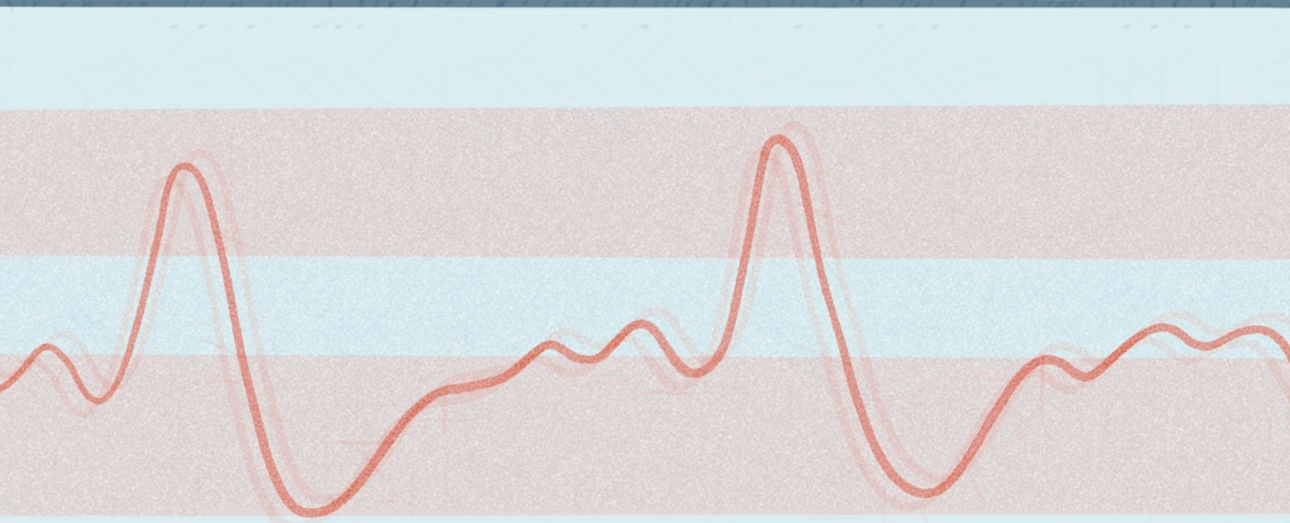
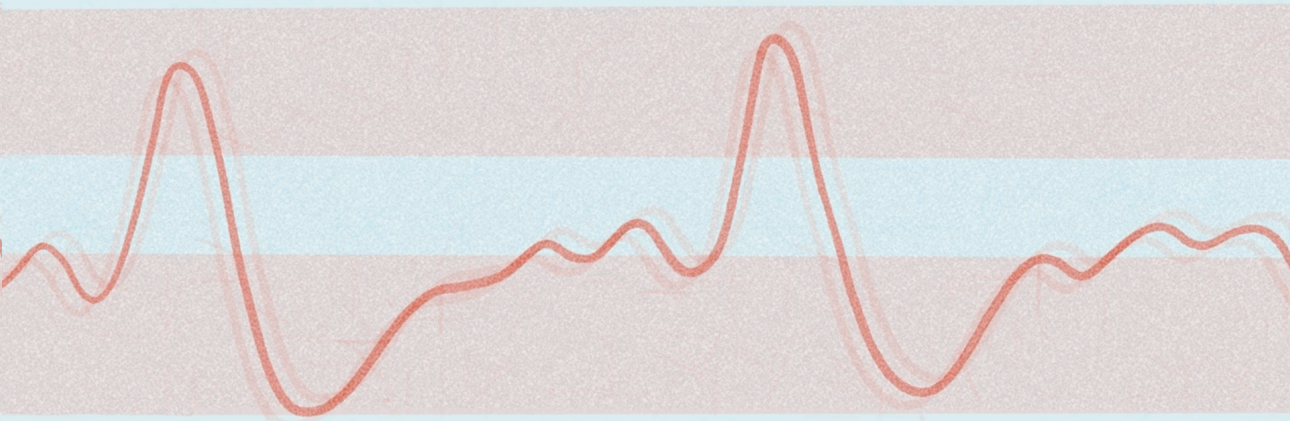
My bad! Subclinical and neurochemical alterations of performance monitoring for self and others

Jansen, M.



Chapter 1

General introduction



"The greatest mistake you can make in life is to be continually fearing you will make one"

Elbert Hubbard

PERFORMANCE MONITORING

Imagine you are driving on the highway to get to your parents' house for a family visit. While doing this, you are continuously monitoring your actions, either consciously or unconsciously. You make sure that you are driving at an appropriate speed, are following the planned route, are respecting road markings, and that you are maintaining enough distance to other vehicles. This way, when you notice that you have accidentally shifted to gear 3 instead of 5 or that you took a wrong turn, you can immediately make the necessary behavioral adjustments, such that you can arrive at your destination in a safe and timely manner.

This example illustrates *performance monitoring*, an essential cognitive process that involves supervising and regulating behavior by evaluating feedback and adjusting subsequent actions. By monitoring our ongoing actions and performance outcomes, we are able to detect errors or mismatches between our intentions or predictions ("when I take this route, I will end up at my parents' house") and our actual experiences ("this is not the route to my parents' house"). By triggering both immediate behavioral adjustments ("taking nearest exit to get back to the correct route") and long-term learning ("remembering to turn left here instead of right on next visit"), this process enables successful adaptive and goal-directed behavior. Hence, performance monitoring is a fundamental process that helps us optimally navigate our complex and ever-changing environment throughout our daily life.

PERFORMANCE MONITORING IN THE BRAIN

Electroencephalography

Errors are important signals telling us that outcomes differ from what was intended or expected. In the early 1990s, a way of measuring the brain's response to errors was discovered, which importantly opened up new avenues of research into the neural mechanisms of performance monitoring. Using electroencephalography (EEG), which is a noninvasive neuroimaging technique used to measure electrical brain activity, a fronto-centrally distributed negative event-related potential (ERP) deflection was found. This so-called *error-related negativity* (ERN) peaks approximately 50 to 100 milliseconds (ms) after the commission of errors (Falkenstein, 1990; Gehring, Goss, Coles, Meyer, & Donchin, 1993), and is followed by a positive deflection referred to as the *error positivity* (Pe) (Figure 1A). The discovery of these error-related ERPs components has substantiated the notion that there is an internal performance-monitoring system in our brain that detects deviations between outcomes and intentions or predictions. Moreover, these components have provided an important means to understanding the underlying neural mechanisms of performance monitoring and their contribution to adaptive and goal-directed behavior.

The ERN and Pe can be most clearly observed in speeded choice reaction time paradigms such as the *Flanker task* (Figure 1B; Eriksen & Eriksen, 1974). In this task, errors are elicited by choosing between different response options in the presence of distracting information and limited response times. Participants see a string of five arrows (or sometimes letters) and are asked to respond with either a left or right button press according to the identity

of the middle arrow. The surrounding arrows can be the same as the middle arrow (congruent trial: e.g., <<<<<), or point in a different direction (incongruent trial: e.g., <<><<). Errors are easily made during incongruent trials because the surrounding arrows elicit opposing response tendencies. Because ERPs are established by measuring and averaging electrical brain activity across multiple events (i.e., erroneous responses), this paradigm provides a useful measure to study brain responses during error detection.

While the functional significance of the ERN has been discussed extensively (see section "WHAT IS THE ERN?" below), the Pe component has received much less attention. Some research has shown that the Pe correlates with the awareness and motivational salience of errors (see e.g., Ullsperger, Danielmeier, & Jocham, 2014). There is also evidence to suggest that the component can be dived in two parts. The 'early' Pe, which closely follows the ERN (150-250 ms after response onset) and has the same frontocentral distribution, has been linked to the more automatic processing of errors. The slower and more parietal 'late' or 'classical' Pe on the other hand, occurs approximately 300-500 ms after an error is made, and is thought to be related to the more conscious or affective processing of errors.

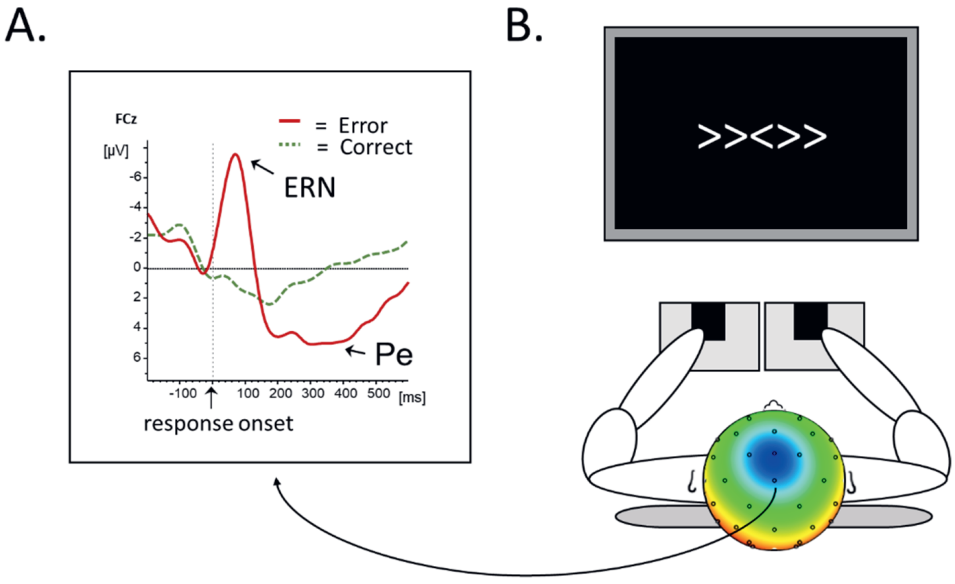


Figure 1. (A) The error-related negativity (ERN) and error positivity (Pe) can be observed as negative- and positive event-related potential (ERP) deflections that are time-locked to response onset. Data from Chapter 2 (de Bruijn, Jansen, & Overgaauw, 2020). (B) Performance-monitoring processes are often studied by recording participant's electrophysiological brain activity during the Flanker task. In this reaction-time task, participants have to respond with a left or right button press according to the identity of the middle arrow. Mistakes are frequently made in this task for trials where the surrounding arrows point in a different direction due to the elicitation of opposing response tendencies, making it an ideal paradigm for the study of errors.

Functional magnetic resonance imaging

Besides EEG, the use of functional magnetic resonance imaging (fMRI) gained popularity in the 2000s as its technology and methods improved. This technique provides an indirect measure of neural activity by capturing changes in blood oxygen level-dependent (BOLD) signals in the brain. Whereas EEG can measure brain activity using millisecond-level timing, fMRI has a much lower temporal resolution, capturing neural changes over seconds to minutes. However, the advantage of this technique is that it can identify spatial patterns of brain activity with much higher resolution than EEG. With the help of paradigms such as the Cannonball task (Figure 2A), fMRI research has shown that correct responses and positive outcomes consistently activate the ventral striatum (VS), while errors and negative feedback specifically activate the anterior insula (AI), and the posterior medial frontal cortex (pMFC), including the anterior midcingulate cortex and the pre-supplementary motor area (Figure 2B; e.g., de Bruijn, de Lange, von Cramon, & Ullsperger, 2009; Koban, Corradi-Dell'Acqua, & Vuilleumier, 2013; Overgaauw, Jansen, & de Bruijn, 2020; Radke, De Lange, Ullsperger, & De Bruijn, 2011). Importantly, electrical source analyses have consistently localized the pMFC as the neural generator of the ERN (Dehaene, Posner, & Tucker, 1994; Holroyd, Dien, & Coles, 1998; Keil, Weisz, Paul-Jordanov, & Wienbruch, 2010; Miltner et al., 2003; Ullsperger & Von Cramon, 2001) and studies using concurrent EEG and fMRI measurements show that this region specifically correlates with single-trial amplitudes of the ERN (Debener et al., 2005; Iannaccone et al., 2015).

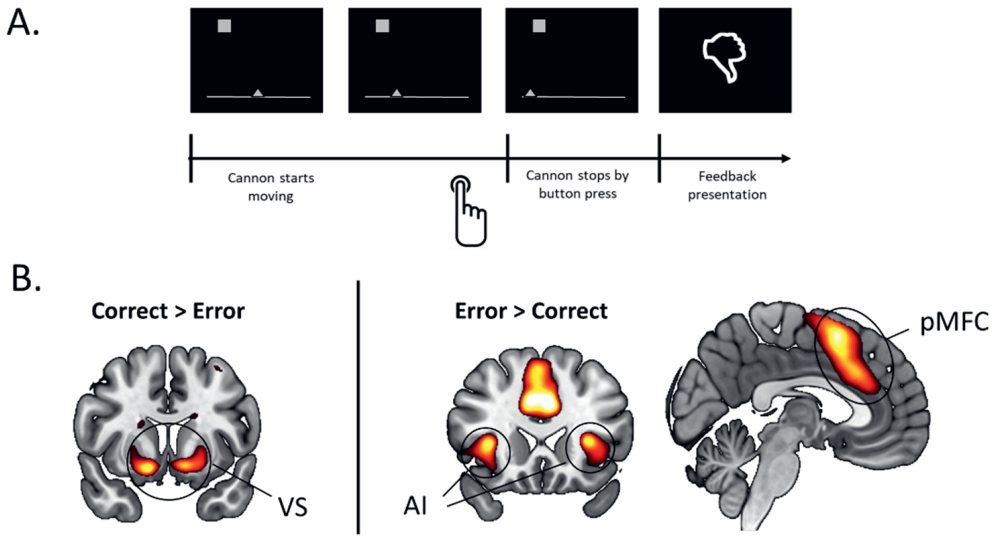


Figure 2. (A) In the Cannonball task, participants have to precisely line up a horizontally moving cannon (triangle) with a stationary target (square), by pressing the button at the right moment. Subsequently, a thumbs up or thumbs down is presented to indicate whether participant's responses were correct or not. (B) Comparing correct responses to errors consistently reveals activation of the ventral striatum (VS), whereas the reverse contrast is associated with activity in the posterior medial frontal cortex (pMFC) and anterior insula (AI). Data from Chapter 7 (Jansen et al., under review).

What is the ERN?

Many theories have been proposed with regard to what exactly the ERN reflects. For example, according to the mismatch hypothesis, one of the earliest theories of the ERN, the component reflects the discrepancy between one's internal representation of the correct action and the actually executed (i.e., erroneous) one (Coles, Scheffers, & Holroyd, 2001). In contrast, the response conflict monitoring theory assumes that the ERN is generated in response to conflicts between competing response tendencies rather than errors per se (Botvinick, Braver, Barch, Carter, & Cohen, 2001). While theories differ with regard to the exact cause of the ERN, there is consensus that the ERN reflects an evaluative signal that indicates the need for adjustments. In support of this, the amplitude of the ERN and error-related pMFC activity have been found to predict subsequent behavioral adjustments and learning (e.g., Debener et al., 2005; Fischer, Danielmeier, Villringer, Klein, & Ullsperger, 2016; Hester, Barre, Murphy, Silk, & Mattingley, 2008; Kerns et al., 2004; Klein et al., 2007; Themanson, Rosen, Pontifex, Hillman, & McAuley, 2012).

Reinforcement learning, prediction errors and dopamine

One particularly influential theory argues that the ERN can be explained by *reinforcement learning* principles conveyed by the mesencephalic *dopamine* system (Holroyd & Coles, 2002). Reinforcement learning provides a framework for how individuals learn to optimize their behavior based on the experienced outcomes of their actions, such as rewards and punishments. The most simple reinforcement learning model (Rescorla, Wagner, Black, & Prokasy, 1972) assumes that we compute internal expectations of the value of certain actions based on past experiences. These expected values are updated when there are differences between expected and actual outcome, which are called *reward prediction errors* and serve as important teaching signals to improve learning and prediction. Importantly, neurophysiological work in animals has shown that dopamine neurons in the midbrain increase firing after better-than-expected outcomes (positive prediction errors) and reduce firing when outcomes are worse than expected (negative prediction errors; Schultz, 2022).

Based on the evidence for dopaminergic encoding of prediction errors, the so-called reinforcement learning theory of the ERN (Holroyd & Coles, 2002) suggests that when errors are committed, the mesencephalic dopamine system conveys a negative prediction-error signal to the pMFC, where it generates the ERN through a disinhibition of apical dendrites of motor neurons. In this theory, the VS acts as an adaptive critic by making predictions about action outcomes and by evaluating whether outcomes are better or worse than predicted. This is in line with findings from computational neuroscience that this region is consistently involved in the signaling of prediction errors (Fouragnan, Retzler, & Philiastides, 2018). These prediction error signals are then carried to the pMFC, which is thought to function as a motor control filter that selects and reinforces the actions that are most successful at carrying out the task at hand. The reinforcement learning theory offers a valuable account of the neural basis of the ERN, and highlights a key role for the neurotransmitter dopamine in the neurobiology of performance monitoring.

ALTERATIONS IN PERFORMANCE MONITORING

A biomarker for psychopathology?

Studying performance monitoring is of critical relevance to the clinical field given that disturbances in these processes may importantly underlie impaired adaptive behavior in psychiatric disorders. For example, diminished performance monitoring may result in excessively impulsive behavior, whereas enhanced performance may be linked to excessively careful or rigid behavior in order to avoid potential harm.

This latter scenario is especially relevant to individuals with *obsessive-compulsive disorder* (OCD). OCD is a complex and burdensome disorder characterized by intrusive and unwanted thoughts, called obsessions, that trigger significant stress or anxiety. These unwanted obsessions are often accompanied by repetitive, ritualistic behaviors or mental acts, called compulsions, which are performed to alleviate the distress caused by the obsessions (Ruscio, Stein, Chiu, & Kessler, 2010). Importantly, research into clinical alterations of performance-monitoring processes began with the suggestion that compulsions in OCD may be the result of persistent high error signals that cannot be corrected by behavioral actions (Pitman, 1987). This theory has been supported by numerous studies showing that individuals with OCD as well as nonclinical samples scoring high on obsessive-compulsive symptoms show enhanced error-related brain signals as indexed by the ERN (see Riesel, 2019 for a meta-analysis and review). Furthermore, based on findings that enhanced ERNs in OCD have been found to be heritable, are independent of symptom reduction or severity, and are also present in unaffected family members, the ERN has even been suggested as a promising *endophenotype* for the disorder (see Riesel, 2019 for a discussion). Endophenotypes are heritable and measurable markers that reflect underlying biological processes associated with a disorder. Importantly, such biomarkers can offer valuable information about the mechanisms that lead to specific disorders, enhance the accuracy of diagnosis, detect those prone to developing a condition, assist in early prevention and intervention, and enable personalized treatment (e.g., Chan & Gottesman, 2008).

Besides being a putative marker for OCD, ERN amplitudes have also been proposed to reflect a more transdiagnostic biomarker that dissociates *internalizing* from externalizing dimensions of psychopathology (Pasion & Barbosa, 2019). ERN amplitudes have been found to be enhanced in internalizing disorders including generalized anxiety disorder, social phobia, and depression, while they seem to be reduced in externalizing disorders such as attention-deficit/hyperactivity disorder, conduct disorder, oppositional defiant disorder, and substance use disorder (Lutz, Kok, & Franken, 2021; Miranda Christine Lutz et al., 2021; Pasion & Barbosa, 2019). This has been suggested to reflect increased threat sensitivity and error aversion in internalizing populations (Weinberg et al., 2016) and the inability to inhibit or change disruptive and maladaptive behavior in externalizing disorders (Miranda C Lutz et al., 2021; Miranda Christine Lutz et al., 2021; Pasion & Barbosa, 2019). With regard to the Pe, there is evidence that this component is altered in externalizing disorders as well (Brazil et al., 2009; Miranda C Lutz et al., 2021).

However, despite the notion that electrophysiological indices of error processing may represent stable, trait-like markers, studies have also indicated that they are modulated by contextual or state-related factors. For example, the ERN is found to be altered by manipulations of the motivational or affective significance of errors, such as when errors are punished (Proudfit, Inzlicht, & Mennin, 2013). This has led to the suggestion that variance in the ERN signal reflect individual differences in one's evaluation of- or sensitivity to endogenous threat (Weinberg et al., 2016). There are also indications that the ERN and Pe are sensitive to differences in state affect such as after mood inductions (see e.g., Dignath, Eder, Steinhauser, & Kiesel, 2020; Paul, Walentowska, Bakic, Dondaine, & Pourtois, 2017). Yet, still little is known about the exact contextual factors that may modulate these ERPs, while such knowledge is critical before the ERN and Pe can be used as biomarkers.

Ovarian hormonal status

One potential modulator of performance monitoring that has received little attention relates to the ovarian hormonal status of women. Ovarian hormones are known to have widespread effects on emotion, cognition, and the brain (e.g., Beltz & Moser, 2020; Dubol et al., 2021; Hamstra, 2021), and interact with many neurochemical systems including the dopaminergic one (Barth, Villringer, & Sacher, 2015) thought to underlie the generation of the ERN (Holroyd & Coles, 2002). Moreover, it has long been recognized that altered levels of ovarian hormones are associated with symptoms of internalizing disorders such as mood and anxiety. For example, women are twice as likely as men to be diagnosed with an internalizing disorder in their reproductive years (Altemus, Sarvaiya, & Epperson, 2014), which points to a role of ovarian hormones in this increased susceptibility of women. Importantly, in women, the main ovarian hormones, estrogen and progesterone, fluctuate as a function of the menstrual cycle. Many women report negative symptoms including heightened anxiety, negative mood and emotional reactivity in the days leading up to menses (luteal phase), which is characterized by high or declining levels of estrogen and progesterone. These symptoms can even become clinically significant, then referred to as premenstrual syndrome or in more severe cases premenstrual dysphoric disorder (Hofmeister & Bodden, 2016). Similarly, many women report negative emotional side effects during the use of oral contraceptives, when synthetic estrogen and progesterone are administered to suppress endogenous ovarian levels (e.g., Poromaa & Segebladh, 2012), and even discontinue contraceptive use because of this (Le Guen, Schantz, Régnier-Loilier, & de La Rochebrochard, 2021). Given this association between ovarian hormones and internalizing symptoms, and the relation between internalizing symptoms and error-related ERPs, one important aim of the current dissertation (**Chapter 5**) was to investigate whether variations in ovarian hormones due to menstrual cycle phase and oral contraceptive use impact amplitudes of these ERPs during a Flanker task as well as *negative affect*, a measure of internalizing symptoms (Crawford & Henry, 2004).

SOCIAL PERFORMANCE MONITORING

The majority of performance-monitoring research has focused on individual or self-oriented contexts, whereby actions and performance outcomes only affect oneself. However, in our everyday life, we are constantly interacting with others. This means that oftentimes our actions do not only affect ourselves, but also affect others around us. Thinking back of our example where we are driving to our parents' house for Christmas, we probably don't want to upset our parents by being late. Additionally, it is likely that we are not in the car by ourselves, but are accompanied by others, such as a partner and/or children, meaning that any mistake we make may also potentially harm them. In this case, knowing that we are responsible for other people's wellbeing, may lead to us to engage in an even more vigilant monitoring of our actions. Consider another example, where we accidentally bump into someone, which leads them to spill their drink. By constantly monitoring our own behavior and how it affects those around us, we can respond in a way that minimizes negative effects of this action, such as apologizing to this other person and buying them a new drink, and paying more attention to where we walk. This helps us prevent harm and allows us to maintain positive relationships with those around us.

Besides helping us to prevent and minimize negative effects of our mistakes for others, constantly monitoring how our actions affects those around us, also importantly allows us to improve other's outcomes. By monitoring the negative and positive consequences that our actions have on others, we can learn to act in a way that benefits others, i.e., in a prosocial manner. For example, if we are at a party and we notice that people are responding positively to our acts of kindness and cooperation, such as offering to refill someone's drink or helping to clean up after the party, this serves as a positive reinforcer, teaching us to repeat such behavior in the future. Performance monitoring thus also enables prosocial behavior, which is crucial for social functioning and building reciprocal relationships (Carlo, 2013).

Whereas some prior studies have investigated the passive social context of monitoring the performance of others (e.g., Brazil et al., 2011; De Bruijn & von Rhein, 2012), little research has been done with regard to how individuals monitor their performance in social scenarios where one's actions directly affect others. Importantly, to get a better perspective on (pro)social- and performance-monitoring processes in daily life, it is essential to investigate how individuals monitor their actions and errors in more interactive social contexts (see e.g, Schilbach, 2016; Schilbach et al., 2013). Moreover, since almost all psychiatric disorders show impairments in effectively engaging in purposeful and successful social interactions (Schilbach, 2016), investigating (pro)social performance-monitoring processes can help us improve our understanding of functional and social impairments in clinical disorders.

SOCIAL PERFORMANCE MONITORING IN THE BRAIN

Social responsibility contexts

Some evidence from fMRI research suggests that social factors such as responsibility for the outcomes of others influence performance-monitoring activity. For instance, increased activity has been observed in the pMFC and AI when errors cause pain to a friend (Koban et al., 2013), and when the person is solely responsible for the error compared to when sharing responsibility with someone else (Cui, Abdelgabar, Keysers, & Gazzola, 2015). In another study, the AI specifically differentiated between observed errors for which participants shared responsibility or not (Cracco, Desmet, & Brass, 2016). Being responsible for the outcomes of a co-actor has also been found to activate brain regions involved in mentalizing (i.e., understanding other people's actions in terms of intentional mental states; Radke et al., 2011). Hence, while heightened feelings of responsibility and affective distress associated with social mistakes may increase performance-monitoring activity, little is known about the impact of social responsibility for harm on ERP indices of performance monitoring. To this end, in **Chapter 2**, we developed a social flanker paradigm, whereby we tested the hypothesis that mistakes that resulted in harmful versus non-harmful consequences for a co-actor, would enhance amplitudes of the ERN and Pe.

Prosocial prediction errors and learning

Researchers have used probabilistic reinforcement learning paradigms such as the *Prosocial-learning task* (Figure 3) to gain more insight into how our brain processes errors in prediction when our actions benefit others. In this task, participants learn through trial and error to select symbols with the highest probability of obtaining reward, with the beneficiary of this reward (oneself, someone else, or no one) differing per condition. By subsequently applying computational reinforcement learning models to participants' performance, trial-by-trial prediction error computations and relevant learning parameters (such as the learning rate, which reflects the speed of learning by calibrating the extent to which prediction errors impact future value estimations of stimuli) are obtained that can be correlated with BOLD responses. Previous studies indicate that "*prosocial learning*" operates using the same reinforcement mechanisms as individual or self-serving learning (Cutler et al., 2021; Lockwood, Apps, Valton, Viding, & Roiser, 2016; Martins, Lockwood, Cutler, Moran, & Paloyelis, 2022; Westhoff, Blankenstein, Schreuders, Crone, & van Duijvenvoorde, 2021). The encoding of prediction errors engages the VS during both forms of learning, yet only prosocial prediction errors were found to activate the subgenual anterior cingulate cortex (sgACC) (Lockwood et al., 2016; Martins et al., 2022), though see Westhoff et al. (2021). This suggests that there might be socially specific neural systems involved when monitoring and learning about outcomes for others (Lockwood, Apps, & Chang, 2020).

Figure 3. The Prosocial-learning task (used in Chapter 6). In this probabilistic reinforcement learning paradigm, participants complete a series of trials where they have to choose between two abstract symbols. Through trial and error, they have to learn which of the symbols has the highest probability of obtaining reward. Participants perform the task in three different conditions: they either play for their own monetary bonus (self-benefitting learning), the bonus of another participant (prosocial learning) or neither's bonus (non-social control condition). Adapted from Lockwood et al. (2016).

ALTERATIONS IN SOCIAL PERFORMANCE MONITORING

Obsessive-compulsive symptoms

Importantly, the way in which social factors such as responsibility for the outcomes of others influence performance-monitoring processes likely depend on individual differences in feelings of responsibility and concern for other's harm. This is especially relevant to those scoring high on obsessive-compulsive symptoms, who are characterized by an excessive sense of responsibility and a fear of making mistakes that might harm others (Hezel & McNally, 2016). For instance, patients with OCD may repeatedly check appliances and locks to prevent accidents like fire or theft from impacting their loved ones. Some may inspect their vehicle repeatedly to ensure they didn't cause harm while driving. Others may engage in repetitive rituals, such as counting, to alleviate the fear that failing to do so may cause harm to someone they care about. The cognitive theory of OCD even posits that an excessive sense of responsibility for causing harm is a key factor in the development of the disorder. This theory suggests that patients misinterpret intrusive thoughts as evidence of their responsibility to prevent harm to others or themselves, leading them to believe that specific actions (compulsions) are necessary to prevent feared outcomes (Salkovskis & Warwick, 1985). Despite these clear social symptoms related to OCD, research on social cognition – referring to all processes or abilities that allow individuals to behave effectively and adaptively in social situations – in this disorder is scarce. In **Chapter 3**, we therefore aimed to review the current evidence for social-cognitive alterations in OCD, and propose that neurocognitive tasks that capture more interactive and implicit aspects of social

cognition, such as social performance-monitoring paradigms, may help improve our understanding of social and daily-life disturbances in this disorder. Subsequently, we aimed to test the hypothesis that individuals scoring high on obsessive-compulsive symptoms might particularly exhibit heightened performance monitoring in social responsibility contexts, when their actions have consequences for others. Hence, in **Chapter 4**, we invited healthy individuals scoring high versus low on obsessive-compulsive symptoms to perform a social flanker task, in which mistakes had negative monetary consequences for oneself, another or no one.

Dopamine and oxytocin

Importantly, alterations in social performance monitoring may also arise as a consequence of neurochemical changes in the brain. As discussed above, the reinforcement learning theory of the ERN indicates a critical role for dopamine in the neural mechanisms underlying performance monitoring. Support for dopaminergic involvement in the encoding of prediction errors and performance monitoring is provided by pharmacological manipulations of this neurochemical system. For example, a study applying computational reinforcement learning models to study brain correlates of prediction errors, found that increasing dopamine using the precursor L-DOPA increased BOLD responses to (positive) prediction errors in the VS (Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006). Moreover, pharmacological evidence suggests that the administration of dopamine agonists increases error-related brain activity and prediction error signaling (e.g., Barnes, O'Connell, Nandam, Dean, & Bellgrove, 2014; De Bruijn, Hulstijn, Verkes, Ruigt, & Sabbe, 2004, 2005; Santesso et al., 2009; Spronk et al., 2016), while dopamine antagonists seem to decrease it (De Bruijn, Sabbe, Hulstijn, Ruigt, & Verkes, 2006; Diederen et al., 2017; Jocham, Klein, & Ullsperger, 2011, 2014; Zirnheld et al., 2004). Associated pharmacological alterations of task performance and learning have been observed as well (Diederen et al., 2017; Guitart-Masip et al., 2014; Pessiglione et al., 2006; Pizzagalli et al., 2008; Santesso et al., 2009; Vo, Seergobin, & MacDonald, 2018; Vo, Seergobin, Morrow, & MacDonald, 2016; Zirnheld et al., 2004). This suggests that altering dopamine levels can modify the neural correlates of performance monitoring and prediction error encoding as well as subsequent behavioral adjustments. However, these pharmacological studies have all focused exclusively on contexts where outcomes solely affect oneself. Hence, relatively little is known about the role of this neurochemical in a social context. Research in rodents suggests that social and non-social prediction errors activate the same dopamine-innervated regions and that dopamine neurons play a role in coding social prediction errors (Manduca et al., 2021). This indicates that enhancing dopamine levels could impact performance monitoring by influencing dopamine-driven prediction errors regardless of whether the individual is acting for themselves or others. However, some studies have also shown that administering L-DOPA led to reduced hyperaltruistic behavior (Crockett et al., 2015) and increased selfish behavior in healthy men (Pedroni, Eisenegger, Hartmann, Fischbacher, & Knoch, 2014). Conversely, blocking dopamine transmission using amisulpride reduced prosocial behavior in women and increased selfish behavior in men (Soutschek et al., 2017). Thus, elevating dopamine levels may also lead to a self-serving bias where personal outcomes become more important than those of others, at least in men.

Furthermore, there is another neurochemical that may be particularly relevant to social performance monitoring. Oxytocin, a hormone and neuropeptide synthesized in the paraventricular and supraoptic nucleus in the hypothalamus (Baskerville & Douglas, 2010), is thought to play an essential role in social behavior and cognition. Initially, it was hypothesized that oxytocin facilitates prosocial behavior (MacDonald & MacDonald, 2010), leading to a surge of pharmacological research examining the potential therapeutic value of oxytocin administration. While this initial hypothesis has been questioned due to inconsistent findings (e.g., Harari-Dahan & Bernstein, 2014; Quintana & Guastella, 2020; Shamay-Tsoory & Abu-Akel, 2016), a more recent theory instead posits that oxytocin increases the salience of social cues (Shamay-Tsoory & Abu-Akel, 2016). In line with this, a study from our lab showed that administration of the neuropeptide oxytocin led to enhanced ERNs for mistakes that negatively affected a co-actor (de Bruijn, Ruissen, & Radke, 2017). Furthermore, Martins et al. (2022) found that a small dose of oxytocin boosted prosocial but not self-benefitting prediction-error encoding in the midbrain and sgACC during a Prosocial-learning task and preserved learning performance in the prosocial condition. These findings indicate that administration of oxytocin may specifically facilitate social performance monitoring.

The social salience theory furthermore suggests that effects of oxytocin are achieved through interactions with the brain's dopaminergic system (Shamay-Tsoory & Abu-Akel, 2016), which is supported by preclinical evidence that oxytocin stimulation can enhance dopamine release during social reward (Hung et al., 2017). However, there is a lack of studies investigating the interaction between dopamine and oxytocin in humans. In **Chapters 6 and 7**, we aimed to gain more insight into the underlying neural mechanisms involved in social performance monitoring and the processing of prosocial prediction errors by pharmacologically manipulating dopamine and oxytocin in a single study, making it possible to establish potential overlap and differences in their effects.

OUTLINE OF THE THESIS

The current thesis consists of two main parts. Part 1 focusses on electrophysiological indices of performance monitoring in a social responsibility context and the role of obsessive-compulsive symptoms herein, whereas Part 2 focusses on neurochemical aspects of individual and social performance monitoring. Figure 4 provides a graphical overview of the topics addressed in this dissertation.

In the first part, we address how the social context of being responsible for the outcomes of another person impact error-related brain activations (Chapter 2), how the study of this specific context is relevant to those with OCD (Chapter 3), and how subclinical obsessive-compulsive symptoms impact performance monitoring in a social responsibility context (Chapter 4). More specifically, in **Chapter 2** we introduce a novel performance-monitoring paradigm whereby mistakes resulted in harmful versus non-harmful consequences for a co-actor, to test the hypothesis that performing in a social responsibility context enhances amplitudes of the ERN. In **Chapter 3**, we present a review in which we summarize the role of social cognition in OCD, and in which we highlight the importance of social

(responsibility) symptoms in this disorder and the relevance of using social performance-monitoring paradigms to gain more insight into social and functional impairments in OCD. In **Chapter 4**, we use a social performance-monitoring paradigm in healthy volunteers to test the hypothesis that being responsible for another person's outcome results in enhanced performance monitoring as reflected in the ERN, particularly for those with high obsessive-compulsive symptoms.

In the second part, we investigate neurochemical aspects of individual and social performance monitoring. Specifically, in **Chapter 5** we investigate how menstrual cycle phase, characterized by fluctuating endogenous levels of ovarian hormones (estrogen and progesterone), and use of the most common type of oral contraceptives, characterized by repressed endogenous- and enhanced exogeneous levels of these hormones, impact individual performance monitoring as indexed by the ERN in women. Chapter 6 and 7 are fMRI studies focusing on the putative modulatory role of neurochemicals oxytocin and dopamine on the neural mechanisms underlying performance monitoring in (pro)social versus individual or self-benefitting contexts in men. In **Chapter 6**, we use the Prosocial-learning task to investigate whether dopamine precursor L-DOPA and oxytocin impact the neural processing of prediction errors and associated learning processes when actions resulted in rewards for oneself versus an anonymous other participant. In **Chapter 7**, we use the Cannonball task to test if these same drugs impact the neural correlates of performance monitoring when mistakes negatively affect the monetary bonus of oneself versus another.

Finally, in **Chapter 8** we summarize and discuss the main findings of this thesis and address important challenges and potential future directions in this research field.

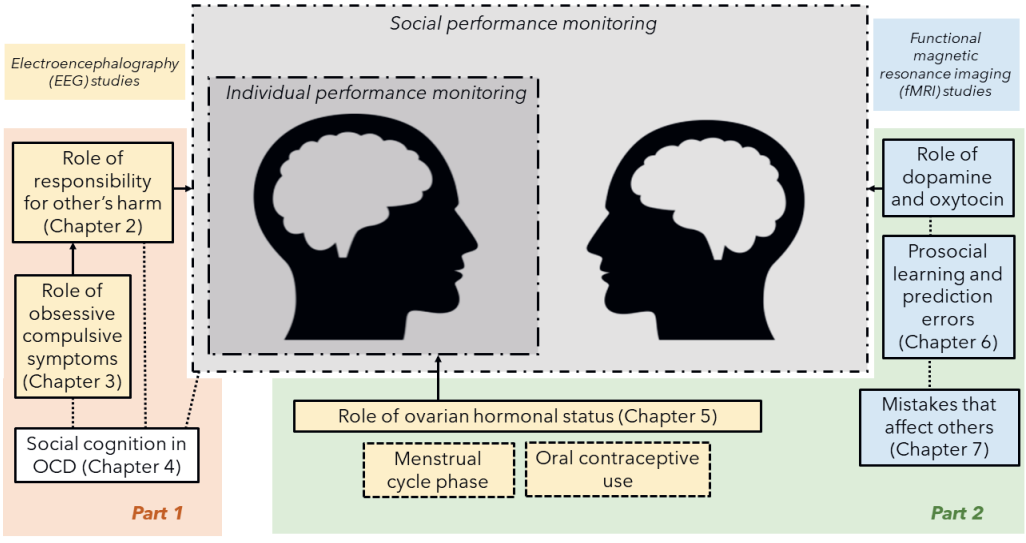


Figure 4. Graphical overview of the topics investigated in this dissertation.

REFERENCES

Barnes, J. J., O'Connell, R. G., Nandam, L. S., Dean, A. J., & Bellgrove, M. A. (2014). Monoaminergic modulation of behavioural and electrophysiological indices of error processing. *Psychopharmacology, 231*(2), 379-392.

Barth, C., Villringer, A., & Sacher, J. (2015). Sex hormones affect neurotransmitters and shape the adult female brain during hormonal transition periods. *Frontiers in neuroscience, 9*, 37.

Baskerville, T. A., & Douglas, A. J. (2010). Dopamine and oxytocin interactions underlying behaviors: potential contributions to behavioral disorders. *CNS neuroscience & therapeutics, 16*(3), e92-e123.

Beltz, A. M., & Moser, J. S. (2020). Ovarian hormones: a long overlooked but critical contributor to cognitive brain structures and function. *Annals of the New York Academy of Sciences, 1464*(1), 156-180.

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.

Brazil, I. A., de Bruijn, E. R., Bulten, B. H., von Borries, A. K. L., van Lankveld, J. J., Buitelaar, J. K., & Verkes, R. J. (2009). Early and late components of error monitoring in violent offenders with psychopathy. *Biological psychiatry, 65*(2), 137-143.

Brazil, I. A., Mars, R. B., Bulten, B. H., Buitelaar, J. K., Verkes, R. J., & De Bruijn, E. R. (2011). A neurophysiological dissociation between monitoring one's own and others' actions in psychopathy. *Biological psychiatry, 69*(7), 693-699.

Carlo, G. (2013). The development and correlates of prosocial moral behaviors. In *Handbook of moral development* (pp. 208-234): Psychology Press.

Chan, R. C., & Gottesman, I. I. (2008). Neurological soft signs as candidate endophenotypes for schizophrenia: a shooting star or a Northern star? *Neuroscience & Biobehavioral Reviews, 32*(5), 957-971.

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.

Cracco, E., Desmet, C., & Brass, M. (2016). When your error becomes my error: anterior insula activation in response to observed errors is modulated by agency. *Social Cognitive and Affective Neuroscience, 11*(3), 357-366.

Crawford, J. R., & Henry, J. D. (2004). The Positive and Negative Affect Schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample. *British journal of clinical psychology, 43*(3), 245-265.

Crockett, M. J., Siegel, J. Z., Kurth-Nelson, Z., Ousdal, O. T., Story, G., Frieband, C., . . . Dolan, R. J. (2015). Dissociable effects of serotonin and dopamine on the valuation of harm in moral decision making. *Current Biology, 25*(14), 1852-1859.

Cui, F., Abdelgabar, A.-R., Keysers, C., & Gazzola, V. (2015). Responsibility modulates pain-matrix activation elicited by the expressions of others in pain. *Neuroimage, 114*, 371-378.

Cutler, J., Wittmann, M. K., Abdurahman, A., Hargitai, L. D., Drew, D., Husain, M., & Lockwood, P. L. (2021). Ageing is associated with disrupted reinforcement learning whilst learning to help others is preserved. *Nature communications, 12*(1), 1-13.

de Bruijn, E. R., de Lange, F. P., von Cramon, D. Y., & Ullsperger, M. (2009). When errors are rewarding. *Journal of Neuroscience, 29*(39), 12183-12186.

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), 151-160.

De Bruijn, E. R., Hulstijn, W., Verkes, R. J., Ruigt, G. S., & Sabbe, B. G. (2005). Altered response evaluation: Monitoring of late responses after administration of {d}-amphetamine. *Journal of Psychophysiology, 19*(4), 311.

de Bruijn, E. R., Jansen, M., & Overgaauw, S. (2020). Enhanced error-related brain activations for mistakes that harm others: ERP evidence from a novel social performance-monitoring paradigm. *Neuroimage, 204*, 116238.

de Bruijn, E. R., Ruissen, M. I., & Radke, S. (2017). Electrophysiological correlates of oxytocin-induced enhancement of social performance monitoring. *Social Cognitive and Affective Neuroscience, 12*(10), 1668-1677.

De Bruijn, E. R., Sabbe, B. G., Hulstijn, W., Ruigt, G. S., & Verkes, R. J. (2006). Effects of antipsychotic and antidepressant drugs on action monitoring in healthy volunteers. *Brain research, 1105*(1), 122-129.

De Bruijn, E. R., & von Rhein, D. T. (2012). Is your error my concern? An event-related potential study on own and observed error detection in cooperation and competition. *Frontiers in neuroscience, 6*, 8.

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.

Dehaene, S., Posner, M. I., & Tucker, D. M. (1994). Localization of a neural system for error detection and compensation. *Psychological science, 5*(5), 303-305.

Diederen, K. M., Ziauddeen, H., Vestergaard, M. D., Spencer, T., Schultz, W., & Fletcher, P. C. (2017). Dopamine modulates adaptive prediction error coding in the human midbrain and striatum. *Journal of Neuroscience, 37*(7), 1708-1720.

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.

Dubol, M., Epperson, C. N., Sacher, J., Pletzer, B., Derntl, B., Lanzenberger, R., . . . Comasco, E. (2021). Neuroimaging the menstrual cycle: a multimodal systematic review. *Frontiers in neuroendocrinology, 60*, 100878.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & psychophysics, 16*(1), 143-149.

Falkenstein, M. (1990). Effects of errors in choice reaction tasks on the ERP under focused and divided attention. *Psychophysiological brain research*.

Fischer, A. G., Danielmeier, C., Villringer, A., Klein, T. A., & Ullsperger, M. (2016). Gender influences on brain responses to errors and post-error adjustments. *Scientific reports, 6*(1), 24435.

Fouragnan, E., Retzler, C., & Philiastides, M. G. (2018). Separate neural representations of prediction error valence and surprise: Evidence from an fMRI meta-analysis. *Human brain mapping, 39*(7), 2887-2906.

Gehring, W. J., Goss, B., Coles, M. G., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological science, 4*(6), 385-390.

Guitart-Masip, M., Economides, M., Huys, Q. J., Frank, M. J., Chowdhury, R., Duzel, E., . . . Dolan, R. J. (2014). Differential, but not opponent, effects of L-DOPA and citalopram on action learning with reward and punishment. *Psychopharmacology, 231*(5), 955-966.

Hamstra, D. (2021). *Mood and the pill*. Leiden University,

Harari-Dahan, O., & Bernstein, A. (2014). A general approach-avoidance hypothesis of oxytocin: accounting for social and non-social effects of oxytocin. *Neuroscience & Biobehavioral Reviews, 47*, 506-519.

Hester, R., Barre, N., Murphy, K., Silk, T. J., & Mattingley, J. B. (2008). Human medial frontal cortex activity predicts learning from errors. *Cerebral Cortex, 18*(8), 1933-1940.

Hezel, D. M., & McNally, R. J. (2016). A theoretical review of cognitive biases and deficits in obsessive–compulsive disorder. *Biological psychology, 121*, 221-232.

Hofmeister, S., & Bodden, S. (2016). Premenstrual syndrome and premenstrual dysphoric disorder. *American family physician, 94*(3), 236-240.

Holroyd, C. B., & Coles, M. G. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychological review, 109*(4), 679.

Holroyd, C. B., Dien, J., & Coles, M. G. (1998). Error-related scalp potentials elicited by hand and foot movements: evidence for an output-independent error-processing system in humans. *Neuroscience letters, 242*(2), 65-68.

Hung, L. W., Neuner, S., Polepalli, J. S., Beier, K. T., Wright, M., Walsh, J. J., . . . Dölen, G. (2017). Gating of social reward by oxytocin in the ventral tegmental area. *Science, 357*(6358), 1406-1411.

Iannaccone, R., Hauser, T. U., Staempfli, P., Walitza, S., Brandeis, D., & Brem, S. (2015). Conflict monitoring and error processing: new insights from simultaneous EEG–fMRI. *Neuroimage, 105*, 395-407.

Jocham, G., Klein, T. A., & Ullsperger, M. (2011). Dopamine-mediated reinforcement learning signals in the striatum and ventromedial prefrontal cortex underlie value-based choices. *Journal of Neuroscience, 31*(5), 1606-1613.

Jocham, G., Klein, T. A., & Ullsperger, M. (2014). Differential modulation of reinforcement learning by D2 dopamine and NMDA glutamate receptor antagonism. *Journal of Neuroscience, 34*(39), 13151-13162.

Keil, J., Weisz, N., Paul-Jordanov, I., & Wienbruch, C. (2010). Localization of the magnetic equivalent of the ERN and induced oscillatory brain activity. *Neuroimage, 51*(1), 404-411.

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

Klein, T. A., Neumann, J., Reuter, M., Hennig, J. r., von Cramon, D. Y., & Ullsperger, M. (2007). Genetically determined differences in learning from errors. *Science, 318*(5856), 1642-1645.

Koban, L., Corradi-Dell'Acqua, C., & Vuilleumier, P. (2013). Integration of error agency and representation of others' pain in the anterior insula. *Journal of Cognitive Neuroscience, 25*(2), 258-272.

Le Guen, M., Schantz, C., Régnier-Loilier, A., & de La Rochebrochard, E. (2021). Reasons for rejecting hormonal contraception in Western countries: A systematic review. *Social science & medicine, 284*, 114247.

Lockwood, P. L., Apps, M. A., & Chang, S. W. (2020). Is there a 'social'brain? Implementations and algorithms. *Trends in cognitive sciences, 24*(10), 802-813.

Lockwood, P. L., Apps, M. A., Valton, V., Viding, E., & Roiser, J. P. (2016). Neurocomputational mechanisms of prosocial learning and links to empathy. *Proceedings of the National Academy of Sciences, 113*(35), 9763-9768.

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

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

MacDonald, K., & MacDonald, T. M. (2010). The peptide that binds: a systematic review of oxytocin and its prosocial effects in humans. *Harvard review of psychiatry, 18*(1), 1-21.

Manduca, A., Carbone, E., Schiavi, S., Cacchione, C., Buzzelli, V., Campolongo, P., & Trezza, V. (2021). The neurochemistry of social reward during development: What have we learned from rodent models? *Journal of Neurochemistry, 157*(5), 1408-1435.

Martins, D., Lockwood, P., Cutler, J., Moran, R., & Paloyelis, Y. (2022). Oxytocin modulates neurocomputational mechanisms underlying prosocial reinforcement learning. *Progress in Neurobiology, 213*, 102253.

Miltner, W. H., Lemke, U., Weiss, T., Holroyd, C., Scheffers, M. K., & Coles, M. G. (2003). Implementation of error-processing in the human anterior cingulate cortex: a source analysis of the magnetic equivalent of the error-related negativity. *Biological psychology, 64*(1-2), 157-166.

Overgaauw, S., Jansen, M., & de Bruijn, E. R. (2020). Self-centered or other-directed: Neural correlates of performance monitoring are dependent on psychopathic traits and social context. *cortex, 129*, 199-210.

Pasion, R., & Barbosa, F. (2019). ERN as a transdiagnostic marker of the internalizing-externalizing spectrum: A dissociable meta-analytic effect. *Neuroscience & Biobehavioral Reviews, 103*, 133-149.

Paul, K., Walentowska, W., Bakic, J., Dondaine, T., & Pourtois, G. (2017). Modulatory effects of happy mood on performance monitoring: Insights from error-related brain potentials. *Cognitive, Affective, & Behavioral Neuroscience, 17*, 106-123.

Pedroni, A., Eisenegger, C., Hartmann, M. N., Fischbacher, U., & Knoch, D. (2014). Dopaminergic stimulation increases selfish behavior in the absence of punishment threat. *Psychopharmacology, 231*(1), 135-141.

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature, 442*(7106), 1042-1045.

Pitman, R. K. (1987). A cybernetic model of obsessive-compulsive psychopathology. *Comprehensive psychiatry, 28*(4), 334-343.

Pizzagalli, D. A., Evins, A. E., Schetter, E. C., Frank, M. J., Pajtas, P. E., Santesso, D. L., & Culhane, M. (2008). Single dose of a dopamine agonist impairs reinforcement learning in humans:

behavioral evidence from a laboratory-based measure of reward responsiveness. *Psychopharmacology, 196*(2), 221-232.
Poromaa, I. S., & Segebladh, B. (2012). Adverse mood symptoms with oral contraceptives. *Acta obstetricia et gynecologica Scandinavica, 91*(4), 420-427.
Proudfit, G. H., Inzlicht, M., & Mennin, D. S. (2013). Anxiety and error monitoring: the importance of motivation and emotion. In (Vol. 7, pp. 636): Frontiers Media SA.
Quintana, D. S., & Guastella, A. J. (2020). An allostatic theory of oxytocin. *Trends in cognitive sciences, 24*(7), 515-528.
Radke, S., De Lange, F., Ullsperger, M., & De Bruijn, E. (2011). Mistakes that affect others: an fMRI study on processing of own errors in a social context. *Experimental brain research, 211*, 405-413.
Rescorla, R. A., Wagner, A. R., Black, A. H., & Prokasy, W. F. (1972). Classical conditioning II: current research and theory.
Riesel, A. (2019). The erring brain: Error-related negativity as an endophenotype for OCD—A review and meta-analysis. *Psychophysiology, 56*(4), e13348.
Ruscio, A. M., Stein, D. J., Chiu, W. T., & Kessler, R. C. (2010). The epidemiology of obsessive-compulsive disorder in the National Comorbidity Survey Replication. *Molecular Psychiatry, 15*(1), 53-63.
Salkovskis, P. M., & Warwick, H. M. (1985). Cognitive therapy of obsessive–compulsive disorder: Treating treatment failures. *Behavioural and Cognitive Psychotherapy, 13*(3), 243-255.
Santesso, D. L., Evins, A. E., Frank, M. J., Schetter, E. C., Bogdan, R., & Pizzagalli, D. A. (2009). Single dose of a dopamine agonist impairs reinforcement learning in humans: Evidence from event-related potentials and computational modeling of striatal-cortical function. *Human brain mapping, 30*(7), 1963-1976.
Schilbach, L. (2016). Towards a second-person neuropsychiatry. *Philosophical Transactions of the Royal Society B: Biological Sciences, 371*(1686), 20150081.
Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience1. *Behavioral and brain sciences, 36*(4), 393-414.
Schultz, W. (2022). Dopamine reward prediction error coding. *Dialogues in clinical neuroscience*.
Shamay-Tsoory, S. G., & Abu-Akel, A. (2016). The social salience hypothesis of oxytocin. *Biological psychiatry, 79*(3), 194-202.
Soutschek, A., Burke, C. J., Raja Beharelle, A., Schreiber, R., Weber, S. C., Karipidis, I. I., . . . Kalenscher, T. (2017). The dopaminergic reward system underpins gender differences in social preferences. *Nature Human Behaviour, 1*(11), 819-827.
Spronk, D. B., Verkes, R. J., Cools, R., Franke, B., Van Wel, J. H., Ramaekers, J. G., & De Bruijn, E. R. (2016). Opposite effects of cannabis and cocaine on performance monitoring. *European Neuropsychopharmacology, 26*(7), 1127-1139.
Themanson, J. R., Rosen, P. J., Pontifex, M. B., Hillman, C. H., & McAuley, E. (2012). Alterations in error-related brain activity and post-error behavior over time. *Brain and cognition, 80*(2), 257-265.
Ullsperger, M., Danielmeier, C., & Jocham, G. (2014). Neurophysiology of performance monitoring and adaptive behavior. *Physiological reviews, 94*(1), 35-79.
Ullsperger, M., & Von Cramon, D. Y. (2001). Subprocesses of performance monitoring: a dissociation of error processing and response competition revealed by event-related fMRI and ERPs. *Neuroimage, 14*(6), 1387-1401.
Vo, A., Seergobin, K. N., & MacDonald, P. A. (2018). Independent effects of age and levodopa on reversal learning in healthy volunteers. *Neurobiology of Aging, 69*, 129-139.
Vo, A., Seergobin, K. N., Morrow, S. A., & MacDonald, P. A. (2016). Levodopa impairs probabilistic reversal learning in healthy young adults. *Psychopharmacology, 233*(14), 2753-2763.
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.
Westhoff, B., Blankenstein, N. E., Schreuders, E., Crone, E. A., & van Duijvenvoorde, A. C. (2021). Increased ventromedial prefrontal cortex activity in adolescence benefits prosocial reinforcement learning. *Developmental cognitive neuroscience, 52*, 101018.
Zirnheld, P. J., Carroll, C. A., Kieffaber, P. D., O'donnell, B. F., Shekhar, A., & Hetrick, W. P. (2004). Haloperidol impairs learning and error-related negativity in humans. *Journal of Cognitive Neuroscience, 16*(6), 1098-1112.