This implies that error processing deficits are not yet detectable at the beginning of a disorder, but rather a result of the disorder. This could also clarify why certain patients show a negative prognosis as opposed to others. This alternative hypothesis seems to be particularly relevant in substance use disorders, as it is known that the prolonged use of

substances has detrimental effects on the brain and cognitive functions (Goldstein & Volkow, 2002, 2011; Leshner, 2003).

To conclude, there is some evidence of early error processing deficits in children in the externalizing spectrum. It remains unclear whether error processing deficits can be considered a vulnerability or consequence of developing externalizing disorders. Since experimental studies addressing causality issues are obviously unethical, more prospective longitudinal cohort studies focused on the ERN and Pe and the development of externalizing problems over time among children in the general population are needed in order to gain insight in developmental aspects and their causal role in problem behaviors.

3) Are there individual differences in ERN/Pe amplitudes within externalizing disorder populations, and what do these individual differences indicate?

Both meta-analyses mentioned previously (Lutz et al., 2021; Pasion & Barbosa, 2019) have shown that there is substantial heterogeneity within clinical groups, indicating that error processing deficits may not be evident for all patients within a clinical disorder and are present in various degrees. There are several possible explanations for the observed variation in the error processing correlates, such as individual differences, presence of comorbidity, severity of symptomatology. First, variability in ERN could be due to individual differences in several domains. For instance, there is evidence that the individual levels of cognitive control in patients with ADHD (Meyer & Hajcak, 2019) and the individual levels of the trait defense reactivity (Weinberg et al., 2012) can modulate the ERN. Several other studies have shed light on additional characteristics that modulate ERN/Pe, such as working memory performance (Miller et al., 2012), fearfulness in toddlerhood (Brooker & Buss, 2014), sensitivity towards rewards and punishment (Boksem et al., 2006; Dikman & Allen, 2000), personality or externalizing traits (Pailing & Segalowitz, 2004; McDonald et al., 2021) and behavioral inhibition (Amodio et al., 2007). Also, it is important to mention that the role of individual differences needs to be studied in large samples. Small sample sizes, typically used in these studies, overestimate effect sizes and have low reproducibility (Button et al., 2013; Larson & Carbine, 2017). An example of this is point is illustrated in the study of Bernoster et al. (2019). In this larger scale study, a clear link between ERN and impulsivity was not found, despite previous reports of this association in smaller scale experiments. The role of individual differences in error processing clearly merits further investigation as it could explain the heterogeneity of the error processing deficits. Another possible explanation for the variation within clinical groups is the presence of comorbid internalizing problems, that often co-occur in for example substance use disorders (Franken et al., 2017; Olvet & Hajcak, 2008; Smith et al., 2017). Or, variability in error processing deficits could be due to the severity of the externalizing symptoms. Both relate to a more dimensional approach of externalizing spectrum, as suggested by Krueger et al. (2007). The high or low levels of externalizing and comorbid problems (irrespective of a particular disorder) are related to more or less pronounced error processing and other executive functioning deficits, can explain possible

variance observed. In turn, studying this link can be insightful for the global functioning of patients. Already, studies have examined the degree of symptomatology and its association with the variation in ERN or Pe. For instance, a smaller ERN has been related to more heavy heroin use (Chen et al., 2013), alcohol use (Campanella et al., 2017; Smith & Mattick, 2013), and nicotine dependence (Luijten et al., 2011). To investigate the comorbidity, symptom severity and global functioning hypotheses, experiments could use ecological momentary assessment applications to asses error processing in combination with comorbid symptoms, symptom severity or functioning ratings in patients. This line of research should be continued and include the Pe component, in order to examine whether and how error processing is related to individual differences, symptom severity, or global functioning in externalizing patients. This knowledge could aid in predicting prognosis of patients and give insight in treatment success.

4) Can the ERN/Pe be indicative of treatment trajectories?

In line with the previous paragraph, discussing the ERN or Pe as an index of symptom severity, error processing could be used as an indicator of status, relapse or treatment success in externalizing disorders (Gorka et al., 2019; Marhe et al., 2013, Steele et al., 2014, and at trend level in Luijten et al., 2016). The study of Gorka et al. (2019) showed that the ERN can possibly be related to pathological stages of patients with alcohol use disorder. In other words, the magnitude of the ERN could differentiate between current, remitted and at-risk in patients with alcohol use disorder. Similarly, although at trend levels, Pe and ERN were related to smoking relapse and resumption in Luijten et al. (2016). Both these studies give insights in how error processing could be directly related to recovery trajectories in addiction disorders. When investigating treatment success, the results of Marhe et al. (2013) and Steele et al. (2014) showed how error processing deficits could be predictive of treatment outcome, by relating error related brain activity to later treatment success. Moreover, the results of Padilla et al. (2011), Schlienz et al. (2013), and Schlienz and Hawk (2017) suggest that ERN is very sensitive to the cessation of alcohol in patients with alcohol disorder. Indeed, distinct activation patterns during error processing of dACC, thalamus, and insula in cognitive processing are found during abstinence in other addiction disorders, such as cocaine-dependency patients, implying that indicators of brain functionality are of predictive value for drug relapse (Connolly et al., 2012; Luo et al., 2013). Indications of the predictive value of error processing for treatment outcome are also found in research in forensic settings. For example, Steele et al. (2015) found that Pe could differentiate between incarcerated males who were or were not subsequently rearrested. In this prospective study, the error related brain activity of adult men was related to later rearrests, information that was gathered in a follow-up. Although several studies show that error processing is potentially predictive of treatment outcome, more evidence is needed to determine its potential in other externalizing disorders.

5) Can error processing be improved by targeted interventions?

At the moment, the improvement of (aspects of) cognitive control, such as error processing, is one of the crucial targets in many studies aiming to treat externalizing disorders. However, in practice it seems rather difficult to improve cognitive control. With exemption of one study (Schoenberg et al., 2014), we are not aware of studies showing that certain treatments can improve error-processing (reflected by improvements in ERN/Pe amplitudes post interventions) specifically, nor that improvements in error processing result in adaptations in behavior. Having said this, there have been successful attempts in the broad area of self-control (Inzlicht et al., 2014) that provide important clues on how to continue the search for effective interventions to improve error processing. Broadly, these studies focus on three types of interventions: cognitive-behavioral training, brain stimulation (Bellaïche et al., 2013; Carmi et al., 2018; Verveer et al., 2021), and meditation techniques (Slagter et al., 2011). Cognitive-behavioral therapy could address elements that are related to cognitive processes. An interesting attempt to test this idea is the study of Schoenberg et al. (2014), where cognitive therapy addressed cognitive flexibility, attention, and behavioral regulation. In this study, they found that mindfulness-based cognitive therapy elevated Pe in adult patients with ADHD. Concerning brain stimulation, current randomized control trials and experiments are exploring the possibility of using brain stimulation to modify cognitive control and neurofeedback to adjust performance monitoring. Non-invasive neurostimulation using transcranial magnetic stimulation (TMS) or transcranial direct current stimulation (tDCS) can alter the activity of targeted brain regions, associated with error processing, such as the ACC. Although research of the use of neurostimulation in externalizing disorders is accumulating and promising (for drug addiction: Song et al., 2019; in ADHD: Soltaninejad et al., 2019), very few studies investigate its effects on cognitive control and even less studies investigate error processing specifically. In healthy volunteers, successful attempts to modulate error processing through tDCS have been reported (Bellaïche et al., 2013). There are preliminary indications that TMS treatment is related to changes in ERN in patients with obsession-compulsive disorder (e.g., Carmi et al., 2018). However, in patients with cocaine addiction that underwent tDCS treatment (Verveer et al., 2021), no changes in the ERN nor craving were found. Last, meditation techniques are another promising avenue to improve cognitive control. Researchers are currently exploring its effects in healthy volunteers (Andreu et al., 2019; Lin et al., 2019; Pozuelos et al., 2019; Quaglia et al., 2019; Saunders et al., 2016; Slagter et al., 2011; Teper et al., 2013) and it seems too early to speculate about the effectiveness in general and for externalizing disorders in particular.

At the moment there are no proven interventions that could indicate that error processing could be improved in patients with externalizing disorders, but cognitive training, particularly like the study of Schoenberg et al. (2014), brain stimulation, and meditation techniques are certainly worth exploring, as they can address cognitive control elements shown in studies with healthy samples. Despite the limitation that improving error processing will only alleviate some of the problems seen in patients with externalizing disorders, future

research should investigate whether improving error processing in turn leads to adapting their behavior.

Conclusion and future directions

Our narrative review shows that error processing and specifically a reduced ERN and Pe are associated with externalizing symptoms, and that empirical evidence is building that deficits in error processing measured at neurophysiological level might be a suitable biomarkers for externalizing conditions. However, many questions remain unanswered. We address several key themes in Table 1, where we give an overview of preliminary evidence and recommendations for future research directions. In this review, we focused on the neurophysiological indices of error processing (ERN and Pe), leaving the discussion on the significance of behavioral indices of error processing untouched. Also, a related ERP, the P300, shares important variance with ERN/Pe, and we would like to shortly touch upon these topics.

Concerning the functional significance of the ERP measures, there are some indications that ERN or Pe is related to task behavior and behavioral adjustment. For instance, the Pe was correlated with post-error slowing (PES; Hajcak et al., 2003; Nieuwenhuis et al., 2001). Additionally, the ERN is associated with post-error accuracy, when mediated by post-error slowing (Beatty et al., 2020). Although there is no consensus yet, several studies (e.g. Cavanagh et al., 2009; Kalfaoğlu et al., 2018) have found relationships between ERN and post-error slowing, indicating that ERN is related to task-related behavior. The authors in Beatty et al. (2020) carefully laid out possible explanations for the inconsistencies around the relationship between behavior and ERN. In addition, it is not clear to what degree the behavioral performance during error processing tasks is affected in patients with externalizing disorders. For example, post-error slowing appears to be affected in patients with ADHD (Balogh & Czobor, 2014) and cocaine use disorder (Franken et al., 2007). On the other hand, there are reports showing a distinct pattern of behavior. For instance, the externalization groups in the study of Gorka et al. (2019) not make more errors in cognitive tasks, nor do they take longer (evidenced in reaction times) in pressing the correct buttons when compared to controls, as in the study of Zhang et al. (2009) in children with ADHD. Studying the significance of behavioral performance during error processing as well as the association of the behavior with the neurophysiological indices is encouraged. Elucidating on this could answer another interesting question: how ecological valid are our lab measures of error-processing when translating to real-life behavior? We recommend researchers to endeavor in elucidating this through e.g., ecological momentary assessment and virtual reality techniques, in order to better test these hypotheses.

Table 1.

Table with key formulations from this review, a short overview of evidence and suggestion for future directions

Key formulations	Preliminary evidence*	Future research directions
Psychometric characteristics of the error- related negativity are clear	Sandre et al., (2020) Klawohn et al., (2020) Riesel et al., (2013) Rietdijk et al., (2014) Clayson (2020)	Report and examine the role of moderators, such as gender and methodological decisions, in ERN/Pe experiments.
Error-related negativity and error positivity is reduced in patients with externalizing disorders	Lutz et al., (submitted) Pasion & Barbosa (2019) Kaiser et al., (in ADHD: 2020) Vallet et al., (2021)	• Extend knowledge of the underlying theories for ERN and functional hypotheses of Pe • Investigate whether error processing is affected in behavioral addiction
Cause or effect of error processing in psychopathology unclear	Meyer et al., (2017) Euser et al., (2013)	Investigating the developmental path of error processing in relation to psychopathology through longitudinal and cross-lagged model designs
Error processing is related to individual differences in the cognitive and personality domains, disorder severity and comorbidity	Cognitive/personality domains: Pailing and Segalowitz (2014) Reward/Punishment sensitivity: Boksem et al., (2006) Symptom severity: Campanell et al., (2017) Comoribity: Franken et al., (2017)	Studying moderating measures to explain heterogeneity in disorders in larger samples • Symptom severity • Relation to comorbidity • Traits and personality • Genes and hormones • Other cognitive measures (e.g. working memory, attentional bias) • Error processing in daily life; using ecological momentary assessments tools
Error processing could be used as an indicator of status, relapse or treatment success	Gorka et al., (2019) Marhe et al., (2013) Steele et al., (2014) Luo et al., (2013) Steele et al. (2015)	Prospective studies using neuroprediction: examining the predictive value of error processing for treatment trajectory and relapse/rearrests rates
Unclear whether and how we can train or stimulate performance monitoring to reduce error processing deficits	Schoenberg et al., (2014) Bellaïche et al., (2013) Verveer et al., (2020)	Through training programs, stimulation techniques or new experimental paradigms, can we improve error processing deficits?

*The authors acknowledge that more studies than mentioned here support the key formulations drawn.

The reduced P300 has been considered a viable predictor for externalizing disorders (see e.g., Patrick et al., 2006). There are however obvious paradigm/component related differences with the reduced ERN/Pe, such as the polarity, latency (ERN) and onset. The latter characteristic differentiates the ERP's the most: the ERN/Pe is triggered by errors (response triggered) whereas the P300 is elicited by a scope of stimuli: affective, oddball etc.. Hence, the P300 reflects broader processes of decision making, leaving the unique feature of ERN and Pe to be specific error processing ERP's. Together these ERP's, among other ERP's such as N200, predict neurocognitive processes and behavior for externalizing disorders. Most studies in this research area consist of relatively small case control samples, in predominantly ADHD and SUD which limits our knowledge on the actual role of error processing. It should be examined whether the variation in error processing found within externalizing disorders is related to differences individuals and their psychopathology. Case-control studies cannot solve questions such as 'is reduced cognitive control a cause or consequence of externalizing disorders?' and 'can we improve error processing in order to help individuals with externalizing problems?'. Future studies should focus on developmental processes in order to clarify a possible causal role, by using longitudinal designs and by exploring the effect of error processing on daily behavior. Also, more studies on the clinical relevance and development of intervention programs to improve error processing in externalizing disorders are needed. We believe that multi-method studies on error processing that are embedded within cognitive neuroscience (MRI, EEG), epidemiology (large prospective cohort study), and the emerging area of ecological momentary assessments (EMA), will be a fruitful new avenue to further explore mechanisms, treatment, and prevention of externalizing disorders.

Chapter 4

Developmental Trajectory of Flanker Performance and its Link to Problem Behavior in 7-to 12-year-old Children

This chapter is submitted as:

Lutz M. C., Kok, R., Koot, S., Van Lier, P. A. C., Buil, M., & Franken, I. H. A. (submitted). Developmental trajectory of flanker performance and its link to problem behavior in 7-to 12-year-old children. Retrieved from https://doi.org/10.31234/osf.io/fj62k

Abstract

This study investigates both the developmental trajectory of flanker task performance in children and the association with the development of teacher-reported problem behavior. Five waves of flanker performance and behavioral and emotional problems were drawn from a large longitudinal sample of elementary school children in the Netherlands (1424 children, ages 7 to 12 years). Latent growth curve modeling (LGM) identified a piecewise decrease in flanker response time: the steepest decline was found from 7 to 9 years old. Boys had lower levels of response time at age 7 than girls. Children showed a linear decrease in behavioral and emotional problems over time. Parallel LGMs revealed that lower levels of initial flanker response time were associated with a stronger decrease in anxiety problems and oppositional defiant problem behavior. A faster decline in response time was associated with a faster decline in depression problems, attention deficit hyperactivity-, and oppositional defiant-related behavior. Results offer insight into the normative development of performance monitoring in childhood and the link between behavioral measures of performance monitoring and behavioral and emotional problems.

Keywords: flanker, conflict and performance monitoring, children, developmental psychopathology, behavioral and emotional problems

Developmental Trajectory of Flanker Performance and its Link to Problem Behavior in 7-to 12-year-old Children

Task performance is often an overlooked marker of performance monitoring in neurophysiological studies, even though it has proven a valuable integrative component when understanding the interplay between the brain and behavior (Sandre et al., 2020; Schroder & Moser, 2014). Performance monitoring, which includes error and conflict monitoring, is a higher-order cognitive function ensuring the ability to reflect on one's actions, detect errors and initiate behavior adjustment (Gehring, Goss, Coles, Meyer, & Donchin, 1993). Like many cognitive abilities, conflict and error monitoring change throughout development (Luna, Marek, Larsen, Tervo-Clemmens, & Chahal, 2015). Differences in the ability to monitor performance have been linked to psychopathology during childhood and adulthood (Meyer, & Hajcak, 2019; Olvet & Hajcak, 2008). However, most studies investigating performance monitoring are cross-sectional, making it impossible to study developmental aspects. It is unclear how flanker performance can be related to the development of behavioral and emotional problems in childhood. Therefore, the current study has two goals: 1) investigate the development of child performance on a modified flanker task and 2) test whether the development of flanker performance is associated with the development of behavioral and emotional problems in a sample of 7-to 12-years old children attending mainstream elementary schools.

Both conflict and error monitoring can be measured through several cognitive paradigms, with the Eriksen flanker task being one of the most used in the field (Eriksen & Eriksen, 1974). The task is simple in its design and instruction, and the cues are non-verbal and easily adaptable to accommodate participants' age or experimental manipulation. It is a forced-choice paradigm in which participants are presented with a string of stimuli and instructed to locate the target symbol accurately and as fast as possible, ignoring the other irrelevant (non-target) symbols (flankers). There are two trial types, congruent and incongruent, where the congruent trial has the same flankers (e.g.,> > > > >), whereas the incongruent contains conflicting flankers (e.g.,> > < > >). This conflicting information requires more complex cognitive processing and is quantified by longer response times on incongruent trials accompanied by more commission errors (Shenhav, Botvinick, & Cohen, 2013).

According to conflict monitoring theory (Botvinick, Braver, Barch, Carter, & Cohen, 2001), dealing with incongruency relies on effectively detecting competing or interfering response options when planning/executing a goal-directed action. Addressing this conflict monitoring system directs attention to current behavior (identification and evaluation of errors) and motivates the adjustment of future behavior (correcting), ultimately reinforcing learning. Conflict monitoring theory is well suited to explain the development of performance monitoring and how behavior and neural maturation are related to cognitive control development (Lo, 2018). Tasks that evoke conflict and error monitoring, such as the flanker task, are associated with specific activation of the anterior cingulate cortex (ACC; Richard Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2003; van Veen & Carter, 2002b). Indeed,

there is a link between level of ACC activity and ongoing task behavior (e.g., Ordaz, Foran, Velanova, & Luna, 2013; Sheth et al., 2012). Although conflict monitoring theory explains behavior in performance monitoring tasks (see Ullsperger et al., 2014), conflict and error monitoring processes are both active during task performance, especially in the brain. This is why other theories, such as the reinforcement learning theory (Holroyd & Coles, 2002) or the adaptive orienting theory of error processing (Wessel, 2018), provide valid theoretical frameworks to explain adaptation behavior following performance monitoring.

The behavioral indices of the flanker task allow for the investigation of performance monitoring on a behavioral level. Using the response times and accuracy per trial type (congruent vs. incongruent), here referred to as the congruency effect, can be an index of conflict monitoring (Botvinick et al., 2001). Another measure that is related to performance monitoring is the speed-accuracy trade-off (SAT; Rabbitt, 1966). Speed-accuracy trade-off (SAT) is the inverse relationship between the speed of response and response accuracy, where an increase in the pace of responding is often at the expense of correctness (Heitz, 2014; Ranger, Kuhn, & Pohl, 2021). According to the most reported SAT score (the inverse efficiency score, Townsend & Ashby, 1983), a larger SAT is observed when an individual is slower in response time while making few errors. A small SAT reflects the focus on faster response time while making more errors (Liesefeld & Janczyk, 2019). Post-error slowing (PES) is a behavioral index of error monitoring, defined as the slowing of response time on a post-error trial. While different accounts explain the PES phenomenon (Danielmeier & Ullsperger, 2011; Dutilh et al., 2012; Rueppel, Mannella, Fitzgerald, & Schroder, 2022), PES generally reflects the behavior adjustment mechanism following error-making and serves as a measure of cognitive control. Both SAT and PES capture individuals' 'strategy' during task performance. Changes in these parameters are the result of adequate performance monitoring, as they are needed to optimize overall performance and the avoidance of future errors. In turn, this ability to adapt behavior after wrongful decision-making is a crucial mechanism for problematic behavior, often observed in different psychological disorders.

Normative performance monitoring development during childhood facilitates the fine-tuning of the abilities to focus attention on relevant stimuli, evaluate wrong decision-making and self-regulate behavior (Denervaud, Hess, Sander, & Pourtois, 2021; Lo, 2018). This is driven by unique changes in the maturation of performance-monitoring brain regions such as ACC (Tamnes, Walhovd, Torstveit, Sells, & Fjell, 2013) and contributing neurotransmitter systems, that change throughout development (Luna et al., 2015). This is why we expect that performance monitoring improves during childhood. Indeed, flanker task performance has been associated with children's age. Cross-sectional studies indicate that older children are in general, more accurate and faster than younger children (Buzzell et al., 2017; Checa, Castellanos, Abundis-Gutiérrez, & Rueda, 2014; Davies, Segalowitz, & Gavin, 2004; Gavin, Lin, & Davies, 2019; Gorday & Meyer, 2018; Overbye et al., 2019). Also, children are slower and commit more errors compared to adults (e.g., Santesso, Segalowitz, & Schmidt, 2006; for a review, see Hämmerer, Müller, & Li, 2014). There are mixed findings considering PES in childhood (Rueppel et al., 2022). Some studies find no age-related changes in PES (Denervaud

et al., 2021; Ladouceur et al., 2007; Taylor, Visser, Fueggle, Bellgrove, & Fox, 2018), some studies report a developmental decrease in PES (Meyer, Weinberg, Klein, & Hajcak, 2012; Smulders, Soetens, & van der Molen, 2016) and others find a developmental increase in PES (Hogan, Vargha-Khadem, Kirkham, & Baldeweg, 2005; Overbye et al., 2019). These discrepancies could be caused by different forms of calculating PES or experimental designs, yet the true meaning of this inconsistency is unknown. Studies usually don't report participants' speed-accuracy trade-offs. A noticeable exception is a study of Ladouceur, Dahl, and Carter (2007), where the sample of 12-year-old children equally valued speed and accuracy during the flanker task performance. Taken together, the above-mentioned studies suggest that children improve in response times and accuracy over time. We expect similar task performance improvement over time in the current study and will explore PES and SAT across 7- to 12-year-olds in the current sample.

Children are vulnerable to developing psychological problems during elementary school (Boer et al., 2021). Although there are other developmental markers that play a role in the etiology of a disorder, performance monitoring markers have been put forth as important neuro-cognitive markers that play a role in the development of emotional and behavioral disorders. It is possible that cognitive markers (such as performance monitoring) can be seen as a risk for psychopathology, that is, cognitive dysfunction is considered transdiagnostic for psychopathology (Abramovitch, Short, & Schweiger, 2021). Yet, well-defined theories that explain how behavioral indices of performance monitoring could explain maladaptive behavior and, in turn emotional and behavioral problems, are lacking. Research relies on the theorical hypotheses that are available to explain neurophysiological indices of performance monitoring. For example, according to the defense reactivity trait hypotheses (Weinberg, Riesel, & Hajcak, 2012), an error can be considered as a threat to an individual's safety, and therefore errors are considered 'bad' and should be avoided (Hajcak & Foti, 2008). Yet, results finding an association between anxiety and behavioral indices of performance monitoring are mixed (Rueppel et al., 2022; Weinberg, et al., 2012). For instance, Meyer, Weinberg, Klein, and Hajcak (2012) found that larger PES and post-error mistakes were more prevalent in children with higher anxiety scores. This was not the case in other studies, where there was no link between cognitive control and anxiety (Ladouceur et al., 2006; Meyer et al., 2013). In youth with major depression disorder diagnosis, no clear association between the response times or accuracy rates and depression scores was observed (Ladouceur et al., 2012). However, young adults who report high levels of depressive symptoms show worse behavioral indices in performance tasks (Compton et al., 2008; Holmes & Pizzagalli, 2007). Several studies show associations between behavioral problems and conflict and error monitoring (Balogh & Czobor, 2016; Meyer et al., 2012; Woltering, Granic, Lamm, & Lewis, 2011) in children. Meta-analytic reviews reveal that individuals (predominantly children) with an attention deficit hyperactivity disorder (ADHD) show a diminished PES, slower response times, and higher error rate than healthy controls (Balogh & Czobor, 2016; Geburek, Rist, Gediga, Stroux, & Pedersen, 2013). In children with conduct or oppositional problem behavior during childhood, studies relating performance monitoring with problem behavior are, to our knowledge, scarce.

An exception is the study by Stieben et al. (2007), which showed that 10-year-old children with externalizing problems did not show post-error slowing. In the study by Woltering et al. (2011), children with high levels of externalizing behavioral problems (based on Child Behavioral Checklist scores) were slower in response times than the comparison group of typically developing children. Given the scant evidence that performance monitoring could be related to problem behavior in children, the studies described here rely on relatively small sample sizes and utilized cross-sectional study designs. Therefore, the second aim of the current study is to contribute to this body of literature, by exploring the level and development of flanker task performance as a predictor of the level and development of anxiety-, depression- related behavior and behavioral problems in a large sample of elementary school children followed throughout elementary school. This study is one of the first longitudinal investigations examining associated changes between behavioral indices of performance monitoring and behavioral and emotional problems throughout childhood.

A few studies found gender differences in flanker task performance in children (Gavin er al., 2019; Torpey, Hajcak, Kim, Kujawa, & Klein, 2012). When controlling for age, boys had a higher error rate and faster response time than girls (Gavin et al., 2019). Studies applying other performance monitoring tasks in young children (e.g., go/nogo tasks) found that gender was associated with response time (Torpey et al., 2012), where girls were slower than boys. In this study, we explore gender differences in flanker task performance, as there might be subtle differences in the initiation level and growth trajectory due to general sex differences in brain morphology and pre-puberal and hormonal changes (Gorday & Meyer, 2018; Ordaz et al., 2013).

Current Study

In the current longitudinal study, we modeled five years of flanker performance drawn from a large sample of mainstream elementary school children followed from age 7 to 12 with the intent to describe flanker performance throughout the elementary school years. This unique sample of repeated measures allows us to explore the development of response time across childhood. Based on previous research on performance monitoring, we expect that children show faster response time as they grow older. Second, we investigate the associated change between the level and development of response times with the level and development of anxiety and depressive related behavior and behavioral (ADHD-symptoms, oppositional defiant and conduct-related) problems. Lastly, we present age-related flanker performance measures (accuracy, congruency effects, PES and SAT's) in this large sample of 7- to 12-year-olds and explore if there are possible gender differences in flanker task performance during these five years as well as gender-specific associations between response times and problem behavior.

Method

Participants

The data were drawn from the project 'Happy Children, Happy adolescents?', a longitudinal elementary school-based study focusing on the interplay between the social-emotional, behavioral, (neuro)cognitive, and bio-psychological development of children in the

Netherlands (Asscheman et al., 2020; Behnsen, Buil, Koot, Huizink, & van Lier, 2018; Tieskens, Buil, Koot, & van Lier, 2021). Participants were recruited from mainstream elementary schools in urban areas in the central part of the Netherlands and rural areas in the Eastern region. For a detailed description of the inclusion of schools, see de Wilde, Koot, and van Lier (2016). Written informed consent from parents was requested for their child(ren) to participate. Each year, we informed the parents, children and teachers about the study and the upcoming data collection. Children, parents and schools were free to refrain from participating at any time of the study. The project and its procedures were approved by the Medical Ethics Committee of the Vrije Universiteit Medical Centre (protocol number NL37788.029.11).

The data used in this study originated from three consecutive age cohorts within the participating schools. Table 1 shows the data structure of the cohorts and descriptive information of the sample extracted from the larger study for data analysis. For cohort 1, data collection started in Grade 1 and continued until the end of elementary school. For cohort 2, data collection began in the second year of kindergarten (please note that the Netherlands has two years of kindergarten) until the end of elementary school. For cohort 3, data collection started in the first year of kindergarten and continued until the end of elementary school. Children's data was included when at least three time points or waves of flanker data were available and when data on age was complete. Half of the schools did not participate in the cognitive tasks in cohort 1 at timepoint 2 (the assessment year 2013) due to logistical complications. We restructured the data according to age so that all children of the same age were included in one measurement wave (i.e., age 7, age 8, etc.). The final sample contained a total of 1424 children (51% boys, nested in 25 schools). Most of the children had a Dutch ethnic background (73%) and came from medium to high social economic status (SES) household (91%).

Procedure

Data were collected annually, usually during spring or summer, and during one or two school days. Children completed a modified flanker task in the afternoon of the testing day, supervised by trained research assistants in a quiet place at the children's schools to ensure privacy and focus. Within the same month, teachers completed an online questionnaire on children's emotional and behavioral problem behavior.

Table 1

Available Sample Size, Age and Gender for Each Year and Cohort.

	Assessment year						
	2012	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**
Cohort 1							
Age mean (sd)	8.1 (.42)	9.2 (.42)	10.1 (.43)	11.1 (.44)	12.1 (.44)		
N	442	255	649	679	632		
Gender (% female)				46.3%			
Cohort 2							
Age mean (sd)	7.1 (.40)	8.1 (.40)	9.0 (.44)	10.0 (.44)	11.0 (.43)	12.0 (.41)	
N	163	258	367	373	363	336	
Gender (% female)				50.1%			
Cohort 3							
Age mean (sd)		7.0 (.38)	7.9 (.39)	8.9 (.38)	9.9 (.40)	10.9 (.38)	12.0 (.37)
N		211	273	309	318	295	284
Gender (% female)				53.5%			

Measures

Flanker Task

The source code for the flanker task can be found on https://osf.io/hmek5/ (van der Jagt & Stoof, 2023). An Eriksen flanker task was modified to accommodate the current sample of children and study design. The flanker task was performed on a tablet computer. The task had four practice items followed by 60 trials. Four different arrows strings (> > < > >, > > > > >, < < > < <, < < < < <) were presented randomly between 1 to 2 seconds on the tablet screen. Participants were instructed to press the left button with the left index finger if the central arrow pointed to the left and the right button with the right index finger if the central arrow pointed to the right. Participants were instructed to press as fast as possible on the correct side. We recorded accuracy and response time from the stimuli onset to button press for congruent (< < < < <, > > > > > 50%) and incongruent (< < > < <, > > < > > 50%) trials. Trials started with a 100 ms cue sign (+), where the central arrow of the string would appear. For cohorts 1 and 2 in 2012, the task contained 56 rather than 60 trials due to a technical error, which is the reason why we report proportion correct instead of number of correct trials as accuracy. A trial was considered invalid when there was no-response within a window of 2000 ms or an invalid range of response time, between 200 and 2000ms. Also, if children finished less than 40 trials, they were excluded from the analysis. This resulted in the exclusion of 33 participants when they were 7 years old, 27 participants that were 8 years old, 35 9-year-old participants, 1 11-year-old

and 2 12-year-olds. Tapping the left and right buttons together was considered an error. Participants who had at least 50% of the trials correct were considered for analyses (excluding 32, 18, 14, 4, 2, and 2 cases for year 7 to 12 respectively).

We summarize means and standard deviations for each trial type's response time and proportion correct (accuracy). In general, faster response times and high accuracy indicate better performance monitoring. Incorrect incongruent response times were subtracted from incorrect congruent trials is here referred to the response time congruency effect, based on a typical flanker effect (Luo & Proctor, 2022). Similarly, the number of incorrect incongruent trials were subtracted from number of incorrect congruent trials to reflect the congruency effect of accuracy. Both are considered behavioral measures of conflict monitoring. We derive post-error and post-correct response times from previous trial response and response times. For post-error slowing, we used the formula for traditional PES (PES_t), which is the reduction in response time following an error, quantified by the following formula: $mRT_{traditional}=mRT_{post\text{-}error} - mRT_{post\text{-}correct}$, where the mean response time (mRT) post-error is subtracted from post-correct (Schroder et al., 2020).

Next, to establish a measure of speed-accuracy trade-off (SAT), we calculated a balanced integration score (BIS; Liesefeld & Janczyk, 2019). The BIS (Liesefeld & Janczyk, Equation 4) includes the standardized response time (ʒRT) of correct trials and the proportion correct ($ʒ_{PC}$): BIS = $ʒ_{pc}$- $ʒ\overline{RT_{i,j}}$. We standardized the response time of correct trials and the proportion correct for each age group, incorporating any variance within this age group. A positive BIS indicates that performance was focused on accuracy rather than speed, in contrast to a negative BIS, which suggests that speed was favored over accuracy. Therefore, a value of 0 indicates that the individual equally valued the speed and choice during the task. We display response times, PES and BIS across ages in Table 2.

Emotional and Behavioral Problem Behavior

Emotional and behavioral problems were assessed using the Problem Behavior at School Interview, teacher report (short version; PBSI, Erasmus Medical Center, 2000; Van Lier et al., 2004). In this 30-item questionnaire, teachers rated items of behavioral and emotional problems on five scales: Attention-Deficit/Hyperactivity Disorder related behavior (AD), oppositional defiant behavior (ODB), conduct problems (CP), anxiety- (ANX) and depression related behavior (DEP). The teacher rated behavior on a 5-point Likert scale, from 0 (never applicable) to 4 (always applicable). AD behavior was measured through 5 items (e.g., the child cannot sit still, is hyperactive), where the Cronbach's alphas (α) for all years ranged between .93 to .94. ODB was covered by 6 items (e.g., child contradicts a lot; range α = .88 to .95). CP was measured through 8 items (e.g., child threatens other children; range α = .90 to .96). ANX was measured using 5 items (e.g., the child is scared at school; range α = .80 to .93) and DEP was measured through 6 items (e.g., child cries or is sad at school; range α = .77 to .88). Higher mean scores indicate more problems. Table 2 presents the means and standard deviation of the PBSI scales.

Gender

Gender was assessed via self-reports and coded as 0 = girl and 1 = boy.

Statistical Analysis

For descriptive statistics, IBM SPSS 28 was used. We used t-tests to examine gender differences for each age group. To investigate the development of flanker response time, we employed latent growth curve models, using the response time of all trial types (congruent and incongruent) as indicators of continuous latent growth factors in Mplus version 8.2 (Muthén & Muthén, 2018). In our latent growth curve models (LGM), the latent intercepts represent the initial mean response times at age 7, and the latent slope represents the mean rate of change across the ages. A positive slope indicates an increase in response time (slower response as children grow older), and a negative slope indicates a decrease in response time (faster responses as children grow older). A robust maximum likelihood estimator (MLR) was employed to deal with (potential) non-normal distributions of the measures. Full information maximum likelihood (FIML) was used to handle missing data. Standard errors were adjusted using a sandwich estimator to account for clustering of children in schools (Williams, 2000).

We fitted separate growth models for all constructs (flanker response time and the PBSI scales: ANX, DEP, ODB, AD, CP) to establish the shape of the growth curve: linear, quadratic, or cubic. Models were evaluated when they were most parsimonious and with three model fit indices: the Comparative Fit Index (CFI), Tucker Lewis Index (TLI, Bentler, & Bonnet, 1980) and the Root Mean Square Error of Approximation (RMSEA; Marsh, Hau, & Wen, 2004). Critical values for CFI ≥ .90, TLI ≥ .95, and RMSEA ≤ .06 were used to determine model fit (Hu & Bentler, 1995). We explored gender differences by estimating the paths of response time freely across gender and then comparing the model fit to the models where regression paths were equal for boys and girls. To identify which model (the freed or constrained model) best fit the data, we applied the Satorra-Bentler chi-square difference test for nested models (Satorra & Bentler, 2001).

Next, five parallel-process LGM's allowed for testing whether level and change in response time were associated with the level and change in behavioral and emotional problems. To this end, the growth parameters of the PBSI scales were regressed on the growth parameters of flanker response times. Possible gender differences in the association between response time and PBSI scales were tested by forcing the parameters to be free across gender and comparing the model to the model where the growth parameters were constrained (unless otherwise specified). A graphical representation of the parallel latent growth model of the flanker response time and problem behavior is visualized in Figure 1, found in the supplementary materials S2, section 1.

Results

Descriptive Statistics

In Table 2, we present the means and standard deviations of different behavioral measures of the flanker task and emotional and behavioral problem behavior between 7 and 12 years old and speed-accuracy trade-offs. See supplementary material S1, tab 1 for gender differences tests in all measures and supplementary material S1 tab 2 for a correlation matrix between the response time and PBSI scales (online available through https://osf.io/xqg86/). Overall, all children scored low on all emotional and behavioral problem scales during all assessment years, indicating low levels of problems. Also, the emotional and behavioral problem scores decreased linearly across the 5 years (supplementary materials S2, section 2). There were random significant correlations between response times and PBSI scales at different ages.

We performed LGM's of proportion correct, response times per trial type and PES to explore the change in these variables over time, since the descriptive information suggests improvement over time. However, all these models had very poor model fit indices and the variances of the slopes were not significant ($p > .112$), suggesting lack of growth (see supplementary materials S2, section 2). For each year, a typical congruency effect in accuracy and response time was observed: there were more errors made on incongruent trials and longer response times for incongruent trials. This means that the typical flanker effect is found at all ages in this sample. T-tests revealed several gender differences at different time points, they can all be viewed in supplementary material S1, tab 1 (online available through https://osf.io/xqg86/). Briefly, at age 8, 9, 11, and 12, the overall percentage correct was larger for girls than boys. At age 8, boys performed faster on all trial types than girls. At other ages, boys and girls performed similarly. Regarding post-error slowing, overall, children showed longer response time after error trials than after correct trials each year. We encountered large kurtosis (ranging from 0.31 to 17.25) for the speed-accuracy trade-off BIS, driven by a few participants who had poor performance (large range in proportion correct). This is why Table 1 shows the means as well as quartiles for BIS. Most participants in each year had a positive BIS, indicating that most children focused on accuracy rather than on speed. The BIS was mostly driven by the reduction of correct trial response time and high accuracy. Independent-sample Mann-Whitney U test revealed a gender difference in BIS only at 12 years of age, where boys had a larger BIS than girls.

Table 2

Means and Standard Deviations for Accuracy, Response Times, Post-Error Slowing, Speed-Accuracy Trade-Offs and PBSI Scales in 7- to 12-Year-Old Children.

	Age					
	7	**8**	**9**	**10**	**11**	**12**
	M (SD)	**M (SD)**	**M (SD)**	**M (SD)**	**M (SD)**	**M (SD)**
N	342	966	917	1336	1308	1250
Accuracy						
Proportion correct	0.93 (.09)	0.95 (.08)	0.96 (.07)	0.96 (.06)	0.96 (.07)	0.97 (.05)
Congruency effect	1.80 (3.74)	1.17 (3.19)	1.06 (3.54)	0.92 (3.02)	1.32 (4.17)	0.80 (2.53)
Response time (in seconds)						
All trial types	1.05 (0.18)	0.93 (0.18)	0.77 (0.14)	0.70 (0.13)	0.65 (0.13)	0.61 (0.11)
Congruency effect	0.06 (0.13)	0.06 (0.12)	0.06 (0.08)	0.05 (0.07)	0.05 (0.07)	0.04 (0.06)
Post-error slowing						
Traditional	0.27 (1.02)	0.29 (1.00)	0.05 (0.26)	0.05 (0.20)	0.04 (0.19)	0.03 (0.18)
Speed-Accuracy trade-off (Balanced integration score)						
Mean (SD)	0.002 (1.34)	0 (1.40)	0.001 (1.43)	-0.005 (1.41)	0.002 (1.49)	0.02 (1.38)
Lower quartile	-0.82	-0.57	-0.49	-0.48	-0.41	-0.48
Median	0.20	0.27	0.29	0.25	0.29	0.26
Higher quartile	1.02	0.94	0.94	0.85	0.87	0.85
Behavioral and Emotional problems: PBSI scores						
AD	1.02 (1.08)	0.91 (1.01)	0.90 (1.00)	0.84 (.95)	0.76 (.94)	0.78 (.93)
OBD	0.74 (.77)	0.69 (.78)	0.72 (.80)	0.66 (.76)	0.62 (.76)	0.65 (.75)
CP	0.42 (.59)	0.38 (.60)	0.38 (.58)	0.35 (.55)	0.30 (.52)	0.34 (.55)
ANX	0.91 (.74)	0.84 (.68)	0.83 (.70)	0.79 (.68)	0.76 (.69)	0.75 (.69)
DEP	0.86 (.71)	0.74 (.69)	0.69 (.67)	0.66 (.64)	0.63 (.64)	0.68 (.67)

Note. AD = Attention-Deficit/Hyperactivity Disorder related behavior; OBD = oppositional defiant behavior; CP = conduct problems; ANX = anxiety; DEP = depression symptoms.

Longitudinal Trajectory of Flanker Response Time

Table 3 presents the model fit indices and model comparisons testing for gender differences in response times. For flanker response time, a quadratic model improved incremental fit over a linear model and was the most parsimonious model over a cubic model. The quadratic model had good fit measures (Table 3) where overall, children responded a little over one second at age 7 on flanker trials (Mean Intercept $\beta = 1.09$, $SE = .019$, $p < 0.001$) and response time decreased over time (Mean Slope $\beta_{linear} = -0.18$, $SE = .011$, $p < 0.001$; Mean Slope $\beta_{quadratric} = -.02$, $SE = .002$, $p < 0.001$). To allow for a better interpretation of the quadratic slope, we segmented the quadratic slopes into two linear slopes (piecewise analyses), which was ultimately more parsimonious over the quadratic slope. Piecewise growth models showed that children's response times decreased at a faster rate between ages 7 and 9 compared to ages 10 and 12 (Mean Slope age 7 to 9 years old $\beta_1 = -.17$, $SE = .01$, $p < 0.001$, Mean Slope age 10 to 12 years old $\beta_2 = -.05$, $SE = .002$, $p < 0.001$). Next, the role of gender in the piecewise flanker response time model was investigated (Table 3). Freeing the intercept across gender significantly improved the model compared to constraining them to be equal for gender (Figure 1), indicating that boys had faster response time at age 7 ($M = 1.04$, $SD = 0.19$) than girls ($M = 1.07$, $SD = 0.16$). Yet, both slopes were similar for boys and girls, suggesting that the rate of decline in response times over time was similar across gender.

Figure 1

Response Time (in Seconds) of All Trail Types From Age 7 to 12-Years.

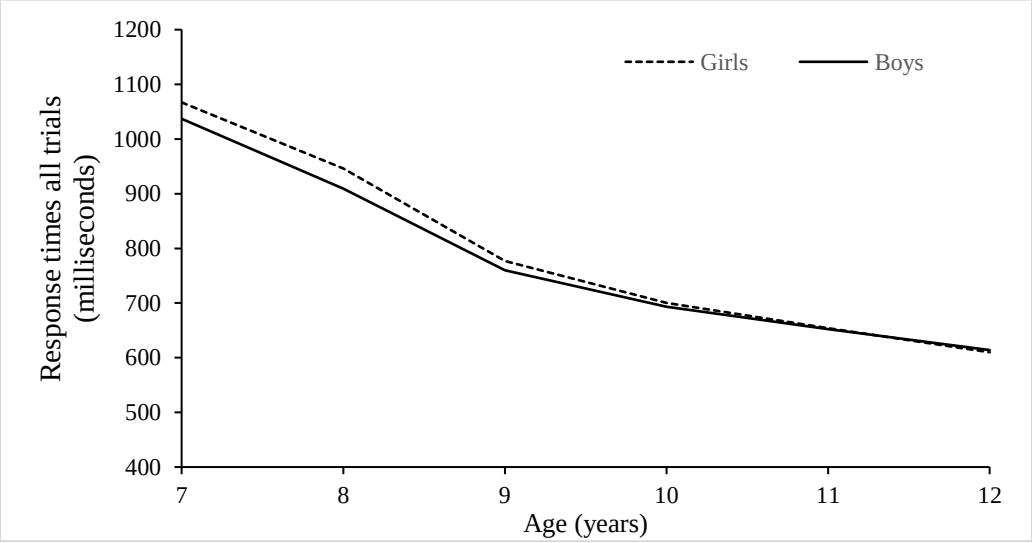


Table 3

Model Fit and Model Comparison for Response Time, Gender Included.

Model	Fit					Difference Tests			
	X^2	*df*	CFI	TFI	RMSEA	Model	ΔX^2	Δ*df*	*p*
1a. Linear model	491.57	16	.58	.61	.15				
1b. Quadratic model	66.01	12	.95	.94	.06	1a vs. 1b	314.62	4	<0.001
1c. Piecewise model	57.86	12	.96	.95	.05				
Gender differences in piecewise model									
2a. Gender free	80.11	24	.96	.95	.06				
2b. Intercept equal	84.75	25	.96	.95	.06	2a vs. 2b	4.23	1	0.04
2c. Slope 1 equal	81.23	25	.96	.95	.06	2a vs. 2c	1.55	1	0.21
2d. Slope 2 equal	81.58	25	.96	.95	.06	2a vs. 2d	2.44	1	0.12

Note. CFI = comparative fit index; TFI = Tucker Lewis Index; RMSEA = root mean square error of approximation.

Associations of Flanker Response time and Problem Behavior.

Model fit and model comparisons testing for gender differences in behavioral and emotional problems initial level (intercept) and development (slope) are shown in Table 2 in the supplementary materials S2, section 3. The regression coefficients for the associations between the growth parameters of response time and the behavioral and emotional problem scales are presented in Table 4. The figures in supplementary materials S2, section 4 are graphical representations of the association models. Note that parameter constraints based on gender differences testing level and growth rate in unconditional models for all constructs were retained in the parallel process LGMs.

For anxiety symptoms, there was a positive association between the initial level of response time at age 7 and the slope of anxiety (*B* = .20, *SE* = .07, *p* = .03, 95% CI of *B* = .03 - .31, β =.25). This suggests that a slower response time at age 7 was associated with a steeper decrease in anxiety scores from age 7 to 12, regardless of gender.

For depression symptoms, there was a positive association between slope of response time at age 7 to 9 years old and the slope of depression scores (*B*= .51, *SE* = .23, *p* = 0.02, CI of *B* = .07 - .95, β =.29). This indicates that a faster decrease in response time during age 7 to 9 years predicted a faster decrease in depression scores between 7 and 12 years for boys and girls.

For ADHD symptoms, the association between the slope of response time at age 7 to 9 and the slope of ADHD was significant and positive for boys (*B* = .63, *SE* = .24, *p* = .01, 95% CI of

B = .16 - 1.11, β =.39), but not significant for girls (*p* = .27). Similarly, the association between the slope of response time and the slope of ADHD between 10 and 12 years old was significant and positive for boys (*B* = 1.44, *SE* = .71, *p* = .04, 95% CI of *B* = .05 - 2.834, β =.28), but not significant for girls (*p* = .49). In other words, for boys a faster decrease in flanker response time was associated with a faster decrease in ADHD-Symptoms.

For OBD symptoms, there was a positive association between the intercept of response time for boys and the slope of OBD scores (*B* = .17, *SE* = .08 *p* = 0.04, 95% CI of *B* = .01 - .33, β =.30), but not for girls (*p* =.46). This suggests that for boys, a slower response time predicted a steeper decline of OBD scores between 7 to 12 years. Also, there was a positive association between the slope of response time between 7 and 9 years and the slope of ODB in boys (*B* = .67, *SE* = .17, *p* < 0.001, 95% CI of *B* = .34 - 1.00, β =.54), but not in girls (*p* = .85). In other words, the faster decline of response time during 7- to 9-year-olds predicted a faster decline in overall OBD scores for boys.

For conduct problems, none of the parameters were significantly linked to each other (table 4). This means that the initiation level and slopes of response time did not predict the level and development of conduct problem symptoms.

Table 4

Association Estimates of Parallel-LGM Models: RT and Behavioral and Emotional Problems.

	RT											
	Intercept				**Slope 1 (age 7 to 9)**				**Slope 2 (age 10 to 12)**			
	Boys		**Girls**		**Boys**		**Girls**		**Boys**		**Girls**	
	B	**SE**	**B**	**SE**	**B**	**SE**	**B**	**SE**	**B**	**SE**	**B**	**SE**
AD												
Intercept	.34	.38	.34	.38								
Slope	.13	.10	.16	.14	.63*	.24	.75	.68	1.44*	.71	-1.45	2.10
OBD												
Intercept	.25	.38	.25	.33								
Slope	.17*	.08	.05	.07	.67**	.17	.12	.64	1.00	.63	.13	2.00
CP												
Intercept	.33	.37	.28	.21								
Slope	-.03	.09	-03	.12	.21	.17	-.63	1.07	-.06	.35	2.33	3.95
ANX												
Intercept	.31	.34	.31	.34								
Slope	.20*	.07	.20*	.07	.65	.35	.65	.35	.53	1.34	.53	1.34
DEP												
Intercept	.74	.39	.16	.31								
Slope	.11	.08	.11	.08	.51*	.23	.51*	.23	.17	.99	.17	.99

Note. *p < .05; **p < .001; RT = response time; AD = Attention-Deficit/Hyperactivity Disorder related behavior; OBD = oppositional defiant behavior; CP = conduct problems; ANX = anxiety; DEP = depression symptoms.

Discussion

The objective of this study was to explore the development of flanker performance and its association with the development of behavioral and emotional problems in children attending mainstream elementary schools. The performance monitoring in children improved over time, evidenced by a decrease in response time on all trial types of a flanker task, where the decrease in response time was most noticeable during age 7 to 9 compared to age 10 to 12. Boys initially had lower response times than girls at age 7, but the pattern of response time development was similar for boys and girls. We found associations between initial level response times and anxiety and opposition deviant behavior. We also found associations between the slope of response time between 7 to 9 and slope of problems relating to ADHD, ODB and depression. Finally, there was an association between the slope of response time between 10 and 12 years and the slope of ADHD.

Although this was not a key aim of this study, this study provides a description of flanker performance measures across childhood. Even when we could not model the shape of growth of the other behavioral indices of the flanker performance, we found that the absolute values at each assessment year suggest improved accuracy and a stable speed-accuracy trade-off. These results, together with the reduction of congruency effects in accuracy and response time, illustrate improved conflict monitoring across age 7 to 12 years old. Post-error slowing (PES) was larger at age 7 and 8, indicating larger response times after error trials. Yet as the children grew older, children had smaller response times after error-making, suggesting improved error monitoring.

The decline in response time and improved accuracy of the children, is in line with other performance monitoring studies across childhood (Davies et al., 2004; Gavin et al., 2019; Overbye et al., 2019). Furthermore, our results fit the general idea of improved cognitive control (Luna et al., 2015) and the diffusion-drift model (DDM) often observed in cognitive force choice paradigms (Dutilh, Forstmann, Vandekerckhove, & Wagenmakers, 2013; Hall, Schreiber, Allen, & Hallquist, 2021; Ratcliff & McKoon, 2008). Briefly, the DDM allows for the translation of task behavior into the components of cognitive processing. In other words: the DDM represents the response time, accuracy and distribution of response times during the decision-making process of two-choice tasks, incorporating the processing of stimuli, noise, and the accumulation of information (Ratcliff & McKoon, 2008). The non-linear trend of response time can be explained by the typical brain maturation during this age group and into adolescence (e.g., Fuhrmann, Madsen, Johansen, Baaré, & Kievit, 2022; Ordaz et al., 2013) and by the development of brain areas (such as the ACC) involved in performance monitoring observed through electrophysiological reports (Boen, Quintana, Ladouceur, & Tamnes, 2022; Davies et al., 2004; Lo, 2018; Overbye et al., 2019). The observed congruency effect (response time and accuracy) across the ages in the current sample corresponds with studies that previously investigated conflict adaption in children (Larson, Clawson, Clayson, & South, 2012; Liu et al., 2018; Mullane, Corkum, Klein, & McLaughlin, 2009; van Meel, Heslenfeld,

Rommelse, Oosterlaan, & Sergeant, 2012). For instance, the response times and accuracy rates of the flanker task performed by the 10-year-olds sample in Lui et al. (2018) and the samples of 6 to 9 years and 10 to 12 years in Van Meel et al. (2012) show similar values as observed in the current study. Considering the behavioral measure of error monitoring, the change in PES around 9 years (and the stable BIS) suggests improved error monitoring at a behavioral level. Although there are various explanations for PES (Danielmeier & Ullsperger, 2011; Dutilh, et al. 2012), it is likely that children in this sample improved at processing the errors (reduction in error response times), improved focus on task performance and increased in error caution. Furthermore, this study identified gender differences in flanker performance across childhood, similarly to Gavin et al. (2019). There is no consensus on the significance of gender differences in performance monitoring tasks in children, however in adults it has been suggested that females are more prone to be distracted by the congruency effect, which is reflected in the PES (Fischer et al., 2016). Together, the flanker behavioral results suggest that a faster rate of performance monitoring and adaptive control development occurs before the age of 9, and as children aged, they became more stable in their ability to monitor their actions and adapt behavior to circumstances.

We have observed several associations between the trajectory of response time and scores on four of the five problem behavior scales. The results indicate that if a child has lower response times at age 7 (that is: better task performance compared to a child with higher response times), then the child shows a steeper decline in anxiety and oppositional defiant behavior scores. Thus, faster task response times in younger children can be predictive of the reduction in problem behavior during childhood. Also, a steeper decline in response time across 7 to 12 years was associated with a steeper decline in ADHD related behavior, particularly for boys. In addition, early (age 7 to 9) fast decline of response time was predictive of a faster decline in depressive (both genders) and behavioral problems (ADHD and ODB for boys only), suggesting that the rate of improved task performance is linked to fast decline in problem behavior. It is noteworthy to mention that several associations were not gender specific, suggesting that the link between task performance and depressive and anxiety problems in children is equally important for both genders during childhood. The lack of gender differences in this link is informative for the use of cognitive control measures when investigating psychopathology. To our knowledge, there are no studies that investigated changes in task performance as a predictor of change of problem symptoms during childhood. However, differences in task behavior between children with and without anxiety and ADHD-related behavior have been frequently reported (e.g., Mullane et al., 2009).

The current study has several limitations. First, a possible limitation regards the current sample, which included primarily Caucasian children from mainstream elementary schools. The included schools were a convenience sample, introducing a possible bias not including unrepresented samples. Also, we do not have records on whether children in these schools had any diagnosed learning or psychological disorders. It is possible that children diagnosed with psychological problems, of which it is known that performance monitoring can

be affected, are included in this sample. Second, the current study used the PBSI, a teacher informed measure on problem behavior observed in the school context. Although scores PBSI could be an accurate substrate of daily maladaptive behavior (especially for ADHD related behavior in schools) and a good substitute for behavioral reports for young children (who might have a reduced ability to self-reflect on their behavior), adding self-report or multi-informant measures can give a more complete few of psychological problems. Last, while there were associations between response time and problem behavior, we cannot draw any conclusions on the directionality of these associations. We based the directionality of the effects on theoretical assumptions, but bidirectional or reverse directions of the associations are possible. Also, the current study investigated performance monitoring on a behavioral level. With the addition of electrophysiological measures, examining the underlying brain processing of the conflict and error monitoring, insight can be given in the complex relation between task behavior and the processes that drive daily maladaptive functioning.

The results of the current study help to pinpoint several future research avenues. Task-related variables of influence, such as response-to-stimulus intervals (Smulders et al., 2016) or trial type expectancy (Gratton et al., 1992) were not investigated here. Also, individual behavioral change or conflict adaptation within the task (between trials), task difficulty and the order of trial types (which determine the level of accumulated conflict) are known to influence the performance in subsequent trials (Larson, Clayson, & Clawson, 2014; Lui et al., 2018). Investigating these behavioral adaptation phenomena over time in children can inform research on the development of cognitive control abilities. With the current knowledge we have on performance monitoring through behavioral studies like the current study and other neurophysiological studies, it is time to investigate how individual differences in trial-to-trial variability in task performance are related to brain changes (e.g., Tamnes et al., 2012) and symptomatology (e.g., Clayson et al., 2022) in longitudinal designs.

Conclusion

Our study is a unique investigation of the developmental trajectory of flanker performance and has explored the associations with problem behavior in elementary school children. Improvement of flanker performance over time was observed, which illustrates the normative development of performance monitoring in 7-to 12-year-old children. We found that flanker performance is associated with problems in anxiety, depression, and ADHD- and oppositional-related behavior during childhood and that there are specific gender differences in these associations. The current study illustrates the significance of behavioral indices of cognitive control development and pinpoints an important link with psychological problems in children.

Supplementary Materials 1:

Available only: https://osf.io/xqg86/.

Supplementary Materials 2:

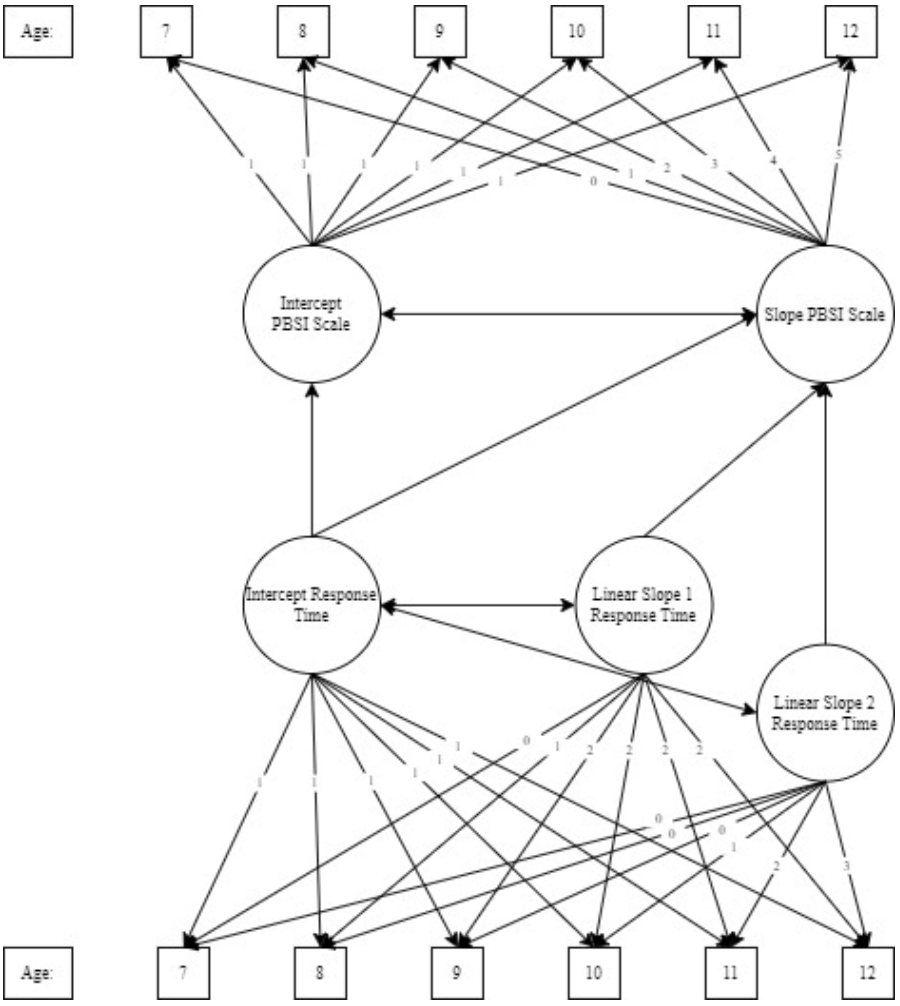
Contents:

Section 1: Graphical representation of parallel latent growth model of response time and emotional or behavioral problems.
Section 2: Model fit for accuracy, response times per trial type and PES.
Section 3: Model fit and model comparisons testing for gender differences of emotional and behavioral problems.
Section 4: Simplified graphical representation of the associations between flanker response times and emotional and behavioral problems.

Section 1

Figure S1

Graphical representation of parallel latent growth model of response time and emotional or behavioral problems.



Section 2:

Table S1.

Model fit for accuracy, response times per trial type and PES.

		Fit					*P*-values for	
Measure	**Model**	X^2	***df***	**CFI**	**TFI**	**RMSEA**	**Variance in Intercept**	**Variance in slope**
Accuracy	Linear	28.52	16	.80	.82	.02	.12	.28
Response time Incongruent trials	Linear	395.33	16	.64	.67	.13	< 0.00	.30
Response time Congruent trials	Linear	450.62	16	.56	.53	.14	< 0.00	.80
Response time Correct trials	Linear	538.98	16	.58	.60	.16	< 0.00	.93
Response time Error trials	Linear	222.17	16	0	-.43	.10	0.03	.65
PES	Linear	45.81	16	0	1.00	.04	.89	.71

Note. CFI = comparative fit index; TFI = Tucker Lewis Index; RMSEA = root mean square error of approximation.

Section 3:

Table S2.

Model fit and model comparisons testing for gender differences of emotional and behavioral problems.

Model	Fit					Difference Tests			
Anxiety	**X^2**	**df**	**CFI**	**TFI**	**RMSEA**	**Model**	**ΔX^2**	**Δdf**	***p***
1a. Linear model	12.72	16	1.00	1.00	0	1a vs. 1b	1.44	4	.83
1b. Quadratic model	12.44	12	1.00	1.00	.01				
Gender differences									
2a. Gender free	28.27	32	1.00	1.00	0				
2b. Intercept equal	28.32	33	1.00	1.00	0	2a vs. 2b	.04	1	.84
2c. Slope equal	28.37	33	1.00	1.00	0	2a vs. 2c	.02	1	.89
Anxiety on Response time models (gender included)									
3a. Constrained model	192.01	128	.97	.97	.03				
3b. Intercept anxiety and intercept RT free	192.20	127	.97	.97	.03	3a vs 3b.	.23	1	.63
3c. Slope anxiety and intercept RT free	192.12	127	.97	.97	.03	3a vs. 3c	.08	1	.72
3d. Slope anxiety and slope s1 RT free	192.44	127	.97	.97	.03	3a vs. 3d	.06	1	.80
3e. Slope anxiety and slope s2 RT free	193.32	127	.97	.97	.03	3a vs. 3e	.02	1	.89
Depression									
1a. Linear model	19.04	16	1.00	1.00	.01				
1b. Quadratic model	14.31	12	1.00	1.00	.01	1a vs. 1b	4.75	4	.31
Gender differences									
2a. Gender free	40.22	32	.99	.99	.02				
2b. Intercept equal	48.50	33	.98	.99	.03	2a vs. 2b	12.69	1	.0004
2c. Slope equal	41.71	33	.99	.99	.02	2a vs. 2c	1.82	1	.18
Table 1 continued			**Fit**				**Difference Tests**		
	X^2	**df**	**CFI**	**TFI**	**RMSEA**	**Model**	**ΔX^2**	**Δdf**	***p***
Depression on Response time models (gender included)									
3a. Constrained model	223.63	127	.96	.96	.03				
3b. Intercept depression and intercept RT free	217.21	126	.96	.96	.03	3a vs 3b.	793.39	1	<0.001
3c. Slope depression and intercept RT free	220.15	126	.96	.96	.03	3a vs. 3c	6.88	1	.01
3d. Slope depression and slope s1 RT free	221.85	126	.96	.96	.03	3a vs. 3d	1.78	1	.18

3e. Slope depression and slope s2 RT free	214.77	126	.96	.96	.03	3a vs. 3e	0.81§	1§	.37§
ADHD									
1a. Linear model	16.58	16	1.00	1.00	.01				
1b. Quadratic model	14.22	12	1.00	1.00	.01	1a vs. 1b	3.32	4	.51
Gender differences									
2a. Gender free	39.09	32	1.00	1.00	.02				
2b. Intercept equal	104.82	33	.95	.95	.06	2a vs. 2b	205.66	1	<0.001
2c. Slope equal	43.84	33	.99	.99	.02	2a vs. 2b	4.26	1	0.04
ADHD on Response time models gender included									
3a. Free model	205.09	122	.97	.97	.03				
3b. Intercept ADHD and intercept RT equal	207.97	123	.97	.97	.03	3a vs 3b.	3.87	1	< 0.05
3c. Slope ADHD and intercept RT equal	206.98	123	.97	.97	.03	3a vs. 3c	2.33	1	.13
3d. Slope ADHD and slope s1 RT equal	205.73	123	.97	.97	.03	3a vs. 3d	.02	1	.89
Table 1 continued			**Fit**				**Difference Tests**		
	X^2	**df**	**CFI**	**TFI**	**RMSEA**	**Model**	**ΔX^2**	**Δdf**	***p***
3e. Slope ADHD and slope s2 RT equal	212.02	123	.97	.97	.03	3a vs. 3e	1.30§	1§	0.25§
CP									
1a. Linear model	34.33	16	0.99	0.99	0.03				
1b. Quadratic model	34.35	12	0.99	0.99	0.04	1a vs. 1b	5.31	4	0.26
Gender differences									
2a. Gender free	60.50	32	0.98	0.98	0.04				
2b. Intercept equal	96.98	33	0.95	0.95	0.05	2a vs. 2b	27.34	1	<0.001
2c. Slope equal	62.65	33	0.98	0.98	0.04	2a vs. 2c	2.42	1	0.12
CP on Response time models gender included									
3a. Free model	235.69	123	.96	.96	.04				
3b. Intercept CP and intercept RT equal	236.14	124	.96	.96	.04	3a vs 3b.	0.01	1	.91
3c. Slope CP and intercept RT equal	235.13	124	.96	.95	.04	3a vs. 3c	.16§	1§	.69§
3d. Slope CP and slope s1 RT equal	239.12	124	.96	.95	.04	3a vs. 3d	1.53§	1§	.22§
3e. Slope CP and slope s2 RT equal	237.62	124	.96	.95	.04	3a vs. 3e	1.95	1	.16
ODD									
1a. Linear model	25.07	16	1.00	1.00	0.02				
1b. Quadratic model	25.74	12	0.99	0.99	0.03	1a vs. 1b	3.04	4	0.55

Gender differences									
2a. Gender free	46.81	32	0.99	0.99	0.03				
2b. Intercept equal	72.89	33	0.97	0.97	0.04	2a vs. 2b	51.28	1	<0.001

Table 1 continued	**Fit**					**Difference Tests**			
	X^2	**df**	**CFI**	**TFI**	**RMSEA**	**Model**	**ΔX^2**	**Δdf**	***p***
2c. Slope equal	47.31	33	.99	.99	.03	2a vs. 2c	.29	1	.59
ODD on Response time models gender included									
3a. Free model	213.24	123	.97	.97	.03				
3b. Intercept ODD and intercept RT equal	212.43	124	.97	.97	.03	3a vs 3b.	.82[§]	1[§]	.37[§]
3c. Slope ODD and intercept RT equal	228.79	124	.97	.97	.03	3a vs. 3c	.73[§]	1[§]	.39[§]
3d. Slope ODD and slope s1 RT equal	213.33	124	.97	.97	.03	3a vs. 3d	.66	1	.42
3e. Slope ODD and slope s2 RT equal	211.17	124	.97	.97	.03	3a vs. 3e	0.02	1	.89

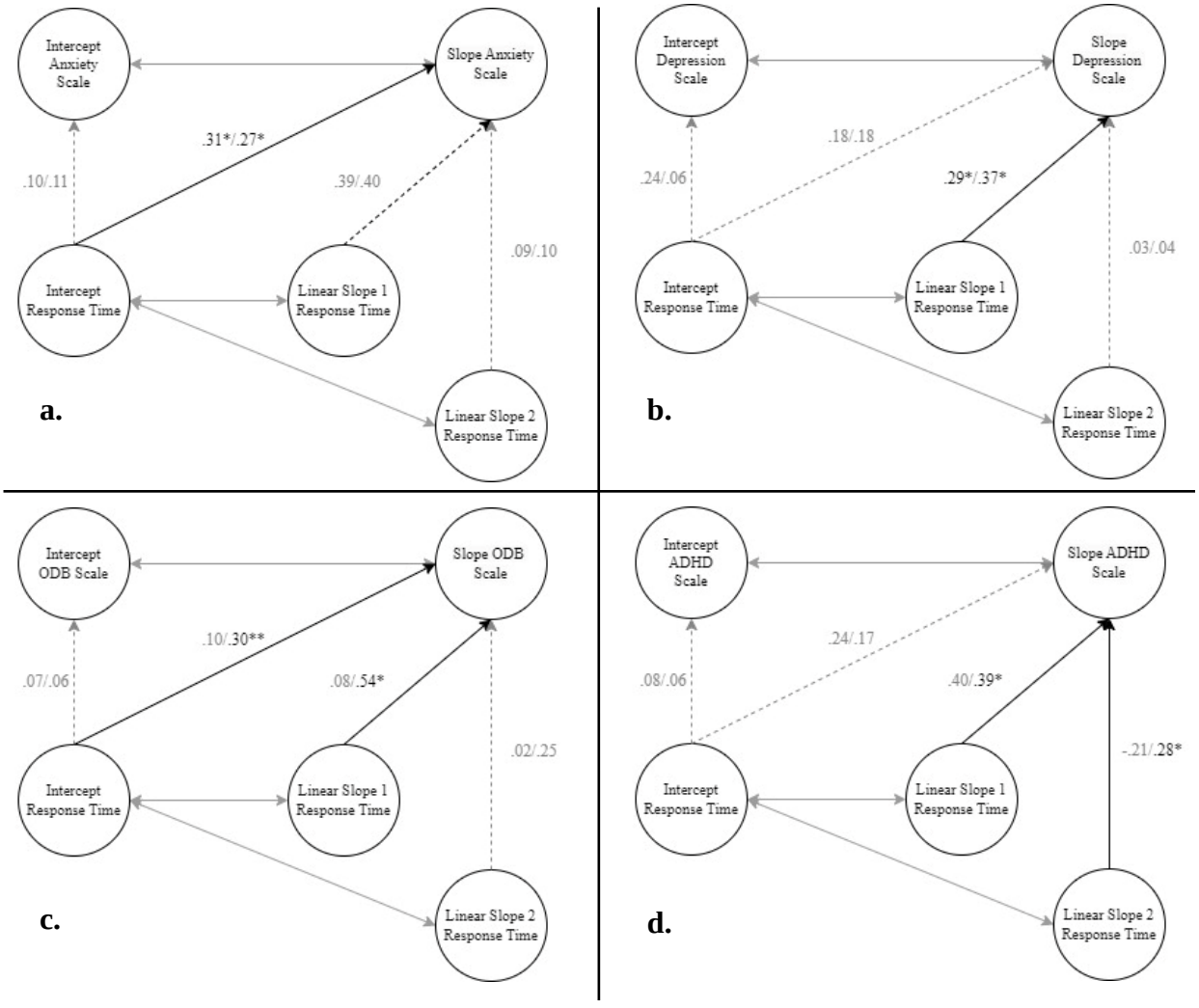
Note. CFI = comparative fit index; TFI = Tucker Lewis Index; RMSEA = root mean square error of approximation.[§] We observed a negative Chi-square value for the Satorra Bentler test. To test if the path was significantly different for boys and girls, we employed the Wald chi-square test of parameter equalities instead. AD = Attention-Deficit/Hyperactivity Disorder related behavior; OBD = oppositional defiant behavior; CP = conduct problems; ANX = anxiety; DEP = depression symptoms.

Section 4

Figures S2

Simplified graphical representation of the associations between flanker response times and

a. PBSI anxiety score.
b. PBSI depression score.
c. PBSI oppositional defiant behavior score.
d. PBSI ADHD score.



Note. Path estimates are standardized regression coefficients. Solid arrow = significant at * p < 0.05 or ** p < 0.001, dashed arrow = insignificant. Gender-specific associations were represented as girls/boys (e.g. .31*/ .27*).

Chapter 5

Investigating the Associations Between Infant Behavioral Inhibition, Social Anxiety, and Social Error Processing during Adulthood

This chapter is based on:

Lutz, M. C., Lakhlani, D., Smith, A. S., Tang, A., Guyer, A., Kok, R., Franken, I. H. A., Fox, N. A., Pine, D. S, & Harrewijn, A. Effects of Infancy Behavioral Inhibition on Social Processing and Risk for Social Anxiety during Adulthood.

Abstract

High infant Behavioral Inhibition (BI) increases risk for social difficulties and anxiety, suggesting long-term consequences for adult socioemotional development. Infant BI and social anxiety also affect error processing. Whether this effect extends to social contexts remains unclear. In this prospective, 30-year longitudinal study, we investigated the long-term association between infant BI and self-reported social anxiety, and the association between infant BI and social anxiety on adult brain and behavioral measures of social error processing. To examine the role of social observation in error processing, participants performed a social flanker task during magnetic resonance imaging. We expected pregenual anterior cingulate activity to be associated with social error processing in participants who scored high on social anxiety and infant temperament. The final sample included data from 24 participants (50% male, mean age 29.50 years). Infant BI was not related to self-reported adult social anxiety scores. We found an interaction effect of temperament and social anxiety on activity in the middle cingulate cortex, temporal and precentral gyrus. We also found a main effect of social vs. alone condition of the social flanker task on activity in middle cingulate cortex. For all regions, activity was stronger during the processing of correct vs. error trials in the social condition. The results suggest that the brain devotes more performance monitoring processing to facilitate better performance during social observation. Despite the small sample size due to attrition over 30 years, we argue that the social flanker task has the potential to investigate social error processing in individuals.

Investigating the Associations Between Infant Behavioral Inhibition, Social Anxiety and Social Error Processing during Adulthood

The temperamental trait behavioral inhibition (BI) is characterized by fearful, shy, timid, or withdrawn behavior in unfamiliar situations early in life (Dilalla et al., 1994; Fox et al., 2001). High observed BI during infancy increases the risk for social difficulties and behavior, such as problems during peer interactions and social withdrawal during childhood (Buzzell et al., 2021; Pérez-Edgar et al., 2011; Sandstrom et al., 2020). Also, there is a relationship between infant BI and difficulties in social interactions during adolescence (Buss et al., 2021; Chronis-Tuscano et al., 2009; McDermott et al., 2009; Nozadi et al., 2018). In addition, approximately 40% of children with BI develop anxiety disorders (Gladstone et al., 2005; Muris et al., 2011; Rapee et al., 2010) and social anxiety in particular (Clauss & Blackford, 2012), making it vital to study who is at risk. Several studies have addressed the potential consequences of infant BI on socioemotional development in adulthood, where and increased risk of internalizing problems was related to high infant BI (Asendorpf et al., 2008; Caspi, 1996; Caspi et al., 2003). In addition, infant BI predicts poorer social and cognitive functioning in adults (Tang et al. 2020). The social and anxiety-related problems that are driven by infant BI suggest that underlying neurocognitive processes may be affected. Therefore, the current study aims to investigate the long-term association between infant BI on self-reported social anxiety, as well as the association between infant BI and social anxiety on brain and behavioral measures of social information processing in adulthood.

Infant BI, Anxiety, and Error Processing

Social anxiety can be defined as heightened distress and fear in response to social evaluation (Rapee & Heimberg, 1997). Social contexts trigger the processing of social and evaluative information, requiring increased vigilance during social interaction. The processing of this information requires performance monitoring, a capacity that is driven by the neural system that is responsible for salience detection presented in social contexts (Henderson et al., 2015). An important component performance monitoring involves the detection and processing of errors, which is measured by the event-related potential error-related negativity (ERN) (Falkenstein et al., 1991; Gehring et al., 1993), or can be observed by the activation of the anterior cingulate cortex (ACC) in the brain (Ridderinkhof et al., 2004; Van Veen & Carter, 2002b). In children and adults, an enhanced neural response to error monitoring is related to anxiety (Meyer, 2017; Moser et al., 2013). In addition, error monitoring appears to be an important moderator in the BI and anxiety relationship (Lahat et al., 2014; Tang et al., 2020; Fox et al., 2023), where a positive association between infant BI and internalizing psychopathology was found in those who had an enhanced ERN. Error processing might be an underlying neurocognitive process that is influenced by BI and anxiety, which in turn, is related to socioemotional development during adulthood. However, this has not yet been examined. Therefore, in the current study, we test the association between infant BI and current social anxiety on error processing in adults.

Social Error Processing in the Brain

Social factors appear to influence error processing in adults, such that the presence of peers increases the pressure to perform well (Geen, 1991; Geen & Bushman, 1989; Zajonc, 1965). Theoretically, biopsychosocial models of social motivation explain how someone, in the presence of others, would be determined to reduce error-making to avoid the risk of social evaluation. Social evaluation is threatening for an individual as this could lead to social exclusion (Baumeister & Leary, 1995) and related consequences such as reduced self-esteem and increase in experience negative emotions (Blackhart et al., 2009). To avoid the risk of social evaluation and its consequences, there is an increase in arousal and fear in the individual during performance monitoring. To measure the effect of the social context during performance monitoring, the social flanker task (Barker et al., 2015; Buzzell et al., 2017; Smith et al., 2019) was developed. It is a modified Eriksen flanker task (Eriksen & Eriksen, 1974) which stimulates the 'social pressure effect' during error monitoring as the participant is instructed that they are being observed by someone during their task performance. In the social flanker task, there is a larger ERN in the condition where the participants thought they were monitored and evaluated during task performance in children (Kim et al., 2005) and adults (Barker et al., 2015). Making errors while they thought they were observed by peers elicited activation in the pregenual part of the anterior cingulate cortex in children with anxiety disorders (pgACC; Smith et al., 2019). On a mechanistic level, this could be interpreted that errors made in a social condition of the social flanker task were more distressing and threatening than in the 'alone' condition.

So far, the social flanker task has not yet been performed in adults to investigate the activation of social error processing by the brain. Also, given that infant BI and social anxiety are related to error processing, it is possible that infant BI and social anxiety are also associated with social error processing. This is because infant BI and anxiety are both associated with a bias to quickly and preferentially process information that is motivationally salient (the social condition). When this bias becomes strong and persistent across development, the fear of performing well in social contexts increases (Henderson et al., 2015). Given this hypothesis, we test if infant BI and self-reported social anxiety are associated with the social error processing in the pgACC during adulthood.

The present study

The aims for the current study are threefold. First, we aim to test whether infant BI is related to self-reported social anxiety in adulthood. Second, we explore the social error processing during brain imaging (MRI) by applying the social flanker task. Third, we aim to test whether infant BI and adult social anxiety are both related to social error processing in adulthood. We preregistered the following hypotheses for this study (Lutz et al., 2022). First, we hypothesized that participants with higher infant BI report more social anxiety symptoms during adulthood (Sandstorm et al., 2020). For the second hypothesis, we expected standard

flanker task effects on a behavioral level (congruency effect; Botvinick et al., 2001) and neural level (activation in the ACC and other performance monitoring-related areas, reported in Smith et al., 2019). We explored possible effects of condition on task performance and whether social anxiety is related to the social flanker task performance. Third, we expected a significant interaction between infant BI and social flanker condition (social vs. alone) on the activation of the pgACC during the processing of errors in incongruent trials. We expected to observe more activation in the pgACC in the social condition than in the alone condition for participants that score high than those who score low on infant BI. Fourth, we hypothesized a significant interaction between current social anxiety symptoms and social flanker condition (social vs. alone) on the pgACC activation during the processing of errors in incongruent trials. We expected to observe more activation in the pgACC in the social condition than in the alone condition for participants that score high vs. low on social anxiety symptoms. Finally, we expected to observe a significant three-way interaction between infant BI, current social anxiety in adulthood, and social flanker condition (social vs. alone) on the activation of the pgACC during the processing of errors in incongruent trials. We expected to observe more activation in the pgACC in participants with both high BI and increased social anxiety symptoms as opposed to participants who score low on both, in the social condition vs. the alone condition during the processing of errors in incongruent trials.

Method

Sample

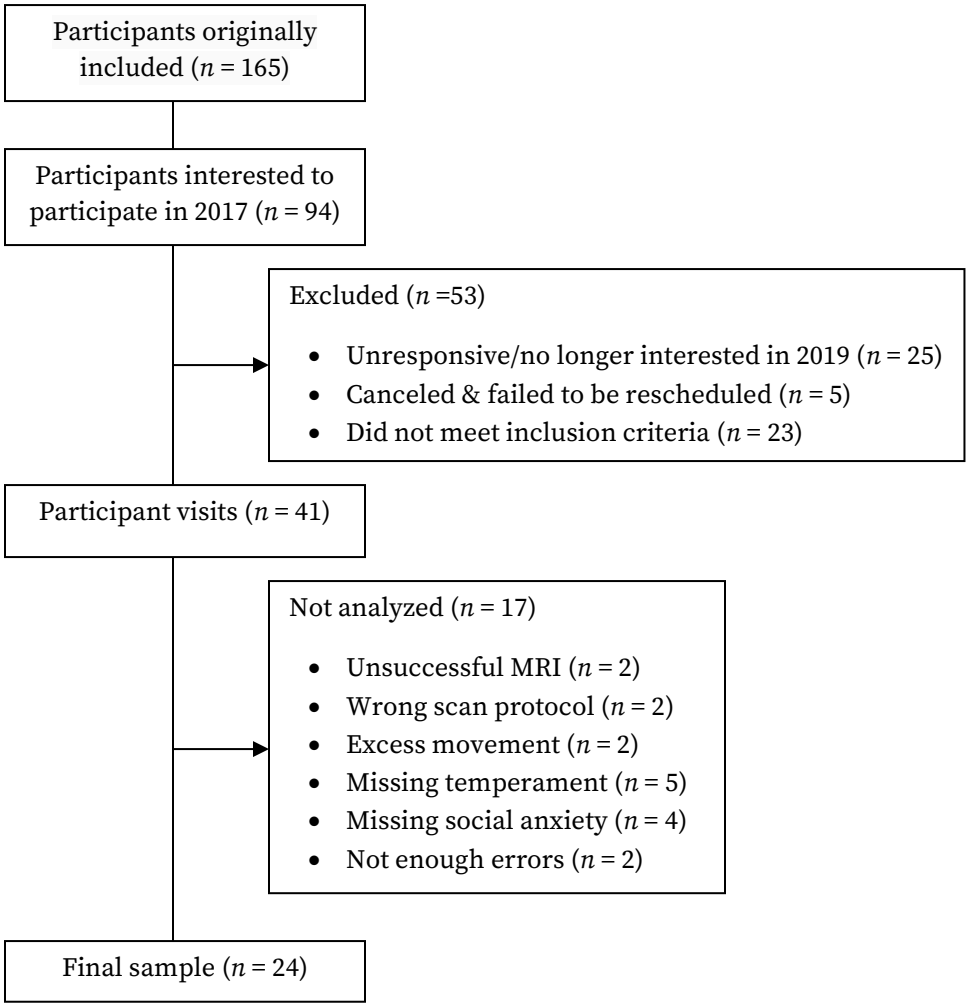
The current study is part of a prospective longitudinal study to investigate the long-term effects of infant temperament on socio-emotional development (Fox et al., 2001). A participant flowchart is presented in Figure 1. Participants (n = 165, 50.1% female) were originally recruited from a cohort of 4-month-old infants screened on motor and emotional reactivity (Calkins et al., 1996; Fox et al., 2001; Kagan et al., 1998) between 1989 and 1993 in the Washington DC area. For more information on the recruitment, in- and exclusion criteria of the infants, see Fox et al. (2001). The exclusion criteria for the current study were contra-indications for the MRI (e.g., pregnancy or metal implants), a family member working at the NIH, IQ below 70, any use of psychotropic medication (e.g., selective serotonin reuptake inhibitor: SSRI's), and diagnosis of psychiatric conditions (psychosis, post-traumatic stress, bipolar, or obsessive-compulsive disorder).

Twenty-four participants (50% female, mean age = 29.48 years old, *SD* = 2.1) were included in the final analysis. All participants were White and almost all were not Latino/Hispanic. Regulations and safety concerns due to the COVID-19 pandemic and the fact that most participants no longer lived near the research facility were reasons for attrition. All participants provided informed written consent prior to completing all mental and physical examinations and MRI scanning. Participants were paid $210 for completing all measurements and compensated for traveling. The Institutional Review Boards of the National Institute for

Mental Health and the University of Maryland approved all procedures (IRB number: 03-M-0186).

Figure 1.

Flowchart of participants.



Procedure

The study consisted of the following components: physical evaluation by a medical doctor, clinical interview by a trained clinician (SCID; Structured Clinical Interview for DSM-IV Axis I disorders; First, Spitzer, Gibbons & Williams, 1995), the vocabulary and matrix reasoning subtest of the WAIS intelligence test (Wechsler Adult Intelligence Scale; Psychological Corporation, 1999), self-report questionnaires, and the MRI scan. The MRI scan consisted of two tasks: the Monetary Incentive Delay task (MID; Knutson et al., 2000; 2005; not relevant for the current study) and the social flanker task. Block randomization was applied for the task in the scanner: half of the participants started with the social flanker, and the other half began

with the MID task. The order of condition (social vs. alone) of the social flanker task was randomly assigned among participants. Participants had the option to visit the NIH and perform all components in one day or have the visit split into two days.

Measures

Infant Temperament: Behavioral Inhibition

Infant BI (Calkins et al., 1996; Kagan et al., 1987) was measured in infants when they were 14 and 24 months (Fox et al., 2001). For the current study, if both measurements were available, they were averaged. If not, we took one of the available BI scores to maximize the sample size. The procedure for determining temperament at 14-month-old and at 24-months-old are briefly explained below. For a full account of the procedure, including the times and setting of the observations, see Fox et al. (2001).

The infant BI score at 14 months is a composite of fearful and avoidant behaviors observed during three episodes in a laboratory setting: a free-play session in an unfamiliar playroom, an introduction to an adult stranger, and an introduction to a novel toy robot. Infants' reactions during these three episodes was recorded and coded using to the following categories: (1) the latency to first touch the toy during free play, (2) the latency to vocalize during free-play, (3) time spent in proximity (within an arm's length) of the mother during free-play, (4) the latency to vocalize to the stranger, (5) the latency to approach the stranger, (6) time spent in proximity to the mother while the stranger presented the infant with a toy, (7) latency to vocalize to the robot, (8) latency to approach the robot, and (9) time spent in proximity to mother during the robot episode.

The infant BI score at 24 months was derived from the reactions during the identical three episodes, and in addition, their willingness to crawl through an inflatable tunnel and the reaction of the infant to an adult stranger in a clown costume. As presented in Fox et al. (2001), infants' behavior was recorded and coded using the following categories: (1) time spent in proximity (within an arm's length) of the mother during free-play, (2) time spent in proximity to the mother while the stranger presented the infant with a toy, (3) time spend in proximity to the mother during the robot episode, (4) time spent in proximity to mother during the tunnel episode (5) latency to approach the stranger and/or touch the toy, (6) latency to approach and/or touch the robot, (7) latency to pass through the tunnel.

For both time points, the summarized scores from the episodes were standardized. Two observers rated each episode, using percentage agreement for reliability for 15% of the data. Pearson's correlations between the paired coders ranged between .85 and 1.0 for the 14-month-olds and .77 to .97 for the 24-month-old infants. We mean-centered the infant BI for the MRI analyses.

Social anxiety: Liebowitz Social Anxiety Scale

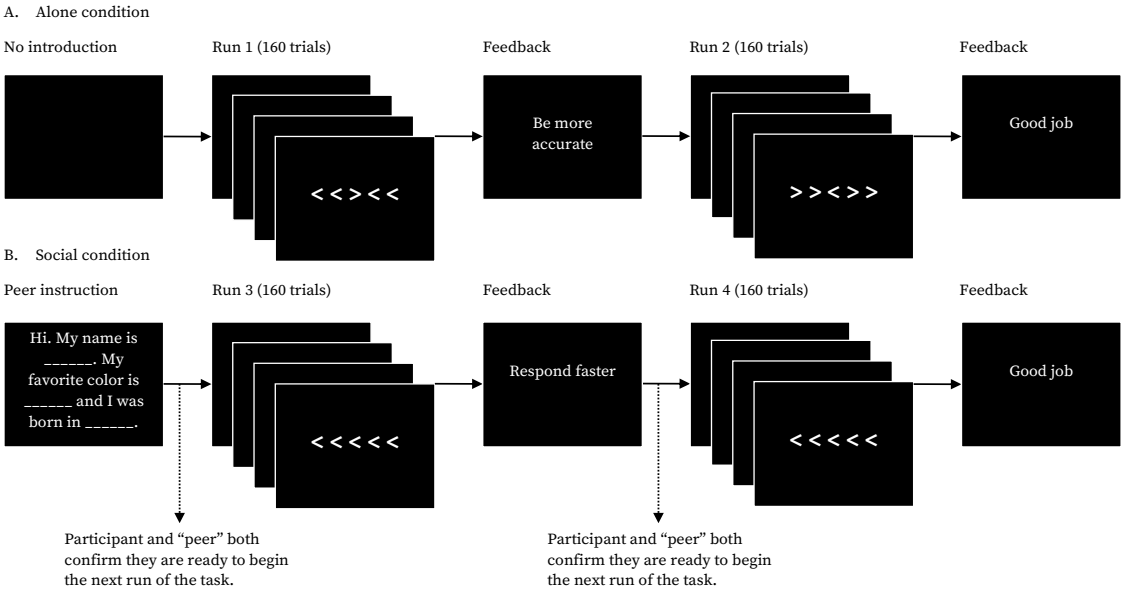
The Liebowitz Social Anxiety Scale (1987) is a self-report questionnaire to measure social anxiety. In this questionnaire, participants had to rate their fear and avoidance in 24 situations on a 4-point Likert scale (0 = none, to 3 = severe). High LSAS scores indicate more social anxiety. The self-report questionnaire has very good psychometric properties, including good test-retest reliability and high internal consistency (α = 0.95 for total score) (Baker et al., 2002; Heimberg et al., 1999). We used the total score by adding all the item scores and then standardizing them among the included participants. For the MRI analyses, we mean-centered the LSAS scores.

Social Error Processing: Social Flanker Task

The social flanker (displayed in Figure 2) is a version of the flanker task (Eriksen & Eriksen, 1974), based on the task reported by Barker et al. (2015; 2018), Buzzell et al. (2017) and Smith et al. (2019). In the current study, the social flanker task is a modified (more trials and shorter response window, described below) version of the task used in Smith et al. (2019) to accommodate the adult sample. In this task, a row of five arrows pointing to either left or right is displayed in each trial. Participants were instructed to indicate the direction of the middle arrow by pressing a button with the corresponding hand. In the congruent trial (50% of the trials), all the arrows point in the same direction, whereas in the incongruent trial (50% of the trials), the middle error is in the opposite direction. After 16 practice trials, participants completed two blocks of trials (160 trials each, approximately a 6-minute run) for the social condition and two blocks for the alone condition (a total of 640 trials). In the *social* condition, the participant believed they were being observed by a same-sex peer, who would make predictions on the participant's performance based on participant information given prior to the scan (first name, age, favorite color). In the alone condition, performance feedback was generated by a computer. Feedback on the performance was given to the participant to maintain a good speed-accuracy trade-off. Specifically, 'good job' was shown for accuracy between 75-90%, 'go faster' was given when accuracy was above 90%, and 'be more accurate' was shown when accuracy was under 75%. In reality, no other participant was present during the social flanker, and pre-recorded audio files facilitated the deception (Smith et al., 2014). All participants received a debriefing about the deception following the scan, including a check question of whether they believed the deception. The social flanker task was performed through E-Prime. All task stimuli were presented in white on a black screen. The flanker stimuli appeared for 200 ms after random amount of time (0 – 300 ms). Response was recorded in a window of 1700 ms. Intertrial interval was 135 ms. We considered the omission of a response as an error.

Figure 2.

Depiction of the social flanker task. The condition was counterbalanced.



Note. Feedback presented on screen was dependent on participants' task performance.

Functional Magnetic Resonance Imaging

fMRI Data Collection. The fMRI images were acquired on a 3T MR750 General Electric scanner (Waukesha, Wisconsin, USA) with a 32-channel head coil at the National Institutes of Health, North Bethesda, MD. For each block, 216 volumes were generated. Standard T2*-weighted echo-planar imaging (EPI) sequence (with 42 interleaved axial slices) was applied to acquire the blood-oxygen-level-dependent (BOLD) signal (repetition time (TR) = 2000ms; time to echo (TE) = 25 ms; flip angle = 60°; FOV = 240 mm; matrix = 96 × 96; in-plane resolution = 3 × 2.5 × 2.5 mm). For the anatomical scans, a T1-weighted magnetization prepared rapid gradient echo was used (MPRAGE; TE = minimum full; TI = 900ms; flip angle = 7°; FOV=256mm; 256 × 256-pixel matrix; in-plane resolution = 1 × 1 × 1 mm).

fMRI Preprocessing. The software program Analysis of Functional NeuroImages (AFNI; Cox, 1996) was used for quality checking, pre-processing, and group-level data analyses. The first and second authors (MCL, DL) performed standard quality checking and pre-processing steps, as described in Smith et al. (2019). Data only from participants with less than 10% TR censored were analyzed. Pre-processing of the MRI data included the inspection of signal spikes, slice-time correction, co-registration, and spatial smoothing (6mm smooth kernel, full width at half maximum). Standard Talaraich space was used for warping, and TRs greater than 1mm of movement were excluded.

For each participant, an individual-level general linear model (GLM) was performed. This included six regressors (time-locked to the onset of each stimulus) for each social flanker condition (social vs. alone) based on their accuracy and trial type: correct-congruent, correct-incongruent, commission error-congruent, commission error-incongruent, omission error-congruent, omission error-incongruent. AFNI's 3dClustSim (Cox et al., 2017) was applied to allow for the computation of a cluster-size threshold for a voxel-wise *p*-value (based on 10,000 Monte-Carlo simulations). Following Eklund et al. (2016), a non-Gaussian auto-correlation was assumed for the smoothing function. The cluster contiguity threshold was determined with a voxel-wise probability threshold of $p < 0.001$ and the family-wise error rate of $\alpha = 0.05$. This resulted in a cluster contiguity threshold of 1161 mm^3 or 43 voxels.

Preregistered Analysis

To test the first hypothesis (whether infant BI is related to social anxiety), we performed a linear regression. For the second hypothesis, testing whether accuracy and response times differ for each trial type and condition of the social flanker task, we performed a 2 by 2 analysis of variance (ANOVA). For the remaining hypotheses, we used a linear mixed-effects model (3dLMEr: Chen et al., 2013), where the brain activity during error vs. correct incongruent trials is the dependent variable, LSAS and BI scores are continuous, between-subject independent variables, and social condition (social vs. alone) is the within-subject independent variable.

Exploratory analysis

We explored the relationship between task performance and current social anxiety. We conducted two repeated-measures ANOVAs, the first examining accuracy and the second, response times. For both analyses, the standardized LSAS score was the continuous between-subjects variable. The two within-subjects variables were the condition (social vs. alone) and trial type (incongruent vs. congruent).

Results

Sample and variable characteristics are presented in Table 1. To answer the first hypothesis, a regression analysis revealed that infant BI and social anxiety in adulthood were not significantly related ($b = 0.034$, $p = .08$).

Table 1.

Descriptives of the sample and all study variables

Demographics (n= 24)	**Mean**	**SD**	**Range**
Age	29.48	2.07	26.7 to 33.59
IQ (n= 10)	116	13.71	99 to 137
Gender	50% male		
Ethnicity	96% Not Latino/Hispanic		
Race	100% White		
Clinical measures			
Standardized infant BI	0.034	2.47	-3.2 to 5.94
LSAS at MRI scan	31.60	26.16	6 to 99
SCID past (n= 18)	Any disorder = 9	No disorder = 9	
SCID present (n= 18)	Yes disorder = 7	No disorder = 11	
Other Information			
Deceived (n= 11)	Yes = 7	No = 2	Not sure = 2

Note. LSAS = Liebowitz Social Anxiety Scale, SCID = Structured Clinical Interview for DSM-IV Axis I disorders.

Social Flanker Task Effects

For the first part of the second hypothesis, we expected standard flanker congruency effects for response times and accuracy. We also explored the effect of condition on response times and accuracy. For response times, no differences between congruent and incongruent trials ($F_{1,88} = 0.57$, $p = .45$), between the social and alone condition ($F_{1,88} = 0.04$, $p = .85$), or the interaction between congruency and condition ($F_{1,88} = 1.04$, $p = .31$) were found. For accuracy, there were differences for incongruent vs. congruent trials ($F_{1,88} = 33.85$, $p < .001$), but not for condition ($F_{1,88} = 0.01$, $p = .92$) or the interaction between congruency and condition ($F_{1,88} = 0.23$, $p = .63$). More errors were made in incongruent trials vs. congruent trials ($t = 5.34$ (23.92) $p < 0.01$).

For the exploratory analyses, to test whether social anxiety was related to task performance, two repeated-measures ANOVA were performed. However, there was no interaction effect of trial type and condition on social anxiety for accuracy ($F_{1,17} = 1.76$, $p = .25$) or response times ($F_{1,17} = 1.32$, $p = .39$).

Table 2.

Behavioral results of the social flanker task.

	Accuracy			
	Congruent		Incongruent	
Condition				
Peer	.98 (.02)		.81 (.16)	
Alone	.97 (.03)		.81 (.14)	
Total	.98 (.02)		.81 (.15)	
	Response time (ms)			
	Congruent		Incongruent	
	Correct	Error	Correct	Error
Condition				
Peer	424.19 (59.16)	430.92 (260.37)	499.45 (74.92)	432.73 (116.84)
Alone	425.37 (56.96)	316.69 (239.94)	503.95 (71.58)	435.73 (109.75)
Total	424.82 (55.82)	425.88 (224.22)	502.44 (69.61)	433.24 (102.83)

Note. Standard deviation is presented in parentheses.

Table 3.

fMRI whole brain results: error vs correct in incongruent trials.

Region	**Voxels**	**Talairach coordinates**		
		x	**y**	**z**
BI x social anxiety x condition	-			
BI x condition	-			
Social anxiety x condition				
Left middle temporal gyrus	75	38.8	38.8	1.2
Left middle cingulate cortex	60	16.2	36.2	26.2
Left middle cingulate cortex	60	18.8	18.8	36.2
Right precentral gyrus and right middle cingulate cortex	47	-21.2	11.2	41.2
BI x social anxiety	-			
Condition				
Right middle cingulate cortex	129	-21.2	31.2	33.8
BI	-			
Social anxiety	-			

Note: Cluster contiguity threshold = 43 voxels. Abbreviations: ME = main effects; BI = behavioral inhibition

fMRI results

Table 3 shows the results of the whole brain analyses. There were no regions identified in the hypothesized three-way interaction between infant BI, social anxiety and condition on brain activity, $p > 0.001$.

Four regions, the left middle temporal gyrus, two regions of the left middle cingulate cortex and the right precentral gyrus/right middle cingulate cortex, emerged for the condition and social anxiety interaction, shown in Figure 4a-d. For all regions, higher anxiety scores were related to more activation during the processing of correct vs. error trials in the social condition. Also, higher social anxiety scores were related to more during the processing of error vs correct trials in the alone condition. The slopes of the associations were, however, not significantly different from 0 (for the social condition: *b*s < -.12, *SE*s < .06, *p*s > .07, for the alone condition *b*s < 2.47, *SE*s < 1.62, *p*s > .14).

There was a main effect of the condition on activity in the right middle cingulate cortex (Figure 3). Participants showed more activation in the correct vs error trials for the social condition compared to the alone condition.

Figure 3

Main effect condition of the social flanker task on activation in the right middle cingulate cortex.

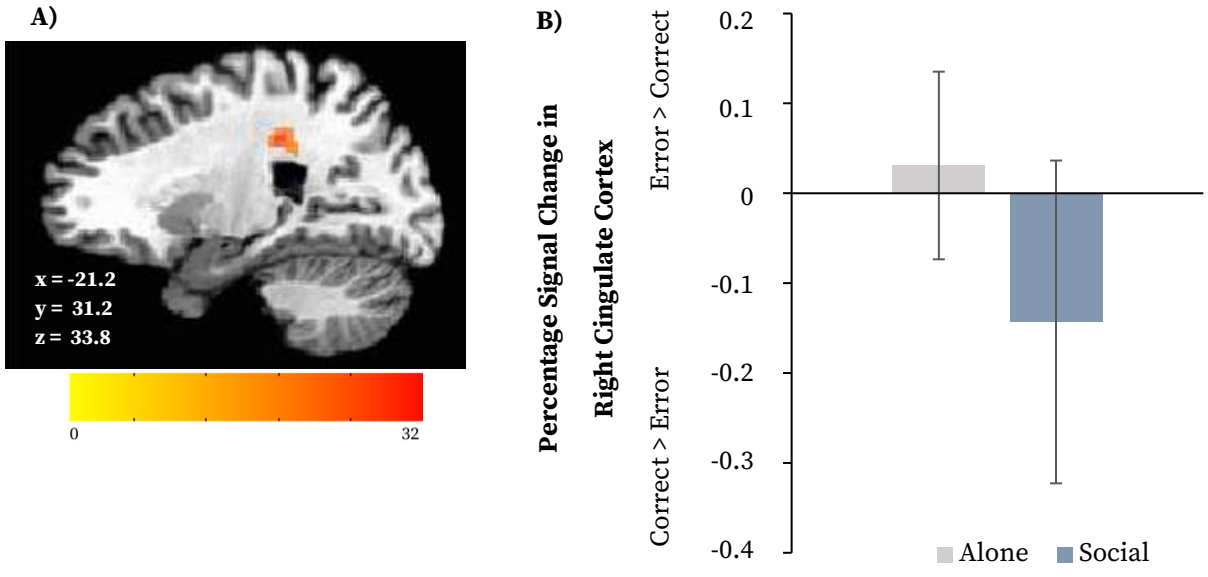
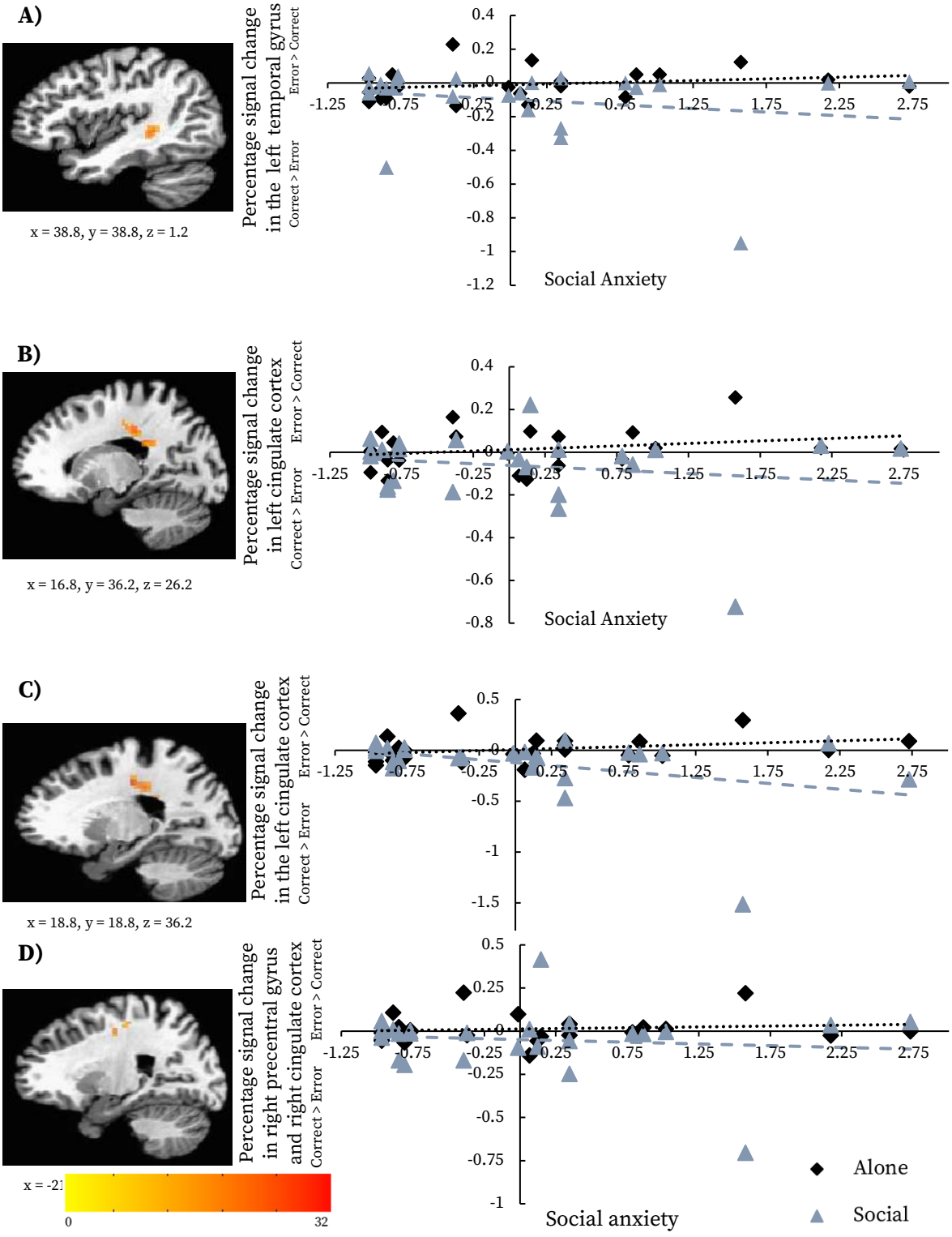


Figure 4

A. Condition by LSAS effect on the left middle temporal gyrus; B. Condition by LSAS effect on the left middle cingulate cortex; C. Condition by LSAS effect on the left middle cingulate cortex; D. Condition by LSAS effect on the right precentral gyrus and right middle cingulate cortex.



Discussion

The purpose of this study was to examine the associations between infant BI, self-reported social anxiety, and social error processing by the brain during adulthood. Because infant BI is an important predictor of the development of social anxiety, we expected that infant BI would be associated with self-reported social anxiety nearly 30 years later in life. However, we did not observe this relationship in the current study. A secondary aim of this study was to investigate social error processing in the brain, assessed by the social flanker task during fMRI. On a behavioral level, we observed a traditional congruency effect on accuracy, but not on response times. At the neural level, we found a main effect of condition on the activity of the right middle cingulate cortex. Next, we investigated the association between infant BI and social anxiety on the social error processing by the brain. We observed an interaction effect of task condition with social anxiety on the brain activity in the left temporal gyrus, left middle cingulate cortex and right precentral gyrus, and right middle cingulate cortex during error processing. Finally, contrary to our expectations, we did not observe an association between infant BI on social error processing by the brain, nor a three-way interaction between infant BI, social anxiety and condition on error processing by the pgACC.

In children with anxiety disorder, a three-way interaction effect of social anxiety on social error processing was observed in the pgACC. The social flanker was originally designed for EEG studies (Barker et al., 2015), with the intention of studying the fast temporal dynamics of error and feedback processing. Although we made accommodations to the task to ensure reliable BOLD signals, it is possible that social error processing is too fast to be detected by fMRI. Furthermore, finding no association between flanker task performance and social anxiety scores is not uncommon in experiments and has been observed previously using this version of the task (Barker et al., 2015; Buzzell et al. 2017; Smith et al., 2019). We also did not find an association between infant BI and social anxiety and infant BI and (social) error processing, despite prior reports (Jarcho et al., 2013; Smith et al., 2019; Tang et al., 2020). A previous study with the same sample (Tang et al., 2020) did find a relation between infant BI and adult social anxiety. More participants were included in this study, not all of them were able to participate in the current fMRI study. Therefore, it is likely that there was insufficient statistical power to detect the relation between infant BI and adult social anxiety in the current sample. Finally, it is possible that infant BI does not influence behavioral performance or error processing directly, but rather indirectly through other cognitive control resources (e.g. attention, self-control) that performance monitoring elicits (Henderson et al., 2015).

In the current study, we found an interaction between condition and self-report social anxiety in activity during error processing in the temporal gyrus and middle cingulate cortex (MCC). Social anxiety was related to more error-related activity during the alone conditions and more correct-related activity during the social conditions was observed for these regions. Activity in the temporal gyrus has previously been associated with the processing of sounds (Mesgarani et al., 2014) and the processing of social concepts and processes (Aldolphs, 2001;

Iacoboni et al., 2004; Satpute & Lieberman, 2006; Zahn et al., 2007). Impairments in the left middle temporal gyrus have been found in patients with social anxiety disorder when compared to healthy controls (e.g., Goldin et al., 2014; Yun et al., 2017). The impairments affected the social-affective communication network and were related to the emotional reactivity and social evaluation in social contexts. Indeed, the heightened orientation and attentional processes that are elicited during communication and social interactions are related to the social fear that is observed in patients with social anxiety (Heeren & McNally, 2016). Thus, considering our findings, the pattern of activity found in the temporal gyrus could suggest heightened attentional control in the social condition, related to processing of potentially socially salient information by the participants. This was reflected in the degree of social anxiety experienced by the participants.

We observed more activity in the right MCC in correct vs. error trials in the social condition compared to the alone condition. Additionally, we observed an interaction between the condition and social anxiety in the left MCC, where higher social anxiety scores were related to more error-related activity in the alone condition and more correct-related activity in the social condition. In general, the MCC is involved in cognitive control activities in healthy individuals (Stevens et al., 2011), which explains the activity during error trials in the alone condition. In addition, the MCC is involved in the processing of social information. For example, the MCC has been implicated in social decision making during social interactions (Apps et al., 2013b). Specifically, the MCC is involved in predicting and monitoring the decisions as well as tracking the outcomes of decisions, for positive unexpected outcomes (Apps et al., 2013a). Furthermore, the gyral surface of the ACC, which is connected to the MCC, appears to be predictive of the connectivity between the dorsomedial prefrontal cortex and the temporoparietal junction, regions that are responsible for social information processing (Balsters et al., 2017; Balsters et al., 2013). The initial finding of MCC activity in correct trials during in the social condition in our study could be indicative of the processing of social decisional making.

The directionality of the condition effect (more activity during correct trials in the social condition) observed here was unexpected. Previous EEG studies showed more error-related brain activity in the social condition (Barker et al., 2015; Buzzell et al., 2017). However, the studies Barker et al. (2018) did not find effect of condition in error and correct related brain activity. The authors state that this finding could be due to an age-related effect on the ERN, as they did find the more error-related activity the social condition for younger vs. older adolescent girls. This observation can be explained by the enhanced social motivation typical for adolescence, corresponding with heightened neural sensitivity to pay attention to social cues (Crone & Dahl, 2012). Likewise, it is possible that the directionality observed in the current study is due to chance (based on the small sample size). A replicating in a larger sample of this design and task can explain the findings of the current study.

On the other hand, the increased activity during correct trials in the social condition could reflect the goal-directed effort during social decision making (Contreras-Huerta et al., 2020). In this case, the increased activity is the response to making the "correct choice" in the presence of peer observation. Another possibility is that the observed activity in the MCC corresponds to correct-related processing. The response event-related potential correct-related negativity is (CRN; Ford, 1999), is studied mostly in EEG studies. Because the CRN is observed during correct choices, in the same latency window and electrode sites (Fz and FCz) as ERN, we assume that neural activity in the ACC drives the CRN. However, the amplitude of the CRN is thought to reflect the reduced certainty of the correct choice (Pailing & Segalowitz, 2004; Scheffers & Coles, 2000) and also appears during error processing (Coles et al., 2001; Vidal et al., 2000). In addition, an increase in the CRN has been attributed to overactive performance monitoring during correct behavior (Bartholow et al., 2005; Endrass, et al., 2010; Hajcak & Simons, 2002). Based on these definitions of the CRN, it is possible that participants were more uncertain about their choices due to the 'social pressure' they experienced during the social condition and responded with increased monitoring of their responses through MCC activity, even though it was a correct choice. It is unclear whether the CRN is affected by social factors or social anxiety, as most CRN studies to date have used non-emotional attentional control tasks (Michael et al., 2021). Only one study (Endrass et al., 2014) found an increased CRN (together with ERN) for patients with social anxiety disorder vs. healthy controls, suggesting that social fear affects both performance monitoring components. In summary, although speculative, the relation between social anxiety and correct-related activity in the MCC might reflect an overreactive performance monitoring system that is triggered during the processing of social decisions.

There are several limitations to this study. First and most evident is the small sample size due to the attrition and eligibility of participants. Attrition is due to the longevity of this study: after 30 years, most of the participants moved out of the area or changed contact information. The small sample size resulted in, what is most likely resulted in underpowered statistical analyses. Eligible participants who were excluded due to medication use, including SSRIs, contributed to the smaller sample size. Although medication use may influence cognitive processes, it raises an interesting point of discussion. It is possible that eligible participants are taking medication because they have developed a disorder such as anxiety, and that temperament played a role in this development. However, the excluded participants did not differ significantly in infant BI from the included participants. To overcome this limitation, the inclusion of patients with a clinical diagnosis and larger sample sizes are recommended. The second limitation is the degree of deception success in the current version of the social flanker task. This version of the social flanker task has not been performed by adults in MRI studies. Although most of the participants (of whom we documented the deception rate) believed they are being observed, some participants reported that they 'could not hear' the peer during the introduction. In future MRI studies, we propose to ensure a successful deception of the social condition. This can be done by 'real-life' introduction of the

peer prior to the experiment or reciprocal evaluation of performance prior to experimentation (e.g., through real-time social interaction tasks; Weinberg et al., 2021).

In this unique 30-year longitudinal neuroimaging study, we administered the social flanker task in adults using MRI, where we found an association between social anxiety and social observation on the correct-related activity. This study provides initial insight into "correct" behavior in social contexts: we therefore encourage future research to further investigate underlying mechanisms that drive this behavior in EEG studies. The current study used prospective design, which may have provided important insights into the process and influences of early markers on later in life processes. This study may serve as model for other prospective research designs, where the availability of clinical, self-report, behavioral, and brain measures could provide for a holistic view on social error processing. For now, we can conclude that social error processing in adults is related to social anxiety and is processed by the middle cingulate cortex of the brain.

Chapter 6

General Discussion

The central focus of the current dissertation was the role of performance monitoring in psychopathology, measured at behavioral and neurophysiological levels. The results of the two longitudinal studies and the synthesis of existing literature in the meta-analysis and narrative provide unique contributions to the research on the role of performance monitoring in the etiology of psychopathology.

Summary of Findings: The Role of Performance Monitoring in Psychopathology

The first central question was to determine the role of error processing in externalizing psychopathology. To this end, two reviews were written examining two error-related brain measures. At the time, there was no meta-analysis that synthesized research on the ERN and Pe nor did they include samples of children with externalizing problems. The results in **Chapter 2** indicate that the ERN and Pe were systematically diminished in studies that primarily examined error processing in adults and children with ADHD and adults with clinical and subclinical addiction. The selected moderators did not contribute to the observed heterogeneity, except for the type of experimental task that was used to assess performance monitoring in the Pe meta-analysis. Studies using the go-nogo task show greater differences in the Pe between externalizing samples and healthy controls than studies using other experimental tasks. Notwithstanding the observed threat of small sample bias (significant for both ERP's), the lack of unpublished studies in the analyses, and the inconclusive *p*-curve analysis for Pe to assess *p*-value reporting, both meta-analyses consistently show reduced error processing in externalizing samples, suggesting impaired anterior cingulate cortex functioning.

In **Chapter 3,** I used the evidence from studies of error processing in externalizing disorders to discuss whether error processing can be considered a biomarker for this spectrum of disorders. Biomarkers are reliable, accurate, discriminative measures of biological substrates sensitive enough to be observed in a heterogeneous disorder sample. They inform researchers and clinicians about etiology, diagnosis, and prognosis. In the review, I show that although EEG correlates of error processing can be considered biomarkers for disorders such as addiction (Luijten et al., 2014), psychopathy (Vallet et al., 2021), and ADHD (Kaiser et al., 2020), there are several disorders from the externalizing spectrum underrepresented (discussed below). There is not only evidence for error processing as a biomarker, but error processing may explain individual differences in other cognitive domains, severity, comorbidity, and treatment status/outcome. Future research is needed to determine whether error processing is a cause or effect of psychopathology and whether brain stimulation can modulate error processing impairments.

Very few studies have access to large enough sample sizes and utilize longitudinal designs that are necessary to examine performance monitoring at the behavioral level in children across development. In **Chapter 4**, I explored flanker response times trajectory and test the association of the response time trajectory with the trajectory of teacher-reported

problem behavior in a unique large sample of Dutch elementary school children with repeated measures of flanker performance. An improved trajectory of response time was observed, evidenced by a piecewise reduction in response times of all trial types. We also observed improved accuracy, reduced post-error slowing, and a strategy of equalizing accuracy and speed as children aged. There were several positive associations observed between the trajectory of response time and the trajectory of problem behavior (excluding conduct problems). A faster response time at age 7, as well as a faster rate of decrease in response time, was positively associated with a corresponding decrease in anxiety, depression, ADHD-related, and oppositional defiant problem behaviors.

In **Chapter 5**, we examined the long-term association between infant behavioral inhibition and self-reported social anxiety in adulthood. Also, we investigated if the two concepts together were associated with brain and behavioral measures of social information processing in adulthood was examined. To study this, we piloted a social version of the flanker task during MRI to identify which areas are involved during social error processing (effect of peer observation vs. alone during performance monitoring) in adults. Results show that the middle cingulate cortex is activated during social error processing. In addition, there was an interaction effect of condition and social anxiety in the left temporal gyrus, left middle cingulate cortex, and right precentral gyrus. The activity in these regions was most pronounced during the correct trials in the social condition. The results indicate specific brain regions to be involved during the increased performance monitoring, in response to 'doing well' in front of the peer. However, there was no relationship between infant BI and adult social anxiety scores, and there was no effect of the social manipulation on the behavioral measures of the flanker task. The results are preliminary, based on a small sample size. Still, pending improvements for the social manipulation, the social flanker task may be a valid paradigm to further investigate social information processing, especially to shed light on how performance monitoring is processed by the brain and how this is related to psychopathology.

Interpretation of Diminished Error-Related Negativity in Externalizing Psychopathology

Traditionally, EEG researchers interpret results by referring to the underlying brain networks driving ERP's. However, truly understanding how the neurophysiological components relate to observable behavior (which is characteristic of the disorder) remains challenging. Occasionally, researchers relate ERP to task behavior or disorder symptoms (serving as proxies for disorder-specific behavior). In addition, the functional significance of ERN and Pe may be explained by a variety of theoretical approaches. Although not a primary outcome of this dissertation, Table 1 provides an overview of propositions that help interpret performance monitoring results from cognitive control studies. I will discuss the behavioral implications of reduced error processing deficits in externalizing psychopathology in the following sections.

Table 1.

Overview of propositions relevant for explaining performance monitoring outcomes in cognitive control studies*

Theory/ Hypothesis	Author	Cognitive Domain	Correlates	Brain region	Tasks	Sample§	Short description	Further reading
Conflict Monitoring Theory	Botvinick et al. 2001; 2004; Yeung et al. 2004	Conflict monitoring Inhibition	ERN FRN N2 PES PIA	ACC PCC PFC	Flanker Go-Nogo Simon Stroop	Healthy Internalizing Externalizing	Explains the detection and processing of conflicting information (discrepancy between error and correct response). The ACC has an evaluative (top-down control) role. ERN reflects the response of conflict on error trials.	Lo, 2018; Yeung et al. 2004; Dignath et al. 2020
Reinforcement Learning Theory	Holroyd & Coles, 2002	Error monitoring Motivation Reward.	ERN Dopamine	Basial ganglia PFC MFC ACC Striatum	Probabilistic learning Task switching	Healthy Addiction ADHD	Error-making calls upon the mesencephalic dopamine system to ensure a negative reinforcement learning signal. Errors serve to 'train' the ACC to direct upon the motor system (also referred to as the response selection model).	Friedman & Robins, 2022
Adaptive Orienting Theory	Wessel, 2018	Error monitoring Orienting Inhibition	PES ERN N2	aMCC	Flanker Go-NoGo Simon Stop Signal	Healthy	Unexpected task events (i.e. errors), trigger several processes, including the inhibition of task behavior and orientation attentional focus to the violation, in order to facilitate adaptive behavior.	Frömer & Shenhav, 2022; Śmigasiewicz et al. 2020
Orienting Theory	Notebaert et al. 2009	Error monitoring	PES	MFC	Flanker Go-NoGo Simon Stop Signal	Healthy	Assumes that errors are infrequent events that capture one's attention.	
Mismatch Theory	Bernstein et al. 1995; Coles et al. 2001; Falkenstein et al. 1991	Conflict & Error monitoring	N2 ERN FRN	aMCC	Any two-choice reaction tasks	Healthy	Servers as the error detection system: the sensitivity to the discrepancy between the decision (error response) and the correct response. Note: the presence of CRN weakens this theory	
Motivational Significance Theory	Gehring et al. 1993; Hajcak & Foti, 2008	Error monitoring Self-regulation	ERN Pe	ACC	Affective manipulations of choice tasks	Healthy Internalizing ADHD	Assumes that errors are motivationally salient and trigger startle/arousal, making errors significant. Also accounts for individual differences.	Hajcak et al. 2005
Predicted Response-Outcome Model	Alexander & Brown, 2010	Error monitoring Self-regulation Reward	ERN Cognitive control SAT N2	mPFC ACC	Change signal Flanker two-armed bandit	Animal Healthy	A cognitive and probabilistic framework that incorporates the response and outcome by the ACC, based on reinforcement learning. Builds forth on the Error likelihood Prediction	Brown & Braver (2005)
Thread Sensitivity	Weinberg et al. 2016	Error monitoring	ERN	ACC	Two-choice reaction tasks, with feedback or punishment	Healthy Anxiety	Errors are motivationally salient and represent aversive and threatening events that trigger a defensive response	Cole et al. 2022; Meyer et al. 2020

Affective Processing Hypothesis	Overbeek et al. 2005	Error monitoring	Pe PES	rACC cACC aIC	Flanker Go-NoGo Oddball Simon Stop Signal	Healthy	The Pe reflects the neuroaffective response of the brain. The response could be seen as "carrying/feeling" for the error	Dignath et al. 2020; Harsay et al., 2012
Error-Awareness Hypothesis	Nieuwenhuis et al. 2001 Ullsperger et al. 2010	Error monitoring	Pe PES Autonomic nervous SAT	ACC aIC rCC	Error signaling Flanker Go-NoGo Simon Antisaccade	Healthy Internalizing Externalizing	Conscious knowledge of mistaken decisions. Pe amplitude reflects either directly error awareness or lead to the processes of becoming aware of the error	Overbeek et al. 2005
Behavior-Adaptation Hypothesis	Gehring et al. 1993 Nieuwenhuis et al. 2001	Error monitoring	Pe PES PIA	ACC CC MFC	Flanker Go-NoGo	Healthy Externalizing	Change in behavior is conditional on the error, evidenced by change in strategy, reaction time or accuracy	Overbeek et al. 2005
Dual mechanism of control	Braver, 2012	Cognitive control	NA	PFC ACC	AX-CPT Stroop Choice reaction	Healthy Internalizing Externalizing	Assumes reactive vs. proactive mechanism of control. Where error detection is a reactive response and conflict monitoring and error likelihood estimation are proactive systems	
Diffusion Drift Model	Radcliff & McKoon, 2008	Decision-making	NA PES	NA	Any two-choice reaction tasks	Healthy	Decision-making involves the accumulation of noisy evidence in favor of one choice over the other, then the evidence for a particular choice 'drifts' over a decision threshold, which determines the final choice.	Dutilh et al. 2012; Desender et al. 2021; Mattes et al. 2022
Accumulation Account	Steinhauser & Yeung, 2012	Error monitoring	SAT Pe	ACC SMA	Discrimination Two-choice reaction	Healthy	The decision process (output) based on sufficient accumulated evidence on error commission (input) leading to being aware of the error.	Lenzoni et al. 2022; Steinhauser & Yeung, 2010
Post-Decisional evidence accumulation	Desender, Ridderinkhof, & Murphy, (2021)	Performance monitoring	Pe PES	rCC aIC ACC	Two-choice reaction	NA	The Pe is the neural marker for the evidence accumulated *after* choices are made. Error detection is "achieved" when enough confidence is reached.	Damasco et al. 2022; Yeung & Summerfield, 2012

Note. There are more cognitive control theories available than presented here. The theories included in the current overview are most relevant to performance monitoring, while I acknowledge they are also relevant for other cognitive processes and brain regions, depending on task manipulations. Also, they differ in perspective and assumptions. Some theories might be inter-related (e.g. Reinforcement and mismatch). [§]Samples stated under this column have been subjected to explaining performance monitoring results. When internalizing and externalizing is stated under the column sample, many disorders in this spectrum are implicated.

Abbreviations: ACC = anterior cingulate cortex; PCC = posterior cingulate cortex; rACC = rostral ACC; cACC = caudal ACC; PFC = prefrontal cortex; MFC = medial frontal cortex; mPFC = medial prefrontal cortex; aMCC = anterior midcingulate cortex; SMA = supplementary motor area; aIC = anterior insula cortex; rCC = rostral cingulate cortex; CC = cingulate cortex; PEA = post-error improvement in accuracy; ERN = error-related negativity; FRN = feedback-related negativity; CRN = correct-related negativity; ADHD = attention-deficit hyperactivity disorder; SAT = speed-accuracy tradeoff.

The ERN is one of the most studied components of performance monitoring and appears to play an important role in both internalizing and externalizing disorders. There are three relevant theories that help explain the implication of reduced ERN for externalizing psychopathology. According to the *Conflict Monitoring Theory* (Botvinick et al., 2004), a conflict is the erroneous response vs. the correct response. The reduced ability to process errors suggests a reduced anterior cingulate response to evaluating conflicting information. In other words, a patient with externalizing disorder may have problems identifying and evaluating correct versus incorrect choices. This explains why patients with externalizing psychopathology exhibit maladaptive behaviors. In the light of the *Reinforcement Learning Theory* (Holroyd & Coles, 2002), where error triggers the dopamine system to ensure learning from the error, reduced error processing suggests that this system is affected in patients with externalizing psychopathology. This helps to explain why patients fail to 'learn' from their mistakes, hindering or prohibiting them from being able to change their behavior. Another relevant framework to explain error processing and subsequent behavior is the *Adaptive Orienting Theory* (Wessel, 2018). The theory suggests that unexpected task events (i.e., errors) elicit a series of processes that orient attention toward the error as well as inhibit current behavior. The reduced ability to improve performance after error making (Franken et al., 2007) and the reduced PES observed in disorders such as addiction (faster response times after errors; Sullivan et al., 2019) provides initial support for this theory. On the other hand, the role of N200 in externalizing samples is still debated (e.g., inconclusive in ADHD: Kaiser et al., 2020; mixed in addiction: Parvaz et al., 2011). In sum, the reduced ERN in externalizing samples suggests a diminished response of the brain to an unexpected or conflicting decision and a failing system of behavioral adaptation following the erroneous decision. All different theoretical perspectives share the involvement of the ACC. Therefore, it is recommended to continue to investigate the link between the ERN, ACC, and behavior adaptation to substantiate the different theoretical perspectives with empirical data.

Chapter 2 and **3** describe considerable heterogeneity of ERN in externalizing samples, which requires further research. There are several potential causes for the observed heterogeneity. First, the heterogeneity in ERN could be due to the heterogeneity of symptoms and behavior within clinical disorders. ERN would thus also be heterogeneous and at the same time disorder specific. For instance, there was a larger ERN reduction for clinical addiction than for subclinical addiction in the meta-analysis. This suggests that the degree of error processing deficits relates to the severity of symptoms for a disorder. To address this, the relationship between the ERN and the degree of disorder-specific symptoms/behavior should be examined in more depth (e.g., Baldwin et al., 2015; Meyer & Hajcak, 2019). Second, other disorder-related variables that influence the ERN, such as age of onset, comorbidity, and medication use, should be examined to explain heterogeneity. Third, the heterogeneity in ERN can be the result of sample variables (e.g., age, gender). Finally, heterogeneity can be the result of the way the ERN is captured. Concerning the last cause, the "absolute" value of the ERN is sensitive to how it is measured (experimental paradigm; Meyer et al., 2013), how the

EEG signal is pre-processed, and how it is calculated (Boen et al., 2021; Clayson, 2020; Klawohn et al., 2020).

Besides error-related ERP's, studying EEG signal oscillations through time-frequency (TF) analyses can give additional insight in the neural underpinnings of error processing. The examination of oscillatory patterns includes the decomposition of the EEG signal into different frequency bands. They are modulated by the cognitive processes or events (therefore, they are referred to as event-related oscillations: ERO's) and inform goal-directed behavior (e.g., Buzzell et al., 2019; Cooper et al., 2019; Valadez & Simons, 2017). Where traditional ERP requires the signal to be synchronous, disregarding any non-phase-locked signal (Luck, 2014), TF distinguishes between amplitude and neural oscillations (including non-phase-locked signals). Also, TF analyses allow for the strength of the signal and the study of phase synchronicity within trials and electrode sites. Theta (Cavanagh & Frank, 2014; Cohen & Donner, 2013) and, to a lesser extent, delta (more reflective of decision-making processes; Fischer et al., 2018), and alpha (Van Driel et al., 2012) are the frequency bands that appear to be characteristic of performance monitoring and error-related brain activity in healthy participants. There are initial reports on performance monitoring-related spectral oscillation differences that have been observed comparing externalizing samples with healthy controls. For instance, reduced power in theta is observed in ADHD (Baijot et al., 20172) and reduced delta oscillations in the acute administration of an opioid (Pfabigan et al., 2021). However, intact performance monitoring is observed in externalizing samples (e.g., Bernat et al., 2011). Currently, our lab is investigating whether there are diminished ERO's in substance use (Stam et al., 2023) and other addictions (Lutz, Franken, & Steele, *in prep*).

Moving forward, additional research is encouraged to further investigate heterogeneity in ERN and spectral oscillations in externalizing samples. This will deepen our fundamental understanding of the underpinnings maladaptive behavior in disorders.

Making Sense of the Error Positivity in Externalizing Disorders.

The late positive potential of error processing Pe was a central measure studied in **Chapters 2** and **3**. Despite the growing interest in this ERP, there is, to this date, still debate on how Pe is linked to behavior. The Pe and ERN are separate ERP's (Dhar, & Pourtois, 2011; Di Gregorio et al., 2018; Endrass et al., 2007; Falkenstein et al., 1991; Nieuwenhuis et al., 2001) and the ERN theories may not apply to Pe (e.g., dopamine does not influence the Pe; De Bruijn et al., 2004). For now, I interpret the results from **Chapter 2** with the available propositions (Desender et al., 2021; Falkenstein, 2004; Overbeek et al, 2005; Steinhauser & Yeung, 2010).

According to the *Error-Awareness Hypothesis,* the Pe reflects the consciousness of the mistake (Godefroid et al., 2016; Nieuwenhuis et al., 2001). There is no consensus on whether the Pe represents error awareness or the signal that leads to error awareness. Yet, several studies point to Pe as a representation of error awareness (Boldt & Yeung, 2015; Hester et al., 2005; Kirschner et al., 2021) or that Pe is larger when errors are perceived (Nieuwenhuis et al.,

2003). A reduced Pe in externalizing disorders indicates that patients are less likely to be conscious of their errors, preventing subsequent corrective behavior. Indeed, EEG and MRI studies confirm that error awareness is reduced in patients with externalizing psychopathology, driven by hypoactivation in the ACC in addiction (Hester et al., 2009; Ridderinkhof et al., 2002) and in ADHD (Geburek et al., 2013). In addition, the anterior insular cortex (AIC; Ullsperger et al., 2010) could play an important role, as error awareness 'activates' the 'salience network' in the AIC through arousal, which subsequently enables task adaptation. Alternatively, the *Affective-Processing Hypothesis* suggests that a reduced Pe constrains the affective appraisal of the event. If this is the case, an individual with externalizing psychopathology fails to process the potential negative consequence and cannot learn or change behavior. It is unclear whether an individual 'doesn't care' about errors or that they simply 'can't feel' for errors. According to the *Behavioral-Adaptation Hypothesis*, the Pe is dependent on the behavioral change after the error (measured by post-error slowing). Given the reduced Pe and PES in patients with ADHD (Balogh & Czobor, 2014) and addiction (Sullivan et al., 2019), this could imply that the signal to "slow down to change behavior" is affected. Last, although not stated as a hypothesis (due to the similarities with the task-related P300), a Pe amplitude can be related to the motivational significance of an error, the activation of the noradrenergic system or the update of the information model of the context of the error. If this is the case, patients with externalizing disorders have a reduced noradrenergic response of the brain toward errors, dampening the alertness and attention triggered by errors. Since P300 has systematically been found to be reduced in patients with externalizing disorders (Patrick et al. 2006; addiction: Euser et al., 2012; Iacono & Malone, 2011; ADHD: Metha, 2020), patients have a reduced capacity to assess the significance of the stimuli presented and to divide appropriate attention processes to focus on task-related stimuli. Finally, a more recent paper suggests that the Pe may reflect the "post-decisional evidence accumulation" (based on the diffusion drift model; Desender et al., 2021). This post-decisional response triggers metacognitive experiences such as error detection and confidence in the decision, which in turn stimulates next-trial adjustment.

An interim conclusion is that a reduced Pe in patients with externalizing psychopathology is indicative of reduced error awareness, evaluation of the error, or reduced initiation of behavioral adaptation following the error. Currently, research on the suggested hypotheses and innovative approaches and their considerations is ongoing (Desender et al., 2021; Overbeek et al., 2014; Wessel, 2012). Therefore, to properly understand the Pe results observed in the field, I encourage the continuation of finding empirical evidence to validate these hypotheses. For now, we must be aware that the absence of theory-driven work in investigating Pe in clinical samples can trigger questionable research practices, such as selective reporting of significant results (for an application in ERN research, see Clayson et al., 2020). This is evidenced by the lack of hypotheses stated in studies as discussed in the supplementary materials of the meta-analysis. Still, we need to gather more information on the Pe in clinical samples to help us understand the framework at hand. Open-science practices

(e.g. pre-registration) may help in this attempt (Paul et al., 2021). Therefore, more research on its neural origin, its psychometric/discriminant properties (Pe vs. P300b or ERN) studies and moderators of Pe is highly recommended.

Under-represented Disorders in the Externalizing Spectrum

Most research on performance monitoring in externalizing psychopathology has been conducted in individuals and patients with ADHD and substance use problems and disorders. There are three externalizing disorders that are currently under-represented in performance monitoring research. The section below discusses why performance monitoring may be informative to the problematic behavior observed in the three disorders: behavioral addiction (BA), and conduct disorders and oppositional defiant disorder, discussed together.

Behavioral Addiction. A promising yet dire avenue that performance monitoring and cognitive control research should turn to is the field of behavioral addiction disorders. There are initial investigations on the ERN and Pe in internet addiction (Zhou et al., 2013), and other 'addictive' behavior, such as food intake (Franken et al., 2018), gaming (Littel et al., 2012), and binge watching (Kilian et al., 2020). Error processing was reduced in individuals with excessive/addictive internet gaming and food intake, as evidenced by a reduced ERN and more error-making than healthy controls. However, other studies show that the ACC is more active in individuals with internet addiction disorder than in healthy controls (Dong et al., 2013), and that error processing is not impaired in individuals who frequently binge watch (Kilian et al.,2020), or game (Luijten et al., 2015). In addition, other brain regions like the precuneus, putamen, and insula may be involved in error processing, as shown in a study on food addiction (Hsu et al., 2017). Clearly, further investigation and replication are needed to establish the role of performance monitoring in BA.

Besides the inconsistency in the initial reports, there are other reasons to investigate cognitive control in BA. First, the observed maladaptive behavior in BA has similarities with the compulsive and obsessive disorder or substance use disorder (repetitive behavior, loss of control, and preoccupation). Since performance monitoring is affected differently in OCD and addiction, it is interesting to see what pattern is specific for BA. In addition, this can inform the notion that shared cognitive dysfunctions contribute to (comorbid) disorders, the philosophy of the Research Domain Criteria (discussed in more detail in a section below). Secondly, the prevalence and incidence of clinical behavioral addictions such as internet gaming disorder and pathological gambling disorder are alarming. Furthermore, it is evident that problematic media use, sex, food, shopping, exercise, and other addictive behavior may have a great impact on individuals and societies. This raises the necessity to "catch up" on how cognitive dysfunction drives maladaptive behavior in BA. Last, BA or problematic addictive behavior affects children and adolescents, providing the opportunity to have a developmental perspective on the role of cognitive control in behavioral addiction (Derevensky, 2019). To

conclude, the knowledge and best practices that researchers have gained from cognitive control research should be applied to BA samples sooner rather than later.

Externalizing disorders in children. Two externalizing disorders are under-represented in performance monitoring research. Conduct disorder is characterized by persistent antisocial behavior that violates the norms and rights of others (American Psychiatric Association, 2022). Oppositional defiant disorder is characterized by hostile, negative, aggressive, and uncooperative behavior towards others (American Academy of Child and Adolescent Psychiatry). Consequences of both disorders affect the individual, including increased risk for substance use and mental health problems, as well as the society: increased risk for criminal behavior. It is therefore essential to improve our understanding of what drives maladaptive behavior and develop possible treatment options (e.g., brain stimulation). For both disorders, deficits in several cognitive control networks and reduced executive functioning have been observed, driven by hypoactivation in the ACC, prefrontal cortex, and caudate (Alegria et al., 2016; Fairchild et al., 2019). Specific performance monitoring studies in children are rare. The studies of Hall et al. (2007) and Woltering et al. (2011) show reduced performance monitoring related stimuli and response ERP's. A meta-analysis with adolescent and adult samples with elevated psychopathy found reduced ERN and Pe (Vallet et al., 2021). In addition, the reduced ERN was driven by impulsive and antisocial traits in individuals with psychopathy. Most interestingly, several studies suggested that ERN/Pe can be indicative for rearrest (e.g., Steele et al., 2015) and treatment trajectories (Woltering et al., 2011.). This demonstrates the potential predictive value of performance monitoring in psychopathology and its course.

Relevance of Performance Monitoring Tasks

Researchers design experimental tasks as a tool to elicit task behavior and brain responses. Relatively simple task adaptations and manipulations to these tasks allow us to test for specific hypotheses, as demonstrated with the social flanker task in **Chapter 5**. Inspired by the work in this dissertation, two important recommendations for future research on performance monitoring tasks are discussed below.

The work presented in this dissertation lays out the groundwork to link affected performance monitoring to maladaptive behavior. As discussed in previous sections, there are several hypotheses for error processing that involve the evaluation of and emotional response to errors. It is based on the idea that errors do not (only) trigger arousal but also a valence-specific effect (Koban & Pourtois, 2014). It has been proposed that the evaluation of errors influences behavior because incorrect actions are considered negative and 'bad' versus correct decisions as positive and 'good' (Aarts et al., 2012; 2013). In addition, a considerable overlap exists in brain regions involved in error, emotional, and social processing, suggesting a potential integrative network that drives behavior (Koban & Pourtois, 2014). Indeed, there is an increase in reports investigating the role of emotional response to errors (for a review in

healthy samples, see Nuñez-Estupiñan et al., 2022). Although this topic is beyond the scope of the current dissertation, moving toward this integrative framework (including the cognitive, emotional, and social propositions) on error processing will aid in defining the role of performance monitoring in psychopathology.

Second, traditional performance monitoring tasks have low ecological validity: they are designed for a controlled laboratory environment and mimic real-life processes. Although task adaptations and innovations are rigorously tested for their face, internal, and construct validity, an effort should be made to increase the ecological and external validity of the tasks and experiments. This can be done by combining methodologies such as ecological momentary assessment (Smith et al., 2019) or the use of virtual/augmented reality (Berger & Davelaar, 2018; Parsons & Courtney, 2016).

The Potential of Behavioral Measures

Traditionally, behavioral measures are reported in neurophysiological studies to 'check' whether the manipulation or task effect was successful. The remaining analyses focus on ERP's, ignoring valuable information that tasks measures can provide, such as individual trial-to-trial changes indicating behavioral changes. In addition, many studies draw conclusions on the directionality between amplitude differences and behavioral performance outcomes without statistically testing whether the association is present. Below, I present four relevant issues and promising avenues on the potential of behavioral measures.

First, testing the relationship between behavioral indices and neurophysiological measures of performance monitoring yields contradicting results. A larger ERN is related to longer response times in consecutive trials (Fischer et al., 2016), and both theta oscillations and ERN were related to post-error adaptation and slowing (Kalfaoğlu et al., 2018; Valadez & Simons, 2017). The association is also found in studies examining single-trial ERN and subsequent-trial behavior (Debener et al., 2005; Steinhauser & Andersen, 2019), single-trial level ERN predicting PES (Beatty et al., 2020) or linking previous trial theta to trial accuracy (Buzzell et al., 2019). On the other hand, other studies did not find any association between task behavior and neural correlates (e.g., Adler, 2021; Buzzell et al., 2017). According to the extensive review by LoTemplio et al. (2023) on the relationship between ERN and task-related behavior (PES and post-error accuracy), the mixed results are due to the involvement of the ACC in many cognitive control processes as well as the way the ERN is measured. Although a meta-analysis confirms that frontal midline theta and PES are positively related (Cavanagh & Shackman, 2015), a debate on whether PES results in behavior adaptation is still ongoing (Buzzell et al., 2017; Kirschner et al., 2021; Notebaert et al., 2009). With the knowledge that both the ERN/PE (**Chapter 2**) and PES are affected in externalizing samples (Balogh & Czobor, 2016; Sullivan et al., 2019), it is possible to hypothesize that they are related. Yet, to my knowledge, this has not been thoroughly investigated in psychopathology, except in anxiety (Cavanagh &

Shackman, 2015). By examining these intricate relationships, we can determine if task measures are indicative of adaptive behavior.

Second, it remains unclear to what degree there is a link between task performance and clinical measures of disorders. For instance, larger response time variability was related to ADHD symptoms (e.g., Epstein et al., 2003). In addiction samples, a meta-analysis (Smith et al., 2014) confirmed that reduced performance on go-nogo and stop-signal tasks is associated with behaviors such as heavy drinking. Examining this link has clinical implications. It can explain the individual differences and within-person differences in disorders, which elucidate on the high heterogeneity observed in psychopathology (Meyer & Hajcak, 2019). Also, several clinical instruments involve the use of performance monitoring tasks and use behavioral indices to support clinical diagnosis (e.g., in the Netherlands: the Amsterdamse Neurologische taken, ANT, de Sonneville, 1996).

Third, there are specific trial-to-trial changes in tasks that influence the overall behavioral results. Although not discussed in this dissertation, studies on healthy participants show that trial-by-trial variables influence task performance and/or neural correlates. These variables include but are not limited to the Gratton effect (conflict adaptation due to sequence of congruency: Gratton et al., 1992; Clayson & Larson, 2013), cumulative error effect (Valadez & Simons, 2017), intertrial intervals (ITI's; Compton et al., 2017; Ehlis et al., 2018), post-error reduction in inference (PERI, Buzzell et al., 2019; Danielmeier & Ullsperger, 2011; Fischer et al., 2018), level of noise (Mattes et al., 2023), and relationships between post-error slowing and post-error accuracy (Schroder et al., 2020). These effects influence the magnitude of the neural correlates and contribute to the heterogeneity of performance monitoring and between-group effects, as observed in the meta-analyses in **Chapter 2**. Although studying trial-by-trial effects requires intensive data handling and analysis, investigating trial-by-trial behavioral change provides unique insight into the mechanisms of behavior adaptation, and should therefore not go unnoticed in future performance monitoring studies. Sharing data and code for (pre-)processing and analysis through online platforms such as GitHub and Open Science Framework will facilitate this endeavor.

Researchers have focused on grant averages and between-group differences, neglecting the structure of ERP data (multiple observations grouped within individuals) and the significance of (intra-)individual variability in ERP's. To deal with the structure of ERP data, multi-level analyses have been proposed (Volspert-Esmond et al., 2018; 2023), incorporating cross-classified data (e.g., individual factors), or hierarchical data (data nested within participants). Exploring intra-individual variability patterns has provided valuable insights (Trenado et al. 2019) into maladaptive behavior observed in conditions, for instance in autism (Magnuson et al., 2020;). It is based on the notion that neural activity is dynamic (spatial and temporal domains) and changes in neural activity within a person can serve as a trait-like representation of behavior (Waschke et al., 2021). Multi-level location-scale modeling has successfully examined intra-individual variability in ERN in patients with depression, general

anxiety, and OCD (Clayson et al., 2022). More specifically, no group differences in ERN were found between the clinical groups, contrary to previous studies. However, between-person differences in the intra-individual variability of correct and error-trials ERN were found. This methodology inspired another study (Lutz et al., 2023 for pre-registration, *in prep*), with the attempt to provide insight into trial-by-trial changes in task behavior and their relationship with all performance monitoring ERP's. The previous- and current trial performance monitoring correlates (means and their variances) N2, P3, ERN, and Pe together were able to predict current trial behavior (response time and accuracy), controlling for previous and current trial type (congruency) and accuracy, within the same participant. The results indicate that the participants 'own' performance monitoring (neural and task level) best predicts their next behavior, an effect that cannot be observed using traditional between-subject approaches.

In conclusion, behavioral indices of performance monitoring provide an understanding of "normal" behavior, but research on how they provide insight into maladaptive behavior remains scarce. Future studies can attempt to examine how task behavior relates to the neural correlates of performance monitoring, clinical symptoms, trial-by-trial changes, and intra-individual variability.

Moving Forward: The Need for Developmental Perspectives and Clinical Applications

A particular challenge in studying cognitive control measures that are primarily driven by brain regions, is due to the fact that the brain is developing continuously during childhood, adolescence, and early adulthood. During adolescence and adulthood, there are significant changes in the mesencephalic dopaminergic system and availability of sites of dopamine release sites (Wahlstrom et al., 2010), which drives the changes observed in the ACC and ERN (Holroyd & Coles, 2002). Researchers have turned their attention to the description of the "typical" brain trajectory of cognitive control measures, which can inform researchers "where" to look for deviations from these trajectories and to test if this is indicative of psychopathology. This is mainly necessary since many mental disorders develop before adulthood (e.g., Olfson et al., 2023), making childhood and adolescence a vulnerable period. A few comprehensive developmental studies on childhood until early adulthood show age-related increases in conflict and error-processing in healthy individuals (Boen et al., 2021; Lo, 2018; Tamnes et al., 2013), yet such developmental or prospective studies with patients are lacking. The developmental perspective in clinical samples allows researchers to test the "changeability" which in turn can inform the progression of disorders. This is why a developmental perspective on performance monitoring in psychopathology is strongly encouraged.

More empirical evidence is needed to see how the separate processes of performance monitoring interactively drive behavior (given the role of the ACC). Rather than studying the cognitive control processes in isolation, different cognitive control processes (e.g., inhibition and working memory; Naim et al., 2023) and other frameworks (threat/reward processing,

attentional biases) all simultaneously contribute to maladaptive behavior and symptoms (Diehl et al., 2018; Ridderinkhof et al., 2004). Research initiatives such as the Research Domain Criteria (RDoC, NIH, Cuthbert, 2022) encourage exploring mental health through a more complete matrix of heuristic topics and the evaluation of domains that contribute to a disorder. For example, a panel of experts agreed that response selection and inhibition constructs from the cognitive control domain contribute most to addiction (Yücel et al., 2018). Several studies conclude that cognitive control deficits are related to disorders (McTeague et al., 2016) such as depression (Dotson et al., 2020), addiction (Brooks et al., 2017), and anxiety (Weinberg et al., 2015).

The studies discussed in **Chapter 3** serve as suggestions for clinical application of performance monitoring research in clinical settings. Future research can be devoted to validating treatment and interventions options that include brain stimulation and improvement in self-monitoring. Non-invasive (deep) brain stimulation (e.g., transcranial magnetic stimulation) shows promising results in modulating the performance monitoring network for some disorders (e.g., in patients with OCD: Balzus et al., 2022) but not in others (e.g., addiction: Verveer, 2020). The next step would be to investigate whether brain stimulation results in the reduction of symptom severity and improved behavior. Finally, a promising avenue includes meditation and mindfulness, which besides other benefits (Im et al., 2021), seem to improve executive functioning (e.g., Teper & Inzlicht, 2013). On a mechanistic level, the emotional acceptance and improved ability to self-monitor increase performance monitoring.

Overall conclusion

In conclusion, this dissertation highlights the role of performance monitoring in psychopathology. Findings indicate that performance monitoring deficits serve as markers for externalizing disorders, as supported by current research. The developmental sample presented here sheds light on the longitudinal trajectory of conflict monitoring in children and its association with adaptive behavior. Additionally, this dissertation offers initial evidence regarding the brain regions involved in social error processing. Through an integrative approach, this dissertation demonstrates how error and conflict processing are implicated in both healthy individuals and psychopathology, employing diverse methods and modalities. The results presented here also provide concrete recommendations for future performance monitoring research and its relevance for clinical practice.

References

Aarts, K., De Houwer, J., & Pourtois, G. (2012). Evidence for the automatic evaluation of self-generated actions. *Cognition, 124*, 117-127 http://dx.doi.org/10.1016/j.cognition.2012.05.009

Aarts, K., De Houwer, J., & Pourtois, G. (2013). Erroneous and correct actions have a different affective valence: Evidence from ERPs. *Emotion, 13*(5), 960-973. https://doi.org/10.1037/a0032808

Aarts, K., & Pourtois, G. (2010). Anxiety not only increases, but also alters early error-monitoring functions. *Cognitive, Affective, & Behavioral Neuroscience, 10*(4), 479-492. https://doi.org/10.3758/CABN.10.4.479

Abramovitch, A., Short, T., & Schweiger, A. (2021). The C Factor: Cognitive dysfunction as a transdiagnostic dimension in psychopathology. *Clinical Psychology Review, 86*, 102007. https://doi.org/10.1016/j.cpr.2021.102007

Adler, P. (2021). The relation between error-related ERP measurements, task performance, and symptoms in forensic samples: A systematic review and meta-analysis [Unpublished master's thesis]. Erasmus University Rotterdam, the Netherlands

Adolphs, R. (2001). The neurobiology of social cognition. *Current opinion in neurobiology, 11*(2), 231-239. https://doi.org/10.1016/S0959-4388(00)00202-6

Albrecht, B., Brandeis, D., Uebel, H., Heinrich, H., Mueller, U. C., Hasselhorn, M., Steinhausen, H-C, Rothenberger, A., & Banaschewski, T. (2008). Action monitoring in boys with Attention-Deficit/Hyperactivity Disorder, their nonaffected siblings, and normal control subjects: Evidence for an endophenotype. *Biological Psychiatry, 64*(7), 615-625. https://doi.org/10.1016/j.biopsych.2007.12.016

Alegria, A. A., Radua, J., & Rubia, K. (2016). Meta-analysis of fMRI studies of disruptive behavior disorders. *American Journal of Psychiatry, 173*(11), 1119-1130. https://doi.org/10.1176/appi.ajp.2016.15081089

American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425787

Amodio, D. M., Master, S. L., Yee, C. M., & Taylor, S. E. (2007). Neurocognitive components of the behavioral inhibition and activation systems: Implications for theories of self-regulation. *Psychophysiology, 45*(1), 11-19. https://doi.org/10.1111/j.1469-8986.2007.00609.x

Andreu, C. I., Palacios, I., Moënne-Loccoz, C., López, V., Franken, I. H. A., Cosmelli, D., & Slagter, H. A. (2019). Enhanced response inhibition and reduced midfrontal theta activity in experienced Vipassana meditators. *Scientific Reports, 9*(1), 1-11. https://doi.org/10.1038/s41598-019-49714-9

Apps, M. A., Lockwood, P. L., Balsters, J. H. (2013b). The role of the midcinulate cortex in monitroing others' decision. *Frontiers in Neuroscience, 7. https://doi.org/10.3389/fnins.2013.00251*

Apps, M. A., Green, R., & Ramnani, N. (2013a). Reinforcement learning signals in the anterior cingulate cortex code for others' false beliefs. *Neuroimage, 64*, 1-9. https://doi.org/10.1016/j.neuroimage.2012.09.010

Arbel, Y., & Donchin, E. (2009). Parsing the componential structure of post-error ERPs: A principal component analysis of ERPs following errors. *Psychophysiology, 46*(6), 1179-1189. https://doi.org/10.1111/j.1469-8986.2009.00857.x

Asendorpf, J. B., Denissen, J. J. A., & van Aken, M. A. G. (2008). Inhibited and aggressive preschool children at 23 years of age: Personality and social transitions into adulthood. *Developmental Psychology, 44*(4), 997-1011. https://doi.org/10.1037/0012-1649.44.4.997

Asscheman, J. S., He, J., Koot, S., Buil, J. M., Krabbendam, L., & van Lier, P. A. (2020). Classroom peer preferences and the development of sharing behavior with friends and others. *International Journal of Behavioral Development, 44*(5), 412-423. https://doi.org/10.1177/0165025420911094

Badre D. (2011). Defining an ontology of cognitive control requires attention to component interactions. *Topics in cognitive science, 3*(2), 217-221. https://doi.org/10.1111/j.1756-8765.2011.01141.x

Baijot, S., Cevallos, C., Zarka, D., Leroy, A., Slama, H., Colin, C., Deconink, N., Dan, B., & Cheron, G. (2017). EEG dynamics of a go/nogo task in children with ADHD. *Brain sciences, 7*(12), 167. https://doi.org/10.3390/brainsci7120167

Baker, S. L., Heinrichs, N., Kim, H. J., & Hofmann, S. G. (2002). The Liebowitz social anxiety scale as a self-report instrument: a preliminary psychometric analysis. *Behaviour research and therapy, 40*(6), 701-715. https://doi.org/10.1016/S0005-7967(01)00060-2

Baldwin, S. A., Larson, M. J., & Clayson, P. E. (2015). The dependability of electrophysiological measurements of performance monitoring in a clinical sample: A generalizability and decision analysis of the ERN and Pe. *Psychophysiology, 52*(6), 790-800. https://doi.org/10.1111/psyp.12401

Balogh, L., & Czobor, P. (2016). Post-error slowing in patients with ADHD: A meta-analysis. *Journal of Attention Disorders, 20*(12), 1004-1016. https://doi.org/10.1177/1087054714528043

Balogh, L., Kakuszi, B., Papp, S., Tombor, L., Bitter, I., & Czobor, P. (2017). Neural correlates of error monitoring in adult attention deficit hyperactivity disorder after failed inhibition in an emotional go/no-go task. *Journal of Neuropsychiatry and Clinical Neuroscience, 29*(4), 326-333. https://doi.org/10.1176/appi.neuropsych.16100183

Balsters, J. H., Apps, M. A. J., Bolis, D., Lehner, R., Gallagher, L., & Wenderoth, N. (2017). Disrupted prediction errors index social deficits in autism spectrum disorder. *Brain, 140*(1), 235-246, https://doi.org/10.1093/brain/aww287

Balsters, J. H., Laird, A. R., Fox, P. T., & Eickhoff, S. B. (2014). Bridging the gap between functional and anatomical features of cortico-cerebellar circuits using meta-analytic connectivity modeling. *Human brain mapping, 35*(7), 3152-3169. https://doi.org/10.1002/hbm.22392

Balzus, L., Klawohn, J., Elsner, B., Schmidt, S., Brandt, S. A., & Kathmann, N. (2022). Non-invasive brain stimulation modulates neural correlates of performance monitoring in patients with obsessive-compulsive disorder. *NeuroImage: Clinical, 35*, 103113. https://doi.org/10.1016/j.nicl.2022.103113

Barker, T. V., Troller-Renfree, S. V., Bowman, L. C., Pine, D. S., & Fox, N. A. (2018). Social influences of error monitoring in adolescent girls. *Psychophysiology, 55*(9), e13089. https://doi.org/10.1111/psyp.13089

Barker, T. V., Troller-Renfree, S., Pine, D. S., & Fox, N. A. (2015). Individual differences in social anxiety affect the salience of errors in social contexts. *Cognitive, Affective, & Behavioral Neuroscience, 15*(4), 723-735. https://doi.org/10.3758/s13415-015-0360-9

Bartholow, B. D., Pearson, M. A., Dickter, C. L., Sher, K. J., Fabiani, M., & Gratton, G. (2005). Strategic control and medial frontal negativity: Beyond errors and response conflict. *Psychophysiology, 42*(1), 33-42. https://doi.org/10.1111/j.1469-8986.2005.00258.x

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychological bulletin, 117*(3), 497. https://doi.org/10.1037/0033-2909.117.3.497

Beatty, P. J., Buzzell, G. A., Roberts, D. M., & McDonald, C. G. (2020). Contrasting time and frequency domains: ERN and induced theta oscillations differentially predict post-error behavior.

Cognitive, Affective, & Behavioral Neuroscience, 20, 636-647. https://doi.org/10.3758/s13415-020-00792-7

Becker, J. B., & Hu, M. (2008). Sex differences in drug abuse. *Frontiers in Neuroendocrinology, 29*(1), 36-47. https://doi.org/10.1016/j.yfrne.2007.07.003

Behnsen, P., Buil, M., Koot, S., Huizink, A., & van Lier, P. (2018). Classroom social experiences in early elementary school relate to diurnal cortisol levels. *Psychoneuroendocrinology, 87*, 1-8. https://doi.org/10.1016/j.psyneuen.2017.09.025

Bellaïche, L., Asthana, M., Ehlis, A.-C., Polak, T., & Herrmann, M. J. (2013). The modulation of error processing in the medial frontal cortex by transcranial direct current stimulation. *Neuroscience Journal, 2013,* 1-10. https://doi.org/10.1155/2013/187692

Bentler, P. M., & Bonnett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*(3), 588-606. https://doi.org/10.1037/0033-2909.88.3.588

Berger, A. M., & Davelaar, E. J. (2018). Frontal alpha oscillations and attentional control: a virtual reality neurofeedback study. *Neuroscience, 378*, 189-197. https://doi.org/10.1016/j.neuroscience.2017.06.007

Bernat, E. M., Nelson, L. D., Steele, V. R., Gehring, W. J., & Patrick, C. J. (2011). Externalizing psychopathology and gain-loss feedback in a simulated gambling task: Dissociable components of brain response revealed by time-frequency analysis. *Journal of Psychopathology and Clinical Science, 120*(2), 352. https://doi.org/10.1037/a0022124

Bernoster, I., de Groot, K., Wieser, M. J. Thurik, R., & Franken, I. H. A. (2019). Birds of a feather flock together: Evidence of prominent correlations within but not between self-report, behavioral and electrophysiological measures of impulsivity. *Biological Psychology, 145*, 112-123. https://doi.org/10.1016/j.biopsycho.2019.04.008

Bernstein, P. S., Scheffers, M. K., & Coles, M. G. H. (1995). "Where did I go wrong?" A psychophysiological analysis of error detection. *Journal of Experimental Psychology. Human Perception and Performance, 21*, 1312-1322. https://doi.org/10.1037//0096-1523.21.6.1312

Beste, C., Domschke, K., Kolev, V., Yordanova, J., Baffa, A., Falkenstein, M., & Konrad, C. (2010). Functional 5-HT1a receptor polymorphism selectively modulates error-specific subprocesses of performance monitoring. *Human Brain Maping, 31*(4), 621-630. https://doi.org/10.1002/hbm.20892

Blackhart, G. C., Nelson, B. C., Knowles, M. L., & Baumeister, R. F. (2009). Rejection elicits emotional reactions but neither causes immediate distress nor lowers self-esteem: A meta-analytic review of 192 studies on social exclusion. *Personality and Social Psychology Review, 13*(4), 269-309. https://doi.org/10.1177/1088868309346065

Bloemen, A.J.P., Oldehinkel, A.J., Laceulle, O.M., Ormel, J., Rommelse, N.N.J., & Hartman, C.A. (2018). The association between executive functioning and psychopathology: General or specific? *Psychological Medicine, 48*(11), 1787-1794. https://doi.org/10.1017/S0033291717003269

Boen, R., Quintana, D. S., Ladouceur, C. D., & Tamnes, C. K. (2022). Age-related differences in the error-related negativity and error positivity in children and adolescents are moderated by sample and methodological characteristics: A meta-analysis. *Psychophysiology, 59*(6), 1-24. https://doi.org/10.1111/psyp.14003

Boer, M., van Dorsselaer, S., de Looze, M., de Roos, S., Brons, H., van den Eijnden, R., Monshouwer, K., Huijnk, W., Ter Bogt, T., Vollebergh, W., & Stevens, G. (2021). HBSC 2021: Gezondheid en Welzijn van Jongeren in Nederland. https://www.trimbos.nl/wp-

content/uploads/2022/09/AF2022-HBSC-2021-Gezondheid-en-welzijn-van-jongeren-in-Nederland.pdf

Boksem, M. A., Tops, M., Wester, A. E., Meijman, T. F., & Lorist, M. M. (2006). Error-related ERP components and individual differences in punishment and reward sensitivity. *Brain Research, 1101*(1), 92-101. https://doi.org/10.1016/j.brainres.2006.05.004

Boldt, A., & Yeung, N. (2015). Shared neural markers of decision confidence and error detection. *Journal of Neuroscience, 35*(8), 3478-3484. https://doi.org/10.1523/JNEUROSCI.0797-14.2015

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108*(3), 624-652. https://doi.org/10.1037/0033-295X.108.3.624

Botvinick, M. M., Cohen, J. D., & Carter, C. C. (2004). Conflict monitoring and anterior cingulate cortex: An update. *Trends in Cognitive Sciences, 8*(12), 539-546. https://doi.org/10.1016/j.tics.2004.10.003

Braver, T. S. (2012). The variable nature of cognitive control: a dual mechanisms framework. *Trends in cognitive sciences, 16*(2), 106-113. https://doi.org/10.1016/j.tics.2011.12.010

Brazil, I. A., de Bruin, E. R. A., Bulten, B. H., von Borries, A. K. L., van Lankveld, J. J. D. M., Buitelaar, J. K., & Verkes, R. J. (2009). Early and late components of error monitoring in violent offenders with psychopathy. *Biological Psychiatry, 65*, 137-143. https://doi.org/10.1016/j.biopsych.2008.08.011

Brooker, R. J., & Buss, K. A. (2014). Toddler fearfulness is linked to individual differences in error-related negativity during preschool. *Developmental Neuropsychology, 39*(1), 1-8. https://doi.org/10.1080/87565641.2013.826661

Brooker, R. J., Buss, K. A., & Dennis, T. A. (2011). Error-monitoring brain activity is associated with affective behaviors in young children. *Developmental Cognitive Neuroscience, 1*(2), 141-152. https://doi.org/10.1016/j.dcn.2010.12.0

Brooks, S. J., Lochner, C., Shoptaw, S., & Stein, D. J. (2017). Using the research domain criteria (RDoC) to conceptualize impulsivity and compulsivity in relation to addiction. *Progress in brain research, 235*, 177-218. https://doi.org/10.1016/bs.pbr.2017.08.002

Brown, J. W., & Braver, T. S. (2005). Learned predictions of error likelihood in the anterior cingulate cortex. *Science, 307*(5712), 1118-1121. https://doi.org/10.1126/science.1105783

Burgio-Murphy, A., Klorman, R., Shaywitz, S. E., Fletcher, J. M., Marchione, K. E., Holahan, J., Stuebin, K. K., Thacher, J. E., & Shaywitz, B. A. (2007). Error-related event-related potentials in children with attention-deficit hyperactivity disorder, oppositional defiant disorder, reading disorder, and math disorder. *Biological Psychology, 75*(1), 75-86. https://doi.org/10.1016/j.biopsycho.2006.12.003

Buss, K. A., Cho, S., Morales, S., McDoniel, M., Webb, A. F., Schwartz, A., Cole, P. M., Dorn, L. D., Gest, S., & Teti, D. M. (2021). Toddler dysregulated fear predicts continued risk for social anxiety symptoms in early adolescence. *Development and Psychopathology, 33*(1), 252-263. https://doi.org/10.1017/S0954579419001743

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience, 14*(5), 365-376. https://doi.org/10.1038/nrn3475

Buzzell, G. A., Barker, T. V., Troller-Renfree, S. V., Bernat, E. M., Bowers, M. E., Morales, S., Bowman, L. C., Henderson, H. A., Pine, D. S., & Fox, N. A. (2019). Adolescent cognitive control, theta oscillations, and social observation. *NeuroImage, 198*, 13-30. https://doi.org/10.1016/j.neuroimage.2019.04.077

Buzzell, G. A., Morales, S., Bowers, M. E., Troller-Renfree, S. V., Chronis-Tuscano, A., Pine, D. S., Henderson, H. A., & Fox, N. A. (2021). Inhibitory control and set shifting describe different pathways from behavioral inhibition to socially anxious behavior. *Developmental Science, 24*(1), e13040. https://doi.org/10.1111/desc.13040

Buzzell, G. A., Richards, J. E., White, L. K., Barker, T. V., Pine, D. S., & Fox, N. A. (2017). Development of the error-monitoring system from ages 9-35: Unique insight provided by MRI-constrained source localization of EEG. *Neuroimage*, *157*, 13-26. https://doi.org/10.1016/j.neuroimage.2017.05.045

Buzzell, G. A., Troller-Renfree, S. V., Barker, T. V., Bowman, L. C., Chronis-Tuscano, A., Henderson, H. A., Kagan, J., Pine, D. S., & Fox, N. A. (2017). A neurobehavioral mechanism linking behaviorally inhibited temperament and later adolescent social anxiety. *Journal of the American Academy of Child & Adolescent Psychiatry, 56*(12), 1097-1105. https://doi.org/10.1016/j.jaac.2017.10.007

Calkins, S. D., Fox, N. A., & Marshall, T. R. (1996). Behavioral and physiological antecedents of inhibited and uninhibited behavior. *Child development, 67*(2), 523-540. https://doi.org/10.1111/j.1467-8624.1996.tb01749.x

Campanella, S., Absil, J., Carbia Sinde, C., Schroder, E., Peigneux, P., Bourguignon, M., Petieau., M., Metens, T., Nouali, M., Goldman, S., Cheron, G., Verbanck, P., & De Tiège, X. (2017). Neural correlates of correct and failed response inhibition in heavy versus light social drinkers: An fMRI study during a go/no-go task by healthy participants. *Brain Imaging and Behavior, 11*(6), 1796-1811. https://doi.org/10.1007/s11682-016-9654-y

Cárdenas, E. F., Hill, K. E., Estes, E., Ravi, S., Molnar Jr, A. E., Humphreys, K. L., & Kujawa, A. (2023). Neural and behavioral indicators of cognitive control in preschoolers with and without prenatal opioid exposure. *Child Neuropsychology*, 1-19. https://doi.org/10.1080/09297049.2023.2196397

Cardinale, E. M., Bezek, J., Morales, S., Filippi, C., Smith, A. R., Haller, S., Valadez, E. A., Harrewijn, A., Philips, D., Chronis-Tuscano, A., Brotman, M. A., Fox, N. A., Pine, D., S., Leibenluft, E. & Kircanski, K. (2023). Cross-sectional and longitudinal associations of anxiety and irritability with adolescents' neural responses to cognitive conflict. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 8*(4), 436-444. https://doi.org/10.1016/j.bpsc.2022.03.007

Carmi, L., Alyagon, U., Barnea-Ygael, N., Zohar, J., Dar, R., & Zangen, A. (2018). Clinical and electrophysiological outcomes of deep TMS over the medial prefrontal and anterior cingulate cortices in OCD patients. *Brain Stimulation, 11*(1), 158-165. https://doi.org/10.1016/j.brs.2017.09.004

Caspi, A. (1996). Behavioral observations at age 3 years predict adult psychiatric disorders: Longitudinal evidence from a birth cohort. *Archives of General Psychiatry, 53*(11), 1033. https://doi.org/10.1001/archpsyc.1996.01830110071009

Caspi, A., & Moffitt, T. E. (2018). All for one and one for all: Mental disorders in one dimension. *American Journal of Psychiatry, 175*(9), 831-844. https://doi.org/10.1176/appi.ajp.2018.17121383

Caspi, A., Harrington, H., Milne, B., Amell, J. W., Theodore, R. F., & Moffitt, T. E. (2003). Children's behavioral styles at age 3 are linked to their adult personality traits at age 26: From age 3 to age 26. *Journal of Personality, 71*(4), 495-514. https://doi.org/10.1111/1467-6494.7104001

Caspi, A., Houts, R. M., Belsky, D. W., Goldman-Mellor, S. J., Harrington, H., Israel, S., Meier, M. H., Ramrakha, S., Shalev, I., Poulton, R., & Moffitt, T. E. (2013). The p Factor. *Clinical Psychological Science, 2*(2), 119-137. https://doi.org/10.1177/2167702613497473

Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in cognitive sciences, 18*(8), 414-421. https://doi.org/10.1016/j.tics.2014.04.012

Cavanagh, J. F., & Shackman, A. J. (2015). Frontal midline theta reflects anxiety and cognitive control: meta-analytic evidence. *Journal of Physiology-Paris, 109*(1-3), 3-15. https://doi.org/10.1016/j.jphysparis.2014.04.003

Cavanagh, J. F., Cohen, M. X., & Allen, J. J. B. (2009). Prelude to and resolution of an error: EEG phase Synchrony reveals cognitive control dynamics during action monitoring. *Journal of Neuroscience, 29*(1), 98-105. https://doi.org/10.1523/JNEUROSCI.4137-08.2009

Cavanagh, J. F., Meyer, A., & Hajcak, G. (2017). Error-specific cognitive control alterations in generalized anxiety disorder. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 2*(5), 413-420, https://doi.org/10.1016/j.bpsc.2017.01.004.

Chang, W-P, Davies, P. L., & Gavin, W. (2009). Error monitoring in college students with attention-deficit/hyperactivity disorder. *Journal of Psychophysiology, 23*(3), 113-125. https://doi.org/10.1027/0269-8803.23.3.113

Checa, P., Castellanos, M. C., Abundis-Gutiérrez, A., & Rueda, M. R. (2014). Development of neural mechanisms of conflict and error processing during childhood: Implications for self-regulation. *Frontiers in Psychology, 5*, 326. https://doi.org/10.3389/fpsyg.2014.00326

Chen, G., Saad, Z. S., Britton, J. C., Pine, D. S., Cox, R. W. (2013). Linear mixed-effects modeling approach to fMRI group analysis. *NeuroImage 73*,176-190. http://doi.org/10.1016/j.neuroimage.2013.01.047

Chen, H., Jiang, H., Guo, Q., Du, J., Wang, J., & Zhao, M. (2013). Case-control study of error-related negativity among males with heroin dependence undergoing rehabilitation. *Shanghai Archives of Psychiatry, 25*(3), 141-148. https://doi.org/10.3969/j.issn.1002-0829.2013.03.003

Chi, P. H., & Deldin, P. J. (2007). Neural evidence for enhanced error detection in major depressive disorder. *American Journal of Psychiatry, 164*, 608-616. https://doi.org/10.1176/appi.ajp.164.4.608

Chronis-Tuscano, A., Degnan, K. A., Pine, D. S., Perez-Edgar, K., Henderson, H. A., Diaz, Y., Raggi, V. L., & Fox, N. A. (2009). Stable early maternal report of behavioral inhibition predicts lifetime social anxiety disorder in adolescence. *Journal of the American Academy of Child & Adolescent Psychiatry, 48*(9), 928-935. https://doi.org/10.1097/CHI.0b013e3181ae09df

Clauss, J. A., & Blackford, J. U. (2012). Behavioral inhibition and risk for developing social anxiety disorder: A meta-analytic study. *Journal of the American Academy of Child & Adolescent Psychiatry, 51*(10), 1066-1075.e1. https://doi.org/10.1016/j.jaac.2012.08.002

Clawson, A., Clayson, P. E., & Larson, M. J. (2013). Cognitive control adjustments and conflict adaptation in major depressive disorder. *Psychophysiology, 50*(8), 711-721. https://doi.org/10.1111/psyp.12066

Clayson, P. E. (2020). Moderators of the internal consistency of error-related negativity scores: A meta-analysis of internal consistency estimates. *Psychophysiology, 57*(8), e13583. https://doi.org/10.1111/psyp.13583

Clayson, P. E., & Larson, M. J. (2013). Psychometric properties of conflict monitoring and conflict adaptation indices: Response time and conflict N 2 event-related potentials. *Psychophysiology, 50*(12), 1209-1219. https://doi.org/10.1111/psyp.12138

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

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

Clayson, P. E., Keil, A., & Larson, M. J. (2022). Open science in human electrophysiology. *International Journal of Psychophysiology, 174*, 43-46. https://doi.org/10.1016/j.ijpsycho.2022.02.002

Clayson, P. E., Rocha, H. A., Baldwin, S. A., Rast, P., & Larson, M. J. (2022). Understanding the error in psychopathology: Notable intraindividual differences in neural variability of performance monitoring. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 7*(6), 555-565. https://doi.org/10.1016/j.bpsc.2021.10.016

Cohen J. *Statistical power analysis for the behavioral sciences.* New York: Academic Press; 1977

Cole, S. L., Cibrian, E., Mirzadegan, I. A., & Meyer, A. (2022). The impact of punishment on error-related brain activity in children. *Developmental Psychobiology, 64*(7), e22318. https://doi.org/10.1002/dev.22318

Coles, M. G., Scheffers, M. K., & Holroyd, C. B. (2001). Why is there an ERN/Ne on correct trials? Response representations, stimulus-related components, and the theory of error-processing. *Biological psychology, 56*(3), 173-189. https://doi.org/10.1016/S0301-0511(01)00076-X

Compton, R. J., Heaton, E., & Ozer, E. (2017). Intertrial interval duration affects error monitoring. *Psychophysiology, 54*(8), 1151-1162. https://doi.org/10.1111/psyp.12877

Compton, R. J., Lin, M., Vargas, G., Carp, J., Fineman, S. L., & Quandt, L. C. (2008). Error detection and posterror behavior in depressed undergraduates. *Emotion, 8*(1), 58–67. https://doi.org/10.1037/1528-3542.8.1.58

Conners, C. K. (1985). The computerized continuous performance test. *Psychopharmacology Bulletin, 21*, 891-892.

Connolly, C. G., Foxe, J. J., Nierenberg, J., Shpaner, M., & Garavan, H. (2012). The neurobiology of cognitive control in successful cocaine abstinence. *Drug and Alcohol Dependence, 121* (1-2), 45-53. https://doi.org/10.1016/j.drugalcdep.2011.08.007

Contreras-Huerta, L. S., Pisauro, M. A., & Apps, M. A. (2020). Effort shapes social cognition and behaviour: A neuro-cognitive framework. *Neuroscience & Biobehavioral Reviews, 118*, 426-439. https://doi.org/10.1016/j.neubiorev.2020.08.003

Conway, C. C., Forbes, M. K., Forbush, K. T., Fried, E. I., Hallquist, M. N., Kotov, R., Mullins-Sweatt, S. N., Shackman, A. J., Skodol, A. E., South, S. C., Sunderland, M., Waszczuk, M. A., Zald, D. H., Afzali, M. H., Bronovalova, M. A., Carragher, N., Docherty, A. R., Jonas, K. G., Krueger, R. F., ... Eaton, N. R. (2019). A hierarchical taxonomy of psychopathology can transform mental health research. *Perspectives on psychological science, 14*(3), 419-436. https://doi.org/10.1177/1745691618810696

Cooper, P. S., Karayanidis, F., McKewen, M., McLellan-Hall, S., Wong, A. S., Skippen, P., & Cavanagh, J. F. (2019). Frontal theta predicts specific cognitive control-induced behavioural changes beyond general reaction time slowing. *Neuroimage, 189*, 130-140. https://doi.org/10.1016/j.neuroimage.2019.01.022

Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research, 29*(3), 162-173. https://doi.org/10.1006/cbmr.1996.0014

Cox, R. W., Chen, G., Glen, D. R., Reynolds, R. C., & Taylor, P. A. (2017). FMRI clustering in AFNI: False-positive rates redux. *Brain Connectivity, 7*(3), 152-171. https://doi.org/10.1089/brain.2016.0475

Crane, N. A., Gorka, S. M., Burkhouse, K. L., Afshar, K., Greenstein, J. E., Aase, D. M., Proescher, E., Schroth, C., Kennedy, A. E., & Phan, K. L. (2018). Neural response to errors is associated with problematic alcohol use over time in combat-exposed returning veterans: An event-related potential study. *Drug and Alcohol Dependence, 183*, 155-161. https://doi.org/10.1016/j.drugalcdep.2017.11.008

Crone, E. A., & Dahl, R. E. (2012). Understanding adolescence as a period of social–affective engagement and goal flexibility. *Nature reviews neuroscience, 13*(9), 636-650. https://doi.org/10.1038/nrn3313

Cuthbert B. N. (2022). Research Domain Criteria (RDoC): Progress and potential. *Current Directions in Psychological Science, 31*(2), 107-114. https://doi.org/10.1177/09637214211051363

Cuthbert, B. (2014) The RDoC framework: facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology. *World Psychiatry, 13*(1), 28-35. https://doi.org/10.1002/wps.20087

Czermainski, F. R., Willhelm, A. R., Santos, Á. Z., Pachado, M. P., & de Almeida, R. M. M. (2017). Assessment of inhibitory control in crack and/or cocaine users: A systematic review. *Trends in Psychiatry and Psychotherapy, 39*(3), 216-225. https://doi.org/10.1590/2237-6089-2016-0043

Czobor, P., Kakuszi, B., Németh, K., Balogh, L., Papp, S., Tombor, L., & Bitter, I. (2017). Electrophysiological indices of aberrant error-processing in adults with ADHD: A new region on interest. *Brain Imaging and Behavior, 11*, 161-162. https://doi.org/10.1007/s11682-016-9610-x

Damaso, K. A., Williams, P. G., & Heathcote, A. (2022). What Happens After a Fast Versus Slow Error, and How Does It Relate to Evidence Accumulation?. *Computational Brain & Behavior, 5*(4), 527-546. https://doi.org/10.1007/s42113-022-00137-2

Danielmeier, C., & Ullsperger, M. (2011). Post-error adjustments. *Frontiers in Psychology, 2,* 1-10. https://doi.org/10.3389/fpsyg.2011.00233

Davies, P. L., Segalowitz, S. J., & Gavin, W. J. (2004). Development of response-monitoring ERPs in 7- to 25-year-olds. *Developmental Neuropsychology, 25*(3), 355-376. https://doi.org/10.1207/s15326942dn2503_6

Davies, P. L., Segalowitz, S. J., Dywan, J., & Pailing, P. E. (2001). Error-negativity and positivity as they relate to other ERP indices of attentional control and stimulus processing. *Biological Psychology, 56*(3), 191-206. https://doi.org/10.1016/S0301-0511(01)00080-1

de Bruijn, E. R., Hulstijn, W., Verkes, R. J., Ruigt, G. S., & Sabbe, B. G. (2004). Drug-induced stimulation and suppression of action monitoring in healthy volunteers. *Psychopharmacology, 177*(1-2), 151-160. https://doi.org/10.1007/s00213-004-1915-6

de Sonneville, L.M.J. (1996). Amsterdamse Neuropsychologische Taken (ANT): Een computergestuurd testprogramma. *Psychologie & Computers, 13*, 94-102.

de Wilde, A., Koot, H. M., & van Lier, P. A. C. (2016). Developmental links between children's working memory and their social relations with teachers and peers in the early school years. *Research on Child and Adolescent Psychopathology, 44*(1), 19-30. https://doi.org/10.1007/s10802-015-0053-4

Debener, S., Ullsperger, M., Siegel, M., Fiehler, K., Von Cramon, D. Y., & Engel, A. K. (2005). Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. *Journal of Neuroscience, 25*(50), 11730-11737. https://doi.org/10.1523/JNEUROSCI.3286-05.2005

Dehaene, S., Posner, M. I., & Tucker, D. M. (1994). Localization of a neural system for error detection and compensation. *Society, 5*(5), 303-305. https://doi.org/10.1111/j.1467-9280.1994.tb00630.x

Denervaud, S., Hess, A., Sander, D., & Pourtois, G. (2021). Children's automatic evaluation of self-generated actions is different from adults. *Developmental Science, 24*, e13045. https://doi.org/10.1111/desc.13045

Derevensky, J. L. (2019). Behavioral addictions: Some developmental considerations. *Current Addiction Reports, 6*, 313-322. https://doi.org/10.1007/s40429-019-00257-z

Desender, K., Ridderinkhof, K. R., & Murphy, P. R. (2021). Understanding neural signals of post-decisional performance monitoring: An integrative review. *Elife, 10*, e67556. https://doi.org/10.7554/eLife.67556

Dhar, M., & Pourtois, G. (2011). Early error detection is generic, but subsequent adaption to errors is not: evidence from ERPs. *Neuropsychologia, 49*(5), 1236-1245. https://doi.org/10.1016/j.neuropsychologia.2011.01.006

Di Gregorio, F., Maier, M. E., & Steinhauser, M. (2018). Errors can elicit an error positivity in the absence of an error negativity: Evidence for independent systems of human error monitoring. *Neuroimage, 172*, 427-436. https://doi.org/10.1016/j.neuroimage.2018.01.081

Diehl, M. M., Lempert, K. M., Parr, A. C., Ballard, I., Steele, V. R., & Smith, D. V. (2018). Toward an integrative perspective on the neural mechanisms underlying persistent maladaptive behaviors. *The European journal of neuroscience, 48*(3), 1870-1883. https://doi.org/10.1111/ejn.14083

Dignath, D., Eder, A. B., Steinhauser, M., & Kiesel, A. (2020). Conflict monitoring and the affective-signaling hypothesis - An integrative review. *Psychonomic Bulletin & Review, 27*, 193-216. https://doi.org/10.3758/s13423-019-01668-9

Dikman, Z. V., & Allen, J. J. (2000). Error monitoring during reward and avoidance learning in high- and low-socialized individuals. *Psychophysiology, 37*(1), 43-54. https://doi.org/10.1111/1469-8986.3710043

Dilalla, L. F., Kagan, J., & Reznick, J. S. (1994). Genetic etiology of behavioral inhibition among 2-year-old children. *Infant Behavior and Development, 17*(4), 405-412. https://doi.org/10.1016/0163-6383(94)90032-9

Dong, G., Lin, X., & Potenza, M. N. (2015). Decreased functional connectivity in an executive control network is related to impaired executive function in Internet gaming disorder. *Progress in Neuropsychopharmacology and Biological Psychiatry, 57*, 76-85. https://doi.org/10.1016/j.pnpbp.2014.10.012

Dong, G., Shen, Y., Huang, J., & Du, X. (2013). Impaired error-monitoring function in people with internet addiction disorder: an event-related fMRI study. *European addiction research, 19*(5), 269-275. https://doi.org/10.1159/000346783

Dong, G., Wang, L., Du, X., & Potenza, M. (2018). Gender-related differences in neural responses to gaming cues before and after gaming: Implications for gender-specific vulnerabilities to Internet gaming disorder. *Social Cognitive and Affective Neuroscience,13*(11), 1203-1214. https://doi.org/10.1093/scan/nsy084

Donoghue, T. & Voytek, B. (2021). Automated meta-analysis of the event-related potential (ERP) literature. *Scientific Reports, 12*, 1867. https://doi.org/10.1038/s41598-022-05939-9

Dotson, V. M., McClintock, S. M., Verhaeghen, P., Kim, J. U., Draheim, A. A., Syzmkowicz, S. M., Gradone, A., M., Bogoian, H. R., & Wit, L. D. (2020). Depression and cognitive control across the lifespan: a systematic review and meta-analysis. *Neuropsychology review, 30*, 461-476. https://doi.org/10.1007/s11065-020-09436-6

Dutilh, G., Forstmann, B. U., Vandekerckhove, J., & Wagenmakers, E. J. (2013). A diffusion model account of age differences in posterror slowing. *Psychology and Aging, 28*(1), 64-76. https://doi.org/10.1037/a0029875

Dutilh, G., Vandekerckhove, J., Forstmann, B. U., Keuleers, E., Brysbaert, M., & Wagenmakers, E. J. (2012). Testing theories of post-error slowing. *Attention, Perception, and Psychophysics, 74*(2), 454-465. https://doi.org/10.3758/s13414-011-0243-2

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*(2). 455-463. https://doi.org/10.1111/j.0006-341x.2000.00455.x

Easdon, C., Izenberg, A., Armilio, M. L., Yu, H., & Alain, C. (2005). Alcohol consumption impairs stimulus- and error-related processing during a Go/No-Go Task. *Cognitive Brain Research, 25*(3), 873-883. https://doi.org/10.1016/j.cogbrainres.2005.09.009

Eaton, N. R., Keyes, K. M., Krueger, R. F., Balsis, S., Skodol, A. E., Markon, K. E., Grant, B. F., & Hasin, D. S. (2012). An invariant dimensional liability model of gender differences in mental disorder prevalence: Evidence from a national sample. *Journal of Psychopathology and Clinical Science, 121*(1), 282-288. https://doi.org/10.1037/a0024780

Efferson, L. M., & Glenn, A. L. (2018). Examining gender differences in the correlates of psychopathy: A systematic review of emotional, cognitive, and morality-related constructs. *Aggressive and Violent Behavior, 41*, 48-61. https://doi.org/10.1016/j.avb.2018.05.009

Egger, M., Smith, G. D., Schneider, M., & Minder, G. (1997). Bias in meta-analysis detected by a simple graphical test. *BMI Clinical Research, 315*(7109), 629-634. https://doi.org/10.1136/bmj.315.7109.629

Ehlis, A. C., Deppermann, S., & Fallgatter, A. J. (2018). Performance monitoring and post-error adjustments in adults with attention-deficit/hyperactivity disorder: an EEG analysis. *Journal of Psychiatry and Neuroscience, 43*(6), 396-406. https://doi.org/10.1503/jpn.170118

Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences, 113*(28), 7900-7905. https://doi.org/10.1073/pnas.1602413113

Endrass, T., Reuter, B., & Kathmann, N. (2007). ERP correlates of conscious error recognition: Aware and unaware errors in an antisaccade task. *European Journal of Neuroscience, 26*(6), 1714-1720. https://doi.org/10.1111/j.1460-9568.2007.05785.x

Endrass, T., Riesel, A., Kathmann, N., & Buhlmann, U. (2014). Performance monitoring in obsessive–compulsive disorder and social anxiety disorder. *Journal of abnormal psychology, 123*(4), 705. https://doi.org/10.1037/ABN0000012

Endrass, T., Schuermann, B., Kaufmann, C., Spielberg, R., Kniesche, R., & Kathmann, N. (2010). Performance monitoring and error significance in patients with obsessive-compulsive disorder. *Biological psychology, 84*(2), 257-263. https://doi.org/10.1016/j.biopsycho.2010.02.002

Epstein, J. N., Erkanli, A., Conners, C. K., Klaric, J., Costello, J. E., & Angold, A. (2003). Relations between continuous performance test performance measures and ADHD behaviors. *Research on Child and Adolescent Psychopathology, 31*, 543-554. https://doi.org/10.1023/A:1025405216339

Erasmus Medical Center. (2000). Problem behavior at school interview. Rotterdam: Department of Child and Adolescent Psychiatry.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Pscyhophysics, 16*, 143-149.

Euser, A. S., Arends, L. R., Evans, B. E., Greaves-Lord, K., Huizink, A. C., & Franken, I. H. A. (2012). The P300 event-related brain potential as a neurobiological endophenotype for substance use disorders: A meta-analytic investigation. *Neuroscience and Biobehavioral Review, 36*(1), 572-603. https://doi.org/10.1016/j.neubiorev.2011.09.002

Euser, A. S., Evans, B. E., Greaves-Lord, K., Huizink, A. C., & Franken, I. H. A. (2013). Diminished error-related brain activity as a promising endophenotype for substance-use disorders: Evidence from high-risk offspring. *Addiction Biology, 18*(6), 970-984. https://doi.org/10.1111/adb.12002

Fairchild, G., Hawes, D.J., Frick, P.J., Copeland, W. E., Odgers, C. L., Franke, B., Freitag, C. M., de Brito, S. A. (2019). Conduct disorder. *Nature Reviews Disease Primers, 5*, 43. https://doi.org/10.1038/s41572-019-0095-y

Falkenstein, M. (2004). ERPcorrelates of erroneous performance. Ullsperger, M., & Falkenstein, M. (Eds.). (2004). *Errors, conflicts, and the brain: current opinions on performance monitoring; proceedings of the conference held in Dortmund, Germany, on July 3 - 5, 200*. Leipzig: Max Planck Institute for Human Cognitive and Brain Sciences.

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

Falkenstein, Michael, Hoormann, J., Christ, S., & Hohnsbein, J. (2000). ERP components on reaction errors and their functional significance: A tutorial. *Biological Psychology, 51*(2-3), 87-107. https://doi.org/10.1016/S0301-0511(99)00031-9

Ficarella, S. C., Rochet, N., & Burle, B. (2019). Becoming aware of subliminal responses: An EEG/EMG study on partial error detection and correction in humans. *Cortex, 120*, 443-456. https://doi.org/10.1016/j.cortex.2019.07.007

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. (1995). The structured clinical interview for DSM-III-R personality disorders (SCID-II). Part I: Description. *Journal of Personality disorders, 9*(2), 83-91.

Fischer, A. G., Danielmeier, C., Villringer, A., Klein, T. A., & Ullspreger, M. (2016). Gender influences on brain responses to errors and post-error adjustments. *Scientific Reports, 6*, 24435. https://doi.org/10.1038/srep24435

Fischer, A. G., Nigbur, R., Klein, T. A., Danielmeier, C., & Ullsperger, M. (2018). Cortical beta power reflects decision dynamics and uncovers multiple facets of post-error adaptation. *Nature Communications 9*, 5038. https://doi.org/10.1038/s41467-018-07456-8

Forbes, M. K., Neo, B., Nezami, O. M., Fried, E. I., Faure, K., Michelsen, B., Twose, M., & Dras, M. (2023). *Elemental psychopathology: Distilling constituent symptoms and patterns of repetition in the diagnostic criteria of the DSM-5*. PsyArXiv. https://doi.org/10.31234/osf.io/u56p2

Ford, J. M. (1999). Schizophrenia: The broken P300 and beyond. *Psychophysiology, 36*(6), 667-682.

Fox, N. A., Henderson, H. A., Rubin, K. H., Calkins, S. D., & Schmidt, L. A. (2001). Continuity and discontinuity of behavioral inhibition and exuberance: Psychophysiological and behavioral influences across the first four years of life. Child development, 72(1), 1-21.

Fox, N. A., Zeytinoglu, S., Valadez, E. A., Buzzell, G. A., Morales, S., & Henderson, H. A. (2023). Annual research review: Developmental pathways linking early behavioral inhibition to later anxiety. *Journal of Child Psychology and Psychiatry, 64*(4), 537-561. https://doi.org/10.1111/jcpp.13702

Franken, I. H. A., Luijten, M., van der Veen, F. M., & van Strien, J. W. (2017). Cognitive control in young heavy drinkers: An ERP study. *Drug and Alcohol Dependence, 175*, 77-83. https://doi.org/10.1016/j.drugalcdep.2017.01.036

Franken, I. H. A., Nijs, I. M. T., Toes, A., & van der Veen, F. M. (2018). Food addiction is associated with impaired performance monitoring. *Biological Psychology, 131*, 49-53. https://doi.org/10.1016/j.biopsycho.2016.07.005

Franken, I. H. A., van Strien, J. W., & Kuijpers, I. (2010). Evidence for a deficit in the salience attribution to errors in smokers. *Drug and Alcohol Dependence, 106*(2-3), 181-185. https://doi.org/10.1016/j.drugalcdep.2009.08.014

Franken, I. H. A., van Strien, J. W., Franzek, E. J., & van de Wetering, B. J. (2007). Error-processing deficits in patients with cocaine dependence. *Biological Psychology, 75*, 45–51. https://doi.org/10.1016/j.biopsycho.2006.11.003

Franken, I. H.A., Luijten, M., van der Veen, F. M., & van Strien, J. W. (2017). Cognitive control in young heavy drinkers: An ERP study. *Drug and Alcohol Dependence, 175*, 77-83. https://doi.org/10.1016/j.drugalcdep.2017.01.036

Fried, E. (2017). Moving forward: how depression heterogeneity hinders progress in treatment and research. *Expert Review of Neurotherapeutics, 17*(5), 423-425, https://doi.org/10.1080/14737175.2017.1307737

Friedman, N. P., & Robbins, T. W. (2022). The role of prefrontal cortex in cognitive control and executive function. *Neuropsychopharmacology, 47*(1), 72-89. https://doi.org/10.1038/s41386-021-01132-0

Frömer, R., & Shenhav, A. (2022). Filling the gaps: Cognitive control as a critical lens for understanding mechanisms of value-based decision-making. *Neuroscience & Biobehavioral Reviews, 134*, 104483. https://doi.org/10.1016/j.neubiorev.2021.12.006

Fuhrmann, D., Madsen, K. S., Johansen, L. B., Baaré, W. F. C., & Kievit, R. A. (2022). The midpoint of cortical thinning between late childhood and early adulthood differs between individuals and brain regions: Evidence from longitudinal modelling in a 12-wave neuroimaging sample. *NeuroImage, 261*, 119507. https://doi.org/10.1016/j.neuroimage.2022.119507

Gavin, W. J., Lin, M.-H., & Davies, P. L. (2019). Developmental trends of performance monitoring measures in 7- to 25-year-olds: Unraveling the complex nature of brain measures. *Psychophysiology, 56*(7). https://doi.org/10.1111/psyp.13365

Geburek, A. J., Rist, F., Gediga, G., Stroux, D., & Pedersen, A. (2013). Electrophysiological indices of error monitoring in juvenile and adult attention deficit hyperactivity disorder (ADHD) - A meta-analytic appraisal. *International Journal of Psychophysiology, 87*(3), 349-362. https://doi.org/10.1016/j.ijpsycho.2012.08.006

Geen, R. G. (1991). Social motivation. *Annual Review of Psychology, 42*, 377-399. https://doi.org/10.1146/annurev.ps.42.020191.002113

Geen, R. G., & Bushman, B. J. (1989). The arousing effects of social presence. *Handbook of social psychophysiology*, 261-281.

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological Science, 4*(6), 385-390. https://doi.org/10.1111/j.1467-9280.1993.tb00586.x

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (2018). The error-related negativity. *Perspectives on Psychological Science, 13*(2), 200-204. https://doi.org/10.1177/1745691617715310

Gilbertson H., Fang, L., Andrzejewski, J.A., & Carlson, J.M. (2021). Dorsal anterior cingulate cortex intrinsic functional connectivity linked to electrocortical measures of error monitoring. *Psychophysiology,58,* e13794. https://doi.org/10.1111/psyp.13794

Gilian, C. M., Fineberg, N. A., & Robbins, T. W. (2017). A trans-diagnostic perspective on obsessive-compulsive disorder. *Psychological Medicine, 47*, 1528-1548. https://doi.org/10.1017/S0033291716002786

Gladstone, G. L., Parker, G. B., Mitchell, P. B., Wilhelm, K. A., & Malhi, G. S. (2005). Relationship between self-reported childhood behavioral inhibition and lifetime anxiety disorders in a clinical sample. *Depression and anxiety, 22*(3), 103-113. https://doi.org/10.1002/da.20082

Global Health Data Exchange. Retrieved from https://ghdx.healthdata.org/

Godefroid, E., Pourtois, G., & Wiersema, J. R. (2016). Joint effects of sensory feedback and interoceptive awareness on conscious error detection: Evidence from event related brain potentials. *Biological psychology*, *114*, 49-60. https://doi.org/10.1016/j.biopsycho.2015.12.005

Goldin, P. R., Ziv, M., Jazaieri, H., Weeks, J., Heimberg, R. G., & Gross, J. J. (2014). Impact of cognitive-behavioral therapy for social anxiety disorder on the neural bases of emotional reactivity to and regulation of social evaluation. *Behaviour research and therapy*, *62*, 97-106. https://doi.org/10.1016/j.brat.2014.08.005

Goldstein, R. Z., & Volkow, N. D. (2002). Drug addiction and its underlying neurobiological basis: Neuroimaging evidence for the involvement of the frontal cortex. *American Journal of Psychiatry, 159*, 1642-1652. https://doi.org/10.1176/appi.ajp.159.10.1642

Goldstein, R. Z., & Volkow, N. D. (2011). Dysfunction of the prefrontal cortex in addiction: Neuroimaging findings and clinical implications. *Nature Reviews Neuroscience, 12*, 652-669. https://doi.org/10.1038/nrn3119

Gonthier, C., Macnamara, B. N., Chow, M., Conway, A. R., & Braver, T. S. (2016). Inducing proactive control shifts in the AX-CPT. *Frontiers in psychology, 7*, 1822. https://doi.org/10.3389/fpsyg.2016.01822

Gorday, J. Y., & Meyer, A. (2018). Linking puberty and error-monitoring: Relationships between self-reported pubertal stages, pubertal hormones, and the error-related negativity in a large sample of children and adolescents. *Developmental Psychobiology, 60*(4), 483-490. https://doi.org/10.1002/dev.21625

Gorka, S. M., Lieberman, L., Kreutzer, K. A., Carrillo, V., Weinberg, A., & Shankman, S. A. (2019). Error-related neural activity and alcohol use disorder: Differences from risk to remission. *Progress in Neuro-Pscychopharmcology and Biological Psychiatry, 92*, 271-278. https://doi.org/10.1016/j.pnpbp.2019.01.011

Gorka, S. M., MacNamara, A., Aase, D. M., Proescher, E., Greenstein, J. E., Walters, R., Passi, H., Babione, J. M., Levy, D. M., Kennedy, A. E., DiGangi, J. A., Rabinak, C. A., Schroth, C., Afshar, K., Fitzgerald, J., Hajcak, G., & Phan, K. L. (2016). Impact of alcohol use disorder comorbidity on defensive reactivity to errors in veterans with posttraumatic stress disorder. *Psychology of Addictive Behaviors, 30*(7), 733-743. https://doi.org/10.1037/adb0000196

Gratton, G., Coles, M. G., & Donchin, E. (1992). Optimizing the use of information: strategic control of activation of responses. *Journal of Experimental Psychology: General, 121*(4), 480.

Groen, Y., Wijers, A. A., Mulder, L. J. M., Waggeveld, B., Minderaa, R. B., & Althaus, M., (2008). Error and feedback processing in children with ADHD and children with autistic spectrum disorder: An EEG event-related potential study. *Clinical Neurophysiology, 119*, 2476-2493.

Groom, M. J. Cahill, J. D., Bates, A. T., Jackson, G. M., Calton, T. G., Liddle, P. F., & Hollis, C. (2010). Electrophysiological indices of abnormal error-processing in adolescents with attention deficits hyperactivity disorder (ADHD). *Journal of Child Psychology and Psychiatry, 51*(1), 66-76. https://doi.org/10.1111/j.1469-7610.2009.02128.x

Groom, M. J., Liddle, E. B., Scerif, G., Liddle, P. F., Batty, M. J., Liotti, M., & Hollis, C. P. (2013). Motivational incentives and methylphenidate enhance electrophysiological correlates of error monitoring in children with attention deficit/hyperactivity disorder. *Journal of Child Psychology and Psychiatry, 54*(8), 836-845. https://doi.org/http://dx.doi.org/10.1111/jcpp.12069

Grützmann, R., Endrass, T., Klawohn, J., & Kathmann, N. (2014). Response accuracy rating modulates ERN and Pe amplitudes. *Biological psychology, 96*, 1-7. https://doi.org/10.1016/j.biopsycho.2013.10.007

Gunther Moor, B., van Leijenhorst, L., Rombouts, S. A., Crone, E. A., & Van der Molen, M. W. (2010). Do you like me? Neural correlates of social evaluation and developmental trajectories. *Social neuroscience, 5*(5-6), 461-482. https://doi.org/10.1080/17470910903526155

Hajcak, G. (2012). What we've learned from mistakes: Insights from error-related brain activity. *Current Directions in Psychological Science, 21*(2), 101-106. https://doi.org/10.1177/0963721412436809

Hajcak, G., & Foti, D. (2008). Errors are aversive: Defensive motivation and the error-related negativity. *Psychological science, 19*(2), 103-108. https://doi.org/10.1111/j.1467-9280.2008.02053.x

Hajcak, G., & Simons, R. F. (2002). Error-related brain activity in obsessive–compulsive undergraduates. *Psychiatry research, 110*(1), 63-72. https://doi.org/10.1016/S0165-1781(02)00034-3

Hajcak, G., Klawohn, J., & Meyer, A. (2019). The utility of event-related potentials in clinical psychology. *Annual Review of Clinical Psychology, 15*, 71-95. https://doi.org/10.1146/annurev-clinpsy-050718-095457

Hajcak, G., McDonald, N., & Simons, R. F. (2003). To err is automatic: Error-related brain potentials, ANS activity, and post-error compensatory behavior. *Psychophysiology, 40*(6), 895-903. https://doi.org/10 .1111/1469-8986.00107

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

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

Hall, N. T., Schreiber, A. M., Allen, T. A., & Hallquist, M. N. (2021). Disentangling cognitive processes in externalizing psychopathology using drift diffusion modeling: Antagonism, but not disinhibition, is associated with poor cognitive control. *Journal of Personality, 89*(5), 970-985. https://doi.org/10.1111/jopy.12628

Ham, T., Leff, A., de Boissezon, X., Joffe, A., & Sharp, D. J. (2013). Cognitive control and the salience network: An investigation of error processing and effective connectivity. *Journal of Neuroscience, 33*(16), 7091-7098. https://doi.org/10.1523/JNEUROSCI.4692-12.2013

Hämmerer, D., Müller, V., & Li, S. C. (2014). Performance monitoring across the lifespan: Still maturing post-conflict regulation in children and declining task-set monitoring in older adults. *Neuroscience and Biobehavioral Reviews, 46*, 105-123. https://doi.org/10.1016/j.neubiorev.2014.06.008

Hanna, G. L., Liu, Y., Rough, H. E., Surapaneni, M., Hanna, B. S., Arnold, P. D., & Gerhing, W. J. (2020). A diagnostic biomarker for pediatric generalized anxiety disorder using the error-related negativity. *Child Psychiatry and Human Development, 51* (5), 827-838. https://doi.org/10.1007/s10578-020-01021-5

Harrer, M., Cuijpers, P., Furukawa, T.A, & Ebert, D. D. (2019). Doing Meta-Analysis in R: A Hand-on Guide. http://doi.org/doi:10.5281/zenodo.2551803

Harsay, H. A., Spaan, M., Wijnen, J. G., & Ridderinkhof, K. R. (2012). Error awareness and salience processing in the oddball task: shared neural mechanisms. Frontiers in Human Neuroscience, 6, 246. https://doi.org/10.3389/fnhum.2012.00246

Hedges, L. V. (1981). Distribution theory of Glass's estimator of effect sizes and related estimators. *Journal of Educational Statistics, 6*(2), 107-128.

Heeren, A., & McNally, R. J. (2016). An integrative network approach to social anxiety disorder: The complex dynamic interplay among attentional bias for threat, attentional control, and symptoms. *Journal of Anxiety Disorders, 42,* 95-104. https://doi.org/10.1016/j.janxdis.2016.06.009

Heimberg, R. G., Horner, K. J., Juster, H. R., Safren, S. A., Brown, E. J., Schneier, F. R., & Liebowitz, M. R. (1999). Psychometric properties of the Liebowitz social anxiety scale. *Psychological medicine, 29*(1), 199-212. https://doi.org/10.1017/S0033291798007879

Heitz, R. P. (2014). The speed-accuracy tradeoff: History, physiology, methodology, and behavior. *Frontiers in Neuroscience, 8*, 1-19. https://doi.org/10.3389/fnins.2014.00150

Henderson, H. A., Pine, D. S., & Fox, N. A. (2015). Behavioral inhibition and developmental risk: A dual-processing perspective. *Neuropsychopharmacology, 40*(1), 207-224. https://doi.org/10.1038/npp.2014.189

Herrmann, M. J., Mader, K., Schreppel, T., Jacob, C., Heine, M., Boreatti-Hümmer, A., Ehlis, A-C., Scheuerpflug, P., Pauli, P., & Fallgatter, A. J. (2010). Neural correlates of performance monitoring in adult patients with attention deficit hyperactivity disorder (ADHD), *The World Journal of Biological Psychiatry*, 11(2-2), 457-464. https://doi.org/10.3109/15622970902977552

Herrmann, M. J., Römmler, J., Ehlis, A. C., Heidrich, A., & Fallgatter, A. J. (2004). Source localization (LORETA) of the error-related-negativity (ERN/Ne) and positivity (Pe). *Cognitive Brain Research, 20*(2), 294-299. https://doi.org/10.1016/j.cogbrainres.2004.02.013

Hester, R., Foxe, J. J., Molholm, S., Shpaner, M., & Garavan, H. (2005). Neural mechanisms involved in error processing: A comparison of errors made with and without awareness. *NeuroImage, 27*(3), 602-608. https://doi.org/10.1016/j.neuroimage.2005.04.035

Hester, R., Nestor, L., & Garavan, H. (2009). Impaired error awareness and anterior cingulate cortex hypoactivity in chronic cannabis users. *Neuropsychopharmacology, 34*(11), 2450-2458. https://doi.org/10.1038/npp.2009.67

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21*(11), 1539-1558. https://doi.org/10.1002/sim.1186

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal, 327*, 557-560. https://doi.org/10.1136/bmj.327.7414.557

Hill, K. E., Oumezaine, B. A., Novak, K. D., Rollock, D., & Foti, D. (2018). Variation in reward and error-related neural measures attributable to age, gender, race and ethnicity. *International Journal of Psychophysiology, 132,* 353-364. https://doi.org/10.1016/j.ijpsycho.2017.12.009

Hogan, A. M., Vargha-Khadem, F., Kirkham, F. J., & Baldeweg, T. (2005). Maturation of action monitoring from adolescence to adulthood: An ERP study. *Developmental Science, 8*(6), 525-534. https://doi.org/10.1111/j.1467-7687.2005.00444.x

Holmes, A. J., & Pizzagalli, D. A. (2007). Task feedback effects on conflict monitoring and executive control: relationship to subclinical measures of depression. *Emotion, 7*(1), 68. https://doi.org/10.1037/1528-3542.7.1.68

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

Holroyd, C. B., & Umemoto, A. (2016). The research domain criteria framework: The case for anterior cingulate cortex. *Neuroscience & Biobehavioral Reviews, 71*, 418-443. https://doi.org/10.1016/j.neubiorev.2016.09.021

Holroyd, C. B., & Yeung, N. (2012). Motivation of extended behaviors by anterior cingulate cortex. Trends in cognitive sciences, 16(2), 122-128. https://doi.org/10.1016/j.tics.2011.12.008

Houston, R. J., & Schlienz, N. J. Event-related potentials as biomarkers of behavior change mechanisms in substance use disorder treatment. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 3*(1), 30-40. https://doi.org/10.1016/j.bpsc.2017.09.006

Hsieh, M. T., Lu, H., Lin, C. I., Sun, T. H., Chen, Y. R., & Cheng, C. H. (2021). Effects of trait anxiety on error processing and post-error adjustments: An event-related potential study with stop-signal task. *Frontiers in Human Neuroscience, 15*, 650838. https://doi.org/10.3389/fnhum.2021.650838

Hsu, J. S., Wang, P. W., Ko, C. H., Hsieh, T. J., Chen, C. Y., & Yen, J. Y. (2017). Altered brain correlates of response inhibition and error processing in females with obesity and sweet food addiction: A functional magnetic imaging study. *Obesity research & clinical practice*, *11*(6), 677-686. http://dx.doi.org/10.1016/j.pscychresns.2015.01.004

Hsu, J., Wang, P., Ko, C., Hsieh, T., Chen, C., & Yen, J. (2017). Altered brain correlates of response inhibition and error processing in females with obesity and sweet food addiction: A functional magnetic imaging study. *Obesity Research & Clinical Practice, 11*(6), 677-686. https://doi.org/10.1016/j.orcp.2017.04.011

Hu, L. T., & Bentler, P. M. (1999). Cuttoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*, 1-55. https://doi.org/10.1080/10705519909540118

Iacoboni, M., Lieberman, M. D., Knowlton, B. J., Molnar-Szakacs, I., Moritz, M., Throop, C. J., & Fiske, A. P. (2004). Watching social interactions produces dorsomedial prefrontal and medial parietal BOLD fMRI signal increases compared to a resting baseline. *Neuroimage*, *21*(3), 1167-1173.https://doi.org/10.1016/j.neuroimage.2003.11.013

Iacono, W. G., & Malone, S. M. (2011). Developmental endophenotypes: Indexing genetic risk for substance abuse with the P300 brain event-related potential. *Child Development Perspectives, 5*, 239-247. https://doi.org/10.1111/j.1750-8606.2011.00205.x

Iacono, W. G., Carlson, S. R., Malone, S. M.,& McGue, M. (2002). P3 Event-related potential amplitude and the risk for disinhibitory disorders in adolescent boys. *Archives of General Psychiatry, 59*, 750-757. https://doi.org/10.1001/archpsyc.59.8.750

Im, S., Stavas, J., Lee, J., Mir, Z., Hazlett-Stevens, H., & Caplovitz, G. (2021). Does mindfulness-based intervention improve cognitive function?: A meta-analysis of controlled studies. *Clinical Psychology Review, 84,* 101972. https://doi.org/10.1016/j.cpr.2021.101972

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

Inzlicht, M., Legault, L., & Teper, R. (2014). Exploring the mechanisms of self-control improvement. *Current Directions in Psychological Science, 23*(4), 302-307. https://doi.org/10.1177/0963721414534256

Jarcho, J. M., Fox, N. A., Pine, D. S., Etkin, A., Leibenluft, E., Shechner, T., & Ernst, M. (2013). The neural correlates of emotion-based cognitive control in adults with early childhood behavioral inhibition. *Biological psychology, 92*(2), 306-314. https://doi.org/10.1016/j.biopsycho.2012.09.008

Jensen, A.R., & Rohwer, W.D. (1966). The Stroop color-word test: A review. *Acta Psychologica, 25* (1): 36-93. https://doi.org/10.1016/0001-6918(66)90004-7

Jonkman, L. M., van Melis, J. J. M., Kemner, C., & Markus, C. R. (2007). Methylphenidate improves deficient error evaluation in children with ADHD: An event-related brain potential study. *Biological Psychology, 76,* 217-229. https://doi.org/10.1016/j.biopsycho.2007.08.004

Kagan, J., Reznick, J. S., & Snidman, N. (1987). The physiology and psychology of behavioral inhibition in children. *Child development*, *58*(6), 1459-1473. https://doi.org/10.2307/1130685

Kagan, J., Snidman, N., & Arcus, D., (1998). Childhood derivatives of high and low reactivity in infancy. *Child Development, 69*, 1483-1493. https://doi.org/10.1111/j.1467-8624.1998.tb06171.x

Kaiser, A., Aggensteiner, P-M., Baumeister, S., Holz, N. E, Banaschewski, T., & Brandeis, D. (2020). Early versus later cognitive event-related potentials (ERPs) in attention deficit/hyperactivity disorder (ADHD): A meta-analysis. *Neuroscience and Biobehavioral Reviews, 112*, 117-134. https://doi.org/10.1016/j.neubiorev.2020.01.019

Kalfaoğlu, Ç., Stafford, T., & Milne, E. (2018). Frontal theta band oscillations predict error correction and posterror slowing in typing. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(1), 69. https://doi.org/10.1037/xhp0000417

Keil, A., Debener, S., Gratton, G., Junghöfer, M., Kappenman, E. S., Luck, S. J., Luu, P., Miller, G. A., & Yee, C. M. (2014). Committee report: publication guidelines and recommendations for studies using electroencephalography and magnetoencephalography. *Psychophysiology*, *51*(1), 1-21. https://doi.org/10.1111/psyp.12147

Kerns, J. G., Cohen, J. D., MacDonald, A. W., III., Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science, 303*(5660), 1023-1026. https://doi.org/10.1126/science.1089910

Kessel, E. M., Meyer, A., Hajcak, G., Dougherty, L. R., Torpey-Newman, D. C., Carlson, G. A., & Klein, D. N. (2016). Transdiagnostic factors and pathways to multifinality: The error-related negativity predicts whether preschool irritability is associated with internalizing versus externalizing symptoms at age 9. *Development and Psychopathology, 28*(4), 913-926. https://doi.org/http://dx.doi.org/10.1017/S0954579416000626

Kilian, C., Bröckel, K. L., Overmeyer, R., Dieterich, R., & Endrass, T. (2020). Neural correlates of response inhibition and performance monitoring in binge watching. *International Journal of Psychophysiology*, *158*, 1-8. https://doi.org/10.1016/j.ijpsycho.2020.09.003

Kim, E. Y., Iwaki, N., Uno, H., & Fujita, T. (2005). Error-related negativity in children: Effect of an observer. *Developmental Neuropsychology, 28*(3), 871-883. https://doi.org/10.1207/s15326942dn2803_7

Kirschner, H., Humann, J., Derrfuss, J., Danielmeier, C., & Ullsperger, M. (2021). Neural and behavioral traces of error awareness. *Cognitive, Affective, & Behavioral Neuroscience*, *21*, 573-591. https://doi.org/10.3758/s13415-020-00838-w

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

Klein, T. A., Ullsperger, M., & Danielmeier, C. (2013). Error awareness and the insula: Links to neurological and psychiatric diseases. *Frontiers in Human Neuroscience, 7,* 14. https://doi.org/10.3389/fnhum.2013.00014

Knutson, B. (2005). Distributed neural representation of expected value. *Journal of Neuroscience, 25*(19), 4806-4812. https://doi.org/10.1523/JNEUROSCI.0642-05.2005

Knutson, B., Westdorp, A., Kaiser, E., & Hommer, D. (2000). FMRI Visualization of brain activity during a monetary incentive delay task. *NeuroImage, 12*(1), 20-27. https://doi.org/10.1006/nimg.2000.0593

Ko, C-H., Hsieh, T-J., Chen, C-Y., Yen, C-F., Chen, C-S., Yen, J-Y., Wang, P-W, & Lui, G-C. (2014). Altered brain activation during response inhibition and error processing in subjects with Internet gaming disorder: A functional magnetic imaging study. *European Archives of Psychiatry and Clinical Neuroscience, 264,* 661-672. https://doi.org/10.1007/s00406-013-0483-3

Koban, L., & Pourtois, G. (2014). Brain systems underlying the affective and social monitoring of actions: an integrative review. *Neuroscience & Biobehavioral Reviews*, *46*, 71-84. https://doi.org/10.1016/j.neubiorev.2014.02.014

Kofler, M. J., Rapport, M. D., Sarver, D. E., Raiker, J. S., Orban, S. A., Friedman, L. M., & Kolomeyer, E. G. (2013). Reaction time variability in ADHD: A meta-analytic review of 319 studies. *Clinical psychology review*, *33*(6), 795-811. https://doi.org/10.1016/j.cpr.2013.06.001

Kotov, R., Krueger, R.F., Watson, D., Achenbach, T.M., Althoff, R.R., Bagby, R.M., Brown, T.A., Carpenter, W.T., Caspi, A., Clark, L.A., Eaton, N.R., Forbes, M.K., Forbush, K.T., Goldberg, D., Hasin, D., Hyman, S.E., Ivanova, M.Y., Lynam, D. R., Markon, K., ... Zimmerman, M. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP). *Journal of Abnormal Psychology, 126* (4), 454-477. doi: 10.1037/abn0000258.

Krueger, R. F., & South, S. C. (2009). Externalizing disorders: Cluster 5 of the proposed meta-structure for DSM-V and ICD-11. *Psychological Medicine, 39*(12), 2061-2070. https://doi.org/10.1017/S0033291709990328

Krueger, R. F., Kotov, R., Watson, D., Forbes, M. K., Eaton, N. R., Ruggero, C. J., Simms, L. J., Widiger, T. A., Achenbach, T. M., Bach, B., Bagby, R. M., Bornovalova, M. A., Carpenter, W. T., Chmielewski, M., Cicero, D. C., Clark, L. A., Conway, C., DeClercq, B., DeYoung, C. G., ... Zimmermann, J. (2018). Progress in achieving quantitative classification of psychopathology. *World psychiatry, 17*(3), 282-293. https://doi.org/10.1002/wps.20566

Krueger, R. F., Markon, K. E., Patrick, C. J., Benning, S. D., & Kramer, M. D. (2007). Linking antisocial behavior, substance use, and personality: an integrative quantitative model of the adult externalizing spectrum. *Journal of Psychopathology and Clinical Science, 116*(4), 645. https://doi.org/10.1037/0021-843X.116.4.645

Krueger, R. F., & Markon, K. E. (2006). Reinterpreting comorbidity: A model-based approach to understanding and classifying psychopathology. *Annual Review in Clinical Psychology, 2*, 111-133. https://doi.org/10.1146/annurev.clinpsy.2.022305.095213.

Ladouceur, C. D. (2016). The error-related negativity: A transdiagnostic marker of sustained threat? *Psychophysiology, 53*(3), 389-392. https://doi.org/10.1111/psyp.12585

Ladouceur, C. D., Dahl, R. E., & Carter, C. S. (2007). Development of action monitoring through adolescence into adulthood: ERP and source localization. *Developmental Science, 10*(6), 874-891. https://doi.org/10.1111/j.1467-7687.2007.00639.x

Ladouceur, C. D., Dahl, R. E., Birmaher, B., Axelson, D. A., & Ryan, N. D. (2006). Increased error-related negativity (ERN) in childhood anxiety disorders: ERP and source localization. *Journal of Child*

Psychology and Psychiatry and Allied Disciplines, 47(10), 1073-1082. https://doi.org/10.1111/j.1469-7610.2006.01654.x

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

Lahat, A., Lamm, C., Chronis-Tuscano, A., Pine, D. S., Henderson, H. A., & Fox, N. A. (2014). Early behavioral inhibition and increased error monitoring predict later social phobia symptoms in childhood. *Journal of the American Academy of Child & Adolescent Psychiatry, 53*(4), 447-455. https://doi.org/10.1016/j.jaac.2013.12.019

Larson, M. J., & Carbine, K. A. (2017). Sample size calculations in human electrophysiology (EEG and ERP) studies: A systematic review and recommendations for increased rigor. *International Journal of Psychophysiology, 111*, 33-41. http://dx.doi.org/10.1016/j.ijpsycho.2016.06.015

Larson, M. J., Baldwin, S. A., Good, D. A., & Fair, J. E. (2010). Temporal stability of the error-related negativity (ERN) and post-error positivity (Pe): The role of number of trials. *Psychophysiology, 47*(6), 1167-1171. https://doi.org/10.1111/j.1469-8986.2010.01022.x

Larson, M. J., Clawson, A., Clayson, P. E., & South, M. (2012). Cognitive control and conflict adaptation similarities in children and adults. *Developmental Neuropsychology, 37*(4), 343-357. https://doi.org/10.1080/87565641.2011.650337

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*, 283-297. http://dx.doi.org/10.1016/j.ijpsycho.2014.06.007

Larson, M. J., Clayson, P. E., Keith, C. M., Hunt, I. J., Hedges, D. W., Nielsen, B. L., & Call, V. R. (2016). Cognitive control adjustments in healthy older and younger adults: Conflict adaptation, the error-related negativity (ERN), and evidence of generalized decline with age. *Biological Psychology, 115*, 50-63. https://doi.org/10.1016/j.biopsycho.2016.01.008

Larson, M. J., South, M. & Clayson, P. E. (2011). Sex differences in error-related performance monitoring. *NeuroReport, 22,* 44-48. https://doi.org/10.1097/WNR.0b013e3283427403

Lee, D., Park, J., Namkoong, K., Kim, I. Y., Jung, Y-C. (2018). Gray matter differences in the anterior cingulate and orbitofrontal cortex of young adults with Internet gaming disorder: Surface-based morphometry. *Journal of Behavioral Addiction, 7*(1), 21-30. https://doi.org/10.1556/2006.7.2018.20

Lenzoni, S., Baker, J., Sumich, A. L., & Mograbi, D. C. (2022). New insights into neural networks of error monitoring and clinical implications: a systematic review of ERP studies in neurological diseases. *Reviews in the Neurosciences, 33*(2), 161-179. https://doi.org/10.1515/revneuro-2021-0054

Leshner, A. (2003). Addiction is a brain disease, and it matters. *Focus, 1*(2), 190-193. https://doi.org/10.1176/foc.1.2.190

Li, C-S. R., Zhang. S., Duann, J-R., Yan, P., Sinha R., & Mazure, C. M. (2009). Gender differences in cognitive control: An extended investigation of the Stop Signal Task. *Brain Imaging and Behavior, 3*, 262-276. https://doi.org/10.1107/s11682-009-9068-1

Liebowitz, M. R. (1987). *Social phobia*. Anxiety, Modern Trends in Psychiatry, 22, pp. 141-73. https://doi.org/10.1159/000414022

Liesefeld, H. R., & Janczyk, M. (2019). Combining speed and accuracy to control for speed-accuracy trade-offs(?). *Behavior Research Methods, 51*(1), 40-60. https://doi.org/10.3758/s13428-018-1076-x

Lin, Y., Eckerle, W. D., Peng, L. W., & Moser, J. S. (2019). On variation in mindfulness training: A multimodal study of brief open monitoring meditation on error monitoring. *Brain Sciences, 9*(9), E226. https://doi.org/10.3390/brainsci9090226

Littel, M., van den Berg, I., Luijten, M., van Rooij, A. J., Keemink, L., & Franken, I. H. A. (2012). Error processing and response inhibition in excessive computer game players: An event-related potential study. *Addiction Biology, 17*(5), 934-947. https://doi.org/10.1111/j.1369-1600.2012.00467.x

Liu, X., Liu, T., Shangguan, F., Sørensen, T. A., Liu, Q., & Shi, J. (2018). Neurodevelopment of conflict adaptation: Evidence from event-related potentials. *Developmental Psychology, 54*(7), 1347-1362. https://doi.org/10.1037/dev0000524

Liu, Y., Hanna, G. L., Carrasco, M., Gehring, W. J., & Fitzgerald, K. D. (2014). Altered relationship between electrophysiological response to errors and gray matter volumes in an extended network for error-processing in pediatric obsessive-compulsive disorder. *Human Brain Mapping, 35*(4), 1143-1153. https://doi.org/10.1002/hbm.22240

Lo, S. L. (2018). A meta-analytic review of the event-related potentials (ERN and N2) in childhood and adolescence: Providing a developmental perspective on the conflict monitoring theory. *Developmental Review, 48*, 82-112. https://doi.org/10.1016/j.dr.2018.03.005

Logan, G. D., Cowan, W. B., & Davis, K. A. (1984). On the ability to inhibit simple and choice reaction time responses: A model and a method. *Journal of experimental psychology: Human perception and performance, 10*(2), 276-291. https://doi.org/10.1037//0096-1523.10.2.276

Loo, S. K., Lenartowicz, A., & Makeig, S. (2016). Use of EEG biomarkers in child psychiatry research - Current state and future directions. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 57*, 4-17. https://doi.org/10.1111/jcpp.12435

LoTemplio, S. B., Lopes, C. L., McDonnell, A. S., Scott, E. E., Payne, B. R., & Strayer, D. L. (2023). Updating the relationship of the Ne/ERN to task-related behavior: A brief review and suggestions for future research. *Frontiers in Human Neuroscience, 17*, 138. https://doi.org/10.3389/fnhum.2023.1150244

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

Luijten, M., Kleinjan, M., & Franken, I.H.A. (2016). Event-related potentials reflecting smoking cue-reactivity and cognitive control as predictors of smoking relapse and resumption. *Psychopharmacology, 233*(15), 2857-2868. https://doi.org/10.1007/s00213-016-4332-8

Luijten, M., Machielsen, M. W., Veltman, D. J., Hester, R., de Haan, L., & Franken, I. H. (2014). Systematic review of ERP and fMRI studies investigating inhibitory control and error processing in people with substance dependence and behavioural addictions. *Journal of Psychiatry and Neuroscience, 39*(3), 149-169. https://doi.org/10.1503/jpn.130052

Luijten, M., Meerkerk, G. J., Franken, I. H., van de Wetering, B. J., & Schoenmakers, T. M. (2015). An fMRI study of cognitive control in problem gamers. *Psychiatry Research: Neuroimaging, 231*(3), 262-268. http://doi.org/10.1016/j.pscychresns.2015.01.004

Luijten, M., Van Meel, C. S., & Franken, I. H. A. (2011). Diminished error processing in smokers during smoking cue exposure. *Pharmacology Biochemistry and Behavior, 97*(3), 514-520. https://doi.org/10.1016/j.pbb.2010.10.012

Luna, B., Marek, S., Larsen, B., Tervo-Clemmens, B., & Chahal, R. (2015). An integrative model of the maturation of cognitive control. *Annual Review of Neuroscience, 38*, 151-170. https://doi.org/10.1146/annurev-neuro-071714-034054

Luo, C., Proctor, R.W. (2022). A diffusion model for the congruency sequence effect. *Psychonomic Bulletin & Review, 29*, 2034-2051. https://doi.org/10.3758/s13423-022-02119-8

Luo, X., Zhang, S., Hu, S., Bednarski, S. R., Erdman, E., Farr, O. M., Hong, K-I., Sinha, R., Mazure, C. G., & Li, C. S. R. (2013). Error processing and gender-shared and-specific neural predictors of relapse in cocaine dependence. *Brain, 136*(4), 1231-1244. https://doi.org/10.1093/brain/awt040

Lutz, M. C., Baldwin, S., Rast, P., Franken, I. H., Larson, M. J., & Clayson, P. E. (2023, February 3). Testing the correspondence between intraindividual variability in performance-monitoring behavior and event-related potentials. Retrieved from osf.io/mzexu. Manuscript available upon request.

Lutz, M. C., Kok, R., Verveer, I., Malbec, M., Koot, S., van Lier, P. A. C., & Franken, I. A. H. (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

Lutz, M.C., Smith, A., Lakhlani, D., Tang, A., Guyer, A., Kok, R., Franken, I. H. A., Fox, N., Pine, D., & Harrewijn, A. (2022). Effects of infant behavioral inhibition on social processing and risk for social anxiety during adulthood. *Open Science Framework*, nwyc9. Retrieved from https://osf.io/nwyc9

Magnuson, J. R., Iarocci, G., Doesburg, S. M., & Moreno, S. (2020). Increased intra-subject variability of reaction times and single-trial event-related potential components in children with autism spectrum disorder. *Autism Research, 13*(2), 221-229. https://doi.org/10.1002/aur.2210

Maij, D. L. R., van de Wetering, B. J. M., & Franken, I. H. A. (2017). Cognitive control in young adults with cannabis use disorder: An event-related brain potential study. *Journal of Psychopharmacology, 31*(8), 1015-1026. https://doi.org/10.1177/0269881117719262

Marhe, R., Van De Wetering, B. J. M., & Franken, I. H. A. (2013). Error-related brain activity predicts cocaine use after treatment at 3-month follow-up. *Biological Psychiatry, 73*(8), 782-788. https://doi.org/10.1016/j.biopsych.2012.12.016

Marquardt, L., Eichele, H., Lundervold, A. J., Haavik, J., & Eichele, T. (2018). Event-related-potential (ERP) correlates of performance monitoring in adults with attention-deficit hyperactivity disorder (ADHD). *Frontiers in Psychology, 9*, 485. https://doi.org/10.3389/fpsyg.2018.00485

Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural equation modeling, 11*(3), 320-341. https://doi.org/10.1207/s15328007sem1103_2

Martel, M. M., Pan, P. M., Hoffmann, M. S., Gadelha, A., do Rosário, M. C., Mari, J. J., Manfro, G. G., Miguel, E. C., Paus, T., Bressan, R. A., Rohde, L. A., & Salum, G. A. (2017). A general psychopathology factor (P factor) in children: Structural model analysis and external validation through familial risk and child global executive function. *Journal of Psychopathology and Clinical Science, 126*(1), 137-148. https://doi.org/10.1037/abn0000205

Mattes, A., Porth, E., & Stahl, J. (2022). Linking neurophysiological processes of action monitoring to post-response speed-accuracy adjustments in a neuro-cognitive diffusion model. *NeuroImage, 247*, 118798. https://doi.org/10.1016/j.neuroimage.2021.118798

Mattes, A., Porth, E., Niessen, E., Kummer, K., Mück, M., & Stahl, J. (2023). Larger error negativity peak amplitudes for accuracy versus speed instructions may reflect more neuro-cognitive alignment, not more intense error processing. *Scientific Reports, 13,* 2259, https://doi.org/10.1038/s41598-023-29434-x

McDermott, J. M., Perez-Edgar, K., Henderson, H. A., Chronis-Tuscano, A., Pine, D. S., & Fox, N. A. (2009). A history of childhood behavioral inhibition and enhanced response monitoring in adolescence are linked to clinical anxiety. *Biological Psychiatry, 65*(5), 445-448. https://doi.org/10.1016/j.biopsych.2008.10.043

McDonald, J. B., Bozzay, M. L., Bresin, K., Verona, E., 2021. Facets of exernalizing psychopathology in relation to inhibitory control and error processing. *International Journal of Psychophysiology, 163,* 79-91. https://doi.org/10.1016/j.ijpsycho.2019.08.007

McLoughlin, G., Albrecht, B., Banaschewski, T., Rotherberger, A., Brandeis, D., Asherson, P. & Kuntsi, J. (2009). Performance monitoring is altered in adults ADHD: A familial event-related potential study. *Neuropsychologica, 47*(14). 3134-3142. https://doi.org/10.1016/j.neuropsychologia.2009.07.013

McTeague, L. M., Goodkind, M. S., & Etkin, A. (2016). Transdiagnostic impairment of cognitive control in mental illness. *Journal of psychiatric research, 83*, 37-46. https://doi.org/10.1016/j.jpsychires.2016.08.001

McTeague, L. M., Huemer, J., Carreon, D. M., Jiang, Y., Eickhoff, S. B., & Etkin, A. (2017). Identification of common neural circuit disruptions in cognitive control across psychiatric disorders. *American Journal of Psychiatry, 174*(7), 676-685. https://doi.org/10.1176/appi.ajp.2017.16040400

Mehta, T., Mannem, N., Yarasi, N. K. & Bollu, P. C. (2020). Biomarkers for ADHD: The present and future directions. *Current Developmental Disorders Reports, 7,* 85092. https://doi.org/10.1007/s40474-020-00196-9

Mesgarani, N., Cheung, C., Johnson, K., & Chang, E. F. (2014). Phonetic feature encoding in human superior temporal gyrus. *Science, 343*(6174), 1006-1010. https://doi.org/10.1126/science.1245994

Meyer, A. (2017). A biomarker of anxiety in children and adolescents: A review focusing on the error-related negativity (ERN) and anxiety across development. *Developmental Cognitive Neuroscience, 27*, 58-68. https://doi.org/10.1016/j.dcn.2017.08.001

Meyer, A., Bress, J. N., Hajcak, G., & Gibb, B. E. (2018). Maternal depression is related to reduced error-related brain activity in child and adolescent offspring. *Journal of Clinical Child & Adolescent Psychology, 47*(2), 324-335. https://doi.org/10.1080/15374416.2016.1138405

Meyer, A., Gibby, B., Wissemann, K., Klawohn, J., Hajcak, G., & Schmidt, N. B. (2020). A brief, computerized intervention targeting error sensitivity reduces the error-related negativity. Cognitive, Affective, & Behavioral Neuroscience, 20, 172-180. https://doi.org/10.3758/s13415-019-00760-w

Meyer, A., & Hajcak, G. (2019). A review examining the relationship between individual differences in the error-related negativity and cognitive control. *International Journal of Psychophysiology, 144*, 7-13. https://doi.org/10.1016/j.ijpsycho.2019.07.005

Meyer, A., & Klein, D. N. (2018). Examining the relationships between error-related brain activity (the ERN) and anxiety disorders versus externalizing disorders in young children: Focusing on cognitive control, fear, and shyness. *Comprehensive Psychiatry, 87*, 112-119. https://doi.org/10.1016/J.COMPPSYCH.2018.09.009

Meyer, A., Hajcak, G., Torpey, D. C., Kujawa, A., Kim, J., Bufferd, S., ... Klein, D. N. (2013). Increased error-related brain activity in six-year-old children with clinical anxiety. *Research on Child and Adolescent Psychopathology, 41*(8), 1257-1266. https://doi.org/10.1007/s10802-013-9762-8

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

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

Meyer. A. (2017). A biomarker of anxiety in children and adolescents: A review focusing on the error-related negativity (ERN) and anxiety across development. *Developmental Cognitive Neuroscience, 27*, 58-68, https://doi.org/10.1016/j.dcn.2017.08.001.

Michael, J. A., Wang, M., Kaur, M., Fitzgerald, P. B., Fitzgibbon, B. M., & Hoy, K. E. (2021). EEG correlates of attentional control in anxiety disorders: A systematic review of error-related negativity and correct-response negativity findings. *Journal of Affective Disorders, 291*, 140-153. https://doi.org/10.1016/j.jad.2021.04.049

Michelini, G., Kitsune, G. L., Cheung, C. H. M., Brandeis, D., Banaschewski, T., Asherson, P., ... Kuntsi, J. (2016). Attention-deficit/hyperactivity disorder remission is linked to better neurophysiological error detection and attention-vigilance processes. *Biological Psychiatry, 80*(12), 923-932. http://dx.doi.org/10.1016/j.biopsych.2016.06.021

Miller, A. E., Watson, J. M. & Strayer, D. L. (2012). Individual differences in working memory capacity predict action monitoring and the error-related negativity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(3), 757-763. https://doi.org/ 10.1037/a0026595.

Miller, G. A., & Rockstroh, B. (2013). Endophenotypes in psychopathology research: Where do we stand? *Annual Review of Clinical Psychology, 9*, 177-213. https://doi.org/10.1146/annurev-clinpsy-050212-185540

Moadab, I., Gilbert, T., Dishion, T. J., & Tucker, D. M. (2010). Frontolimbic activity in a frustrating task: Covariation between patterns of coping and individual differences in externalizing and internalizing symptoms. *Development and Psychopathology, 22*(2), 391-404. https://doi.org/10.1017/S0954579410000131

Moeller, S. J., Froböse, M. J., Konova, A. B., Misryrlis, M., Parvaz, M. A., Goldstein, R. Z., & Alia-Klein, N. (2014a). Common and distinct neural correlates of inhibitory dysregulation: Stroop fMRI study of cocaine addiction and intermittent explosive disorder. *Journal of Psychiatric Research, 58*, 55-62. https://doi.org/10.1016/j.jpsychires.2014.07.016

Moeller, S. J., Konova, A. B., Parvaz, M. A., Tomasi, D., Lande, R.D., Fort, C. & Goldstein, R. Z. (2014b). Functional, structural, and emotional correlates of impaired insight in cocaine addiction. *JAMA Psychiatry, 71*(1), 61-70. https://doi.org/10.1001/jamapsychiatry.2013.2833

Monoach, D. S., & Agam, Y. (2013). Neural markers of errors as endophenotypes in neuropsychiatric disorders. *Frontiers in Human Neuroscience, 7,* 350. https://doi.org/10.3389/fnhum.2013.00350

Morales, S., Bowers, M. E., Leach, S. C., Buzzell, G. A., Fifer, W., Elliott, A. J., & Fox, N. A. (2022). Time-frequency dynamics of error monitoring in childhood: An EEG study. *Developmental Psychobiology, 64*, e22215.https://doi.org/10.1002/dev.22215

Moran, T. P., Schroder, H. S., Kneip, C., & Moser, J. S. (2017). Meta-analysis and psychophysiology: A tutorial using depression and action-monitoring event-related potentials. *International Journal of Psychophysiology, 111*, 17-32. https://doi.org/10.1016/j.ijpsycho.2016.07.001

Morgan, A. B., & Lilienfeld, S. O. (2000). A meta-analytic review of the relation between antisocial behavior and neuropsychological measures of executive function. *Clinical Psychology Review, 20*, 113-136. https://doi.org/10.1016/S0272-7358(98)00096-8

Morie, K. P., De Sanctis, P., Garavan, H., & Foxe, J. J. (2014). Executive dysfunction and reward dysfunction: A high-density electrical mapping study in cocaine abusers. *Neuropharmacology, 85*, 397-407. https://doi.org/10.1016/j.neuropharm.2014.05.016

Moser, J. S., Moran, T. I. M. P., Kneip, C., Schroder, H. S., & Larson, M. J. (2016). Sex moderates the association between symptoms of anxiety, but not obsessive compulsive disorder, and error-monitoring brain activity: A meta-analytic review. *Psychophysiology, 53*, 21-29. https://doi.org/10.1111/psyp.12509

Moser, J. S., Moran, T. P., Schroder, H. S., Donnellan, M. B., & Yeung, N. (2013). On the relationship between anxiety and error monitoring: a meta-analysis and conceptual framework. *Frontiers in Human Neuroscience, 7*, 466. https://doi.org/10.3389/fnhum.2013.00466

Mulder, M. J., Bos, D., Weusten, J. M., van Belle, J., van Dijk, S. C., Simen, P., van England, H., & Durston, S. (2010). Basic impairments in regulating the speed-accuracy tradeoff predict symptoms of attention-deficit/hyperactivity disorder. *Biological psychiatry, 68*(12), 1114-1119. https://doi.org/10.1016/j.biopsych.2010.07.031.

Mullane, J. C., Corkum, P. v., Klein, R. M., & McLaughlin, E. (2009). Interference control in children with and without ADHD: A systematic review of flanker and simon task performance. *Child Neuropsychology, 15*(4), 321-342. https://doi.org/10.1080/09297040802348028

Munro, G. E. S., Dywan, J., Harris, G. T., McKee, S., Unsal, A., & Segalowitz, S. J. (2007). ERN varies with degree of psychopathy in an emotional discrimination task. *Biological Psychology, 76,* 31-42. https://doi.org/10.1016/j.biopsycho.2007.05.004

Muris, P., van Brakel, A. M., Arntz, A., & Schouten, E. (2011). Behavioral inhibition as a risk factor for the development of childhood anxiety disorders: A longitudinal study. *Journal of child and family studies, 20*, 157-170. https://doi.org/10.1007/s10826-010-9365-8

Muthén & Muthén (2018). Mplus Editor (version 8.2) [Software]. *Los Angeles, CA: Muthen & Muthen.*

Naim, R., Jeong, A., Lee, K., Laky, Z. E., Pereira, F., Zheng, C., & Brotman, M. (2023). Examining domains of executive functions in children with externalizing versus internalizing problems in the ABCD data. *Biological Psychiatry*, *93*(9), S133. https://doi.org/10.1016/j.biopsych.2023.02.337

Niessen, E., Fink, G. R., Hoffmann, H. E., Weiss, P. H., & Stahl, J. (2017). Error detection across the adult lifespan: Electrophysiological evidence for age-related deficits. *NeuroImage, 152*, 517-529. https://doi.org/10.1016/j.neuroimage.2017.03.015.

Nieuwenhuis, S., Ridderinkhof, K. R., Blom, J., Band, G. P., & Kok, A. (2001). Error-related brain potentials are differentially related to awareness of response errors: evidence from an antisaccade task. *Psychophysiology*, *38*(5), 752-760. https://doi.org/10.1111/1469-8986.3850752

Notebaert, W., Houtman, F., Van Opstal, F., Gevers, W., Fias, W., & Verguts, T. (2009). Post-error slowing: an orienting account. *Cognition*, *111*(2), 275-279. https://doi.org/10.1016/j.cognition.2009.02.002

Novianti, P. W., Roes, K. C. B., & van der Tweel, I. (2014). Estimation of between-trial variance in sequential meta-analyses: A simulation study. *Contemporary Clinical Trials, 37*(1), 129-138. https://doi.org/10.1016/j.cct.2013.11.012

Nozadi, S. S., White, L. K., Degnan, K. A., & Fox, N. A. (2018). Longitudinal relations between behavioral inhibition and social information processing: Moderating role of maternal supportive reactions to children's emotions. *Social Development, 27*(3), 571-585. https://doi.org/10.1111/sode.12286

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

Olfson, M., Wall, M. M., Wang, S., & Blanco, C. (2023). Prevalence and correlates of mental disorders in children aged 9 and 10 years: results from the ABCD study. *Journal of the American Academy of Child and Adolescent Psychiatry*, S0890-8567(23)00178-8. https://doi.org/10.1016/j.jaac.2023.04.005

Olvet, D. M., & Hajcak, G. (2008). The error-related negativity (ERN) and psychopathology: Toward an endophenotype. *Clinical Psychological Review, 28*(8), 1343-1354. https://doi.org/10.1016/j.cpr.2008.07.003

Olvet, D. M., & Hajcak, G. (2009). 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

Ordaz, S. J., Foran, W., Velanova, K., & Luna, B. (2013). Longitudinal growth curves of brain function underlying inhibitory control through adolescence. *Journal of Neuroscience, 33*(46), 18109-18124. https://doi.org/10.1523/JNEUROSCI.1741-13.2013

Orr, C., & Hester, R. (2012). Error-related anterior cingulate cortex activity and the prediction of conscious error awareness. *Frontiers in human neuroscience, 6*, 177. https://doi.org/10.3389/fnhum.2012.00177

Orwin, R. G. (1983). A fail-safe N for effect size in meta-analysis. *Journal of educational statistics, 8*(2), 157-159.

Overbeek, T. J. M., Nieuwenhuis, S., & Ridderinkhof, K. R. (2005). Dissociable components of error processing: On the functional significance of the Pe vis-à-vis the ERN/Ne. *Journal of Psychophysiology 19*(4), 319-329. https://doi.org/10.1027/0269-8803.19.4.319

Overbye, K., Walhovd, K. B., Paus, T., Fjell, A. M., Huster, R. J., & Tamnes, C. K. (2019). Error processing in the adolescent brain: Age-related differences in electrophysiology, behavioral adaptation, and brain morphology. *Developmental Cognitive Neuroscience, 38*, 100665. https://doi.org/10.1016/j.dcn.2019.100665

Padilla, M. L., Colrain, I. M., Sullivan, E. V., Mayer, B. Z., Turlington, S. R., Hoffman, L. R., Wagstaff, A., E., & Pfefferbaum, A. (2011). Electrophysiological evidence of enhanced performance monitoring in recently abstinent alcoholic men. *Psychopharmacology, 213*(1), 81-91. https://doi.org/10.1007/s00213-010-2018-1

Pailing, P. E., & Segalowitz, S. J. (2004). The error-related negativity as a state and trait measure: Motivation, personality, and ERPs in response to errors. *Psychophysiology, 41*(1), 84-95. https://doi.org/10.1111/1469-8986.00124

Park, M., Jung, M. H., Lee, J., Choi, A. R., Chung, S. J., Kim, B., Kim, D., J., & Choi, J. (2020). Neurophysiological and cognitive correlates of error processing deficits in Internet Gaming Disorder. *Cerebral Cortex, 30*(9), 4914-4921. https://doi.org/10.1093/cercor/bhaa083

Parvaz, M. A., Alia-Klein, N., Woicik, P. A., Volkow, N. D., & Goldstein, R. Z. (2011). Neuroimaging for drug addiction and related behaviors. *Reviews in the Neurosciences, 22*(6), 609-624. https://doi.org/10.1515/RNS.2011.055

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*, 133-149. https://doi.org/10.1016/j.neubiorev.2019.06.013

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

Paul, M., Govaart, G. H., & Schettino, A. (2021). Making ERP research more transparent: Guidelines for preregistration. *International Journal of Psychophysiology, 154*, 52-63. https://doi.org/10.1016/j.ijpsycho.2021.02.016

Pavlov, Y, G., Adamain, N., Appelhoff, S., Arvaneh, M., Benwell, C. S. Y., Beste, C., Bland, A. R., Bradford, D. E., Bublatzky, F., Busch, N. A., Clayson, P. E., Cruse, D., Czeszumki, A., Dreber, A., Dumas, G., Ehinger, B., Ganis, G., He, X., Hinojosa, J. A., Huber-Huber, C., ... Mushtaq, F. (2021). #EEGManyLabs: Investigating the replicability of influential EEG experiments. *Cortex, 144*, 213-229. https://doi.org/10.1016/j.cortex.2021.03.013

Pérez-Edgar, K., Reeb-Sutherland, B. C., McDermott, J. M., White, L. K., Henderson, H. A., Degnan, K. A., Hane, A. A., Pine, D. S., & Fox, N. A. (2011). Attention biases to threat link behavioral inhibition to social withdrawal over time in very young children. *Research on Child and Adolescent Psychology, 39*, 885-895. https://doi.org/10.1007/s10802-011-9495-5

Pfabigan, D. M., Rütgen, M., Kroll, S. L., Riečanský, I., & Lamm, C. (2021). The administration of the opioid buprenorphine decreases motivational error signals. *Psychoneuroendocrinology, 128*, 105199. https://doi.org/10.1016/j.psyneuen.2021.105199

Pourtois, G., Vocat, R., N'Diaye, K., Spinelli, L., Seeck, M., & Vuilleumier, P. (2010). Errors recruit both cognitive and emotional monitoring systems: Simultaneous intracranial recordings in the dorsal anterior cingulate gyrus and amygdala combined with fMRI. *Neuropsychologia, 48*(4), 1144-1159. https://doi.org/10.1016/j.neuropsychologia.2009.12.020

Pozuelos, J., P., Mead, B. R., Reuda, M. R., & Malinowski, P. (2019). Chapter 6: Short-term mindful breath awareness training improves inhibitory control and response monitoring. *Progress in Brain Research, 244,* 137-163. https://doi.org/10.1016/bs.pbr.2018.10.019

Quaglia, J. T., Zeidan, F., Grossenbacher, P.G., Freeman, S.P., Braun, S.E., Martelli, A., Goodman, R. J., & Brown, K. W. (2019). Brief mindfulness training enhances cognitive control in socioemotional contexts: Behavioral and neural evidence. *PLOS ONE 14*(7): e0219862. https://doi.org/10.1371/journal.pone.0219862

Rabbitt, P. M. A. (1966). Errors and error correction in choice-response tasks. *Experimental Psychology, 71*(2), 264-272.

Ranger, J., Kuhn, J. T., & Pohl, S. (2021). Effects of motivation on the accuracy and speed of responding in tests: The speed-accuracy tradeoff revisited. *Measurement, 19*(1), 15-38. https://doi.org/10.1080/15366367.2020.1750934

Rapee, R. M., & Heimberg, R. G. (1997). A cognitive-behavioral model of anxiety in social phobia. *Behaviour research and therapy, 35*(8), 741-756. https://doi.org/10.1016/S0005-7967(97)00022-3

Rapee, R. M., Kennedy, S. J., Ingram, M., Edwards, S. L., & Sweeney, L. (2010). Altering the trajectory of anxiety in at-risk young children. *American Journal of Psychiatry, 167*(12), 1518-1525. https://doi.org/10.1176/appi.ajp.2010.09111619

Rass, O., Fridberg, D. J., & O'Donnell, B. F. (2014). Neural correlates of performance monitoring in daily and intermittent smokers. *Clinical Neurophysiology, 125*(7), 1417-1426. https://doi.org/10.1016/j.clinph.2013.12.001

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation, 20*(4), 873-922. https://doi.org/10.1162/neco.2008.12-06-420

Ridderinkhof, K. R., de Vlugt, Y., Bramlage, A., Spaan, M., Elton, M., Snel, J., & Band, G. P. (2002). Alcohol consumption impairs detection of performance errors in mediofrontal cortex. *Science, 298*(5601), 2209-2211. https://doi.org/10.1126/science.1076929

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

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science, 306*(5695), 443-447. https://doi.org/10.1126/science.1100301

Ridderinkhof, K. R., Van Den Wildenberg, W. P., Segalowitz, S. J., & Carter, C. S. (2004). Neurocognitive mechanisms of cognitive control: the role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain and Cognition, 56*(2), 129-140. https://doi.org/10.1016/j.bandc.2004.09.016

Riesel, A. (2019). The erring brain: Error-related negativity as an endophenotype for OCD - A review and meta-analysis. *Psychophysiology, 56*(4), 1-22. https://doi.org/10.1111/psyp.13348

Riesel, A., Klawohn, J., Grutzmann, R., Kaufmann, C., Heinzel, S., Bey, K., Lennertz, L., Wagnger, M., & Kathmann, N. (2019). Error-related brain activity as a transdiagnostic endophenotype for obsessive-compulsive disorder, anxiety and substance use disorder. Psychological Medicine, 49(7), 1207-1217. https://doi.org/10.1017/S0033291719000199

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

Riesel, A., Weinberg, A., Endrass, T., Meyer, A., & Hajcak, G. (2013). The ERN is the ERN is the ERN? Convergent validity of error-related brain activity across different tasks. *Biological Psychology, 93*(3), 377-385. https://doi.org/10.1016/j.biopsycho.2013.04.007

Rietdijk, W. J. R., Franken, I. H. A., & Thurik, A. R. (2014). Internal consistency of event-related potentials associated with cognitive control: N2/P3 and ERN/Pe. *PLoS ONE, 9*(7), e102672. https://doi.org/10.1371/journal.pone.0102672

Rosburg, T., Deuring, G., Boillat, C., Lemoine, P., Falkenstein, M., Graf, M., & Mager, R. (2018). Inhibition and attentional control in pedophilic child sexual offenders - An event-related potential study. *Clinical Neurophysiology, 129*(9), 1990-1998. https://doi.org/10.1016/j.clinph.2018.06.029

Rosvold, H. E., Mirsky, A. F., Sarason, I., Bransome Jr, E. D., & Beck, L. H. (1956). A continuous performance test of brain damage. *Journal of Consulting Psychology, 20*(5), 343. https://psycnet.apa.org/doi/10.1037/h0043220

Rudo-Hutt, A. S. (2015). Electroencephalography and externalizing behavior: A meta-analysis. *Biological Psychology, 105*, 1-19. https://doi.org/10.1016/j.biopsycho.2014.12.005

Rueppel, M., Mannella, K. A., Fitzgerald, K. D., & Schroder, H. S. (2022). Post-error slowing in anxiety and obsessive-compulsive disorders. *Cognitive, Affective and Behavioral Neuroscience, 22*(3), 610-624. https://doi.org/10.3758/s13415-021-00976-9

Ruigrok, A. N. V., Salimi-Korshidi, G., Lai, M-G., Baron-Cohen, S., Lombardo, M. V. Trait, R. J., & Suckling, J. (2014). A meta-analysis of sex differences in human brain structure. *Neuroscience and Biobehavioral Reviews, 39*, 34-50. https://doi.org/10.1016/Journalneubiorev.2013.12.004

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

Sandstrom, A., Uher, R., & Pavlova, B. (2020). Prospective association between childhood behavioral inhibition and anxiety: A Meta-Analysis. *Research on Child and Adolescent Psychopathology, 48*(1), 57-66. https://doi.org/10.1007/s10802-019-00588-5

Santesso, D. L., Segalowitz, S. J., & Schmidt, L. A. (2006). Error-related electrocortical responses in 10-year-old children and young adults. *Developmental Science, 9*(5), 473-481. https://doi.org/10.1111/j.1467-7687.2006.00514.x

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika, 66*(4), 507-514.

Satpute, A. B., & Lieberman, M. D. (2006). Integrating automatic and controlled processes into neurocognitive models of social cognition. *Brain Research, 1079*(1), 86-97. https://doi.org/10.1016/j.brainres.2006.01.005

Saunders, B., Rodrigo, A.H. & Inzlicht, M. (2016). Mindful awareness of feelings increases neural performance monitoring. *Cognitive, Affective and Behavioral Neuroscience, 16*, 93-105. https://doi.org/10.3758/s13415-015-0375-2

Scheffers, M. K., & Coles, M. G. H. (2000). Performance monitoring in a confusing world: Error-related brain activity, judgments of response accuracy, and types of errors. *Journal of Experimental Psychology: Human Perception and Performance, 26*(1), 141-151. https://doi.org/10.1037/0096-1523.26.1.141

Schellekens, A. F., De Bruijn, E. R., Van Lankveld, C. A., Hulstijn, W., Buitelaar, J. K., De Jong, C. A., & Verkes, R. J. (2010). Alcohol dependence and anxiety increase error-related brain activity. *Addiction, 105*(11), 1928-1934. https://doi.org/10.1111/j.1360-0443.2010.03065.x

Schlienz, N. J., & Hawk, L. W. J. (2017). Probing the behavioral and neurophysiological effects of acute smoking abstinence on drug and nondrug reinforcement during a cognitive ask. *Nicotine & Tobacco Research, 19*(6), 729-737. https://doi.org/10.1093/ntr/ntw277

Schlienz, N. J., Hawk, L. W., & Rosch, K. S. (2013). The effects of acute abstinence from smoking and performance-based rewards on performance monitoring. *Psychopharmacology, 229*(4), 701-711. https://doi.org/10.1007%2Fs00213-013-3131-8

Schoenberg, P. L. A., Hepark, S., Kan, C., Barendregt, H. P., Buitelaar, J. K., & Speckens, A. E. M. (2014). Effects of mindfulness-based cognitive control therapy on neurophysiological correlates of performance monitoring in adult attention-deficit/hyperactivity disorder. *Clinical Neurophysiology, 125*, 1407-1416. http://dx.doi.org/10.1016/Journalclinph.2013.11.031

Schrijvers, D., De Bruijn, E. R., Maas, Y. J., Vancoillie, P., Hulstijn, W., & Sabbe, B. G. (2009). Action monitoring and depressive symptom reduction in major depressive disorder. *International Journal of Psychophysiology, 71*(3), 218-224. https://doi.org/10.1016/Journalijpsycho.2008.09.005

Schroder, H. S., & Moser, J. S. (2014). Improving the study of error monitoring with consideration of behavioral performance measures. *Frontiers in Human Neuroscience, 8*. https://doi.org/10.3389/fnhum.2014.00178

Schroder, H. S., Nickels, S., Cardenas, E., Breiger, M., Perlo, S., & Pizzagalli, D. A. (2020). Optimizing assessments of post-error slowing: A neurobehavioral investigation of a flanker task. *Psychophysiology, 57*(2), 1-17. https://doi.org/10.1111/psyp.13473

Schwarzer, G. (2007) Meta: An R package for meta-analysis. *R News, 7*, 40-45.

Segalowitz, S. J., & Dywan, J. (2009). Individual differences and developmental change in the ERN response: implications for models of ACC function. *Psychological research, 73*(6), 857-870. https://doi.org/10.1007/s00426-008-0193-z

Senderecka, M., Grabowska, A., Szewczyk, J., Gerc, K., & Chmylak, R. (2012). Response inhibition of children with ADHD in the stop-signal task: An event-related potential study. *International Journal of Psychophysiology, 85*(1), 93-105. 10.1016/j.ijpsycho.2011.05.007

Shackman, A. J., Salomons, T. V., Slagter, H. A., Fox, A. S., Winter, J. J., & Davidson, R. J. (2011). The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nature Reviews Neuroscience, 12*(3), 154-167. https://doi.org/10.1038/nrn2994

Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron, 79*(2), 217-240. https://doi.org/10.1016/j.neuron.2013.07.007

Sherman, L. E., Rosenbaum, G. M., Smith, A. R., Botdorf, M. A., Fettich, K., Patrianakos, J. L., McCloskey, M., Steinberg, L. D., & Chein, J. M. (2019). The interactive effects of peers and alcohol on functional brain connectivity in young adults. *NeuroImage, 197*, 264-272. https://doi.org/10.1016/j.neuroimage.2019.04.003

Sheth, S. A., Mian, M. K., Patel, S. R., Asaad, W. F., Williams, Z. M., Dougherty, D. D., ... Eskandar, E. N. (2012). Human dorsal anterior cingulate cortex neurons mediate ongoing behavioural adaptation. *Nature, 488*(7410), 218-221. https://doi.org/10.1038/nature11239

Shiels, A. N., Reardon, K. W., Brandes, C. M., & Tackett, J. L. (2019). The *p* factor in children: Relationships with executive functions and effortful control. *Journal of Research in Personality, 82*, 103853. https://doi.org/10.1016/j.jrp.2019.103853

Shiels, K., & Hawk, L. W. (2010). Self-regulation in ADHD: The role of error processing. *Clinical Psychology Review, 30*(8), 951-961. https://doi.org/10.1016/j.cpr.2010.06.010

Siddaway, A. P., Wood, A. M., & Hedges, L. V. (2019). How to do a systematic review: A best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annual review of psychology, 70*, 747-770. https://doi.org/10.1146/annurev-psych-010418-102803

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014a). P-curve: a key to the file-drawer. *Journal of experimental psychology: General, 143*(2), 534. https://doi.org/10.1037/a0033242

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014b). P-curve and effect size: Correcting for publication bias using only significant results. *Perspectives on Psychological Science, 9*(6), 666-681. https://doi.org/10.1177/1745691614553988

Simonsohn, U., Simmons, J. P., & Nelson, L. D. (2015). Better P-curves: Making P-curve analysis more robust to errors, fraud, and ambitious P-hacking, a Reply to Ulrich and Miller. *Journal of Experimental Psychology: General, 144*(6), 1146–1152. https://doi.org/10.1037/xge0000104

Slagter, H. A., Davidson, R. J., Lutz, A. (2011). Mental training as a tool in the neuroscientific study of and cognitive plasticity. *Frontiers in Human Neuroscience, 5*, 17. https://doi.org/10.3389/fnhum.2011.00017

Śmigasiewicz, K., Ambrosi, S., Blaye, A., & Burle, B. (2020). Inhibiting errors while they are produced: Direct evidence for error monitoring and inhibitory control in children. *Developmental Cognitive Neuroscience, 41*, 100742. https://doi.org/10.1016/j.dcn.2019.100742

Smith, A. R., Chein, J., & Steinberg, L. (2014). Peers increase adolescent risk taking even when the probabilities of negative outcomes are known. *Developmental Psychology, 50*(5), 1564–1568. https://psycnet.apa.org/doi/10.1037/a0035696

Smith, A. R., Kircanski, K., Brotman, M. A., Do, Q. B., Subar, A. R., Silk, J. S., Engel, S., Crosby, R. D., Harrewijn, A., White, L. K., Haller, S. P., Cardinale, E. M., Buzzell, G. A., Barker, T., Leibenluft, E., & Pine, D. S. (2019). Advancing clinical neuroscience through enhanced tools: Pediatric social anxiety as an example. *Depression and Anxiety, 36*(8), 701-711. https://doi.org/10.1002/da.22937

Smith, J. L., & Mattick, R. P. (2013). Evidence of deficits in behavioural inhibition and performance monitoring in young female heavy drinkers. *Drug and Alcohol Dependence, 133*(2), 398-404. https://doi.org/10.1016/j.drugalcdep.2013.06.020

Smith, J. L., Mattick, R. P., & Sufani, C. (2017). Error detection and behavioural inhibition in young heavy drinkers. *Drug and Alcohol Dependence, 171*, 20-30. https://doi.org/10.1016/j.drugalcdep.2016.11.016

Smith, J. L., Mattick, R. P., Jamadar, S. D., & Iredale, J. M. (2014). Deficits in behavioural inhibition in substance abuse and addiction: A meta-analysis. *Drug and alcohol dependence, 145*, 1-33. https://doi.org/10.1016/j.drugalcdep.2014.08.009

Smulders, S. F. A., Soetens, E., & van der Molen, M. W. (2016). What happens when children encounter an error? *Brain and Cognition, 104*, 34-47. https://doi.org/10.1016/j.bandc.2016.02.004

Sokhadze, E., Stewart, C., Hollifield, M., & Tasman, A. (2008). Event-related potential study of executive dysfunctions in a speeded reaction task in cocaine addiction. *Journal of Neurotherapy, 12*(4), 185-204. https://doi.org/10.1080/10874200802502144

Soltaninejad, Z., Nejati, V., & Ekhtiari, H. (2019). Effect of anodal and cathodal transcranial direct current stimulation on DLPFC on modulation of inhibitory control in ADHD. *Journal of Atten. Disord. 23*(4), 325-332. https://doi.org/10.1177/1087054715618792

Song, S., Zilverstand, A., Gui, W., Li, H-J., & Zhou, X. (2019). Effects of single-session versus multi-session non-invasive brain stimulation on craving and consumption in individuals with drug addiction, eating disorders or obesity: A meta-analysis. *Brain Stimulation, 12*, 606-618. https://doi.org/10.1016/j.brs.2018.12.975

Stam, H., Franken, I. H. A., van der Veen, F., & Steele, V. R. (2023, June 9). Neurological and behavioral markers of addiction problems. https://doi.org/10.17605/OSF.IO/4FQEA

Steele, V. R., Claus, E. D., Aharoni, E., Vincent, G. M., Calhoun, V. D., & Kiehl, K. A. (2015). Multimodal imaging measures predict rearrest. *Frontiers in Human Neuroscience, 9*, 425. https://doi.org/10.3389/fnhum.2015.00425

Steele, V. R., Fink, B. C., Maurer, J. M., Arbabshirani, M. R., Wilber, C. H., Jaffe, A. J., Sidz, A., Pearlson, G. D., Calhoun, V. D., Clark, V. P., & Kiehl, K. A. (2014). Brain potentials measured during a Go/NoGo task predict completion of substance abuse treatment. *Biological Psychiatry, 76*(1), 75-83. https://doi.org/10.1016/j.biopsych.2013.09.030

Steinhauser, M., & Andersen, S. K. (2019). Rapid adaptive adjustments of selective attention following errors revealed by the time course of steady-state visual evoked potentials. *Neuroimage, 186*, 83-92. https://doi.org/10.1016/j.neuroimage.2018.10.059

Steinhauser, M., & Yeung, N. (2010). Decision processes in human performance monitoring. *Journal of Neuroscience, 30*(46), 15643-15653. https://doi.org/10.1523/JNEUROSCI.1899-10.2010

Steinhauser, M., & Yeung, N. (2012). Error awareness as evidence accumulation: effects of speed-accuracy trade-off on error signaling. Frontiers in human neuroscience, 6, 240. https://doi.org/10.3389/fnhum.2012.00240

Stevens, F. L., Hurley, R. A., & Taber, K. H. (2011). Anterior cingulate cortex: Unique role in cognition and emotion. *The Journal of neuropsychiatry and clinical neurosciences, 23*(2), 121-125. https://doi.org/10.1176/jnp.23.2.jnp121.

Stieben, J., Lewis, M. D., Granic, I., Zelazo, P. D., Segalowitz, S., & Pepler, D. (2007). Neurophysiological mechanisms of emotion regulation for subtypes of externalizing children. *Development and Psychopathology, 19*(2), 455-480. https://doi.org/10.1017/S0954579407070228

Sullivan, R. M., Perlman, G., & Moeller, S. J. (2019). Meta-analysis of aberrant post-error slowing in substance use disorder: Implications for behavioral adaptation and self-control. *European Journal of Neuroscience, 50*(3), 2467-2476. https://doi.org/10.1111/ejn.14229

Tamnes, C. K., Fjell, A. M., Westlye, L. T., Østby, Y., & Walhovd, K. B. (2012). Becoming consistent: Developmental reductions in intra-individual variability in reaction time are related to white matter integrity. *Journal of Neuroscience, 32*(3), 972-982. https://doi.org/10.1532/JNEUROSCI.4779-11.2012

Tamnes, C. K., Walhovd, K. B., Torstveit, M., Sells, V. T., & Fjell, A. M. (2013). Performance monitoring in children and adolescents: A review of developmental changes in the error-related negativity and brain maturation. *Developmental Cognitive Neuroscience, 6*, 1-13. https://doi.org/10.1016/j.dcn.2013.05.001

Tang, A., Crawford, H., Morales, S., Degnan, K. A., Pine, D. S., & Fox, N. A. (2020). Infant behavioral inhibition predicts personality and social outcomes three decades later. *Proceedings of the National Academy of Sciences, 117*(18), 9800-9807. https://doi.org/10.1073/pnas.1917376117

Taylor, J. B., Visser, T. A. W., Fueggle, S. N., Bellgrove, M. A., & Fox, A. M. (2018). The error-related negativity (ERN) is an electrophysiological marker of motor impulsiveness on the Barratt Impulsiveness Scale (BIS-11) during adolescence. *Developmental Cognitive Neuroscience, 30*, 77-86. https://doi.org/10.1016/j.dcn.2018.01.003

Taylor, S. F., Stern, E. R., & Gehring, W. J. (2007). Neural systems for error monitoring: recent findings and theoretical perspectives. *The Neuroscientist, 13*(2), 160-172. https://doi.org/10.1177/1073858406298184

Teper, R., & Inzlicht, M. (2013). Meditation, mindfulness and executive control: the importance of emotional acceptance and brain-based performance monitoring. *Social cognitive and affective neuroscience*, *8*(1), 85-92. https://doi.org/10.1093/scan/nss045

Teper, R., Segal, Z. V., & Inzlicht, M. (2013). Inside the mindful mind: How mindfulness enhances emotion regulation through improvements in executive control. *Current Directions Psychological Science, 22*(6) 449-454. https://doi.org/10.1177/0963721413495869

The Psychological Corporation. (1999). *Manual for the wechsler abbreviated scale of intelligence*. San Antonio, TX: Author.

Tieskens, J. M., Buil, J. M., Koot, S., & van Lier, P. A. (2021). Developmental associations between risk-taking and anxiety symptoms across ages 8–12 years. *Child Development, 92*(6), 2563-2576. https://doi.org/10.1111/cdev.13644

Tops, M. & Boksem, M. A. (2011). Cortisol involvement in mechanisms of behavioral inhibition. *Pyschophysiology, 48*(5), 723-732. https://doi.org/10.1111/j.1469-8986.2010.01131.x

Torpey, D. C., Hajcak, G., Kim, J., Kujawa, A., & Klein, D. N. (2012). Electrocortical and behavioral measures of response monitoring in young children during a Go/No-Go task. *Developmental Psychobiology, 54*(2), 139-150. https://doi.org/10.1002/dev.20590

Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modelling of elementary psychological processes*. New York, NY: Cambridge University Press.

Trenado, C., González-Ramírez, A., Lizárraga-Cortés, V., Pedroarena Leal, N., Manjarrez, E., & Ruge, D. (2019). The potential of trial-by-trial variabilities of ongoing-EEG, evoked potentials, event related potentials and fMRI as diagnostic markers for neuropsychiatric disorders. *Frontiers in Neuroscience, 12*, 850. https://doi.org/10.3389/fnins.2018.00850

Tseng, W. L., Deveney, C. M., Stoddard, J., Kircanski, K., Frackman, A. E., Yi, J. Y., Hsu, D., Moroney, E., Machlin, L., Danahue, L., Roule, A., Perhamus, G., Reynolds, R. C., Roberson-Nay, R.,

Hettema, J. M., Tobwin, K. E., Stringaris, A., Pine, D. S., Brotman, M. A., & Leibenluft, E. (2019). Brain mechanisms of attention orienting following frustration: Associations with irritability and age in youths. *American Journal of Psychiatry, 176*(1), 67-76. https://doi.org/10.1176/appi.ajp.2018.18040491

Ullsperger, M., Danielmeier, C., & Jocham, G. (2014). Neurophysiology of performance monitoring and adaptive behavior. *Physiological Reviews, 94*(1), 35-79. https://doi.org/10.1152/physrev.00041.2012

Ullsperger, M., Harsay, H. A., Wessel, J. R., & Ridderinkhof, K. R. (2010). Conscious perception of errors and its relation to the anterior insula. *Brain Structure and Function, 214*, 629-643. https://doi.org/10.1007/s00429-010-0261-1

Valadez, E. A., & Simons, R. F. (2018). The power of frontal midline theta and post-error slowing to predict performance recovery: Evidence for compensatory mechanisms. *Psychophysiology, 55*(4), e13010. https://doi.org/10.1111/psyp.13010

Vallet, W., Neige, C., Mouchet-Mages, S., Brunelin, J., & Grondin, S. (2021). Response-locked component of error monitoring in psychopathy: A systematic review and meta-analysis of error-related negativity/positivity. *Neuroscience and Biobehavioral Reviews, 123*, 104-119. https://doi.org/10.1016/j.neubiorev.2021.01.004 .

Van Aert, R. C., Wicherts, J. M., & van Assen, M. A. (2016). Conducting meta-analyses based on p values: Reservations and recommendations for applying p-uniform and p-curve. *Perspectives on Psychological Science, 11*(5), 713-729. https://doi.org/10.1177/1745691616650874

Van De Voorde, S., Roeyers, H., & Wiersema, J. R. (2010). Error monitoring in children with ADHD or reading disorder: An event-related potential study. *Biological Psychology, 84*(2), 176-185. https://doi.org/10.1016/j.biopsycho.2010.01.011

Van der Borght, L., Braem, S., Stevens, M., & Notebaert, W. (2016). Keep calm and be patient: The influence of anxiety and time on post-error adaptations. *Acta Psychologica, 164*, 34-38. https://doi.org/10.1016/j.actpsy.2015.12.007.

Van der Jagt, A., & Stoof, C. (2023, May 11). Corsi - an ipad app for motoric and psychological tests. https://doi.org/10.17605/OSF.IO/HMEK5

Van der Molen, M. J., Poppelaars, E. S., Van Hartingsveldt, C. T., Harrewijn, A., Gunther Moor, B., & Westenberg, P. M. (2014). Fear of negative evaluation modulates electrocortical and behavioral responses when anticipating social evaluative feedback. *Frontiers in Human Neuroscience, 7*, 936. https://doi.org/10.3389/fnhum.2013.00936

van Driel, J., Ridderinkhof, K. R., & Cohen, M. X. (2012). Not all errors are alike: theta and alpha EEG dynamics relate to differences in error-processing dynamics. *Journal of Neuroscience, 32*(47), 16795-16806. https://doi.org/10.1523%2FJNEUROSCI.0802-12.2012

Van Lier, P. A. C., Muthén, B. O., van der Sar, R. M., & Crijnen, A. A. M. (2004). Preventing disruptive behavior in elementary school children: Impact of a universal classroom-based intervention. *Journal of Consulting and Clinical Psychology, 72*(3), 467-478. https://doi.org/10.1037/0022-006X.72.3.467

Van Meel, C. S., Heslenfeld, D. J., Rommelse, N. N., Oosterlaan, J., & Sergeant, J. A. (2012). Developmental trajectories of neural mechanisms supporting conflict and error processing in middle childhood. *Developmental Neuropsychology, 37*(4), 358-378. https://doi.org/10.1080/87565641.2011.653062

Van Veen, V., & Carter, C. S. (2002a). The anterior cingulate as a conflict monitor: fMRI and ERP studies. *Physiology and Behavior, 77*(4), 477-482. https://doi.org/10.1016/S0031-9384(02)00930-7

Van Veen, V., & Carter, C. S. (2002b). The tinning of action-monitoring processes in the anterior cingulate cortex. *Journal of Cognitive Neuroscience, 14*(4), 593-602. https://doi.org/10.1162/08989290260045837

Verbruggen, F., & Logan, G. D. (2008). Response inhibition in the stop-signal paradigm. *Trends in Cognitive Sciences, 12*(11), 418-424. https://doi.org/10.1016/j.tics.2008.07.005

Veroniki, A. A., Jackson, D., Viechtbauer, W., Bender, R., Bowden, J., Knapp, G., Kuss, O., Higgins, J. PT., Langan, D., & Salanti, G. (2016). Methods to estimate the between-study variance and its uncertainty in meta-analysis. *Research Synthesis Methods, 7*, 55-79. https://doi.org/10.1002/jrsm.1164

Verveer, I. (2020). *Bridging the Gap Between Neurocognitive Insights and the Addiction Clinic: The Effects and Underlying Mechanisms of Transcranial Direct Current Stimulation in Substance Use Disorder*.[Doctoral dissertation]. Erasmus University Rotterdam.

Verveer, I., Hill, A. T., Franken, I. H. A., Yücel, M., Van Dongen, J. D. M., & Segrave, R. (2021). Modulation of control: Can HD-tDCS targeting the dACC reduce impulsivity? *Brain Research, 1756*(1), 147282. https://doi.org/10.1016/j.brainres.2021.147282

Vidal, F., Hasbroucq, T., Grapperon, J., & Bonnet, M. (2000). Is the 'error negativity' specific to errors? *Biological psychology, 51*(2-3), 109-128. https://doi.org/10.1016/S0301-0511(99)00032-0

Viechtbauer, W. (2005). Bias and efficiency of meta-analytic variance estimators in the random-effects model. *Journal of Educational and Behavioral Statistics, 30*(3), 261-293. https://doi.org/10.3102/10769986030003261

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software 36*(3). https://doi.org/10.18637/jss.v036.i03

Viechtbauer, W., & Cheung, M. W. L. (2010). Outlier and influence diagnostics for meta-analysis. *Research synthesis methods, 1*(2), 112-125. https://doi.org/10.1002/jrsm.11

Vilà-Balló, A., Hdez-Lafuente, P., Rostan, C., Cunillera, T., & Rodriguez-Fornells, A. (2014). Neurophysiological correlates of error monitoring and inhibitory processing in juvenile violent offenders. *Biological Psychology, 102*, 141-152. http://dx.doi.org/10.1016/j.biopsycho.2014.07.021

Vocat, R., Pourtois, G., & Vuilleumier, P. (2008). Unavoidable errors: A spatio-temporal analysis of time-course and neural sources of evoked potentials associated with error processing in a speeded task. *Neuropsychologia, 46*(10), 2545-2555. https://doi.org/10.1016/j.neuro psych ologia.2008.04.006

Volkow, N. D., Koob, G. F., & McLellan, A. T. (2016). Neurobiologic advances from the brain disease model of addiction. *New England Journal of Medicine, 374*(4), 363-371. https://doi.org/10.1056/NEJMra1511480

Volpert-Esmond, H. I., Merkle, E. C., Levsen, M. P., Ito, T. A., & Bartholow, B. D. (2018). Using trial-level data and multilevel modeling to investigate within-task change in event-related potentials. *Psychophysiology, 55*(5), e13044. https://doi.org/10.1111/psyp.13044

Volpert-Esmond, H. I., Page-Gould, E., & Bartholow, B. D. (2021). Using multilevel models for the analysis of event-related potentials. *International Journal of Psychophysiology, 162*, 145-156. https://doi.org/10.1016/j.ijpsycho.2021.02.006

Wahlstrom, D., White, T., & Luciana, M. (2010). Neurobehavioral evidence for changes in dopamine system activity during adolescence. *Neuroscience & Biobehavioral Reviews, 34*(5), 631-648. https://doi.org/10.1016/j.neubiorev.2009.12.007

Wang, L., Gu, Y., Zhao, G., & Chen, A. (2020). Error-related negativity and error awareness in a Go/No-go task. *Scientific reports, 10*(1), 4026. https://doi.org/10.1038/s41598-020-60693-0

Waschke, L., Kloosterman, N. A., Obleser, J., & Garrett, D. D. (2021). Behavior needs neural variability. *Neuron, 109*(5), 751-766. https://doi.org/10.1016/j.neuron.2021.01.023

Wechsler D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI)*. San Antonio, TX: Psychological Corporation.

Weinberg, A., Dieterich, R., & Riesel, A. (2015). Error-related brain activity in the age of RDoC: A review of the literature. *International Journal of Psychophysiology, 98*. 276-299. https://doi.org/10.1016/j.ijpsycho.2015.02.029

Weinberg, A., Ethridge, P., Pegg, S., Freeman, C., Kujawa, A., & Dirks, M. A. (2021). Neural responses to social acceptance predict behavioral adjustments following peer feedback in the context of a real-time social interaction task. Psychophysiology, 58(3), e13748. https://doi.org/10.1111/psyp.13748

Weinberg, A., Liu, H., & Shankman, S. A. (2016). Blunted neural response to errors as a trait marker of melancholic depression. *Biological Psychology, 113*, 100-107. https://doi.org/10.1016/j.biopsycho.2015.11.012

Weinberg, A., Meyer, A., Hale-Rude, E., Perlman, G., Kotov, R., Klein, D. N., & Hajcak, G. (2016). Error-related negativity (ERN) and sustained threat: Conceptual framework and empirical evaluation in an adolescent sample. *Psychophysiology, 53*(3), 372-385. https://doi.org/10.1111/psyp.12538

Weinberg, A., Riesel, A., & Hacjak, G. (2012). Integrating multiple perspectives on error-related brain activity: The ERN as a neural indicator of trait defensive reactivity. *Motivation and Emotion, 36*, 84-100. https://doi.org/10.1007/s11031-011-9269-y

Wessel, J. R. (2012). Error awareness and the error-related negativity: evaluating the first decade of evidence. *Frontiers in Human Neurosciencec, 6*, 88. https://doi.org/10.3389/fnhum.2012.00088

Wessel, J. R. (2018). An adaptive orienting theory of error processing. *Psychophysiology*, *55*(3), 1-21. https://doi.org/10.1111/psyp.13041

Wessel, J. R., Danielmeier, C., Morton, J. B., & Ullsperger, M. (2012). Surprise and error: common neuronal architecture for the processing of errors and novelty. *Journal of Neuroscience, 32*(22), 7528-7537. https://doi.org/10.1523/JNEUROSCI.6352-11.2012

Wiersema, J. R., van der Meere, J. J., & Roeyers, H. (2005). ERP correlates of impaired error monitoring in children with ADHD. *Journal of Neural Transmission, 112*(10), 1417-1430. https://doi.org/10.1007/s00702-005-0276-6

Wiersema, J. R., Van Der Meere, J. J., & Roeyers, H. (2009). ERP correlates of error monitoring in adult ADHD. *Journal of Neural Transmission, 116*(3), 371-379. https://doi.org/10.1007/s00702-008-0165-x

Wiggins, J. L., Brotman, M. A., Adleman, N. E., Kim, P., Oakes, A. H., Reynolds, R. C., Chen, G., Pine, D. S., & Leibenluft, E. (2016). Neural correlates of irritability in disruptive mood dysregulation and bipolar disorders. *American Journal of Psychiatry, 173*(7), 722-730. https://doi.org/10.1176/appi.ajp.2015.15060833

Wild-Wall, N., Oades, R. D., Schmidt-Wessels, M., Christiansen, H., & Falkenstein, M. (2009). Neural activity associated with executive functions in adolescents with attention-deficit/hyperactivity disorder (ADHD). *International Journal of Psychophysiology, 74*, 19-27. https://doi.org/10.1016/j.ijpsycho.2009.06.003

Williams, R. L. (2000). A note on robust variance estimation for cluster-correlated data. *Biometrics, 56*(2), 645-646. https://doi.org/10.1111/j.0006-341X.2000.00645.x

Woltering, S., Granic, I., Lamm, C., & Lewis, M. D. (2011). Neural changes associated with treatment outcome in children with externalizing problems. *Biological Psychiatry, 70*(9), 873-879. https://doi.org/10.1016/j.biopsych.2011.05.029

Xing, L., Yuan, K., Bi, Y., Yin, J., Cai, C., Feng, D., Li, Y., Song, M., Wang, H., Y, D., Xue, T., Jin, C., Quin, W., & Tian, J. (2014). Reduced fiber integrity and cognitive control in adolescents with internet gaming disorder. *Brain Research, 1586*, 109-117. https://doi.org/10.1016/j.brainres.2014.08.044

Xue, J. M., Lin, P-Z., Sun, J-W., & Cao, F-L. (2017). Disrupted executive function and aggression in individuals with a history of adverse childhood experiences: An event-related potential study. *Journal of Nervous and Mental Disease, 205*(12), 942-951. https://doi.org/10.1097/NMD.0000000000000743

Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychological review, 111*(4), 931. https://psycnet.apa.org/doi/10.1037/0033-295X.111.4.931

Yücel, M., Oldenhof, E., Ahmed, S. H., Belin, D., Billieux, J., Bowden-Jones, H., Carter, A., Chamberlain, S. R., Clark, L., Connor, J., Daglish, M., Dom, G., Dannon, P., Duka, T., Fernandez-Sarrano, M. J., Field, M., Franken, I., Goldstein, R. Z., Gonzalez, R., Goudriaan, A., E., ... Verdejo-Garcia, A. (2019). A transdiagnostic dimensional approach towards a neuropsychological assessment for addiction: an international Delphi consensus study. *Addiction, 114*(6), 1095-1109. https://doi.org/10.1111/add.14424

Yun, J. Y., Kim, J. C., Ku, J., Shin, J. E., Kim, J. J., & Choi, S. H. (2017). The left middle temporal gyrus in the middle of an impaired social-affective communication network in social anxiety disorder. *Journal of affective disorders, 214*, 53–59. https://doi.org/10.1016/j.jad.2017.01.043

Zahn, R., Moll, J., Krueger, F., Huey, E. D., Garrido, G., & Grafman, J. (2007). Social concepts are represented in the superior anterior temporal cortex. *Proceedings of the National Academy of Sciences, 104*(15), 6430-6435. https://doi.org/10.1073/pnas.0607061104

Zajonc, R. B. (1965). Social Facilitation: A solution is suggested for an old unresolved social psychological problem. *Science, 149*(3681), 269-274. https://doi.org/10.1126/science.149.3681.269

Zhang, J-S., Wang, Y., Cai, R-G., & Yan, C-H. (2009). The brain regulation mechanism of error monitoring in impulsive children with ADHD – An analysis of error related potentials. *Neuroscience Letters, 460*, 11-15. https://doi.org/10.1016/j.neulet.2009.05.027

Zhou, Y., Lin, F., Du, Y., Zhao, Z., Xu, J., & Lei, H. (2011). Gray matter abnormalities in Internet addiction: A voxel-based morphometry study. *European Journal of Radiology, 79*(1), 92-95. https://doi.org/10.1016/j.ejrad.2009.10.025

Zhou, Z., Li, C., & Zhu, H. (2013). An error-related negativity potential investigation of response monitoring function in individuals with Internet addiction disorder. *Frontiers in Behavioral Neuroscience, 7*, 131. https://doi.org/10.3389/fnbeh.2013.00131

Zijlmans, J., Bevaart, F., van Duin, L., Luijks, M. J. A., Popma, A., & Marhe, R. (2019). Error-related brain activity in relation to psychopathic traits in multi-problem young adults: An ERP study. *Biological Psychology, 144*, 46-53. https://doi.org/10.1016/j.biopsycho.2019.03.014

Zwetsloot, P. P., Van Der Naald, M., Sena, E. S., Howells, D. W., IntHout, J., De Groot, J. A., Chamuleau, S. Aj., MacLeod, M. R., & Wever, K. E. (2017). Standardized mean differences cause funnel plot distortion in publication bias assessments. *eLife, 6*, e24260. https://doi.org/10.7554/eLife.24260.

About the Author

Miranda C. Lutz was born on November 10, 1990 in Leiderdorp, the Netherlands. After completing her diploma program at the International Baccalaureate Organization in Oegstgeest in 2019, she completed a bachelor's degree in Pedagogical Sciences at the University of Amsterdam and two master's degrees in Education and Child Studies at the University of Leiden. During her master studies she developed her research interest: working with EEG and MRI measures to better understand psychopathology. For one year, she worked full-time as a research assistant in the lab of Prof. Andrea Evers, PhD, and as a research trainee with Kim de Jong, PhD. In 2017, she started her PhD at Erasmus University Rotterdam, funded by the Vital Cities and Citizen grant, under the supervision of Prof. Ingmar Franken, PhD and Prof. Pol van Lier, PhD. She supported the large data collection of the longitudinal study 'Happy child, Happy adolescent'. In 2022, she was on a research visit to the National Institute of Mental Health, to work closely with Daniel Pine, MD and Prof. Nathan Fox, PhD (University of Maryland) on the longitudinal study of the effects of behavioral inhibition. Miranda has coordinated and taught several courses and supervised several undergraduate and graduate students in psychology. She has conducted several workshops on meta-analysis for colleagues and students. In addition, she has established several national and international collaborations, organized and co-hosted several events for researchers (Research Days for the Dutch Society of Developmental Psychology, VNOP and Dutch Peer Relation Research Day, hosted in Rotterdam). Miranda has dedicated herself to improve the well-being and working conditions of PhD students in her department, for which she initiated the DPECS PhD representatives. Her efforts were rewarded with the 'Best PhD colleague 2020' award from her department's Erasmus Graduate School. Miranda is an advocate of open science and transparent practices in academia, sharing her experiences on social media and publishing code and data available on OSF. While working on her Ph.D., Miranda continues to teach swimming lessons to children at the local swimming pool in her spare time. Miranda will continue to work for DPECS as an Assistant Professor under the supervision of Prof. Matthias Wieser, PhD. She lives with her husband Lars and two daughters (Nora Maya, born 2018 and Isa Elyn, born 2021) in Oegstgeest.

PhD Portfolio

Publications

De Jong, K., Conijn, J. M., Gallagher, R. A.V., Reshetnikova, A. S., Heij, M. & Lutz, M. C. (2021). Using progress Feedback to improve outcomes and reduce dropout, treatment duration, and deterioration: A multilevel meta-analysis. *Clinical Psychological Review, 85*, 102002. https://doi.org/10.1016/j.cpr.2021.102002

Lutz, M. C., Kok, R., & Franken, I. H. A. (2021). Event-related potential (ERP) measures of error processing as biomarkers of externalizing disorders: A narrative review. *International Journal of Psychophysiology, 166*, 151-159. https://doi.org/10.1016/j.ijpsycho.2021.06.002

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

Submitted/Pre-registered papers

Aslanidou, A., Malbec, M., Lutz, M.C., & Wieser, M. J.: Review on steady-state EEG: ability to differentiate between stimuli. Preregistration: https://osf.io/nq745/.

Calga, C.*, Troost, L.*, Ultanir, D.*, Lutz, M.C., & Schulte, M. (2023). TMS. Submitted to *Frontiers for Young Minds.* *shared first authorship.

Horoz, N., van Atteveldt, N., van Lenthe, F., Groeniger, J. O., Houweling, T. A., Koot, H. M., **Lutz, M. C.**, & Buil, M. (2023, May 26). Are Household- and School-level Parental Education Associated with Academic Self-Concept Development in Elementary School? https://doi.org/10.17605/OSF.IO/5BTP4

Lutz M. C., Buil, M., Kok, R., Koot, S., Franken, I. H. A., & Van Lier, P. A. C. Developmental trajectory of flanker performance and its link to problem behavior in 7-to 12-year-old children. Retrieved from: https://osf.io/10.31234/osf.io/fj62k

Lutz, M. C., Davidsen, A. H., & de Jong, K. Validity study of the Dutch translation of the Assessment for Signal Clients questionnaire. Revied & Re-submitted in *Clinical Psychology and Psychotherapy*

Lutz, M. C., Lakhlani, D., Tang, A., Smith, A. S., Kok, R., Franken, I. H. A., Guyer, A., Fox, N. A., Pine, D. S, & Harrewijn, A. Effects of Infancy Behavioral Inhibition on Social Processing and Risk for Social Anxiety during Adulthood. Pre-registration: https://osf.io/nwyc9

Lutz, M.C., Rast, P., Baldwin, S. A., Franken, I. H. A., Larson, M. J., & Clayson, P. Testing the Correspondence Between Intraindividual Variability in Performance-Monitoring Behavior and Event-Related Potentials. Pre-registration under embargo: https://osf.io/mzexu

Presentations/talks

- 23.04.21, Open Science Rotterdam. Publication bias and P-curve analysis for Meta-Analysis. https://tinyurl.com/4zuwtfya

- 13.06.22, CPDD Minneapolis, USA. Poster M36 entitled: Diminished Error-Related Negativity and Error Positivity in Adults With Addiction Problems and Disorders: A Meta-Analysis on Error Processing
- 20.09.22, Rotterdam, the Netherlands. Master course (Addiction) lecture entitled: Behavioral addictions: are we in control of our behavior?
- 28.09.22, SPR Vancouver, Canada. Poster session 1 entitled: Relationship between within-person differences in error-related negativity and error positivity and correct-trial response-time means and variances in healthy participants

Courses

- EGSH Communicating your research: Lessons from Bitescience
- EGSH Data analysis with R
- EGSH Professionalism and integrity in research
- EGSH Searching, finding and managing your literature
- EGSH Shut up and write
- EGSH Multi-level modeling I
- EGSH Multi-level modeling II
- EGSH Making your research proposal work for you
- EGSH Research Synthesis & Meta-analysis
- EGSH Self-presentation: focus, structure
- EARA Advanced Structural Equational Modeling
- EPOS &Helmholtz workshop: Inhibition Winterschool

Awards | Recognition

2022 Travel grants: Van der Gaag beurs, KNAW; €4575 Erasmus Trust Fund, €2500; SPR Family Support grant: $400

2021 Award for best Ph.D. Colleague 2020 from the Graduate School Awards for Ph.D. Excellence, Graduate School of Social Sciences and Humanities, EUR

2020 Seed funding for Reducing externalizing behavior in adolescents using co-creation, in collaboration with Michelle Achterberg, Ph.D. & Lysanne te Brinke, Ph.D., €1500.

Academic Activities

Events	Organized and hosted: Graduate Research Days VNOP-CAS 2018-2020; PRO 2021
Committee	Founder of departmental PhD representatives, Erasmus University Rotterdam
Peer-Review	Neuroscience & Biobehavioral Reviews; European Addiction Research

Teaching

- Bachelor and Master thesis supervision
- Coordination and teacher for three workgroups in the master Kind & Jeugd Psychologie
- Teacher for third-year workgroups and advanced research trainee program in the bachelor Psychology