



**Universiteit  
Leiden**  
The Netherlands

## **Enhanced error-related brain activations for mistakes that harm others: ERP evidence from a novel social performance-monitoring paradigm**

Bruijn, E.R.A. de; Jansen, M.; Overgaauw, S.

### **Citation**

Bruijn, E. R. A. de, Jansen, M., & Overgaauw, S. (2019). Enhanced error-related brain activations for mistakes that harm others: ERP evidence from a novel social performance-monitoring paradigm. *Neuroimage*, 204, 116238. doi:10.1016/j.neuroimage.2019.116238

Version: Publisher's Version

License: [Creative Commons CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/)

Downloaded from: <https://hdl.handle.net/1887/138521>

**Note:** To cite this publication please use the final published version (if applicable).



# Enhanced error-related brain activations for mistakes that harm others: ERP evidence from a novel social performance-monitoring paradigm

Ellen R.A. de Bruijn<sup>a,b,\*</sup>, Myrthe Jansen<sup>a,b</sup>, Sandy Overgaauw<sup>a,b</sup>

<sup>a</sup> Leiden University, Department of Clinical Psychology, Leiden, the Netherlands

<sup>b</sup> Leiden Institute for Brain and Cognition (LIBC), Leiden, the Netherlands

## ARTICLE INFO

### Keywords:

Error-related negativity (ERN)  
Event-related potential (ERP)  
Harmful mistakes  
Responsibility  
Social performance monitoring

## ABSTRACT

Our mistakes often have negative consequences for ourselves, but may also harm the people around us. Continuous monitoring of our performance is therefore crucial for both our own and others' well-being. Here, we investigated how modulations in responsibility for other's harm affects electrophysiological correlates of performance-monitoring, viz. the error-related negativity (ERN) and error positivity (Pe). Healthy participants (N = 27) performed a novel social performance-monitoring paradigm in two responsibility contexts. Mistakes made in the harmful context resulted in a negative consequence for a co-actor, i.e., hearing a loud aversive sound, while errors in the non-harmful context were followed by a soft non-aversive sound. Although participants themselves did not receive auditory feedback in either context, they did experience harmful mistakes as more distressing and reported higher effort to perform well in the harmful context. ERN amplitudes were enhanced for harmful compared to non-harmful mistakes. Pe amplitudes were unaffected. The present study shows that performing in a potentially harmful social context amplifies early automatic performance-monitoring processes and increases the impact of the resulting harmful mistakes. These outcomes not only further our theoretical knowledge of social performance monitoring, but also demonstrate a novel and useful paradigm to investigate aberrant responsibility attitudes in various clinical populations.

Knowledge about how our actions may affect others is essential for behaving in a socially appropriate manner. This especially holds for mistakes, as errors made in a social context not only affect ourselves, but they may additionally harm other people. These so-called *social* or *harmful mistakes* are therefore often associated with enhanced feelings of responsibility (de Bruijn et al., 2017; Koban et al., 2013). Consequently, during social interactions it is important to monitor our ongoing performance and to regulate our actions in a way that is aimed at optimizing the interaction (de Bruijn et al., 2012a, 2017). Studies on performance monitoring have mainly focused on non-social settings (for recent reviews see e.g., Gehring et al., 2018; Ullsperger et al., 2014a), but researchers have now also started investigating human performance-monitoring processes in different social contexts (for a review see Koban and Pourtois, 2014). For example, it has been demonstrated that (partly) overlapping neural mechanisms are involved in monitoring own versus other's errors (van Schie et al., 2004; de Bruijn et al., 2009, 2012b; Shane et al., 2008). However, differences may also exist depending on the nature of the social interaction. Monitoring other's performance in cooperative versus competitive contexts may

result, for example, in differential involvement of both monitoring- (Newman-Norlund et al., 2009) and reward-related brain areas (de Bruijn et al., 2009). Although monitoring of errors made by others is crucial for social processes such as observational learning (see e.g., Brazil et al., 2011), the disadvantage of focusing on these types of errors is that the interactive nature of human social behavior is not taken into account. As a result, this type of research that mainly employs a so-called spectator account (Schilbach et al., 2013; Schilbach, 2016) not directly advances our understanding of performance-monitoring processes related to knowledge about how our actions affect others. The aim of the current study is to address this issue by directly comparing mistakes made in harmful versus non-harmful social contexts.

Studies using event-related potentials (ERPs) have identified several components specific for error-detection processes. The most well-known ERP component is the error-related negativity (ERN), a sharp negative deflection peaking around 50–70 ms after an erroneous response (Falkenstein et al., 1990; Gehring et al., 1993). This negative peak is followed by positive deflections, which can be classified into the early error positivity (Pe) and the late Pe (see e.g., de Bruijn et al., 2017; Ullsperger

\* Corresponding author. Department of Clinical Psychology, Wassenaarseweg 52, 2333, AK, Leiden, the Netherlands.

E-mail address: [edebruijn@fsw.leidenuniv.nl](mailto:edebruijn@fsw.leidenuniv.nl) (E.R.A. de Bruijn).

<https://doi.org/10.1016/j.neuroimage.2019.116238>

Received 4 March 2019; Received in revised form 18 July 2019; Accepted 30 September 2019

Available online 1 October 2019

1053-8119/© 2019 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

et al., 2014b). The ERN is thought to result from dopamine-based prediction errors enabling flexible adaptive behavior by triggering short-term behavioral adjustments (see e.g., Debener et al., 2005). The Pe is assumed to be more specifically involved in conscious affective processing of mistakes (Ullsperger et al., 2014a) and/or subjective confidence of one's actions (Boldt and Yeung, 2015). Functional magnetic resonance imaging (fMRI) studies have revealed an emotional and cognitive control network involving posterior medial frontal cortex (pmFC; including anterior midcingulate cortex and pre-supplementary motor area), insular cortex, and (ventral) striatum (see e.g., Debener et al., 2005; de Bruijn et al., 2009; Ullsperger et al., 2014a). Activations in pmFC are increased for erroneous compared to correct actions, while striatum activity is more pronounced for positive compared to negative outcomes. Using simultaneous EEG/fMRI recordings and single-trial analyses, Debener et al. demonstrated that ERN amplitudes were associated with both increased pmFC activations and prolonged reaction times following an erroneous response. This increase in reaction time can be interpreted as error-related adaptive behavior – i.e., taking more time to increase the likelihood of responding correctly on the next trial – and is known as post-error slowing (Dutilh et al., 2012; Rabbitt, 1966; but see Notebaert et al., 2009 for an alternative explanation in terms of re-orienting). In short, people adjust their behavior to achieve the most optimal outcome using various neural mechanisms. However, the efficiency in detecting and monitoring errors is related to individual differences in personality traits, functioning of the autonomic nervous system, and the experience of emotions (Segalowitz and Dywan, 2009).

Indeed, modulations of the ERN have been reported in many studies and the component seems particularly sensitive to individual trait differences and distress associated with the mistake. For example, ERNs are larger when error significance is enhanced, e.g., when an error is punished (e.g., Endrass et al., 2010; Meyer and Gawlowska, 2017; Riesel et al., 2012; Potts, 2011) or while being observed and/or evaluated by another person (Hajcak et al., 2005; Masaki et al., 2017; Voegler et al., 2018). The role of the individual salience of errors also fits well with the many observations of enhanced ERN amplitudes in patients suffering from anxiety-related disorders, such as obsessive-compulsive disorder (for recent reviews see Endrass and Ullsperger, 2014; Perera et al., 2019).

As stated above, most research so far has either focused on performance monitoring in non-social contexts or on processes such as error observation that do not require social interactive behavior. As a result, only little is known about monitoring of performance that may actually affect others. Fortunately, a few recent studies have started shedding light on the latter. Using fMRI, Koban et al. (2013), for example, showed that socially harmful mistakes (i.e., mistakes that caused pain in others) more strongly activate the cingulate-insula network compared to non-harmful errors. Yu et al. (2014) demonstrated enhanced activity in the same network when only the participant in the scanner made a mistake and was hence solely responsible for causing pain in another person compared to the situation in which both participants responded incorrectly and thus shared the responsibility. The authors interpreted these effects in terms of interpersonal guilt, a negative emotional state experienced when inflicting harm on others. In a study from our lab, we have demonstrated that performing in a high responsibility context, i.e., when actions additionally had consequences for a co-actor, was associated with recruitment of dorsal medial prefrontal cortex (dMPFC) an area involved in social-reasoning processes such as sharing or inferring other's states (i.e., *mentalizing*) (Radke et al., 2011). Enhanced activation in this area has also been reported during the observation of a co-actor receiving painful shocks when the participant had full versus shared responsibility for causing the pain (Cui et al., 2015).

However, to our knowledge, only one previous study from our own lab investigated how differences in responsibility for other's harm may affect early automatic error-related ERP components such as the ERN and Pe. In this recent pharmacological study, we demonstrated oxytocin-induced enhancements of the ERN for social compared to non-social mistakes (de Bruijn et al., 2017). Contrary to our expectations though,

this enhancement for socially harmful mistakes was not present in the placebo condition. We speculated that this unexpected finding was the result of the indirect and subtle responsibility manipulation used. More specifically and unlike the existing fMRI studies, the co-actor did not observe the participant's performance and was thus unaware of his/her mistakes. In addition, the participants' mistakes did not directly affect the outcome of the co-actor, but rather affected the (somewhat vague) long-term possibility of winning an additional joint prize after data collection for the entire study would have been completed. The advantage of this method was that it provided a lot of experimental control (i.e., the non-social and social condition only differed with respect to instructions provided), but it may also have reduced the impact of the manipulation. So although previous studies have shown the involvement of social cognitive as well as performance-monitoring related neural mechanisms when performing in high-responsibility contexts, it is unknown if error-related ERP components such as the ERN and Pe are also differently modulated by harmful versus non-harmful mistakes in a social context.

The current study aims at answering this question by having participants perform a social flankers task in both a non-harmful and a harmful condition. Social mistakes were manipulated by providing aversive noise blasts over headphones to a co-actor for each mistake made by the participant in the harmful condition. In the non-harmful condition however, the co-actor would hear a non-aversive soft sound following each mistake. Importantly, participants did not receive these sounds themselves and both performance and feedback was thus similar for them in both conditions. We hypothesized that mistakes in the harmful condition would be associated with enhanced error significance and thus larger ERN and Pe amplitudes compared to the non-harmful condition.

## 1. Materials and methods

### 1.1. Participants

Twenty-nine healthy volunteers participated in our study. Data of two participants had to be removed because of insufficient number of error trials to analyze (<6 error trials, see Olvet and Hajcak, 2009). The data of the remaining 27 healthy volunteers (university students; 17 females, mean age = 22.0 years. SD = 3.7 years) were analyzed. All participants were recruited using the online Leiden University Research Participation System. Exclusion criteria were a diagnosis of psychiatric disorders or use of antidepressant medication. Participants received course credit or financial compensation for their time. All participants provided written informed consent. The research was approved by the local ethical committee of the institute of Psychology and in accordance with the latest version of the Declaration of Helsinki.

### 1.2. Procedure

Two participants were invited to the lab. The experimenter explained that one of them would perform the flanker task while the other person's task was to count their mistakes based on the soft or loud noise he/she would hear over their headphones. We told them that we were both interested in the cognitive abilities of the person performing the flanker task, and in the effects of the interference of aversive sounds on the person counting the mistakes. The participant who counted the mistakes was actually a same-gender confederate who was invited to the lab to make it believable for the actual participant that their mistakes had negative consequences for another person.

Before starting the task, participants completed 40 practice trials "out loud" while the confederate was sitting next to them. During this phase, both the participant and the confederate heard the noise through the speakers whenever the participant made a mistake. This was done in order to make the participant aware of the averseness of the loud noise and the impact it had on the other person. Additionally, we asked both the participant and the confederate to rate the (un)pleasantness of both

the soft and the loud noise on a scale from 1 (very pleasant) to 7 (very unpleasant) as a manipulation check. The participants could not see each other's ratings, but the experimenter always intentionally mentioned that the confederate had chosen for the highest level of unpleasantness in case of the loud noise.

During the task, both participants wore headphones. We explained to the participants that even though they would not actually hear the noise, it would help them to stay focused during task performance. A setup with two computer displays was used. The displays were positioned on two separate tables divided by a screen (see Fig. 1). Both participants were in the same room but were unable to see or hear each other. After completion of the study, participants were debriefed. The experimenter informed them that the other person was a confederate who did not actually hear noise blasts following mistakes of the participant.

### 1.3. Task

We used a novel social variant of the Flanker task (Eriksen and Eriksen, 1974), the so-called error-responsibility task (ERT), in which participants had to respond to the central arrow (< or >) of a string of 5 arrows by a left or right button press. The central arrow can either be the same as the surrounding (i.e., flanking) arrows (congruent trial: <<<<< or >>>>>) or different (incongruent trial: <<<<< or >>>>>). Each trial started with the presentation of a fixation (1000 ms) followed by a blank screen (250 ms). Next, the stimulus was presented for 100 ms. After this, a blank screen was presented during which participants could respond (900 ms). Participants were instructed to respond as fast and accurate as possible.

The ERT consisted of two conditions of 416 trials each including 50% congruent and 50% incongruent trials presented in a random order. A short break was introduced halfway in each condition. The order of the two conditions was counterbalanced across participants. In the "Non-harmful" condition, mistakes made by the participant resulted in the generation of a soft and not unpleasant noise for the other person, while in the "Harmful" condition mistakes resulted in a loud and aversive noise blast delivered over the headphones to the other participant. Note that there were no direct observable consequences for the performing participant in either condition.

During each break, subjective levels of anxiety, frustration, desperation, boredom, and effort were measured using visual analogue scales (VAS) where we asked participants to rate (on a continuous scale of

0–12) how they felt during performance of the previous block of trials. The total task duration was around 30–40 min, including the time spent on answering these questions.

### 1.4. Exit questions

After performing the ERT, participants filled out an exit-questionnaire. Participants were asked to indicate how upset they were when they made a mistake resulting in a soft/loud noise for the other person (1 = not at all, 7 = very much). They also had to indicate more generally how they experienced the task regarding pleasantness (1 = very unpleasant, 7 = very pleasant), difficulty (1 = very difficult, 7 = very easy), motivation (1 = very low, 7 = very high), and how often they were thinking about the consequences for the other person when they made a mistake (1 = never, 7 = always). Finally, they were asked to indicate how much connection they felt with the other person using a pictorial measure of closeness (Aron et al., 1992; 1 = not at all, 7 = very much). The results for the questions directly related to the manipulation are reported below, while the outcomes for the more general task-experience related questions are reported in the [supplementary material](#).

### 1.5. Questionnaires

As part of our standard procedure, we also collected data from several relevant trait questionnaires, such as empathy and psychopathic traits. This was, however, not the main purpose of the study and our sample size is relatively small to investigate possible individual differences in these traits. However, for transparency and completeness, the outcomes of the correlation analyses with these questionnaires are reported in the [supplementary material](#).

### 1.6. EEG data collection and analyses

EEG data were recorded from 31 scalp electrodes located according to an extended version of the international 10–20 system (5 midline electrodes: Fz, FCz, Cz, Pz, Oz; 26 Lateral electrodes: AF3/4, F3/4, F7/8, FC1/2, FC5/6, C3/4, CP1/2, CP5/6, P3/4, P7/8, T7/8; PO3/4, O1/2). Bipolar vertical (below and above the left eye) and horizontal electro-oculograms (EOGs) were recorded. Monopolar recordings were referenced to the common mode sensor (CMS) and drift was corrected with a driven right leg (DRL) electrode (for details see <http://www.biosemi.com/faq/cms&drl.html>). Signals were DC amplified and digitized with a BioSemi ActiveTwo system at a sampling rate of 512 Hz. EEG data was further analyzed offline using Brain Vision Analyzer 2.0 (BVA; Brain Products, Munich, Germany). All signals were re-referenced to the average of the left and right mastoids, filtered with a band-pass filter between 0.02 and 20 Hz and with a notch filter of 50 Hz, followed by a lenient artifact rejection to remove large artifacts. Eye movements were then corrected using the automatic independent component analyses (ICA) for ocular corrections implemented in BVA, using a slope algorithm for blink detection, 512 ICA steps, and an infomax restricted ICA (mean number of components removed was 7). The ICA was followed by a stricter artifact rejection on the main electrodes of interest (Fz, FCz, Cz, and Pz) with the following settings: maximal amplitude difference of 100  $\mu$ V in 200 ms intervals, minimal allowed amplitude: -75  $\mu$ V, maximal allowed amplitude: 75  $\mu$ V. Mean number of trials in averages: Correct Beep = 155, Correct Loud = 160, Error Beep = 31, Error Loud = 25. Significantly more correct trials than error trials entered the averages ( $p < .001$ ), but neither the main effect of condition ( $p = .412$ ) nor the interaction between the two was significant ( $p = .107$ ).

Response-locked ERPs were baseline corrected relative to a 200 ms pre-stimulus baseline and averaged for correct and incorrect responses to incongruent stimuli and for each subject separately. ERN amplitude was determined on these subject averages by subtracting the most negative peak in the 0–150 ms time window after response onset from the most

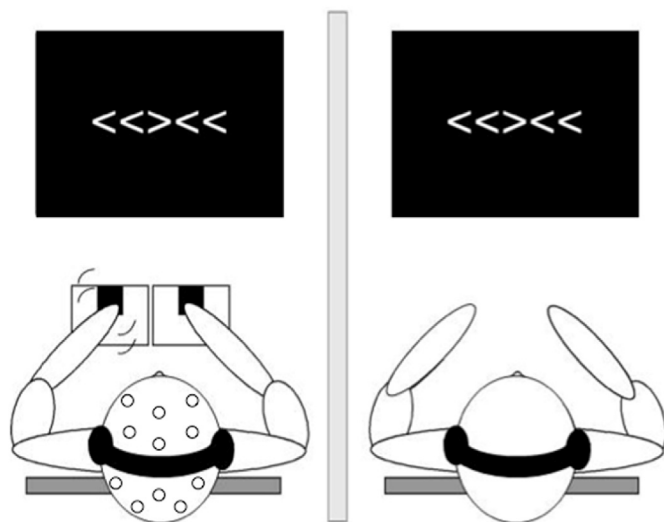


Fig. 1. Experimental setup of the error responsibility task. The left participant is performing the Flankers task, while the right participant is counting their mistakes based on the auditory feedback presented over the headphones.

positive peak in the time window starting 80 ms before and ending 80 ms after response onset at electrodes Fz, FCz, and Cz covering the typical frontocentral distribution of the ERN (see e.g., de Bruijn et al., 2017; Falkenstein et al., 1990; Gehring et al., 1993). Additionally, the ERN was quantified as the mean amplitude in an area of  $\pm 20$  ms around the most negative peak on erroneous responses (cf. Endrass et al., 2014; Riesel et al., 2017, 2019). To limit the influence of baseline fluctuations – visible in the grand average waveforms – a  $-50$  to  $0$  ms baseline correction before determining the mean amplitude was employed. Pe amplitudes were determined for both the early and the late component (Ullsperger et al., 2014a). The early Pe was defined as the most positive peak in the  $150$ – $250$  ms post-response time window and the late Pe was determined as the mean amplitude in the  $300$ – $500$  ms post-response time window at Fz, FCz, Cz, and Pz (see de Bruijn et al., 2017). Note that analyses on stimulus-locked ERP components (separately for correct congruent and incongruent stimuli) are reported in the [supplementary material](#).

### 1.7. Statistical analyses

For the behavioral data, we focused on reaction times, percentage of erroneous responses, and post-error slowing (Rabbitt, 1966). First, all trials with too fast ( $<150$  ms), too slow ( $>800$  ms) or no responses were removed from the dataset (2.76% of all trials). Post-error slowing (PES) was defined in a so-called robust way that has been shown to be less sensitive to global performance fluctuations (Dutilh et al., 2012). PES robust is quantified as a single-trial value of PES by comparing correct trials preceding (pre-error) and following an error (post-error). Only error trials that were both preceded and followed by at least one correct response were included. Mean individual reaction times and error rates were entered into repeated measures general linear models (RM-GLMs) with the possible within-subject factors Congruency (2 levels: congruent, incongruent), Correctness (2 levels: correct, error), Post-error slowing (2 levels: pre-error, post-error), and Context (2 levels: non-harmful, harmful). As erroneous responses to congruent trials are rare and thus better not included in reaction-time analyses, we used two separate statistical models to investigate the presence of standard flankers effects (see e.g., de Bruijn et al., 2017). The first one included only correct reaction times and the factors Congruency and Context. The second ANOVA included only responses to incongruent trials and the factors Correctness and Context. For the ERP data, individual mean amplitudes were entered into RM-GLMs with the possible within-subject factors Correctness, Condition, and Electrode (3 or 4 levels: Fz, FCz, Cz, Pz). Follow-up tests for significant main effects or interactions were Bonferroni corrected for multiple comparisons.

## 2. Results

### 2.1. Manipulation checks

Mean ratings of the exit questions are presented in Table 1. On a scale of 1–7, ranging from not unpleasant at all to very unpleasant, participants rated the soft noise in the non-harmful condition significantly as less unpleasant ( $M = 1.63$ ) than the loud noise in the harmful condition ( $M = 5.74$ ;  $t(26) = -23.96$ ,  $p < .001$ ). Also, the exit questionnaires indicated that participants felt on average more upset when making a mistake in the harmful ( $M = 5.00$ ) compared to the non-harmful condition ( $M = 2.74$ ;  $t(26) = -9.33$ ,  $p < .001$ ). The VAS scales – with scores ranging from 1 to 12 – showed that participants reported more effort to perform well in the harmful ( $M = 9.84$ ,  $SD = 1.81$ ) than in the non-harmful condition ( $M = 9.18$ ,  $SD = 1.82$ ;  $t(26) = -4.18$ ,  $p < .001$ ). Levels of desperation ( $2.19$  vs  $1.47$ ;  $p = .034$ ) and boredom ( $4.30$  vs  $5.05$ ;  $p = .048$ ) were non-significant after Bonferroni correction. Anxiety ( $2.38$  vs  $1.70$ ;  $p = .064$ ) and frustration ( $3.67$  vs  $2.97$ ;  $p = .080$ ) scores did not show significant differences between the two conditions either.

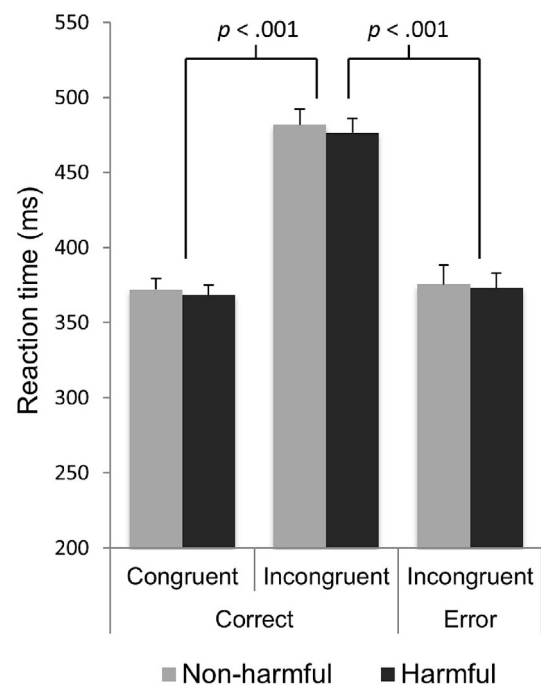
**Table 1**

Mean ratings (standard deviations in parentheses) for the exit questions asked at the end of the experiment.

| Exit Question                                                                                                     | Mean (SD)   |
|-------------------------------------------------------------------------------------------------------------------|-------------|
| Noise in the non-harmful condition (1 = not unpleasant at all, 7 = very unpleasant)                               | 1.63 (.63)  |
| Noise in the harmful condition (1 = not unpleasant at all, 7 = very unpleasant)                                   | 5.74 (.94)  |
| Upset with mistakes in the non-harmful condition (1 = not at all, 7 = very much)                                  | 2.74 (1.48) |
| Upset with mistakes in the harmful condition (1 = not at all, 7 = very much)                                      | 5.00 (1.54) |
| Pleasantness of task in general (1 = very unpleasant, 7 = very pleasant)                                          | 3.78 (1.22) |
| Difficulty of task in general (1 = very difficult, 7 = very easy)                                                 | 4.15 (1.32) |
| General motivation during task (1 = not motivated, 7 = very motivated)                                            | 6.00 (.88)  |
| How often were you thinking about the consequences of your mistakes for the other person? (1 = never, 7 = always) | 3.78 (1.76) |
| Self-other integration (1 = minimal, 7 = maximal)                                                                 | 3.37 (1.69) |

### 2.2. Behavioral results

Reaction-time analyses revealed the expected Flanker effects (see Fig. 2). The analysis on correct RTs only, revealed a significant main effect of Congruency,  $F(1,26) = 685.20$ ,  $p < .001$ ,  $\eta_p^2 = .96$ , with faster responses to correct congruent trials (370 ms) compared to correct incongruent ones (479 ms). The analysis on incongruent trials only, demonstrated a significant main effect of Correctness,  $F(1,26) = 384.31$ ,  $p < .001$ ,  $\eta_p^2 = .94$ , with erroneous responses to incongruent (374 ms) being faster than correct responses to incongruent trials. In both analyses, the main effect of Context was not significant (both  $F_s < 1$ ) and neither did the interactions between the factors and Context reach significance (both  $F_s < 1$ ). With respect to error rates, more errors were made on incongruent trials (14.99%) compared to congruent ones (0.53%;  $F(1,26) = 118.69$ ,  $p < .001$ ,  $\eta_p^2 = .82$ ). Neither the main effect of Context,  $F(1,26) = 2.52$ ,  $p = .135$ ,  $\eta_p^2 = .09$ , nor the interaction between



**Fig. 2.** Mean reaction times for correct and erroneous responses to congruent and incongruent trials in both contexts. Error bars represent standard errors of the mean.



Congruency and Context,  $F(1,26) = 2.30$ ,  $p = .142$ ,  $\eta_p^2 = .08$ , reached significance.

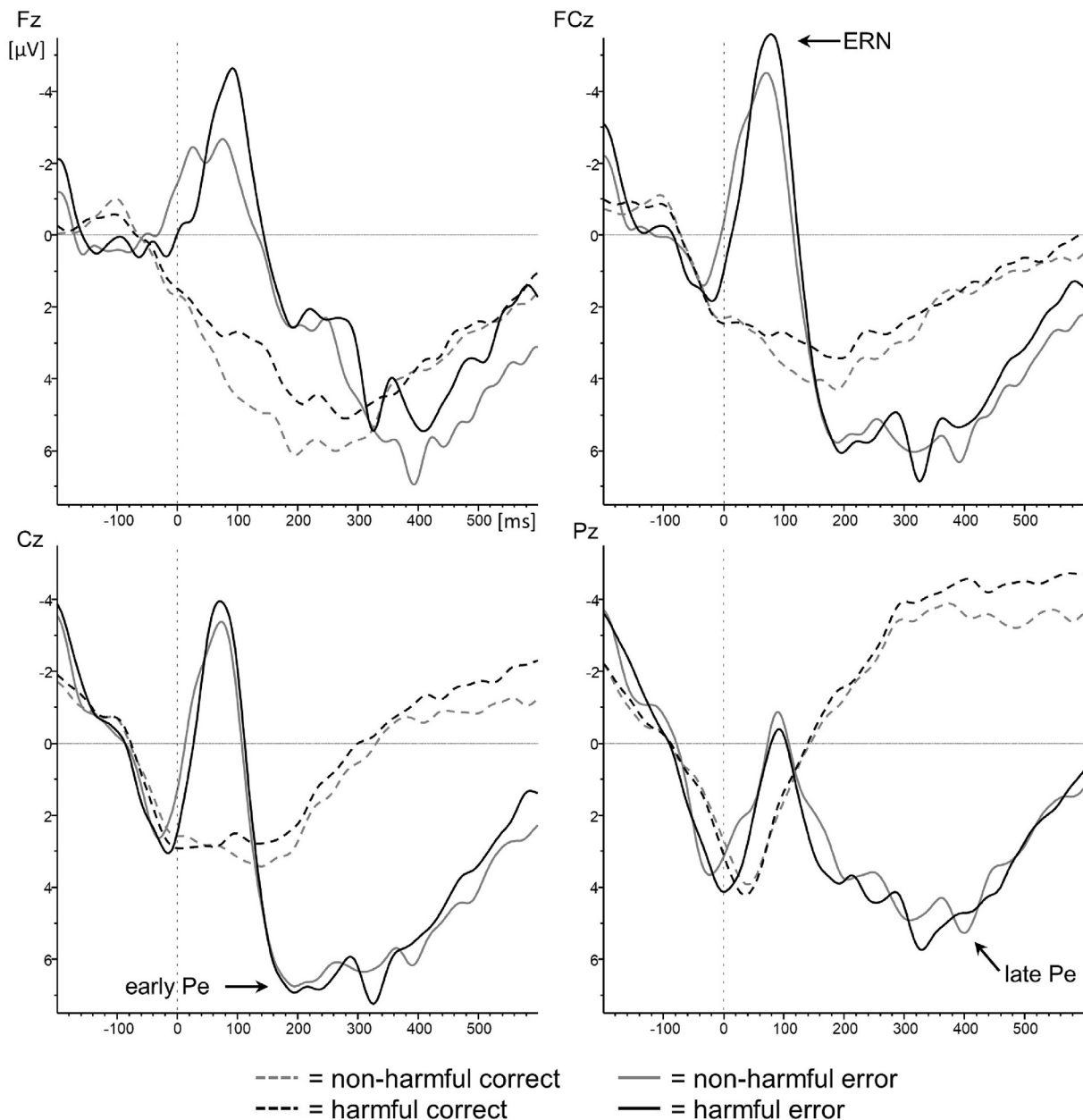
The analyses on adaptive behavior following errors revealed a main effect of post-error slowing,  $F(1,26) = 59.93$ ,  $p < .001$ ,  $\eta_p^2 = .70$ , with slower reaction times for correct responses following a mistake (427 ms) than before a mistake (395 ms). Again, neither the main effect of Context nor the interaction between the two was significant (both  $F_s < 1$ ).

### 2.3. ERN results

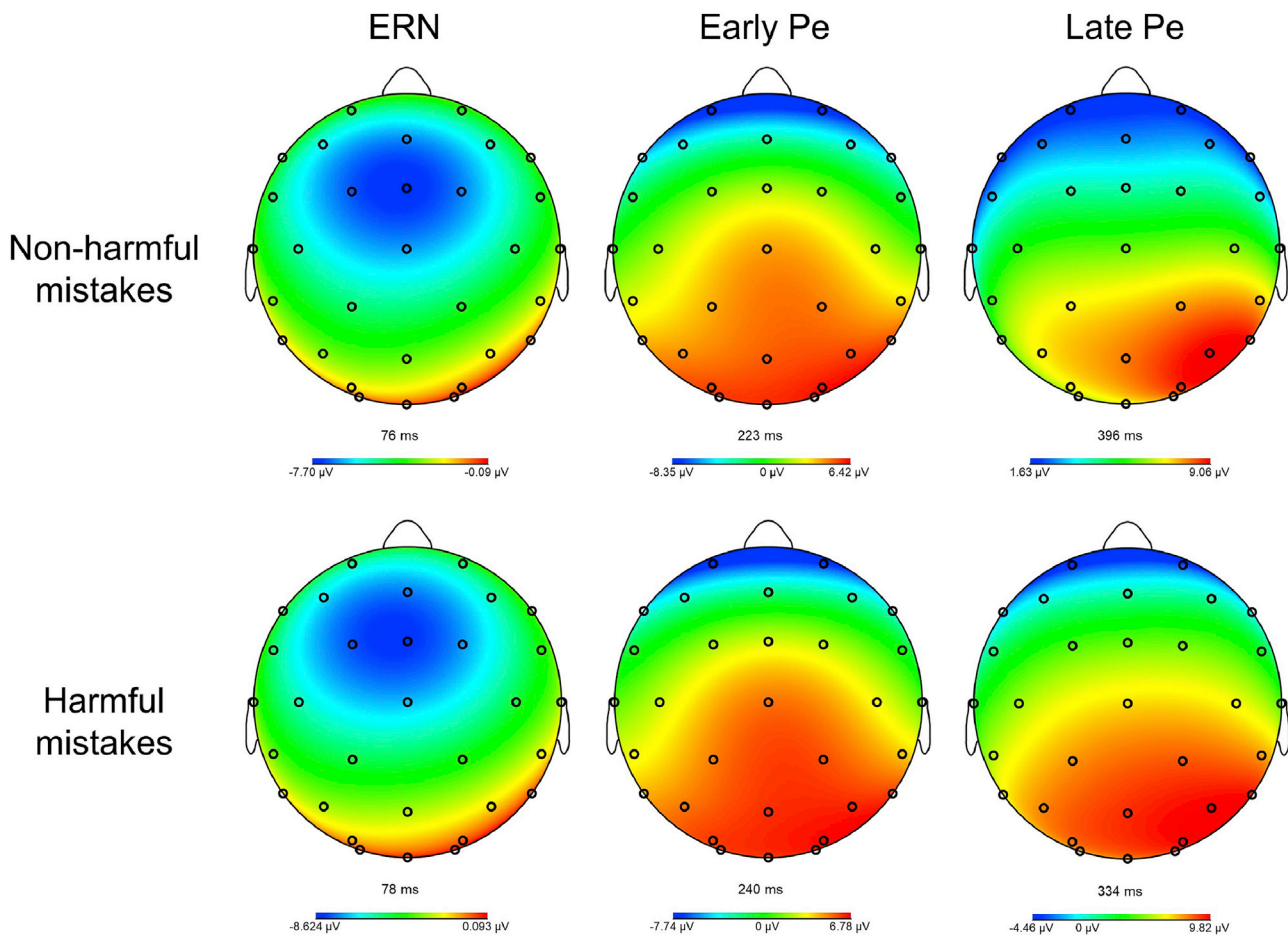
The grand average ERP waveforms for the different conditions at midline electrodes are depicted in Fig. 3A. The analyses on ERN amplitude revealed a main effect of Correctness,  $F(1,26) = 50.82$ ,  $p < .001$ ,  $\eta_p^2 = .66$ , with larger ERN amplitudes following erroneous (−10.42  $\mu\text{V}$ ) compared to correct responses (−4.58  $\mu\text{V}$ ). Neither the main effect of Electrode,  $F(1,26) = 2.06$ ,  $p = .149$ ,  $\eta_p^2 = .14$ , nor the main effect of Context was significant,  $F(1,26) = 2.72$ ,  $p = .111$ ,  $\eta_p^2 = .10$ . However,

both the interaction between Correctness and Context,  $F(1,26) = 4.44$ ,  $p = .045$ ,  $\eta_p^2 = .15$ , and the three-way interaction between Electrode, Context, and Correctness were significant,  $F(2,25) = 3.73$ ,  $p = .038$ ,  $\eta_p^2 = .23$ . None of the remaining two-way interactions were significant (both  $F_s < 1.66$ ,  $p_s > .21$ ).

Follow-up analyses showed that the effect of Context was significant for erroneous responses,  $F(1,26) = 4.94$ ,  $p = .035$ ,  $\eta_p^2 = .16$ , with increased ERN amplitudes in the harmful condition (−10.91  $\mu\text{V}$ ) compared to the non-harmful condition (−9.92  $\mu\text{V}$ ). The effect of Context was not significant for correct responses ( $F < 1$ ). The main effect of Electrode was not significant for the analyses on erroneous responses only,  $F(2,25) = 2.30$ ,  $p = .121$ ,  $\eta_p^2 = .16$ , but numerically ERN amplitudes were maximal at Fz (−10.86  $\mu\text{V}$ ) and FCz (−10.73  $\mu\text{V}$ ) compared to Cz (−9.67  $\mu\text{V}$ ) reflecting the frontocentral topographical distribution of the ERN (see also Fig. 4). In line with this, the differences in ERN amplitude between the two contexts were largest at Fz (1.14  $\mu\text{V}$ ) and FCz (1.15  $\mu\text{V}$ ) compared to Cz (0.70  $\mu\text{V}$ ).



**Fig. 3.** Response-locked grand average ERP waveforms for the non-harmful (grey lines) and the harmful context (black lines). Correct (dashed lines) and incorrect responses (solid lines) are depicted. Response onset is at 0 ms.



**Fig. 4.** Topographical distributions of the ERP components of interest derived from the peak onsets at the difference waveforms (error minus correct) for both the non-harmful and the harmful context.

Please note that the additional analysis using an area around the most negative peak revealed a similar pattern as the peak-to-peak ERN quantifications, with a significant main effect of Context,  $F(1,26) = 4.27$ ,  $p = .049$ ,  $\eta_p^2 = .14$ , reflecting enhanced ERN amplitudes in the harmful ( $-7.23 \mu\text{V}$ ) compared to the non-harmful condition ( $-6.20 \mu\text{V}$ ).

#### 2.4. Early Pe results

The analyses demonstrated a main effect of Electrode,  $F(3,24) = 16.25$ ,  $p < .001$ ,  $\eta_p^2 = .67$  and a main effect of Correctness  $F(1,26) = 10.74$ ,  $p = .003$ ,  $\eta_p^2 = .29$ . The interaction between the two was also significant,  $F(3,24) = 9.34$ ,  $p < .001$ ,  $\eta_p^2 = .54$ . Follow-up tests showed that the effect of Correctness was significant at all electrodes (all  $ps < .011$ ), except for Fz ( $p = .425$ ), reflecting a more centroparietal distribution of the early Pe component compared to the ERN (see Fig. 4). In line with this, numerically the largest difference between correct and incorrect waveforms was seen at Pz ( $5.55 \mu\text{V}$ ) in comparison to Cz ( $5.27 \mu\text{V}$ ), FCz ( $3.25 \mu\text{V}$ ), and Fz ( $1.18 \mu\text{V}$ ). The main effect of Context was not significant ( $F < 1$ ). Neither did any of the remaining two-way and three-way interactions reach significance (all  $Fs < 2.18$ , all  $ps > .15$ ).

#### 2.5. Late Pe results

The late Pe analyses demonstrated a similar pattern as the early Pe, with significant main effects of Electrode,  $F(3,24) = 6.45$ ,  $p = .002$ ,  $\eta_p^2 = .45$  and Correctness,  $F(1,26) = 51.24$ ,  $p < .001$ ,  $\eta_p^2 = .66$ . The interaction between the two was also significant,  $F(3,24) = 9.59$ ,  $p < .001$ ,  $\eta_p^2 = .55$ . Follow-up tests showed that the correctness effect was not significant at

Fz ( $p = .26$ ), but was significant at the other three locations (all  $ps < .001$ ). Again, numerically the largest difference between correct and incorrect waveforms was observed at Pz ( $8.36 \mu\text{V}$ ; Cz =  $6.18 \mu\text{V}$ ; FCz =  $3.59 \mu\text{V}$ ; Fz =  $1.28 \mu\text{V}$ ; see Fig. 4). The main effect of Context was not significant ( $F < 1$ ). Neither did any of the remaining two-way and three-way interactions reach significance (all  $Fs < 1.27$ , all  $ps > .13$ ). Please note that analyses on stimulus-locked ERP components (N1, N2, and P3) did not reveal any effects of the manipulation (see [supplementary material](#)).

#### 2.6. ERP-self report correlations

Exploratory spearman correlation analyses showed that ERN amplitudes at Cz correlated significantly with the self-reported effort participants put in to perform well (as measured with the VAS) in the harmful condition ( $\rho = -.442$ ,  $p = .021$ ). ERN amplitudes at FCz ( $\rho = .502$ ,  $p = .008$ ) and Cz ( $\rho = .436$ ,  $p = .023$ ) also correlated significantly with experienced boredom in the harmful condition. Please note that the ERN is a negative ERP component. Negative correlation coefficients thus reflect larger ERN amplitudes to be associated with higher scores on the questionnaires. For a report of all correlation analyses, see [supplementary material](#).

### 3. Discussion

The aim of the current study was to investigate performance-monitoring processes in harmful and non-harmful social conditions using ERPs. Our manipulation check showed that participants rated the

loud noise more aversive than the soft noise and they felt more upset when making mistakes that harmed others. They also reported higher effort to perform well in harmful contexts than in non-harmful ones. These results indicate that our manipulation was successful and that the error responsibility paradigm thus created an effective harmful and non-harmful social context. Behaviorally, all expected standard flanker task effects were present. Congruency effects were observed for both reaction times and error rates and erroneous responses were faster than correct ones. Also, participants slowed down following a mistake, reflecting post-error slowing (Rabbitt, 1966). These behavioral effects were of similar size in both conditions and also overall reaction times were comparable for the harmful and non-harmful situations. At the electrophysiological level, the results showed that ERN amplitudes at frontocentral electrode locations were larger for mistakes that harmed others compared to non-harmful mistakes. Pe amplitudes, however, were not different for the two conditions.

The absence of manipulation-induced behavioral effects is not very unexpected given previous research. Performance-monitoring studies using (variants of) a flanker task often fail to report behavioral effects, because of the strict instructions that emphasize both speed and accuracy (cf. de Bruijn et al., 2017). As a minimum number of errors are required for the analyses, only a small reaction-time window is available during which responses can be given, thus limiting variability in reaction times. Also, the aim is to keep error frequencies as comparable as possible between the different conditions, as it is known that performance differences, such as dissimilar error rates may affect ERN amplitudes (see e.g., Fischer et al., 2017; Gehring et al., 1993). We can therefore conclude that the currently found effects on ERN amplitude are not confounded by more general differences in performance.

Note that the emphasis on both speed and accuracy may also limit the occurrence of post-error slowing, as this is usually more pronounced when accuracy is emphasized over speed (Danielmeier and Ullsperger, 2011). As a consequence, ERN effects in the absence of effects on behavioral adaptations are also quite common in the literature. For example, studies comparing clinical populations to healthy controls often report this pattern (see e.g., Endrass et al., 2014; Riesel et al., 2019) and almost all previous pharmacological performance-monitoring studies have demonstrated ERN effects in the absence of effects on post-error slowing (see e.g., Barnes et al., 2014; de Bruijn et al., 2004, 2006, 2017; Riba et al., 2005; Spronk et al., 2014, 2016; Zirnheld et al., 2006). Also, an alternative account of post-error slowing states that slowing down following a mistake does not reflect adaptive behavior, but rather a re-orienting process after an infrequent unexpected event (Notebaert et al., 2009). The discussion about what post-error slowing precisely reflects is thus still ongoing in the literature. We are therefore cautious in interpreting the absence of effects of context on this form of behavioral adaptation, e.g., in terms of compensatory mechanisms.

The ERN analyses showed the expected pattern of enhanced ERN amplitudes for erroneous compared to correct responses. Also, the typical topographical distribution of the component was observed with maximum amplitudes at frontocentral electrodes (see e.g., Gehring et al., 2018 for a recent review). In line with our hypotheses the data revealed enhanced ERNs for harmful compared to non-harmful mistakes. This outcome is consistent with the idea that error significance or subjective salience is an important determinant of ERN amplitude and that the component is sensitive to distress (Bartholow et al., 2005). This is particularly interesting as –unlike previous studies– there was no consequence associated with errors for participants themselves in neither condition. Based on the outcomes of the questionnaires, however, it is plausible to assume that making mistakes in the high responsibility context is associated with enhanced distress. The neural generator of the ERN, the anterior midcingulate cortex, is known to not only respond to errors, but to multiple aversive signals such as negative affect, cognitive conflict, and pain (Shackman et al., 2011). The area has therefore been argued to play a central role in adaptive behavior by signaling aversive information. The ERN, originating from this area, may be amplified by

aversive states that indicate a greater need for cognitive control, such as distress (Nash et al., 2014). The present study thus shows that performing in a potentially harmful social context may represent an aversive state that amplifies ERN amplitudes and increases the impact of the resulting harmful mistakes. Social saliency, here expressed in the level of possible harm inflicted on others, is thus sufficient for modulating early performance-monitoring processes reflected in the ERN.

A theoretical explanation of the ERN that also allows mathematical formalization holds that the component reflects a (reward) prediction error that triggers behavioral adjustments, such as error-correction processes or error-prevention strategies (for a recent overview on the physiological principles of performance monitoring, see Ullsperger, Danielmeier and Jocham, 2014b). A prediction error is elicited when expected and actual outcomes differ and scales (negatively) with expectancy. Consistent with this idea, surprising outcomes are associated with greater responses in areas in pmFC including anterior midcingulate cortex (Ullsperger et al., 2014b). More evidence for a role of prediction errors in performance monitoring comes from studies that have demonstrated prediction errors to scale with the ERN elicited by unexpected feedback (the so-called feedback-related negativity or FRN; see e.g., Chase et al., 2011; Fischer and Ullsperger, 2013; Walsh and Anderson, 2013). We propose that the currently self-reported higher effort to perform well in the harmful condition may be related to subjective expectancy and hence the magnitude of prediction errors generated. Specifically, participants' expectations may be altered in high-effort contexts rendering mistakes more surprising than in low-effort situations, which may thus result in relatively larger prediction errors. Although dedicated studies preferably using a combination of single-trial analyses and computational modeling are needed to confirm this explanation, the current data do provide indirect support in the form of the significant correlation between ERN amplitudes and self-reported effort.

While the ERN has mainly been interpreted in terms of fast automatic and unconscious error detection (see e.g., Ullsperger et al., 2014a; Nieuwenhuis et al., 2001), the Pe has been linked to affective conscious processing of the error (see e.g., Ullsperger et al., 2014a; Overbeek et al., 2005). An increase in Pe amplitude for harmful mistakes would thus have fitted the interpretation of subjective saliency and “feeling bad” about inflicting harm on others. Previously, we reported *reduced* late Pe amplitudes for a social context, where mistakes additionally affected a co-actor compared to an individual context where mistakes only had negative consequences for the participant (de Bruijn et al., 2017). However, the main difference with the current study is that participants now were solely responsible for the consequences for the co-actor whereas previously they shared responsibility for their joint outcome. Accordingly, the earlier reported finding of reduced Pe amplitudes could be explained by diffusion of responsibility and social loafing (see e.g., Gilovich et al., 2005; Karau and Williams, 1993; Latané et al., 1979), which is absent in the current paradigm. The presently found ERN effects in the absence of Pe modulations thus suggest that socially harmful mistakes specifically affect early performance-monitoring processes and not, or at least to a lesser extent, later, more affective processes related to error detection.

The current findings are also relevant from a clinical perspective, as differences in the magnitude of the concern people have for others can be observed in various disorders. Enhanced feelings of guilt or responsibility for harm is, for example, observed in OCD (Mancini and Gangemi, 2017; Salkovskis et al., 2000). Patients often worry that harm may come to others because of something they do or fail to do (Hezel and McNally, 2016). Although patients are aware that these behaviors do not have a realistic relation to the possible harm (e.g., taking a long route to the supermarket to prevent a fire in one's parents apartment), they do feel the need to perform this behavior to reduce the distress associated with the thought and/or the likelihood of the event taking place. Conversely, individuals that score high on psychopathic traits are characterized by reduced feelings of guilt and a lack of remorse (Cleckley, 1982; Hare and McPherson, 1984; Prado et al., 2016), which may promote antisocial



behavior (Blair, 1995; Cima et al., 2009). Also, psychopathic individuals often do not take responsibility for their actions and show decreased concern for how their behavior may negatively affect others (Hare, 1980). Studying performance monitoring in socially harmful contexts may thus provide important insight into the often reported but still poorly understood altered performance-monitoring processes in these disorders. For example, enhanced ERN amplitudes have been repeatedly demonstrated in OCD and have therefore been proposed to reflect an endophenotype of the disorder (Endrass and Ullsperger, 2014; Riesel et al., 2015). However, excessive performance monitoring is not exclusive to OCD, but is also observed in anxiety and depressive disorders and may thus represent a more general biomarker (Olvet and Hajcak, 2008; Weinberg et al., 2012). Investigating processing of harmful mistakes from a social perspective in these various disorders may thus facilitate formulating a more specific endophenotype thereby advancing our understanding of disturbed performance monitoring in these clinical populations.

In addition, future studies could focus in more detail on the role of individual differences in relevant personality traits such as empathy or perspective taking. Note that we presently did not design our study to specifically investigate this issue (e.g., the sample size was too small, no pre-selection on the basis of these traits took place, etc.), but the current data did not provide evidence for a central role of these traits either (see supplementary results). Alternatively, one could argue that the impact of the manipulation was simply not strong enough for these patterns to emerge. Manipulations, such as establishing prior cooperative or competitive states (cf. Ruissen and de Bruijn, 2016) may enhance the impact of the manipulation and the role that certain traits play in these processes. Such manipulations may then help in disentangling the exact role of task-induced affective processes or personality traits in modulations of social performance monitoring.

Another interesting option for future research is to further investigate the absence of a behavioral equivalent of the currently found context effects in our ERP measures. Using a probabilistic learning paradigm, for example, one could establish whether participants also learn faster from errors or negative feedback when negative consequences for another person are involved. Support for the possible influence of social manipulations on performance monitoring in the context of learning comes from a recent study by Voegler et al. (2019) in patients with social anxiety disorder. The results demonstrated reduced learning from negative feedback in these patients as well as modulations of the feedback-related negativity specifically under social observation. In short, we emphasize the importance of employing experimental designs that incorporate social manipulations to increase our understanding of the complex interplay of social and personality factors in performance monitoring as well as its clinical relevance.

To conclude, using a novel social performance-monitoring paradigm, we demonstrated enhanced early performance monitoring as reflected in the ERN for harmful compared to non-harmful mistakes. This finding not only extends existing fMRI studies that demonstrated the involvement of both social-cognitive and performance-monitoring related neural mechanisms, but it also reveals that social saliency in itself is sufficient to modulate early, automatic performance-monitoring processes. The present outcomes are consistent with theories that propose that the distress associated with errors scales with ERN amplitudes and that the ERN reflects prediction errors that are used to improve performance through adaptive behavior. As a result, the current study not only furthers our theoretical and fundamental knowledge of performance monitoring, but also opens up new research avenues into alterations in these processes often observed in clinical disorders characterized by aberrant responsibility attitudes.

## Acknowledgments

This work was supported by a personal grant from the Netherlands Organization for Scientific Research to E.D.B (NWO; VIDI grant nr. 452-

12-005). The authors would like to thank Sandor de Munck and Margit Ruissen for their help in data collection.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.neuroimage.2019.116238>.

## References

- Aron, A., Aron, E.N., Smollan, D., 1992. Inclusion of other in the self scale and the structure of interpersonal closeness. *J. Personal. Soc. Psychol.* 63 (4), 596.
- 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.
- 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, 33–42.
- Blair, R., 1995. A cognitive developmental approach to morality: investigating the psychopath. *Cognition* 57, 1–29.
- Boldt, A., Yeung, N., 2015. Shared neural markers of decision confidence and error detection. *J. Neurosci.* 35, 3478–3484.
- Brazil, I.A., Mars, R.B., Bulten, B.H., Buitelaar, J.K., Verkes, R.J., De Bruijn, E.R., 2011. Aneurophysiological dissociation between monitoring one's own and others' actions in psychopathy. *Biol. Psychiatry* 69, 693–699.
- Chase, H.W., Swanson, R., Durham, L., Benham, L., Cools, R., 2011. Feedback-related negativity codes prediction error but not behavioral adjustment during probabilistic reversal learning. *J. Cogn. Neurosci.* 23, 936–946.
- Cima, M., Tonnaer, F., Hauser, M.D., 2009. Psychopaths know right from wrong but don't care. *Soc. Cogn. Affect. Neurosci.* 5, 59–67.
- Cleckley, H.M., 1982. *The Mask of Sanity*. New American library, New York.
- Cui, F., Abdelgabar, A.R., Keyesers, C., Gazzola, V., 2015. Responsibility modulates pain-matrix activation elicited by the expressions of others in pain. *Neuroimage* 114, 371–378.
- Danielmeier, C., Ullsperger, M., 2011. Post-error adjustments. *Front. Psychol.* 2, 233.
- de Bruijn, E.R., de Lange, F.P., von Cramon, D.Y., Ullsperger, M., 2009. When errors are rewarding. *J. Neurosci.* 29, 12183–12186.
- de Bruijn, H., Hulstijn, W., Verkes, R.J., Ruijt, G.S., Sabbe, B.G., 2004. Drug-induced stimulation and suppression of action monitoring in healthy volunteers. *Psychopharmacology (Berl)* 177, 151–160.
- de Bruijn, E.R.A., Mars, R.B., Bekkering, H., Coles, M.G.H., 2012. Your mistake is my mistake... or is it? Behavioral adjustments following own and observed actions in cooperative and competitive contexts. *Q. J. Exp. Psychol.* 65, 317–325.
- de Bruijn, E.R.A., Ruissen, M.I., Radke, S., 2017. Electrophysiological correlates of oxytocin-induced enhancement of social performance monitoring. *Soc. Cogn. Affect. Neurosci.* 12, 1668–1677.
- de Bruijn, Sabbe, B.G., Hulstijn, W., Ruijt, G.S., Verkes, R.J., 2006. Effects of antipsychotic and antidepressant drugs on action monitoring in healthy volunteers. *Brain Res.* 1105, 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. *Front. Neurosci.* 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. *J. Neurosci.* 25, 11730–11737.
- Dutilh, G., van Ravenzwaaij, D., Nieuwenhuis, S., van der Maas, H.L.J., Forstmann, B.U., Wagenmakers, E.J., 2012. How to measure post-error slowing, a confound and a simple solution. *J. Math. Psychol.* 56, 208–216.
- Endrass, T., Schuermann, B., Kaufmann, C., Spielberg, R., Kniesche, R., Kathmann, N., 2010. Performance monitoring and error significance in patients with obsessive-compulsive disorder. *Biol. Psychol.* 84, 257–263.
- Endrass, T., Riesel, A., Kathmann, N., Buhlmann, U., 2014. Performance monitoring in obsessive-compulsive disorder and social anxiety disorder. *J. Abnorm. Psychol.* 123 (4), 705–714.
- Endrass, T., Ullsperger, M., 2014. Specificity of performance monitoring changes in obsessive-compulsive disorder. *Neurosci. Biobehav. Rev.* 46, 124–138.
- Eriksen, B.A., Eriksen, C.W., 1974. Effects of noise letters upon the identification of a target letter in a nonsearch task. *Percept. Psychophys.* 16, 143–149.
- Falkenstein, M., Hohnsbein, J., Hoormann, J., Blanke, L., 1990. Effects of errors in choice reaction tasks on the ERP under focused and divided attention. In: Brunia, C.H.M., Gaillard, A.W.K., Kok, A. (Eds.), *Psychophysiological Brain Research*. Tilburg University Press, pp. 192–195.
- Fischer, A.G., Klein, T.A., Ullsperger, M., 2017. Comparing the error-related negativity across groups: the impact of error- and trial-number differences. *Psychophysiology* 54, 998–1009.
- Fischer, A.G., Ullsperger, M., 2013. Real and fictive outcomes are processed differently but converge on a common adaptive mechanism. *Neuron* 79, 1243–1255.
- Gehring, W.J., Goss, B., Coles, M.G.H., Meyer, D.E., Donchin, E., 2018. The error-related negativity. *Perspect. Psychol. Sci.* 13, 200–204.
- Gehring, W.J., Goss, B., Coles, M.G.H., Meyer, D.E., Donchin, E., 1993. A neural system for error detection and compensation. *Psychol. Sci.* 4, 385–390.

- Gilovich, T., Keltner, D., Nisbett, R., 2005. *Social Psychology*, first ed. W.W. Norton & Company, New York, NY.
- Hajcak, G., Moser, J.S., Yeung, N., Simons, R.F., 2005. On the ERN and the significance of errors. *Psychophysiology* 42, 151–160.
- Hare, R.D., McPherson, I.M., 1984. Violent and aggressive behaviour by criminal psychopaths. *Int. J. Law Psychiatry* 7, 35–50.
- Hare, R.D., 1980. A research scale for the assessment of psychopathy in criminal populations. *Personal. Individ. Differ.* 1, 111–119.
- Hezel, D.M., McNally, R.J., 2016. A Theoretical review of cognitive biases and deficits in obsessive-compulsive disorder. *Biol. Psychol.* 121, 221–232.
- Karau, S.J., Williams, K.D., 1993. Social loafing, A meta-analytic review and theoretical integration. *J. Personal. Soc. Psychol.* 65, 681–706.
- Koban, L., Pourtois, G., 2014. Brain systems underlying the affective and social monitoring of actions: an integrative review. *Neurosci. Biobehav. Rev.* 46, 71–84.
- Koban, L., Corradi-Dell'Acqua, C., Vuilleumier, P., 2013. Integration of error agency and representation of others' pain in the anterior insula. *J. Cogn. Neurosci.* 25, 258–272.
- Latané, B., Williams, K., Harkins, S., 1979. Many hands make light the work, the causes and consequences of social loafing. *J. Personal. Soc. Psychol.* 37, 822–832.
- Mancini, F., Gangemi, A., 2017. Obsessive patients and deontological guilt: a review. *Psychopathol. Rev.* 4, 155–168.
- Masaki, H., Maruo, Y., Meyer, A., Hajcak, G., 2017. Neural correlates of choking under pressure: athletes high in sports anxiety monitor errors more when performance is being evaluated. *Dev. Neuropsychol.* 42, 104–112.
- Meyer, A., Gawlowska, M., 2017. Evidence for specificity of the impact of punishment on error-related brain activity in high versus low trait anxious individuals. *Int. J. Psychophysiol.* 120, 157–163.
- Nash, K., Prentice, M., Hirsh, J., McGregor, I., Inzlicht, M., 2014. Muted neural response to distress among securely attached people. *Soc. Cogn. Affect. Neurosci.* 9, 1239–1245.
- Newman-Norlund, R.D., Ganesh, S., van Schie, H.T., De Bruijn, E.R.A., Bekkering, H., 2009. Self-identification and empathy modulate error-related brain activity during the observation of penalty shots between friend and foe. *Soc. Cogn. Affect. Neurosci.* 4, 10–22.
- 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, 752–760.
- Notebaert, W., Houtman, F., Opstal, F.V., Gevers, W., Fias, W., Verguts, T., 2009. Post-error slowing: an orienting account. *Cognition* 111, 275–279.
- Olvet, D.M., Hajcak, G., 2008. The error-related negativity (ERN) and psychopathology: toward an endophenotype. *Clin. Psychol. Rev.* 28, 1343–1354.
- Olvet, D.M., Hajcak, G., 2009. The stability of error-related brain activity with increasing trials. *Psychophysiology* 46, 957–961.
- Overbeek, T.J.M., Nieuwenhuis, S., Ridderinkhof, K.R., 2005. Dissociable components of error processing, on the functional significance of the Pe vis-a-vis the ERN/Ne. *J. Psychophysiol.* 19, 319–329.
- Perera, M.P.N., Bailey, N.W., Herring, S.E., Fitzgerald, P.B., 2019. Electrophysiology of obsessive compulsive disorder: a systematic review of the electroencephalographic literature. *J. Anxiety Disord.* 62, 1–14.
- Potts, G.F., 2011. Impact of reward and punishment motivation on behavior monitoring as indexed by the error-related negativity. *Int. J. Psychophysiol.* 81, 324–331.
- Prado, C.E., Treeby, M.S., Crowe, S.F., 2016. Examining the relationships between sub-clinical psychopathic traits with shame, guilt and externalization response tendencies to everyday transgressions. *J. Forensic Psychiatry Psychol.* 27, 569–585.
- Rabbitt, P.M., 1966. Errors and error correction in choice-response tasks. *J. Exp. Psychol.* 71, 264–272.
- Radke, S., de Lange, F.P., Ullsperger, M., de Bruijn, E.R.A., 2011. Mistakes that affect others: an fMRI study on processing of own errors in a social context. *Exp. Brain Res.* 211, 405–413.
- Riba, J., Rodríguez-Fornells, A., Morte, A., Münte, T.F., Barbanoj, M.J., 2005. Noradrenergic stimulation enhances human action monitoring. *J. Neurosci.* 25 (17), 4370–4374.
- Riesel, A., Endrass, T., Auerbach, L.A., Kathmann, N., 2015. Overactive performance monitoring as an endophenotype for obsessive-compulsive disorder: evidence from a treatment study. *Am. J. Psychiatry* 172, 665–673.
- Riesel, A., Weinberg, A., Endrass, T., Kathmann, N., Hajcak, G., 2012. Punishment has a lasting impact on error-related brain activity. *Psychophysiology* 49, 239–247.
- Riesel, A., Goldhahn, S., Kathmann, N., 2017. Hyperactive performance monitoring as a transdiagnostic marker: results from health anxiety in comparison to obsessive-compulsive disorder. *Neuropsychologia* 96, 1–8.
- Riesel, A., Klawohn, J., Grützmann, R., Kaufmann, C., Heinzl, S., Bey, K., Lennertz, L., Wagner, M., Kathmann, N., 2019. Error-related brain activity as a transdiagnostic endophenotype for obsessive-compulsive disorder, anxiety and substance use disorder. *Psychol. Med.* 1–11.
- Ruissen, M.I., de Bruijn, E.R.A., 2016. Competitive game play attenuates self-other integration during joint task performance. *Front. Psychol.* 7, 274.
- Salkovskis, P.M., Wroe, A.L., Gledhill, A., Morrison, N., Forrester, E., Richards, C., Reynolds, M., Thorpe, S., 2000. Responsibility attitudes and interpretations are characteristic of obsessive compulsive disorder. *Behav. Res. Ther.* 38, 347–372.
- Schilbach, L., 2016. Toward a second-person neuropsychiatry. *Phil. Trans. R. Soc. B* 371, 1686.
- Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., Vogeley, K., 2013. Toward a second-person neuroscience. *Behav. Brain Sci.* 36, 393–414.
- Segalowitz, S.J., Dywan, J., 2009. Individual differences and developmental change in the ERN response: implications for models of ACC function. *Psychol. Res.* 73, 857–870.
- 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. *Nat. Rev. Neurosci.* 12, 154–167.
- Shane, M.S., Stevens, M., Harenski, C.L., Kiehl, K.A., 2008. Neural correlates of the processing of another's mistakes: a possible underpinning for social and observational learning. *Neuroimage* 42, 450–459.
- Ullsperger, M., Fischer, A.G., Nigbur, R., Endrass, T., 2014a. Neural mechanisms and temporal dynamics of performance monitoring. *Trends Cogn. Sci.* 18, 259–267.
- Sprong, D.B., Dumont, G.J.H., Verkes, R.J., De Bruijn, 2014. The acute effects of MDMA and ethanol administration on electrophysiological correlates of performance monitoring in healthy volunteers. *Psychopharmacology* 231 (14), 2877–2888.
- Sprong, D.B., Verkes, R.J., Cools, R., Franke, B., Van Wel, Ramaekers, J.G., et al., 2016. Opposite effects of cannabis and cocaine on performance monitoring. *Eur. Neuropsychopharmacol.* 26, 1127–1139.
- Ullsperger, M., Danielmeier, C., Jocham, G., 2014b. Neurophysiology of performance monitoring and adaptive behavior. *Physiol. Rev.* 94, 35–97.
- van Schie, H.T., Mars, R.B., Coles, M.G., Bekkering, H., 2004. Modulation of activity in medial frontal and motor cortices during error observation. *Nat. Neurosci.* 7, 549–554.
- Voegler, R., Peterburs, J., Lemke, H., Ocklenburg, S., Liepelt, R., Straube, T., 2018. Electrophysiological correlates of performance monitoring under social observation in patients with social anxiety disorder and healthy controls. *Biol. Psychol.* 132, 71–80.
- Voegler, R., Peterburs, J., Bellebaum, C., Straube, T., 2019. Modulation of feedback processing by social context in social anxiety disorder (SAD)-an event-related potentials (ERPs) study. *Sci. Rep.* 9, 4795.
- Weinberg, A., Riesel, A., Hajcak, G., 2012. Integrating multiple perspectives on error-related brain activity: the ERN as a neural indicator of trait defensive reactivity. *Motiv. Emot.* 36, 84–100.
- Walsh, M.M., Anderson, J.R., 2013. Electrophysiological responses to feedback during the application of Abstract rules. *J. Cogn. Neurosci.* 25, 1986–2002.
- Yu, H., Hu, J., Hu, L., Zhou, X., 2014. The voice of conscience: neural bases of interpersonal guilt and compensation. *Soc. Cogn. Affect. Neurosci.* 9, 1150–1158.
- Zirnheld, P.J., Carroll, C.A., Kieffaber, P.D., O'Donnell, B.F., Shekhar, A., Hetrick, W.P., 2006. Haloperidol impairs learning and error-related negativity in humans. *J. Cogn. Neurosci.* 16, 1098–1112.