

ORIGINAL ARTICLE OPEN ACCESS

Right but Still Lousy: Correct Responses With an Unfavorable Outcome Elicit an Error Positivity

Peter Löschner  | Marco Steinhauser

Department of Psychology, Catholic University of Eichstätt-Ingolstadt, Eichstätt, Bavaria, Germany

Correspondence: Peter Löschner (peter.loeschner@ku.de)

Received: 26 November 2024 | **Revised:** 18 July 2025 | **Accepted:** 8 September 2025

Funding: This research was supported by a grant from the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) to Marco Steinhauser (Grant STE 1708/4-2).

Keywords: error monitoring | error positivity | error-related negativity | event-related potentials

ABSTRACT

The error positivity (Pe) is a neural correlate of performance monitoring that is observed after errors in choice tasks. The results of previous studies suggest that the Pe reflects a monitoring process that goes beyond the mere distinction between correct and incorrect responses. Here, we investigated the idea that the Pe represents a higher-order error signal reflecting an inference-based outcome evaluation. To this end, we created a multistage task whose overall outcome depended on the correctness of each individual stage and was revealed not until the last stage. This implied that the final response could lead to an unfavorable outcome even if it was objectively correct. Our results replicated the general finding that a Pe occurs immediately after errors within each stage. Crucially, we also obtained a Pe after correct responses associated with an unfavorable outcome at the final stage. Moreover, a pattern classifier trained to decode this higher-order Pe successfully decoded the Pe for incorrect responses. These results suggest that the Pe represents an evaluative process that infers the outcome by integrating multiple error signals and taking context into account.

1 | Introduction

To efficiently adapt behavior, monitoring processes are required that detect negative outcomes and adjust behavior accordingly. In experimental tasks with a single behavioral response, this monitoring mainly involves distinguishing between correct and incorrect responses. In such scenarios, event-related potentials (ERPs) such as the error-related negativity (Ne/ERN; Falkenstein et al. 1991; Gehring et al. 1993) and the error positivity (Pe; Overbeek et al. 2005) are studied as they represent distinct processes of error monitoring. However, everyday tasks often consist of several behavioral responses distributed over multiple stages, and the overall outcome of such a multistage task can be determined not until the last stage is finished. In such a scenario, not only the correctness of single responses but the outcome of the whole multistage task must be evaluated.

Here, we hypothesize that the Pe represents a higher-order error signal that indicates the overall outcome in a multistage task. To investigate this, we created a multistage task in which the overall outcome can be unfavorable even if the last response is correct and tested whether this response leads to a Pe.

The Pe is a slow parietal positive deflection that peaks between 200 and 400 ms post-response (Falkenstein et al. 1991; Overbeek et al. 2005). It has been linked to error awareness (Endrass et al. 2007; Nieuwenhuis et al. 2001; O'Connell et al. 2007), leading to the theory that the Pe reflects the accumulation of evidence that an error has been made (Steinhauser and Yeung 2010, 2012; Ullsperger et al. 2010). The Pe has also been associated with decision confidence (Boldt and Yeung 2015), motivational processes (Drizinsky et al. 2016; Schroder et al. 2014) or negative affect (Falkenstein

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Psychophysiology* published by Wiley Periodicals LLC on behalf of Society for Psychophysiological Research.

et al. 2000; Rodeback et al. 2020). In contrast, the Ne/ERN is an earlier fronto-central negativity observed immediately following erroneous responses, peaking between 50 and 100 ms post-response (Falkenstein et al. 2000; Ullsperger, Danielmeier, and Jocham 2014). Theoretical accounts suggest that the Ne/ERN reflects a mismatch between the given and actually correct response (Coles et al. 2001), a post-response conflict (Yeung et al. 2004) or a prediction error (Holroyd and Coles 2002). In each of these accounts, the Ne/ERN represents a rapid error signal indicating that the immediately preceding response is incorrect.

While the Pe and Ne/ERN have initially been interpreted as reflecting different stages of the same error monitoring process, there is increasing evidence that they represent dissociable monitoring systems that evaluate behavior in different ways. Di Gregorio et al. (2018) showed that, in a situation in which participants can infer that an error has occurred without knowing the correct response, a Pe can occur in the absence of an Ne/ERN (see also Hewig et al. 2011). Their results suggest that, while the Ne/ERN relies on a comparison of correct and actual responses, the Pe represents an inference-based evaluation that takes context information into account. In another study (Pfister et al. 2016) participants could intentionally commit errors by breaking task rules. These intentional errors did not elicit an Ne/ERN but a Pe-like positivity, which might reflect that intentional errors are interpreted as a negative outcome (resulting in a Pe) even though they are not incorrect as they do not deviate from an intention (resulting in no Ne/ERN). Finally, dual-tasking experiments revealed that a Pe can even occur after a correct response (Steinhauser and Steinhauser 2021). When two tasks were performed in close succession, an error in the first task elicited a Pe following the correct response in the second task. This was interpreted as reflecting a temporal deferment of error monitoring if both tasks overlap in time.

The results described above suggest that the Pe represents an outcome evaluation that goes beyond the mere distinction between correct and incorrect as reflected by the Ne/ERN. Here, we propose that the Pe represents a higher-order error signal that is based on an inference-based outcome evaluation which takes context information into account. To investigate this idea, we consider a multistage task that consists of a series of decisions, but whose overall outcome depends on the correctness of each individual stage and is revealed not until the last stage. Such a task requires error monitoring on two hierarchical levels: the level of individual responses and the level of the overall outcome. Crucially, we assume that the Ne/ERN indicates error monitoring on the response level, whereas the Pe represents a higher-order signal that tracks the overall outcome. To test this, we created a situation within our multistage task in which the response at the last stage is correct, whereas the overall outcome is suboptimal. For this situation, we predicted that the correct response will elicit a Pe (but no Ne/ERN) indicating that the overall outcome is unfavorable. Such a result would not only demonstrate that the Pe relates to a higher-order error signal reflecting the outcome of a task (rather than the correctness of a response), it would provide further evidence for a dissociation between Ne/ERN and Pe (Di Gregorio et al. 2018).

1.1 | The Present Study

In the present study, we created an experimental paradigm inspired by the game “Rock, Paper, Scissors”. In the original game, players repeatedly form one of three hand gestures representing a rock, a paper, and a scissor. To determine which player has won the round, the three gestures establish an intransitive set of relations with rock winning against scissor, scissor winning against paper, and paper winning against rock. If both players show the same gestures, the round is counted as draw. In our version, this game is turned into a two-stage decision in which the final outcome depended on both decision stages (Figure 1B). Initially, participants are shown the gesture they are playing against (the cue, e.g., the rock). At Stage 1, participants have to choose from two pairs of gestures: The winning pair contains the cue together with a winning gesture (e.g., the rock and the paper). The losing pair contains the cue together with a losing gesture (e.g., the rock and the scissor). At Stage 2, participants have to choose a gesture from the previously chosen Stage-1 pair. The final outcome depends on whether the chosen gesture wins or loses against the cue gesture, or whether cue and chosen gesture are the same (draw).

The structure of this two-stage task implies that the final outcome depends on the decisions at both stages (Figure 1A): If the winning pair was chosen at Stage 1, the final outcome can be either a win or a draw. But if the losing pair was chosen at Stage 1, the final outcome can be a draw or a loss. To maximize the outcome, the *correct* choices would be the winning pair at Stage 1 and the winning gesture (if choosing from a winning pair) or the gesture resulting in a draw (if choosing from a losing pair) at Stage 2. As we were interested in deviations from optimal outcomes; however, we enforced *errors* by inducing speed pressure. An error occurred if the losing pair was chosen at Stage 1, or if the losing gesture (if choosing from a losing pair) or the gesture resulting in a draw (if choosing from a winning pair) was chosen at Stage 2. In either of these cases, we expected to observe markers of error monitoring such as the Ne/ERN and Pe after these incorrect choices. However, crucial for the present purpose is the scenario in which a Stage-1 error is followed by a correct response at Stage 2. In this case, the losing pair is chosen at Stage 1, which implies that the best possible final outcome at Stage 2 is a draw. This creates a situation in which a correct Stage-2 response leads to an unfavorable outcome. The question is whether the Stage-2 response in this situation is associated with a Pe. Such a finding would demonstrate that the Pe reflects a higher-order error signal that indicates unfavorable outcomes even if the response was correct.

2 | Method

2.1 | Participants

A priori power analysis revealed that a sample size of 24 participants was required to allow the crucial within-subjects t-test testing for the presence of a Ne/ERN or Pe to detect medium effect sizes of 0.52 with a statistical power of 0.8 (G*Power; Faul et al. 2007, 2009). After initial data from 24 participants were collected, one participant had to be excluded because no trials remained for analyzing Stage-2 errors after EEG preprocessing. Another participant had to be excluded because less than

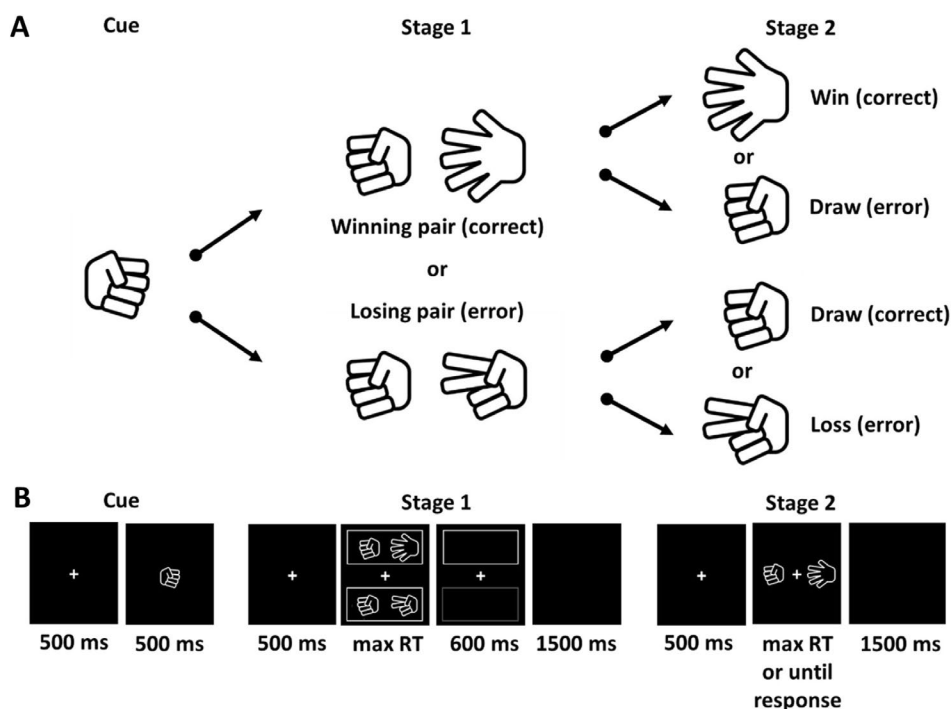


FIGURE 1 | Overview of an exemplary trial. (A) All possible decisions within the trial. (B) Time course of the trial. First, a cue is shown indicating the gesture of the opponent (here the rock). On Stage 1, participants have to choose between the winning pair (containing the cue gesture and the winning gesture, here the rock and the paper) and the losing pair (containing the cue gesture and the losing gesture, here the rock and the scissor). On Stage 2, participants have to choose from the two gestures of the pair chosen in Stage 1. Participants can make a correct response or an error at Stage 2 leading to an overall win, a loss, or a draw. max RT = maximum response time of the respective stage.

6 trials remained for analyzing higher-order errors (Pontifex et al. 2010). These two were replaced with new participants. This resulted in a sample of 24 usable participants (4 male, 0 diverse; age $M = 23.3$ years, $SD = 4.17$; all right-handed). All participants were recruited at the Catholic University of Eichstätt-Ingolstadt and received either course credit or payment (10.50 € per hour) in addition to a performance-dependent monetary incentive of up to 4.80 €. All participants gave informed consent, and the study was approved by the ethics committee of the Catholic University of Eichstätt-Ingolstadt.

2.2 | Stimuli

Stimuli were three hand gestures representing the options of the game “rock, paper, scissors,” which were either presented as pairs of options with the fingers on the left (Stage 1 in Figure 1) or as cues with the fingers on the right (Cue in Figure 1). On the screen, the rock gesture had a height of 1.30° and a width of 1.22° , the paper gesture had a height of 1.72° and a width of 1.60° , and the scissor gesture had a height of 1.30° and a width of 1.60° . All gestures were drawn with white lines on a black background. Stimuli were presented on an Acer (XB253Q) flatscreen with a frame rate of 144 Hz.

2.3 | Task and Procedure

Each trial consisted of two stages and each stage required a decision by the participants (Figure 1). The trial started with a central fixation cross displayed for 500 ms. This was followed by a gesture

representing the cue for 500 ms. After another fixation cross for 500 ms, the first-stage stimulus appeared, which consisted of two pairs of gestures, each within a white frame above and below the fixation cross. One of the pairs was the same gesture as the cue together with the gesture that would win against the cue (the winning pair). The other pair was the same gesture as the cue together with the gesture that would lose against the cue (the losing pair). The position of gestures (left, right) as well as pairs (up, down) was random and balanced within each block. Participants had to make their Stage-1 decision by pressing the arrow-up or arrow-down key with the index finger of their dominant hand. After the response, only the two frames were shown for 600 ms (whereas the gestures disappeared) and the frame not chosen was darkened. Then, a black screen was presented for 1500 ms. Stage 2 started with a fixation cross for 500 ms followed by the pair of gestures chosen in Stage 1 presented left and right of a fixation cross. The position of the two gestures was randomized. Participants made their Stage-2 decision by pressing the arrow-left or arrow-right key with the same finger as for the Stage-1 decision. After the response, a black screen was depicted for 1500 ms until the next trial started. Participants were instructed to respond as quickly and as accurately as possible and to relocate their index finger in the middle of the arrow keys after each response.

To ensure an adequate error rate, adaptive response windows were employed at each stage to induce time pressure. If a maximum response time (RT) has passed without a response, a feedback screen (“Zu spät!”, engl. too late) was presented immediately for 1000 ms (and no frames were depicted at Stage 1). If participants pressed a key prior to the presentation of a stimulus at any stage, feedback was provided for a duration of 1000 ms

("Zu früh gedrückt!", engl. pressed too early) and the next trial started. The maximum RT was adjusted separately for each stage and depended on the error rate of the preceding 20 trials at this stage. The targeted error rate ranged between 15% and 20% for Stage 1 and between 5% to 10% for Stage 2. If the error rate fell below 15% at Stage 1 or 5% at Stage 2, the maximum RT for the respective stage was decreased by 30 ms. If the error rate exceeded 20% at Stage 1 or 10% at Stage 2, the maximum RT for the respective stage was increased by 30 ms. If a response was too late, then the maximum RT was additionally increased by 60 ms. The starting value for each stage was 1500 ms. The maximum RT could not exceed 1500 ms. Additionally, to prevent too many too late responses, the lower limit for the maximum RT was the 90th percentile of the participant's RT of all preceding trials. This adaptive response window led to a maximum RT of 1053 ms (SE = 56 ms) for Stage 1 and 950 ms (SE = 56 ms) for Stage 2.

To incentivize correct and quick responding, a monetary incentive system was implemented. In each trial, participants received points based on their final outcome, and the score of accumulated points was exchanged for money at the end of the experiment (10 points = 1 ct.). Ten points were added for a win, and five points were subtracted for a loss. No points were added or subtracted for a draw. A failure to respond within the maximum RT led to an additional loss of 30 points. The current score was displayed for 4000 ms after every 12th trial, followed by a black screen for 500 ms before the next trial started.

The experiment comprised 10 test blocks with 48 trials each, resulting in a total of 480 trials. Oral as well as written feedback emphasizing the need to respond either faster or more accurately was administered after each block. Feedback was triggered if the error rate exceeded predefined thresholds set at 15%–20% for Stage 1 and 5%–10% for Stage 2. Prior to these test blocks, participants conducted four practice blocks to familiarize participants with the task and the monetary incentive system. The first practice block comprised 12 trials in which only Stage 2 without a response window was practiced. The second practice block comprised 24 trials with both stages but again without response windows. The third and fourth practice blocks comprised 48 trials, in which the response windows and the incentives were introduced. After the Stage-2 response, participants received immediate feedback on the awarded points in the current trials ("10", "0", "-5", or "-30"), which was presented 1000 ms after the response for 1000 ms, followed by a black screen for 500 ms. After these practice blocks, participants were informed that, in the test blocks, there was no trial-wise feedback about the outcome but that a score was presented every 12 trials. This score was reset to zero at the beginning of the test blocks.

2.4 | EEG Data Acquisition

64 Ag-AgCl electrodes were used in a 10-10 system to acquire EEG-Data with a BioSemi Active-Two system (BioSemi, Amsterdam, The Netherlands; channels Fp1, AF7, AF3, F1, F3, F5, F7, FT7, FC5, FC3, FC1, C1, C3, C5, T7, TP7, CP5, CP3, CP1, P1, P3, P5, P7, P9, PO7, PO3, O1, Iz, Oz, POz, Pz, CPz, Fpz, Fp2,

AF8, AF4, AFz, Fz, F2, F4, F6, F8, FT8, FC6, FC4, FC2, FCz, Cz, C2, C4, C6, T8, TP8, CP6, CP4, CP2, P2, P4, P6, P8, P10, PO8, PO4, O2 and the left and right mastoid). For reference and ground electrodes, the Common Mode Sense (CMS) and Driven Right Leg (DRL) electrodes were used, which were placed centrally and close to POz. Additionally, the vertical and horizontal electrooculogram (EOG) was recorded with BioSemi FLAT Active electrodes (Ag-AgCl) above and below the right eye and on the outer canthi of both eyes. Both EEG and EOG were recorded at a sampling rate of 1024 Hz, and no online filtering was applied. All electrodes were off-line re-referenced to averaged mastoids.

2.5 | Data Analysis

2.5.1 | Behavioral Data Analysis

Each response was assigned to the condition Stage (Stage 1, Stage 2) and classified as a correct response or error. At Stage 1, the response was correct if the winning pair was chosen and an error if the losing pair was chosen. At Stage 2, the response was correct if the winning gesture (if choosing from a winning pair) or the gesture that results in a draw (if choosing from a losing pair) was chosen. An error occurred at Stage 2 if the losing gesture (if choosing from a losing pair) or the gesture resulting in a draw (if choosing from a winning pair) was chosen. A correct Stage-2 response that was preceded by a Stage-1 error was defined as a higher-order error. All trials were excluded whose RTs deviated more than three standard deviations from the RT mean of each condition and participant. Furthermore, trials featuring responses either too early ($M = 0.31\%$, $SE = 0.10\%$) or too late ($M = 4.97\%$, $SE = 0.45\%$) were excluded from subsequent analysis. Prior to statistical testing of error rates, arcsine transformations were applied (Winer et al. 1991). We first compared error rates of the two stages using a paired t-test. Then, RTs were subjected to a two-way repeated measures ANOVA with the variables Stage (Stage 1 vs. Stage 2) and Correctness (correct vs. error). To address post-error adjustments, the error rate at Stage 2 and the RTs of correct Stage-2 responses were compared between Stage-2 responses following a correct Stage-1 response and those following a Stage-1 error using paired t-tests.

2.5.2 | EEG Preprocessing

All analyses were conducted using custom scripts in MATLAB v9.10 (The Mathworks, Natick, MA) and EEGLAB v2019.1 (Delorme and Makeig 2004). Continuous EEG data was downsampled to 512 Hz, low-pass filtered (40 Hz) and then high-pass filtered (0.1 Hz) with a least-squares FIR filter (-6 dB half-amplitude cutoff). Epochs ranging from -500 ms before to 1000 ms after each response (Stage 1 and Stage 2) were extracted and baseline-corrected using a baseline interval from -150 to -50 ms before the response. Bad electrodes were identified using EEGLAB's channel rejection routine (pop_rejchan.m) with a joint probability and a kurtosis criterion (both with a threshold of 5) and were interpolated using spherical spline interpolation. On average, 4.39% of electrodes for Stage 1 ($SE = 0.62\%$) and 3.79% for Stage 2 ($SE = 0.56\%$) were

interpolated. To correct eye blinks and muscular artifacts, an infomax-based Independent Component Analysis (Bell and Sejnowski 1995) was conducted on a copy of the dataset to which an additional 1 Hz high-pass filter was applied. Independent components (ICs) reflecting ocular and motor artifacts were then identified using the ICLabel toolbox (Pion-Tonachini et al. 2019) and removed from the original dataset. On average, 10.0 ICs were removed per participant for Stage 1 ($SE = 0.86$) and 11.92 for Stage 2 ($SE = 1.03$). Of these, a mean of 4.46 ICs in Stage 1 ($SE = 0.73$) and 4.33 in Stage 2 ($SE = 0.70$) represented motor noise which might reflect that participants had to move their fingers to the respective keys. Retaining these ICs would have led to a more than two times higher number of excluded trials in the next preprocessing step. In this next step, epochs whose channel activity exceeded $\pm 100 \mu V$ or whose joint probability deviated more than 5 standard deviations from the mean within the time range of interest (-150 to 400 ms) were removed. On average, 9.90% of trials in Stage 1 ($SE = 1.54\%$) and 10.55% of trials for Stage 2 ($SE = 2.21\%$) were removed in this way. Finally, epochs were averaged separately for Stage-1 responses and Stage-2 responses and separately for the different conditions as described below. No trials with errors in both stages were analyzed because these trials were too infrequent ($M = 6.75$, $SE = 0.96$). The number of trials in the remaining conditions ranged between 13.9 ($SE = 1.99$) for Stage-2 errors following correct Stage-1 responses and 332.6 ($SE = 9.96$) trials for correct Stage-2 responses following correct Stage-1 responses. Note that, in the former condition, four participants had less than 6 trials, which has been recommended as a minimum to analyze the Ne/ERN (Pontifex et al. 2010). Therefore, after the above-described replacement of two participants, we additionally conducted all analyses described below without participants with less than 6 trials. These analyses replicated all effects from the full sample if not stated otherwise. Because the recommendation of Pontifex et al. (2010) is valid for the Ne/ERN while our component of interest was the Pe, we additionally applied the method of Luck et al. (2021) to check whether small trial numbers in some participants influenced our results. That is, we calculated standardized measurement errors (SME) for each component (Ne/ERN, Pe) and condition for the time windows used in the respective analyses. For Stage-2 errors, SMEs would have been improved when one participant was excluded for the Ne/ERN and two participants were excluded for the Pe. When conducting our analyses (including the MVPA) without these participants, almost all results remained the same (with one exception noted in the results section). We therefore retained these participants in the reported analyses.

2.5.3 | ERP Data Analysis

The Pe and Ne/ERN were defined as the difference in mean amplitudes between errors and correct trials (for a discussion of this definition, see Discussion section). As participants had to move their finger to the respective response keys and ERP data were locked to the keypress rather than to the movement onset, we had to deal with a temporal offset in response-locked ERP components. This was relevant mainly for the Ne/ERN but less so for the more long-going Pe. We therefore did not predefine a time window for the Ne/ERN but used a cluster-based permutation

test within a broader time period at electrode FCz. In contrast, the Pe was quantified within the frequently used time window of 200 ms to 400 ms (Endrass et al. 2007; Overbeek et al. 2005; Ullsperger, Fischer, et al. 2014) at electrode POz (Steinhauser and Yeung 2012; Steinhauser and Steinhauser 2021). Please note that the Pe is often quantified at more anterior electrodes such as Pz or CPz. To make sure that choosing POz does not inappropriately bias our data, we applied all Pe analysis in the present study also for data at electrodes Pz and CPz but obtained the same results for all analysis.

First, we investigated whether a typical Ne/ERN and the Pe were present for Stage-1 errors and Stage-2 errors. Note that because trials with errors in both stages were excluded, Stage-1 and Stage-2 errors refer to trials with a correct response in the other stage. For the Ne/ERN, we used a cluster-based permutation test to compare correct and incorrect responses at each stage. Using the Mass Univariate ERP Toolbox (Groppe et al. 2011), we identified clusters and calculated 100,000 permutations within the time range of -50 to 200 ms at FCz. A cluster inclusion threshold as well as an output threshold of $p = 0.05$ was used. For the Pe, we performed within-subject t-tests comparing errors with correct trials at each stage. We then analyzed both components in higher-order errors by comparing correct Stage-2 responses following Stage-1 errors with those following correct Stage-1 responses using the same methods as described above. We finally compared all three error types (Stage-1 error, Stage-2 errors, higher-order errors) by subjecting Ne/ERN and Pe amplitudes to a one-way repeated measures ANOVA with error type as the variable. Pairwise differences were tested using paired t-tests. Ne/ERN amplitudes were quantified in the time windows identified by the permutation tests in the previous step. As no significant cluster was found for higher-order errors, the time window for Stage-2 errors was used.

2.5.4 | Multivariate Pattern Analysis (MVPA)

An MVPA was used to investigate whether the error-related brain activity found for higher-order errors shows a similar spatial distribution as that for Stage-2 errors. To this end, we used a rationale that has been applied in previous studies (Boldt and Yeung 2015; Di Gregorio et al. 2018; Steinhauser and Yeung 2010, 2012). First, a logistic regression classifier was trained to discriminate between correct Stage-2 responses following a Stage-1 error (corresponding to a high-order error) and correct Stage-2 responses following correct Stage-1 responses using the linear integration method (Parra et al. 2002, 2005). Classifier input was the mean single-trial amplitude from all 64 electrodes in a specific time window. To trace the time course of classifier performance, classifiers were trained for consecutive but overlapping time windows in the interval between -100 and 500 ms with a width of 50 ms and a step size of 25 ms. The training set comprised all correct Stage-2 responses following a Stage-1 error (the higher-order error trials) as well as half of the correct Stage-2 responses following correct Stage-1 responses (the all-correct trials) as the other half was used for a later analysis step. For this, we divided the correct Stage-1 responses into trials with odd and even trial numbers and used trials with odd numbers for training. To balance the number of higher-order errors and all-correct trials, the set of higher-order error trials

(which were less frequent in all participants) was duplicated until the difference in trial numbers between the two conditions was smaller than the number of higher-order trials. Then, higher-order trials were randomly drawn and added to the training set until trial numbers in both conditions were balanced. A leave-one-out cross-validation procedure was used to evaluate the classifiers' performance. Each trial was classified after training with a training set from which this trial (and its duplicates) had been removed. The removed trials were replaced using a similar method as described in the previous step.

Classifier performance was measured by an Az score with a value of 0.5 indicating classification on a chance level and a value of 1 indicating perfect classification. A permutation test was used to test whether Az scores differed from chance level. For each participant, we randomized the assignment of trials to the two categories (higher-order error vs. all correct) and calculated an Az score. Repeating this procedure 1000 times resulted in a distribution of Az scores, and the 5%-percentile from this distribution was used as a critical Az score. Inference testing for each classifier was done by comparing the mean Az score (averaged across participants) to the respective mean critical Az score. To visualize the spatial distribution of the discriminating component, we provide a topography of coupling coefficients representing the activity at each electrode site that correlates with the discriminating component (Parra et al. 2002, 2005).

As the time range in which our classifiers showed significant classification performance strongly overlaps with the time period of the Pe, we averaged across all significant classifiers and concluded that the resulting classifier can be viewed as a decoder of the Pe-like component elicited by higher-order errors (which was also confirmed by the topography of the discriminating component). In a second step, we then tested whether this classifier can also distinguish between Stage-2 errors and correct Stage-2 responses in the Pe time range. This would provide evidence that similar ERP components (here: a Pe) are involved in both types of errors. To this end, we applied the classifier to Stage-2 errors (following correct Stage-1 responses) and those all-correct trials that were not chosen for the training set (trials with even numbers). Prediction values were calculated for each trial and time point within the epoch. These prediction values can be interpreted as reflecting how strongly the spatial distribution of brain activity at this time point resembles the Pe-like activity that is decoded by the classifier. Prediction values were averaged separately for Stage-2 errors and correct Stage-2 trials and compared on a group level using a paired t-test.

3 | Results

3.1 | Behavioral Data

Error rates and RTs for Stage 1 and Stage 2 are depicted in Table 1. Participants made more errors in Stage 1 than in Stage 2, $t(23) = 43.06$, $p < 0.001$, $d = 0.65$, which is in accordance with the adaptive RT criterion aiming for more errors within Stage 1. The analysis of RTs revealed a significant main effect for Stage, $F(1,23) = 32.26$, $p < 0.001$, $\eta_p^2 = 0.58$, but no significant effect of Correctness, $F(1,23) = 0.22$, $p = 0.641$, $\eta_p^2 < 0.01$, and no significant interaction, $F(1,23) = 2.60$, $p = 0.121$, $\eta_p^2 = 0.10$. These results indicate faster RTs for Stage 2 ($M = 446$ ms, $SE = 12$ ms) than for Stage 1 ($M = 502$ ms, $SE = 14$ ms), again reflecting differences in the adaptive RT criterions. We next investigated whether Stage-2 performance was affected by the correctness of the Stage-1 response (Table 1). Participants made significantly more errors in Stage 2 after an erroneous Stage-1 response than after a correct Stage-1 response, $t(23) = 34.34$, $p < 0.001$, $d = 0.60$. However, Stage-2 RT was not affected by the correctness of the Stage-1 response, $t(23) = 1.37$, $p = 0.183$, $d = 0.57$, indicating no post-error slowing.

3.2 | ERP Data

Response-locked ERPs used to analyze the Pe and Ne/ERN amplitudes for Stage 1 and Stage 2 are depicted in Figure 2. For Stage 1, we found the typical Pe pattern of a positive posterior deflection in errors ($M = 1.55 \mu V$, $SE = 1.33 \mu V$) relative to correct trials ($M = -4.73 \mu V$, $SE = 1.35 \mu V$) in the time range of 200–400 ms at electrode POz, $t(23) = 5.52$, $p < 0.001$, $d = 0.94$ (right panel in Figure 2A). For the Ne/ERN, the cluster-based permutation revealed significantly more negative pronounced errors ($M = 1.53 \mu V$, $SE = 0.65 \mu V$) compared to correct trials ($M = 4.52 \mu V$, $SE = 0.54 \mu V$) in a time window of –47 to 78 ms ($p < 0.001$) at electrode FCz (left panel in Figure 2A). Unexpectedly, the spatial distribution of this Ne/ERN was not restricted to frontocentral electrodes but spanned across both anterior and posterior electrode sites.

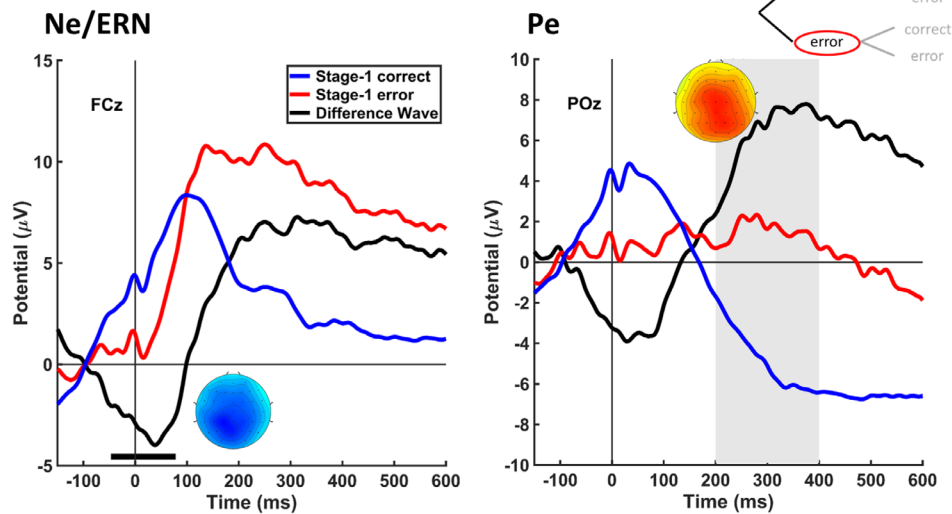
For Stage 2, we also found an increased positivity for errors ($M = 0.28 \mu V$, $SE = 1.60 \mu V$) compared to correct trials ($M = -9.98 \mu V$, $SE = 1.24 \mu V$) at electrode POz indicating a Pe, $t(23) = 8.40$, $p < 0.001$, $d = 1.44$ (right panel in Figure 2B). For the Ne/ERN, a negative cluster in the time window of –37 to 39 ms ($p = 0.030$) was identified at electrode FCz which indicated a more negative activity for errors ($M = -1.53 \mu V$, $SE = 0.90 \mu V$) than for correct trials ($M = 1.83 \mu V$, $SE = 0.70 \mu V$). In contrast to

TABLE 1 | Error rates and response times for Stage 1 and Stage 2.

	Error rates (%)	Correct RT (ms)	Error RT (ms)
Stage 1	15.08 (1.15)	506 (6)	500 (10)
Stage 2	5.60 (1.5)	446 (7)	461 (10)
Stage 2 post correct	4.45 (1.40)	446 (7)	447 (6)
Stage 2 post error	14.05 (1.40)	470 (14)	450 (10)

Note: Within-subjects standard errors (Morey 2008) are provided in parentheses. Abbreviation: RT = response time.

A - Stage 1



B - Stage 2

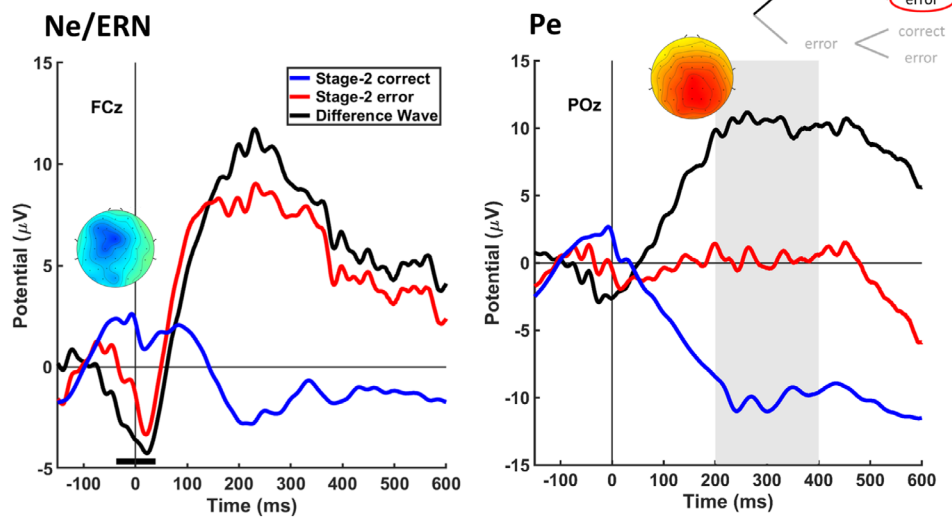


FIGURE 2 | Error-related brain activity for Stage-1 errors and Stage-2 errors. Panels show Stage-1 locked (A) and Stage-2 locked (B) ERPs and difference waves (errors minus corrects) at electrode FCz for analyzing the Ne/ERN and at POz for analyzing the Pe. Black horizontal lines indicate the cluster representing the Ne/ERN as revealed by the permutation test. Gray areas indicate the time interval for statistical testing of the Pe. Scalp topographies indicate the spatial distribution of the difference wave in the time interval representing the respective component.

Stage-1 errors, the Ne/ERN for Stage-2 errors had a clear fronto-central distribution.

After demonstrating that errors at each stage elicit the typical error-related brain activity, we investigated whether also a correct Stage-2 response leads to a Pe if this correct response was associated with an unfavorable outcome due to a Stage-1 error. To this end, we compared Stage-2 correct responses with a preceding Stage-1 error with those with a preceding correct Stage-1 response (see Figure 3). As hypothesized, we found such a higher-order Pe as reflected by a more positive deflection in correct Stage-2 responses following a Stage-1 error ($M = -6.20 \mu V$, $SE = 1.61 \mu V$) than in those following a correct Stage-1 response ($M = -9.98 \mu V$, $SE = 1.24 \mu V$) at electrode POz, $t(23) = 4.85$, $p < 0.001$, $d = 0.53$ (right panel in Figure 3). In contrast, the cluster-based permutation test at electrode FCz did not

find a negative cluster representing an Ne/ERN (left panel in Figure 3).

We then compared the three conditions (Stage-1 errors, Stage-2 errors, correct Stage-2 responses following Stage-1 errors) by considering the difference waves representing the respective Pes and Ne/ERNs (Figure 4). A one-way ANOVA for the Pe revealed that the three conditions differed significantly, $F(2, 46) = 16.22$, $p < 0.001$, $\eta_p^2 = 0.41$. The Pe for Stage-2 errors was significantly larger ($M = 10.25 \mu V$, $SE = 1.22 \mu V$) than the Pe for Stage-1 errors ($M = 6.28 \mu V$, $SE = 1.14 \mu V$), $t(23) = 4.00$, $p < 0.001$, $d = 1.06$. Moreover, the higher-order Pe (correct Stage-2 responses following Stage-1 errors) was significantly smaller ($M = 3.77 \mu V$, $SE = 0.77 \mu V$) compared to both, the Pe for Stage-1 errors, $t(23) = -2.11$, $p = 0.046$, $d = 0.52$, and the Pe for Stage-2 errors, $t(23) = -5.20$, $p < 0.001$, $d = 1.27$. Please note that the former comparison between

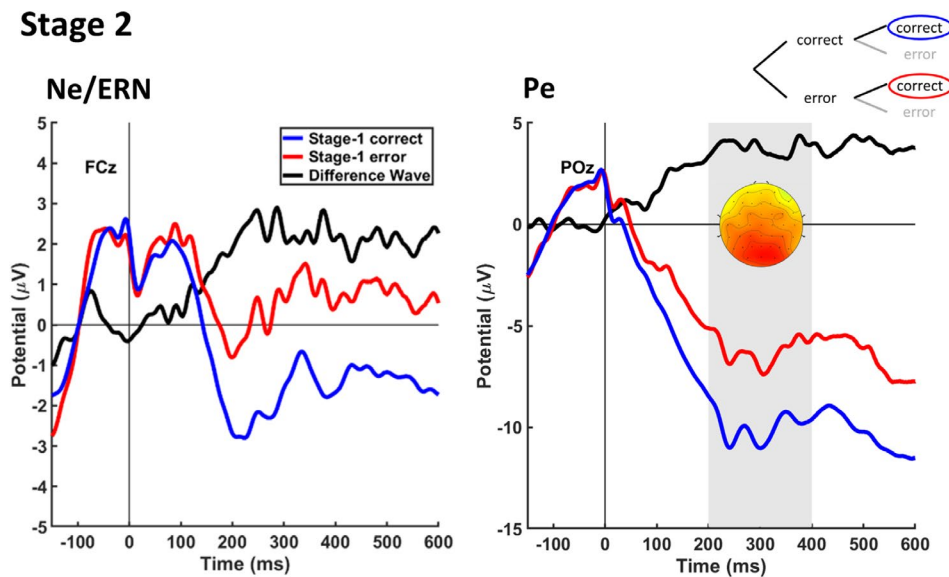


FIGURE 3 | Error-related brain activity reflecting higher-order errors at Stage 2. Panels show Stage-2-locked ERPs and difference waves (error minus corrects) at electrode FCz for analyzing the Ne/ERN and at POz for analyzing the Pe. The gray area indicates the time interval for statistical testing of the Pe. The scalp topography indicates the spatial distribution of the difference wave in the time interval of the Pe.

Difference Waves

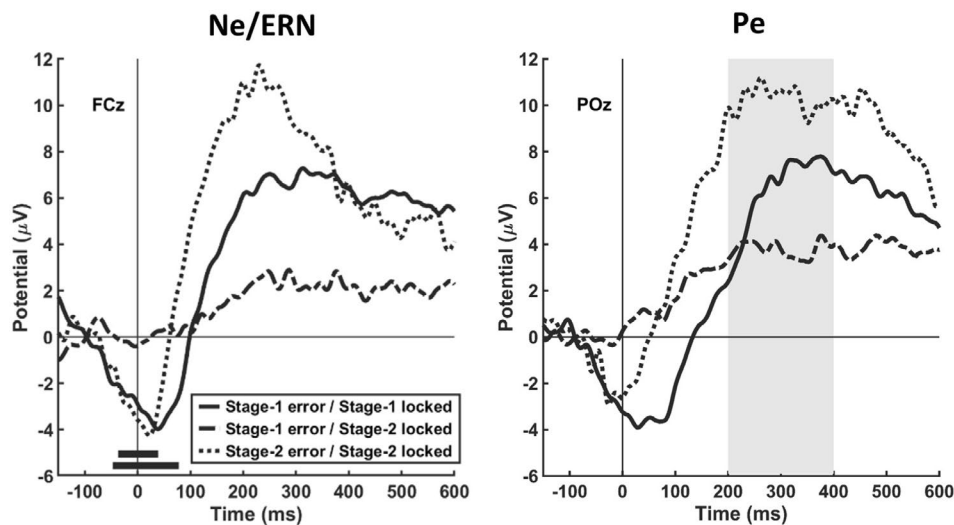


FIGURE 4 | Difference waves (errors minus corrects) representing the Ne/ERN at electrode FCz and the Pe at electrode POz for Stage-1 errors in Stage-1-locked data as well as Stage-2 errors and higher-order errors in Stage-2-locked data. The black horizontal lines indicate the time windows found by the permutation tests for the Ne/ERN. The gray area indicates the time interval for statistical testing for the Pe.

higher-order Pe and Pe for Stage-1 errors did not reach significance, $t(21) = -1.62$, $p = 0.120$, $d = 0.35$, when the two participants with low SMEs were removed from the sample (see Section 2.5.2). To compare the Ne/ERN for the three conditions, we used the time windows associated with the respective clusters for Stage-1 and Stage-2 errors. For correct Stage-2 responses following Stage-1 errors, we used the time window associated with the cluster for Stage-2 errors. The one-way ANOVA again revealed that the three conditions differed significantly, $F(2, 46) = 5.68$, $p = 0.013$, $\eta_p^2 = 0.12$. We found no significant difference between Ne/ERNs for Stage-1 errors ($M = -2.99 \mu V$, $SE = 0.58 \mu V$) and Stage-2 errors ($M = -3.36 \mu V$, $SE = 1.09 \mu V$), $t(23) = 0.38$, $p = 0.710$, $d = 0.16$ (Figure 4). However, the Ne/ERN for correct Stage-2 responses

following Stage-1 errors ($M = -0.18 \mu V$, $SE = 0.68 \mu V$) was significantly smaller than the Ne/ERN for Stage-2 errors, $t(23) = 2.48$, $p = 0.021$, $d = 0.51$, as well as the Ne/ERN for Stage-1 errors, $t(23) = 3.66$, $p < 0.001$, $d = 1.53$ (Figure 4).

3.3 | Multivariate Pattern Analysis

An MVPA was used to demonstrate that the higher-order Pe involves activity with a similar spatial distribution as the Pe on error trials, which would provide evidence that the positivity following higher-order errors is indeed a Pe. We trained a classifier to discriminate between correct Stage-2 trials that follow Stage-1 errors

and those that follow a correct Stage-1 response. Az scores representing classifier performance for consecutive time windows are depicted in Figure 5A. Permutation tests revealed that classifier performance is significantly above chance for time windows between 100 and 400 ms, which overlaps with the time window of the Pe. Also, the spatial distribution of the discriminating component in this time window reveals a pattern similar to a Pe.

In a second step, we investigated whether this classifier can decode Stage-2 errors. To this end, we used the classifier to distinguish between Stage-2 errors and correct Stage-2 responses (both following a correct Stage 1). This resulted in prediction values for each time window and each trial type, which are depicted in Figure 5B. We expected higher prediction values for Stage-2 errors than for correct Stage-1 responses if the Pe on Stage-2 errors and the higher-order Pe share similar spatial features. In line with this expectation, we found a significant difference in prediction values in the time window of the Pe (200–400 ms), $t(23) = 3.03$, $p = 0.006$, $d = 0.618$. We further tested whether this effect is specific for the Pe time window or whether a similar difference can be found for the Ne/ERN time window. In the Ne/ERN time window previously obtained for Stage-2 errors (–25 to 25 ms), no significant difference between prediction values was observed, $t(23) = 0.70$, $p = 0.494$, $d = 0.143$.

Please note that the prediction values are the output of a logistic regression model. Accordingly, values above 0.5 indicate that a trial is more likely to be an error (vs. correct) while values below 0.5 indicate that a trial is more likely to be correct (vs. an error). The results in Figure 5B thus indicate that Stage-2 errors are not identified as errors but rather as less correct trials. This could reflect that the classifier is somewhat biased towards correct trials because trials of this category were used for training (although we did not use the same correct trials for training and prediction). However, this does not change our interpretation. A classifier that is trained to decode the higher-order Pe predicts

differences between correct and Stage-2 errors. And therefore, activity with a similar spatial distribution underlies the difference between correct and higher-order errors (for which the classifier was trained) and the difference between correct and Stage-2 errors (to which the classifier was applied).

4 | Discussion

The present study investigated the idea that error monitoring evaluates the overall outcome of a multistage task beyond the outcome of individual stages, and that this overall outcome is reflected by the Pe. This assumption was tested in a two-stage task in which the overall outcome depended on the performance at each individual stage. Crucially, if an incorrect response was provided at Stage 1, even a correct response at Stage 2 led to an unfavorable outcome because a draw is less favorable than a win as it does not result in monetary gain. Our results show that this case led to a posterior positivity after the correct Stage-2 response, which we interpreted as a higher-order Pe. This interpretation received support from an MVPA demonstrating that a pattern classifier trained on this higher-order Pe could decode the Pe observed after incorrect Stage-2 responses. In conclusion, the evidence presented here supports the idea of a higher-order error monitoring system which evaluates the overall outcome in a multistage task, and which is reflected by a Pe.

Our results suggest that the Pe reflects more than the accumulated evidence for an error in the immediately preceding response. The evidence accumulation account (Steinhauser and Yeung 2010; Ullsperger et al. 2010) states that the Pe represents the accumulated evidence that an error has been committed. While this can explain the Pe following incorrect Stage-1 and Stage-2 responses as well as corresponding results from single-stage tasks (Murphy et al. 2015; Navarro-Cebrian et al. 2016; Steinhauser and Yeung 2010, 2012; Ullsperger et al. 2010), it is insufficient to account for the higher-order Pe following correct

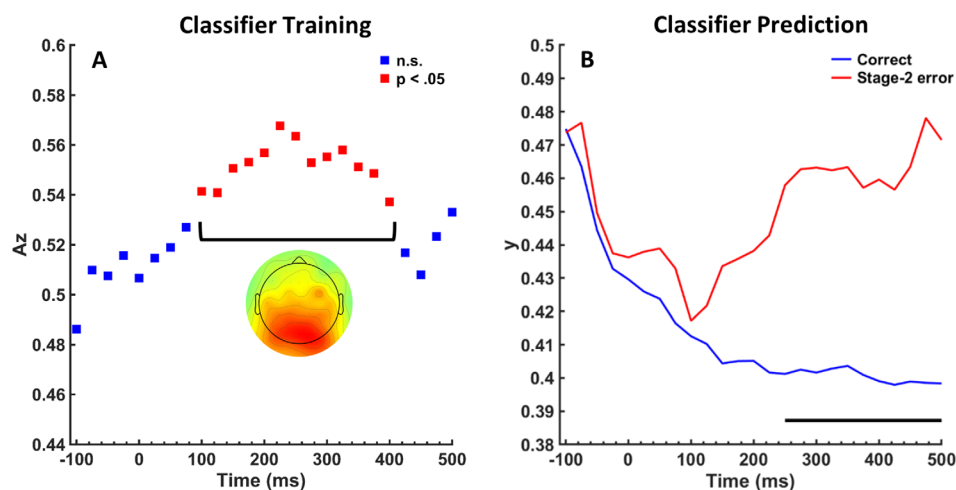


FIGURE 5 | Multivariate pattern analysis. (A) Discrimination performance (Az) after training of a multivariate pattern classifier that distinguishes between correct Stage-2 responses following Stage-1 errors (higher-order errors) and Stage-2 responses following correct Stage-1 responses. Red squares indicate time windows with a significant discrimination performance as revealed by a permutation test. The topography shows the discriminating component averaged across the significant time windows (200–500 ms). (B) Prediction values (y) after applying the classifier to Stage-2-errors and correct Stage-2-responses (both following correct Stage-1 responses). The black bar shows the time window in which prediction values of both conditions differ significantly ($p < 0.05$).

Stage-2 responses. Our results are also difficult to explain by the idea that the Pe reflects the confidence of having made an error (Boldt and Yeung 2015). Even though no confidence judgments were assessed in the present study, there is no reason to assume that participants' confidence at Stage 2 is impaired because of an incorrect Stage-1 response, as there was enough time between stages to complete post-decisional processing of Stage 1. We thus interpret the higher-order Pe as reflecting the outcome of the whole multistage task. More specifically, the Pe for correct Stage-2 responses indicates that the overall outcome of the trial is unfavorable (as it corresponds to a draw), even though the Stage-2 response is correct (as it corresponds to the better of two options). We call this Pe a higher-order Pe because it represents an evaluative signal that takes the whole multistage task into account rather than a single response.

The higher-order Pe is smaller than the Pe for Stage-1 errors and Stage-2 errors. Relative to the pre-response baseline, the amplitude on higher-order errors in the Pe time window is even negative, and a positivity is obtained only because we measured the Pe as the difference between correct and higher-order error trials. Given that only some studies in the literature use this difference approach (e.g., Boldt and Yeung 2015; Di Gregorio et al. 2016; Falkenstein et al. 1990, 1991) whereas others use the absolute amplitude for error trials (e.g., Endrass et al. 2010; Hewig et al. 2011), the question emerges as to whether the higher-order Pe defined as a difference meets the characteristics of a true Pe. We believe that defining the Pe as a difference between correct and error trials leads to a better estimate of late error-related brain activity because the time course of response-locked activity on correct and error trials is strongly affected by stimulus-locked components such as the P300. The immediate decline of posterior activity after the (correct and incorrect) response at each stage in Figures 2 and 3 suggests that the response coincides with the peak of the P300 in our data and that the decline reflects the decrease of the P300 after this peak. As a consequence, a valid quantification of the Pe is possible only if this P300 time course is subtracted by using the difference between correct and error trials. In the [Supporting Information](#), we provide evidence for this view by analyzing epochs comprising both stages with a pre-stimulus baseline, in which the influence of the stimulus-locked P300 can be clearly identified. In these analyses, we also show that the higher-order Pe as well as all other reported forms of error-related brain activity are replicated with a pre-stimulus baseline. This additionally indicates that our results cannot be accounted for by effects during the pre-response period.

A further question is whether the higher-order Pe and the Pe following an incorrect response reflect the same or different mechanisms. From the latter perspective, one might argue that the two types of Pe reflect two levels of error monitoring, the level of individual responses and the level of the multistage task. While the former takes place after each response, the latter additionally occurs at the end of a trial. The higher-order monitoring process could integrate individual error signals from each stage with task rules specifying the overall outcome (e.g., "error at Stage 1 and correct response at Stage 2 results in a draw") to provide evidence for an unfavorable outcome reflected by the higher-order Pe. Such an inference-based metacognitive process could fulfill functions such as identifying errors that are not characterized by a single incorrect response, such as order errors

in dual tasking (e.g., Steinhauser et al. 2021). It could also underlie the ability to report the source of an unfavorable outcome (Di Gregorio et al. 2016). Solving the so-called credit assignment problem (Minsky 1961), that is, the problem of assigning an outcome to its underlying source, is crucial for learning from errors and is particularly relevant in the context of multistage tasks (e.g., Fu and Anderson 2008; Stolyarova 2018; Walsh and Anderson 2014; Wurm et al. 2024).

Alternatively, the higher-order Pe and the Pe following incorrect responses could reflect the same mechanism. It is conceivable that both represent a neural signal indicating that the current response led to an unfavorable outcome, irrespective of whether this is due to an error in a single response or a specific response pattern in a multistage task. The higher-order Pe in our paradigm could reflect the influence of context on the evidence accumulation process. A correct Stage-1 response leads to a win context, which implies that only an incorrect Stage-2 response entails an unfavorable outcome. However, an error at Stage 1 leads to a loss context, and hence, even a correct Stage-2 response is associated with an unfavorable outcome. In this case, the context itself delivers the evidence for an unfavorable outcome, thus inducing a Pe. A necessary implication of such a single-mechanism explanation is that any Pe reflects a higher-order monitoring system because it takes contextual information into account and thus operates on a level beyond that of individual responses.

This single-mechanism explanation of the Pe receives support from our decoder analysis which suggests that the higher-order Pe and Pe after incorrect responses share similar spatial features. However, another observation might speak against this view. Trials with a "Stage-1 correct/Stage-2 error" pattern and trials with a "Stage-1 error/Stage-2 correct" pattern both lead to the same outcome (i.e., a draw). However, the Pe amplitude in the former case (i.e., after Stage-2 errors) is substantially larger than the Pe amplitude in the latter case (i.e., after correct Stage-2 responses). This appears to be more compatible with the idea of two types of Pe: The Stage-2 response on "Stage-1 correct/Stage-2 error" trials leads to both, a Pe due to the incorrect response and a higher-order Pe, whereas the Stage-2 response on "Stage-1 error/Stage-2 correct" trials leads only to a higher-order Pe. This putatively additive pattern of the two types of Pe might imply that the two underlying monitoring processes are independent but coincide in time. Further studies are necessary to determine whether the two types of Pe are based on a single mechanism or reflect two different levels of monitoring. This could be achieved by source modeling to identify whether the two types of Pe are generated by different networks of neural generators.

The design of the present study implies that decisions at the second stage are made either in a gain context (correct responses lead to a win, incorrect responses lead to a draw) or a loss context (incorrect responses lead to loss, correct responses lead to a draw). Several previous studies contrasted conditions with a gain, loss, or neutral context (neither gain nor loss) to investigate how motivational context influences error monitoring (Endrass et al. 2010; Maruo et al. 2016; Overmeyer et al. 2023; Potts 2011). In these studies, context was manipulated by the experimenter rather than being a consequence of an own preceding decision. While motivational context showed mixed effects on the Ne/ERN, the

majority of studies found no significant effects on the Pe (Maruo et al. 2016; Overmeyer et al. 2023; Potts 2011). Only a single study reported an increased Pe for loss context versus neutral context (Endrass et al. 2010). Crucially, none of these studies reported evidence that context influenced Pe activity on correct trials. This suggests that context alone cannot account for the observation of the higher-order Pe. It appears necessary that the unfavorable outcome was caused by the participants' own incorrect response at Stage 1. This highlights an important aspect, namely that the higher-order Pe is still a signature of performance monitoring rather than a pure index of behavioral outcome.

The present study reports a Pe that is not caused by an error in the current response (Stage 2) but by an error in the preceding response (Stage 1). A similar pattern has been observed in dual-tasking experiments (Steinhauser and Steinhauser 2021) in which participants responded to two independent tasks whose stimuli were separated either by a short interval (300 ms) or a long interval (1200 ms). With a short interval, it was observed that the Pe for errors in the first task was reduced, but instead a Pe emerged after the correct response in the second task. This was interpreted as reflecting an adaptive mechanism that shifts the Pe away from the initially incorrect task towards the end of the entire dual-task to prevent interference between error monitoring and task processing (Buzzell et al. 2017; Houtman and Notebaert 2013; Jentsch and Dudschig 2009). This interpretation received further support from the finding that the Pe to the (erroneous) first and (correct) second task response correlated negatively. It is unlikely that the higher-order Pe in the present data is due to a similar mechanism. Steinhauser and Steinhauser (2021) observed this effect only for temporally overlapping tasks but not if stimuli were separated by a long interval. In the present study, the stimuli of the two stages are separated by about 2600 ms, which is even longer than the long interval in their study. However, both studies demonstrate that a Pe can occur following a correct response, albeit for different reasons.

Remarkably, the Ne/ERN did not show the pattern of a higher-order error monitoring signal as it only occurred immediately after incorrect responses. This is consistent with theories of the Ne/ERN such as conflict monitoring theory (Yeung et al. 2004) or reinforcement learning theory (Holroyd and Coles 2002). Conflict monitoring predicts a Ne/ERN only if a response is followed by a corrective tendency which should never occur for correct Stage-2 responses even if they are associated with an unfavorable outcome. Reinforcement learning theory assumes that a Ne/ERN is elicited if a response indicates an outcome that is worse than expected. The scenario in which we observed a higher-order Pe (i.e., a correct Stage-2 response following an error at Stage 1) causes an unfavorable outcome (draw), which is not worse than expected as the Stage-1 error leads to a losing context in which this draw is the best possible outcome. Hence, the fact that the higher-order Pe is not accompanied by a higher-order Ne/ERN is fully compatible with these theories. It also confirms previous findings that a Pe can occur in the absence of a Ne/ERN, and hence, that these components reflect independent systems of error monitoring (Di Gregorio et al. 2018). A less expected outcome is that the Ne/ERN at Stage 1 (but not at Stage 2) is also observable at posterior electrodes. This is not due to the choice of the pre-response

baseline as an Ne/ERN-like negative deflection is also observable in the stimulus-locked data with a pre-stimulus baseline (see [Supporting Information](#)). One possibility is that this is related to the specific motor requirements in our paradigm, which could be addressed in future studies.

The present results demonstrate that a full understanding of neural signatures like the Pe cannot be achieved if error processing is exclusively investigated in single-stage tasks like the flanker task. More complex scenarios like our multistage task can reveal aspects of error monitoring that remain unnoticed in these simple tasks. This implies that it could be valuable to turn to more ecologically valid contexts. Real-world decision-making often involves multistage processes where an error in one stage can influence a final outcome (e.g., medical diagnoses, financial decisions). Recently, several studies addressed error monitoring in sports (Denul et al. 2025), car driving (Zhang et al. 2015) or immersive VR environments (Arake et al. 2022). Focusing on such settings could clarify the role of the higher-order Pe in adaptive behavior and outcome evaluation. The lack of ecological validity of single-stage choice tasks used in laboratory settings could also be the reason why brain activity in these tasks is only weakly related to everyday behavior (e.g., Saunders et al. 2022). Focusing on more complex tasks should be accompanied by involving more representative samples. Gender imbalances in typical studies on error monitoring—and also in the present study—could impair the generalizability of results as gender has been shown to influence neural signatures of error monitoring (Fischer et al. 2016; Larson et al. 2011).

5 | Conclusion

The present study demonstrates that a Pe can occur if a response is associated with an unfavorable outcome even if this response was correct. This suggests that the Pe reflects an inference-based outcome evaluation that takes place at a higher level of error monitoring that either integrates multiple error signals in a multistage task or takes context into account. In contrast, the Ne/ERN appears to operate on a lower level of error monitoring that mainly evaluates the correctness of an immediately preceding response.

Author Contributions

Peter Löschner: conceptualization, formal analysis, investigation, methodology, visualization, writing – original draft. **Marco Steinhauser:** conceptualization, funding acquisition, methodology, supervision, writing – review and editing.

Acknowledgments

We thank Celina Watschke, Valentin Sonnleitner, and Elias Staatz for their assistance in data collection. The data and analysis scripts of this study are available from the corresponding author upon request. Open Access funding enabled and organized by Projekt DEAL.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

- Arake, M., H. Ohta, A. Tsuruhara, et al. 2022. "Measuring Task-Related Brain Activity With Event-Related Potentials in Dynamic Task Scenario With Immersive Virtual Reality Environment." *Frontiers in Behavioral Neuroscience* 16: 779926. <https://doi.org/10.3389/fnbeh.2022.779926>.
- Bell, A. J., and T. J. Sejnowski. 1995. "An Information-Maximization Approach to Blind Separation and Blind Deconvolution." *Neural Computation* 7, no. 6: 1129–1159. <https://doi.org/10.1162/neco.1995.7.6.1129>.
- Boldt, A., and N. Yeung. 2015. "Shared Neural Markers of Decision Confidence and Error Detection." *Journal of Neuroscience* 35, no. 8: 3478–3484. <https://doi.org/10.1523/JNEUROSCI.0797-14.2015>.
- Buzzell, G. A., P. J. Beatty, N. A. Paquette, D. M. Roberts, and C. G. McDonald. 2017. "Error-Induced Blindness: Error Detection Leads to Impaired Sensory Processing and Lower Accuracy at Short Response-Stimulus Intervals." *Journal of Neuroscience* 37, no. 11: 2895–2903. <https://doi.org/10.1523/JNEUROSCI.1202-16.2017>.
- Coles, M. G. H., M. K. Scheffers, and C. B. Holroyd. 2001. "Why Is There an ERN/ne on Correct Trials? Response Representations, Stimulus-Related Components, and the Theory of Error-Processing." *Biological Psychology* 56, no. 3: 173–189. [https://doi.org/10.1016/S0301-0511\(01\)00076-X](https://doi.org/10.1016/S0301-0511(01)00076-X).
- Delorme, A., and S. Makeig. 2004. "EEGLAB: An Open Source Toolbox for Analysis of Single-Trial EEG Dynamics Including Independent Component Analysis." *Journal of Neuroscience Methods* 134, no. 1: 9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009>.
- Denul, A., G. Pourtois, T. Loeys, and W. Notebaert. 2025. "Social Post-Error Adaptations Across Four NBA Basketball Seasons." *Scientific Reports* 15, no. 1: 17919. <https://doi.org/10.1038/s41598-025-02006-x>.
- Di Gregorio, F., M. E. Maier, and M. Steinhauser. 2018. "Errors Can Elicit an Error Positivity in the Absence of an Error Negativity: Evidence for Independent Systems of Human Error Monitoring." *NeuroImage* 172: 427–436. <https://doi.org/10.1016/j.neuroimage.2018.01.081>.
- Di Gregorio, F., M. Steinhauser, and M. E. Maier. 2016. "Error-Related Brain Activity and Error Awareness in an Error Classification Paradigm." *NeuroImage* 139: 202–210. <https://doi.org/10.1016/j.neuroimage.2016.05.074>.
- Drizinsky, J., J. Zülch, H. Gibbons, and J. Stahl. 2016. "How Personal Standards Perfectionism and Evaluative Concerns Perfectionism Affect the Error Positivity and Post-Error Behavior With Varying Stimulus Visibility." *Cognitive, Affective, & Behavioral Neuroscience* 16, no. 5: 876–887. <https://doi.org/10.3758/s13415-016-0438-z>.
- Endrass, T., B. Reuter, and N. Kathmann. 2007. "ERP Correlates of Conscious Error Recognition: Aware and Unaware Errors in an Antisaccade Task." *European Journal of Neuroscience* 26, no. 6: 1714–1720. <https://doi.org/10.1111/j.1460-9568.2007.05785.x>.
- Endrass, T., B. Schuermann, C. Kaufmann, R. Spielberg, R. Kniesche, and N. Kathmann. 2010. "Performance Monitoring and Error Significance in Patients With Obsessive-Compulsive Disorder." *Biological Psychology* 84, no. 2: 257–263. <https://doi.org/10.1016/j.biopsycho.2010.02.002>.
- Falkenstein, M., J. Hohnsbein, J. Hoormann, and L. Blanke. 1990. "Effects of Errors in Choice Reaction Tasks on the ERP Under Focused and Divided Attention." In *Psychophysiological Brain Research*, edited by C. H. M. Brunia, A. W. K. Gaillard, and A. Kok, 192–195. Tilburg University Press.
- Falkenstein, M., J. Hohnsbein, J. Hoormann, and L. Blanke. 1991. "Effects of Crossmodal Divided Attention on Late ERP Components. II. Error Processing in Choice Reaction Tasks." *Electroencephalography and Clinical Neurophysiology* 78, no. 6: 447–455. [https://doi.org/10.1016/0013-4694\(91\)90062-9](https://doi.org/10.1016/0013-4694(91)90062-9).
- Falkenstein, M., J. Hoormann, S. Christ, and J. Hohnsbein. 2000. "ERP Components on Reaction Errors and Their Functional Significance: A Tutorial." *Biological Psychology* 51, no. 2-3: 87–107. [https://doi.org/10.1016/S0301-0511\(99\)00031-9](https://doi.org/10.1016/S0301-0511(99)00031-9).
- Faul, F., E. Erdfelder, A. Buchner, and A.-G. Lang. 2009. "Statistical Power Analyses Using G*Power 3.1: Tests for Correlation and Regression Analyses." *Behavior Research Methods* 41, no. 4: 1149–1160. <https://doi.org/10.3758/BRM.41.4.1149>.
- Faul, F., E. Erdfelder, A.-G. Lang, and A. Buchner. 2007. "G*Power 3: A Flexible Statistical Power Analysis Program for the Social, Behavioral, and Biomedical Sciences." *Behavior Research Methods* 39, no. 2: 175–191. <https://doi.org/10.3758/BF03193146>.
- Fischer, A. G., C. Danielmeier, A. Villringer, T. A. Klein, and M. Ullsperger. 2016. "Gender Influences on Brain Responses to Errors and Post-Error Adjustments." *Scientific Reports* 6, no. 1: 24435. <https://doi.org/10.1038/srep24435>.
- Fu, W.-T., and J. R. Anderson. 2008. "Solving the Credit Assignment Problem: Explicit and Implicit Learning of Action Sequences With Probabilistic Outcomes." *Psychological Research* 72, no. 3: 321–330. <https://doi.org/10.1007/s00426-007-0113-7>.
- Gehring, W. J., B. Goss, M. G. H. Coles, D. E. Meyer, and E. Donchin. 1993. "A Neural System for Error Detection and Compensation." *Psychological Science* 4, no. 6: 385–390. <https://doi.org/10.1111/j.1467-9280.1993.tb00586.x>.
- Groppe, D. M., T. P. Urbach, and M. Kutas. 2011. "Mass Univariate Analysis of Event-Related Brain Potentials/Fields II: Simulation Studies." *Psychophysiology* 48, no. 12: 1726–1737. <https://doi.org/10.1111/j.1469-8986.2011.01272.x>.
- Hewig, J., M. G. H. Coles, R. H. Trippe, H. Hecht, and W. H. R. Miltner. 2011. "Dissociation of Pe and ERN/ne in the Conscious Recognition of an Error." *Psychophysiology* 48, no. 10: 1390–1396. <https://doi.org/10.1111/j.1469-8986.2011.01209.x>.
- Holroyd, C. B., and M. G. H. Coles. 2002. "The Neural Basis of Human Error Processing: Reinforcement Learning, Dopamine, and the Error-Related Negativity." *Psychological Review* 109, no. 4: 679–709. <https://doi.org/10.1037/0033-295X.109.4.679>.
- Houtman, F., and W. Notebaert. 2013. "Blinded by an Error." *Cognition* 128, no. 2: 228–236. <https://doi.org/10.1016/j.cognition.2013.04.003>.
- Jentzsch, I., and C. Dudschig. 2009. "Short Article: Why Do We Slow Down After an Error? Mechanisms Underlying the Effects of Posterror Slowing." *Quarterly Journal of Experimental Psychology* 62, no. 2: 209–218. <https://doi.org/10.1080/17470210802240655>.
- Larson, M. J., M. South, and P. E. Clayson. 2011. "Sex Differences in Error-Related Performance Monitoring." *Neuroreport* 22, no. 1: 44–48. <https://doi.org/10.1097/WNR.0b013e3283427403>.
- Luck, S. J., A. X. Stewart, A. M. Simmons, and M. Rhemtulla. 2021. "Standardized Measurement Error: A Universal Metric of Data Quality for Averaged Event-Related Potentials." *Psychophysiology* 58, no. 6: e13793. <https://doi.org/10.1111/psyp.13793>.
- Maruo, Y., A. Schacht, W. Sommer, and H. Masaki. 2016. "Impacts of Motivational Valence on the Error-Related Negativity Elicited by Full and Partial Errors." *Biological Psychology* 114: 108–116. <https://doi.org/10.1016/j.biopsycho.2015.12.004>.
- Minsky, M. 1961. "Steps Toward Artificial Intelligence." *Proceedings of the IRE* 49, no. 1: 8–30. <https://doi.org/10.1109/JRPROC.1961.287775>.
- Morey, R. D. 2008. "Confidence Intervals From Normalized Data: A Correction to Cousineau (2005)." *Tutorial in Quantitative Methods for Psychology* 4, no. 2: 61–64. <https://doi.org/10.20982/tqmp.04.2.p061>.
- Murphy, P. R., I. H. Robertson, S. Harty, and R. G. O'Connell. 2015. "Neural Evidence Accumulation Persists After Choice to Inform Metacognitive Judgments." *eLife* 4: 1–23. <https://doi.org/10.7554/eLife.11946>.

- Navarro-Cebrian, A., R. T. Knight, and A. S. Kayser. 2016. "Frontal Monitoring and Parietal Evidence: Mechanisms of Error Correction." *Journal of Cognitive Neuroscience* 28, no. 8: 1166–1177. https://doi.org/10.1162/jocn_a_00962.
- Nieuwenhuis, S., K. R. Ridderinkhof, J. Blom, G. P. H. Band, and A. Kok. 2001. "Error-Related Brain Potentials Are Differentially Related to Awareness of Response Errors: Evidence From an Antisaccade Task." *Psychophysiology* 38, no. 5: S0048577201001111. <https://doi.org/10.1017/S0048577201001111>.
- O'Connell, R. G., P. M. Dockree, M. A. Bellgrove, et al. 2007. "The Role of Cingulate Cortex in the Detection of Errors With and Without Awareness: A High-Density Electrical Mapping Study." *European Journal of Neuroscience* 25, no. 8: 2571–2579. <https://doi.org/10.1111/j.1460-9568.2007.05477.x>.
- Overbeek, T. J. M., S. Nieuwenhuis, and K. R. Ridderinkhof. 2005. "Dissociable Components of Error Processing: On the Functional Significance of the Pe Vis-à-Vis the ERN/ne." *Journal of Psychophysiology* 19, no. 4: 319–329. <https://doi.org/10.1027/0269-8803.19.4.319>.
- Overmeyer, R., H. Kirschner, A. G. Fischer, and T. Endrass. 2023. "Unraveling the Influence of Trial-Based Motivational Changes on Performance Monitoring Stages in a Flanker Task." *Scientific Reports* 13, no. 1: 19180. <https://doi.org/10.1038/s41598-023-45526-0>.
- Parra, L. C., C. Alvino, A. Tang, et al. 2002. "Linear Spatial Integration for Single-Trial Detection in Encephalography." *NeuroImage* 17, no. 1: 223–230. <https://doi.org/10.1006/nimg.2002.1212>.
- Parra, L. C., C. D. Spence, A. D. Gerson, and P. Sajda. 2005. "Recipes for the Linear Analysis of EEG." *NeuroImage* 28, no. 2: 326–341. <https://doi.org/10.1016/j.neuroimage.2005.05.032>.
- Pfister, R., R. Wirth, K. A. Schwarz, A. Foerster, M. Steinhauser, and W. Kunde. 2016. "The Electrophysiological Signature of Deliberate Rule Violations." *Psychophysiology* 53, no. 12: 1870–1877. <https://doi.org/10.1111/psyp.12771>.
- Pion-Tonachini, L., K. Kreutz-Delgado, and S. Makeig. 2019. "ICLabel: An Automated Electroencephalographic Independent Component Classifier, Dataset, and Website." *NeuroImage* 198: 181–197. <https://doi.org/10.1016/j.neuroimage.2019.05.026>.
- Pontifex, M. B., M. R. Scudder, M. L. Brown, et al. 2010. "On the Number of Trials Necessary for Stabilization of Error-Related Brain Activity Across the Life Span." *Psychophysiology* 47, no. 4: 767–773. <https://doi.org/10.1111/j.1469-8986.2010.00974.x>.
- Potts, G. F. 2011. "Impact of Reward and Punishment Motivation on Behavior Monitoring as Indexed by the Error-Related Negativity." *International Journal of Psychophysiology* 81, no. 3: 324–331. <https://doi.org/10.1016/j.ijpsycho.2011.07.020>.
- Rodeback, R. E., A. Hedges-Muncy, I. J. Hunt, K. A. Carbine, P. R. Steffen, and M. J. Larson. 2020. "The Association Between Experimentally Induced Stress, Performance Monitoring, and Response Inhibition: An Event-Related Potential (ERP) Analysis." *Frontiers in Human Neuroscience* 14: 1–16. <https://doi.org/10.3389/fnhum.2020.00189>.
- Saunders, B., M. Milyavskaya, and M. Inzlicht. 2022. "Longitudinal Evidence That Event Related Potential Measures of Self-Regulation Do Not Predict Everyday Goal Pursuit." *Nature Communications* 13, no. 1: 3201. <https://doi.org/10.1038/s41467-022-30786-7>.
- Schroder, H. S., T. P. Moran, M. B. Donnellan, and J. S. Moser. 2014. "Mindset Induction Effects on Cognitive Control: A Neurobehavioral Investigation." *Biological Psychology* 103, no. 1: 27–37. <https://doi.org/10.1016/j.biopsycho.2014.08.004>.
- Steinhauser, M., and N. Yeung. 2010. "Decision Processes in Human Performance Monitoring." *Journal of Neuroscience* 30, no. 46: 15643–15653. <https://doi.org/10.1523/JNEUROSCI.1899-10.2010>.
- Steinhauser, M., and N. Yeung. 2012. "Error Awareness as Evidence Accumulation: Effects of Speed-Accuracy Trade-Off on Error Signaling." *Frontiers in Human Neuroscience* 6: 1–12. <https://doi.org/10.3389/fnhum.2012.00240>.
- Steinhauser, R., S. Kübler, M. Steinhauser, and T. Schubert. 2021. "Neural Correlates of Task-Order Preparation in Dual Tasks: An EEG Study." *Journal of Cognitive Neuroscience* 33, no. 10: 1–16. https://doi.org/10.1162/jocn_a_01752.
- Steinhauser, R., and M. Steinhauser. 2021. "Adaptive Rescheduling of Error Monitoring in Multitasking." *NeuroImage* 232: 117888. <https://doi.org/10.1016/j.neuroimage.2021.117888>.
- Stolyarova, A. 2018. "Solving the Credit Assignment Problem With the Prefrontal Cortex." *Frontiers in Neuroscience* 12: 182. <https://doi.org/10.3389/fnins.2018.00182>.
- Ullsperger, M., C. Danielmeier, and G. Jocham. 2014. "Neurophysiology of Performance Monitoring and Adaptive Behavior." *Physiological Reviews* 94, no. 1: 35–79. <https://doi.org/10.1152/physrev.00041.2012>.
- Ullsperger, M., A. G. Fischer, R. Nigbur, and T. Endrass. 2014. "Neural Mechanisms and Temporal Dynamics of Performance Monitoring." *Trends in Cognitive Sciences* 18, no. 5: 259–267. <https://doi.org/10.1016/j.tics.2014.02.009>.
- Ullsperger, M., H. A. Harsay, J. R. Wessel, and K. R. Ridderinkhof. 2010. "Conscious Perception of Errors and Its Relation to the Anterior Insula." *Brain Structure and Function* 214, no. 5–6: 629–643. <https://doi.org/10.1007/s00429-010-0261-1>.
- Walsh, M. M., and J. R. Anderson. 2014. "Navigating Complex Decision Spaces: Problems and Paradigms in Sequential Choice." *Psychological Bulletin* 140, no. 2: 466–486. <https://doi.org/10.1037/a0033455>.
- Winer, B. J., D. R. Brown, and K. M. Michels. 1991. *Statistical Principles in Experimental Design*. 3rd ed. McGraw-Hill.
- Wurm, F., B. Ernst, and M. Steinhauser. 2024. "Surprise-Minimization as a Solution to the Structural Credit Assignment Problem." *PLoS Computational Biology* 20: e1012175. <https://doi.org/10.1371/journal.pcbi.1012175>.
- Yeung, N., M. M. Botvinick, and J. D. Cohen. 2004. "The Neural Basis of Error Detection: Conflict Monitoring and the Error Related Negativity." *Psychological Review* 11, no. 4: 931–959. <https://doi.org/10.1037/0033-295X.111.4.931>.
- Zhang, H., R. Chavarriaga, Z. Khaliliardali, L. Gheorghe, I. Iturrate, and J. d. R. Millán. 2015. "EEG-Based Decoding of Error-Related Brain Activity in a Real-World Driving Task." *Journal of Neural Engineering* 12, no. 6: 066028. <https://doi.org/10.1088/1741-2560/12/6/066028>.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Figure S1:** Error-related brain activity for correct trials, Stage-1 errors (with correct Stage-2 response) and Stage-2-errors (with correct Stage-1 response) at electrode FCz. (A) Stage-1-locked waveforms in a time window comprising both stages. (B) Stage-2-locked waveforms. C and D: Difference waves between error trials and correct trials. Boxplots show the distribution of response times at the respective stage. **Figure S2:** Error-related brain activity for correct trials, Stage-1 errors (with correct Stage-2 response), and Stage-2-errors (with correct Stage-1 response) at electrode Pz. (A) Stage-1-locked waveforms in a time window comprising both stages. (B) Stage-2-locked waveforms. C and D: Difference waves between error trials and correct trials. Boxplots show the distribution of response times at the respective stage. **Figure S3:** Error-related brain activity for correct trials, Stage-1 errors (with correct Stage-2 response) and Stage-2-errors (with correct Stage-1 response) at electrode POz. (A) Stage-1-locked waveforms in a time window comprising both stages. (B) Stage-2-locked waveforms. C and D: Difference waves between error trials and correct trials. Boxplots show the distribution of response times at the respective stage.