Adaptive rescheduling of error monitoring in multitasking

Abbreviated title: Error monitoring in multitasking

Robert Steinhauser*[1], Marco Steinhauser[1]

[1] Department of Psychology, Catholic University of Eichstätt-Ingolstadt, Ostenstraße 25, 85072 Eichstätt, Germany

Corresponding author: Robert Steinhauser, Catholic University of Eichstätt-Ingolstadt, Ostenstraße 25, 85072 Eichstätt, Germany. Phone: +49 8421 9321120. E-mail: robert.steinhauser@ku.de.

33 pages, 8 figures, 0 tables

Word counts: Abstract: 150; Main text: 6393; Methods: 2196

Abstract

The concurrent execution of temporally overlapping tasks leads to considerable interference between the subtasks, which is overcome by a neural mechanism that actively defers processing of the secondary task. In this study, we recorded error-related scalp potentials to investigate whether such a mechanism can also be found in error monitoring. Participants worked on a multitasking paradigm with partially overlapping execution of two tasks (T1, T2). The error positivity, a correlate of conscious error evaluation, was reduced after T1 errors but occurred after a correct T2-response instead. MVPA-based and regression-based single-trial analysis revealed a trial-wise trade-off between immediate and postponed error processing. Two follow-up experiments confirmed this finding, showed that this result is not due to credit-assignment errors in which a T1 error is falsely attributed to T2, and verified the link to conscious error detection. These findings demonstrate a remarkable flexibility of the human brain when adapting to multitasking situations.

(150 words)

Keywords: multitasking, dual-tasking, error awareness, Pe, ERN

Introduction

Adaptive human behavior requires an error monitoring system that constantly checks whether an error has occurred and initiates adjustments to cognition and behavior accordingly. Numerous studies could show that such a system exists in the human brain, detecting errors fast and reliably (Ullsperger, Danielmeier, & Jocham, 2014). Most studies, however, investigated error monitoring when single tasks were executed in isolation. In contrast, everyday behavior is characterized by the concurrent execution of temporally overlapping tasks. As dual-tasking is associated with decrements in task performance (Pashler, 1994; Tombu & Jolicœur, 2003), it is plausible to assume that also error monitoring suffers under these conditions. Indeed, recent evidence suggested that dual-tasking leads to specific impairments to the neural correlates of error monitoring (Klawohn, Endrass, Preuss, Riesel, & Kathmann, 2016; Weißbecker-Klaus, Ullsperger, Freude, & Schapkin, 2016). In the present study, we investigate how the brain adapts to these dual-tasking conditions. By considering error-related brain activity in event-related potentials (ERPs), we ask how monitoring processes are reorganized to maintain reliable error detection when the brain is confronted with two temporally overlapping tasks.

Close temporal succession of more than one task leads to considerable interference between the subtasks, resulting in performance decrements compared to the execution of single tasks (Jersild, 1927; Telford, 1931). As dual-task interference arises predominantly in central, decision-related processes, the brain adapts to this interference by serializing these processes, granting attentional resources to only one task representation at a time (Meyer & Kieras, 1997; Pashler, 1994). Neuroscientific studies suggested that interference as well as the resulting serialization is linked to a fronto-parietal network that is able to grant access to consciousness to only one task representation, by synchronizing distinct brain regions in a self-amplifying process (Dehaene, Kerszberg, & Changeux, 1998; Del Cul, Baillet, &

Dehaene, 2007; Gross et al., 2004; Hesselmann, Flandin, & Dehaene, 2011; Sergent, Baillet, & Dehaene, 2005; Sergent & Dehaene, 2004).

A number of electrophysiological studies has shown that the serialized access to consciousness manifests in the P300 component (Dell'Acqua, Jolicoeur, Vespignani, & Toffanin, 2005; Hesselmann et al., 2011; Sergent & Dehaene, 2004; Sigman & Dehaene, 2008). Sergent and colleagues (2005) observed a significantly reduced P300 on stimuli that were not perceived consciously due to the continuous processing of a preceding stimulus, whereas earlier visual potentials were equally elicited by consciously perceived and unperceived stimuli. Similarly, Sigman and Dehaene (2008) found in a dual-tasking paradigm that neural correlates of early attentional processes were executed in parallel for both tasks, whereas the P300 for the second task (T2) was delayed until response selection for the first task (T1), represented by its own P300, had been completed (see also Dell'Acqua et al., 2005; Hesselmann et al., 2011). Recently, Marti, King, & Dehaene (2015) provided evidence that this serialization of task selection is the result of a competition between task-related processes when they strive for conscious access. As only one task representation at a time can get access to consciousness, neural representations of the two tasks repel each other in a way that initially, only processing of T1 is enabled, and that of T2 is actively delayed.

It has recently been suggested that also processes involved in the detection and evaluation of performance errors suffer from dual-tasking interference. Research has robustly identified two ERPs associated with distinct stages of error monitoring, which are executed in a fixed temporal cascade immediately following the erroneous response (Ullsperger, Fischer, Nigbur, & Endrass, 2014) – the error-related negativity (Ne/ERN, Falkenstein et al., 1991; Gehring et al., 1993) and the error positivity (Pe, Overbeek et al., 2005). The Ne/ERN, an early fronto-central negativity on error trials, has been assumed to reflect response conflict (Yeung, Botvinick, & Cohen, 2004) or prediction errors (Holroyd & Coles, 2002) that

accompany performance errors. The Pe, on the other hand, is a later positive deflection over parietal electrodes, which is supposed to reflect conscious error detection (Boldt & Yeung, 2015; Nieuwenhuis, Ridderinkhof, Blom, Band, & Kok, 2001; Overbeek et al., 2005; M. Steinhauser & Yeung, 2010), likely resulting from the process of accumulating evidence that an error has occurred (M. Steinhauser & Yeung, 2010). It has recently been shown that the Pe and conscious error awareness can occur even if the earlier Ne/ERN is impaired or absent (Di Gregorio, Maier, & Steinhauser, 2018; Maier, Di Gregorio, Muricchio, & Di Pellegrino, 2015a). Moreover, particularly the Pe interferes with subsequent stimulus processing (Buzzell et al. 2017, but see Beatty et al., 2018) or is reduced when two temporally overlapping tasks have to be executed (Weißbecker-Klaus et al., 2016). In contrast, the Ne/ERN appears to be sensitive to working memory load even if no concurrent task execution is required (Klawohn et al., 2016; Maier & Steinhauser, 2017; Moser, Moran, Schroder, Donnellan, & Yeung, 2013). The observation that mainly the Pe, and thus conscious error processing, suffers from dual-tasking interference is in accordance with the aforementioned studies on dual-tasking. The Pe exhibits striking physiological similarities with the P300, which, as outlined above, is particularly affected by dual-tasking interference. It shares the time course and scalp topography of the P300 (Leuthold & Sommer, 1999; Overbeek et al., 2005) and has been linked to joint activation in bilateral prefrontal and parietal brain regions (Hester, Foxe, Molholm, Shpaner, & Garavan, 2005), equivalently to the synchronization of brain areas that was previously suggested to form the basis of the P300 (Dehaene et al., 1998; Sergent & Dehaene, 2004; Sigman & Dehaene, 2008).

Whereas the similarity between the Pe and P300 can explain why specifically the Pe suffers from dual-tasking interference, little is known how the brain maintains the ability to consciously detect errors under dual-tasking. One solution to this problem would be to serialize not only task processing but also the accompanying error monitoring processes

(Hochman & Meiran, 2005). Processing a T1 error could lead to an additional deferment of T2 processing, an effect that has indeed been demonstrated in behavioral studies (Jentzsch & Dudschig, 2009; M. Steinhauser, Ernst, & Ibald, 2017). However, reduced error signaling rates when an erroneous response is quickly followed by another task (Rabbitt, 2002) and the observation of a reduced Pe in dual-tasking (Weißbecker-Klaus et al., 2016) suggest that such a *serialization* cannot fully protect error monitoring from dual-task interference. Consequently, it might be necessary to actively defer conscious error detection to the end of a dual-task scenario. In this case, processing a T1 error would in some or all trials not occur until the response to T2. Such an *adaptive rescheduling* of conscious error processing would imply that the accumulation of evidence for an error that underlies the Pe can be detached from early error signals represented by the Ne/ERN.

The present study investigated this adaptive rescheduling hypothesis, which proposes that the brain adapts to dual-task interference in error monitoring by deferring conscious error detection to after completion of the dual task. To this end, we focused on T1 errors but analyzed ERPs following both the T1-response and the T2-response. We contrasted trials with a sufficiently long interval between the two task stimuli that allowed serial task execution (1200 ms) with trials that required overlapping task execution due to a considerably shorter interval (300 ms). In serial task execution, we expected to see the typical Ne/ERN and Pe after the erroneous T1-response. Crucially, in overlapping task execution, we predicted that the Pe after the T1-response would be reduced or even absent while a Pe would now be obtained after the correct T2-response. Such a result would demonstrate that the decision about the correctness of T1 is deferred in some or all trials, respectively, until the completion of T2 and extend previous accounts on an active deferment of component processes in dual-tasking (Hesselmann et al., 2011; Sigman & Dehaene, 2008; R. Steinhauser & Steinhauser, 2018) to the area of performance monitoring.

Results

Experiment 1

24 healthy adult participants worked on a variant of the psychological refractory period (PRP) paradigm, which is commonly used to investigate mutual interference between subtasks in dual-tasking situations (Pashler, 1994; Tombu & Jolicœur, 2003). The details of the experimental paradigm can be found in Figure 1.

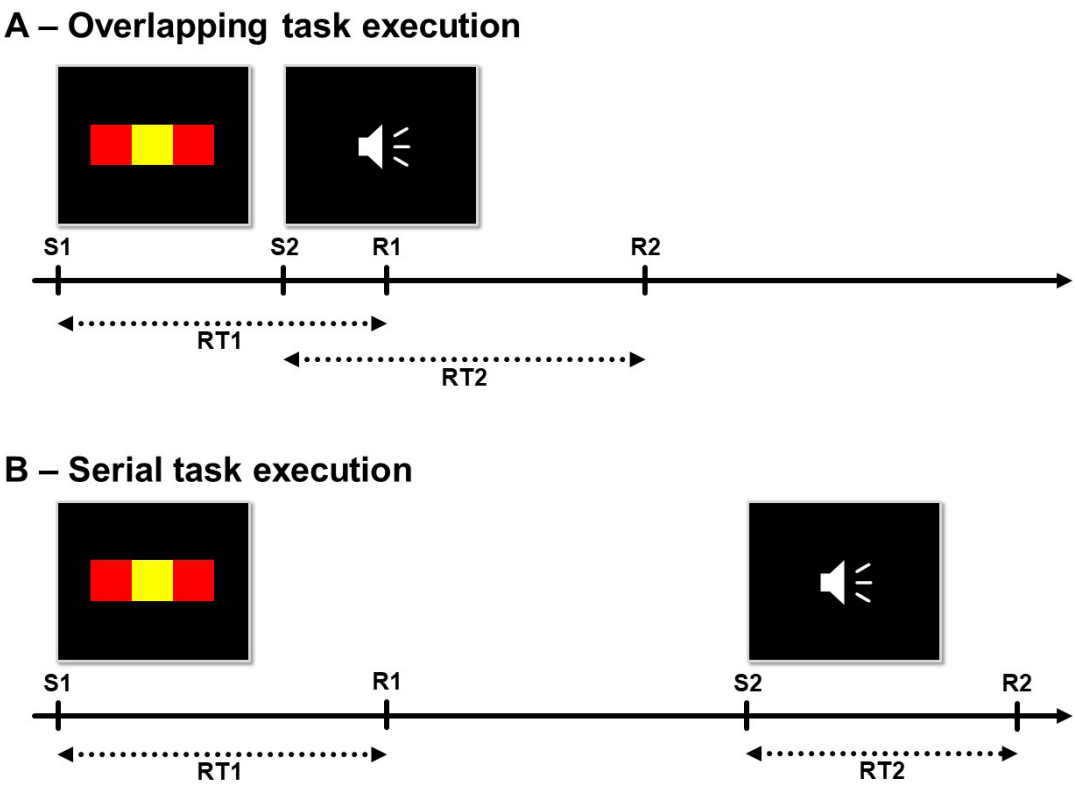

Figure 1. Time course of two example trials. On each trial, participants had to respond to two tasks that were separated by an unpredictable stimulus onset asynchrony of 300 ms (overlapping task execution, A) or 1200 ms (serial task execution, B). A three-choice color variant of the flanker task, in which the color of the central square had to be indicated (red, yellow, blue), served as Task 1. A two-choice pitch discrimination task was presented subsequently as Task 2, with an auditory sine tone stimulus of 400 Hz or 900 Hz. Participants were instructed to respond as fast as possible to both tasks, in the given order. S1 = Stimulus 1. S2 = Stimulus 2. R1 = Response 1. R2 = Response 2. RT1 = response time to Task 1. RT2 = response time to Task 2.

Behavior. RTs and error rates are depicted in Figure 2. To verify that our paradigm creates a dual-tasking scenario with overlapping task execution, we first examined whether two typical effects of dual-tasking can be observed in this dataset: First, the so-called PRP effect refers to the observation that RTs to T2 increase with a decreasing stimulus onset asynchrony (SOA), and thus indicates a form of dual-task cost. This effect is typically explained by the idea that, with more overlap between tasks, T2 execution is delayed (Pashler, 1994) or suffers from depleted resources (Tombu and Jolicœur, 2003). Second, it has recently been shown that T1 errors lead to increased RTs to T2, and this post-error slowing is larger for short SOAs than for long SOAs (Steinhauser et al., 2017). This phenomenon presumably reflects interference between T1 error monitoring and execution of T2, which again is larger with short SOAs.

To analyze both effects in our data, we subjected RTs for T1 and T2 separately to repeated measures ANOVAs on the variables SOA (short vs. long) and T1 Correctness (correct vs. error). Characteristic of PRP paradigms, no significant effects were found for RTs to T1, Fs(1, 23) < 2.70, ps > .11, $\eta_p^2 < .11$. In contrast, RTs to T2 were considerably slower on trials with short SOA than with long SOA, $F(1, 23) = 83.4$, $p < .001$, $\eta_p^2 = .78$, reflecting the typical PRP effect. In addition, RTs to T2 were slower when the preceding T1-response was an error, $F(1,23) = 56.8$, $p < .001$, $\eta_p^2 = .71$, and as expected, this post-error slowing on T2 was increased with short SOA, $F(1, 23) = 25.9$, $p < .01$, $\eta_p^2 = .53$. Thus, our data replicate well-known signatures of dual-task interference, which demonstrates that the short-SOA condition induced overlapping task execution. In addition, an analysis of inter-response intervals (IRIs, the time between the T1 response and the T2 response) in the short-SOA condition ruled out that participants had grouped their responses. While IRIs were considerable longer on T1 errors (587 ms) compared to corrects (381 ms) due to the above mentioned post-error slowing, $t(23) = 8.35$, $p < .001$, $d = 1.25$, IRIs in both conditions were

far beyond the commonly used threshold for response grouping of 50 – 100 ms (Hommel, 1998; Pashler & Johnston, 1989; Welford, 1952).

Error rates for T1 were high enough for an analysis of error-related brain activity. A mean T1 error rate of 8.31 % resulted in an average number of 98.3 trials with T1 errors but correct T2 responses per participant. A repeated measures ANOVA on the variables SOA (short vs. long) and Task (T1 vs. T2) yielded a significant interaction, $F(1, 23) = 8.06$, $p = .009$, $\eta_p^2 = .26$, indicating that, whereas T1 errors were equally frequent in short-SOA and long-SOA trials, T2 errors were more frequent in short-SOA trials as compared to long-SOA trials, $t(23) = 2.69$, $p = .013$, $d = .50$. This again demonstrates increased interference in the short-SOA condition.

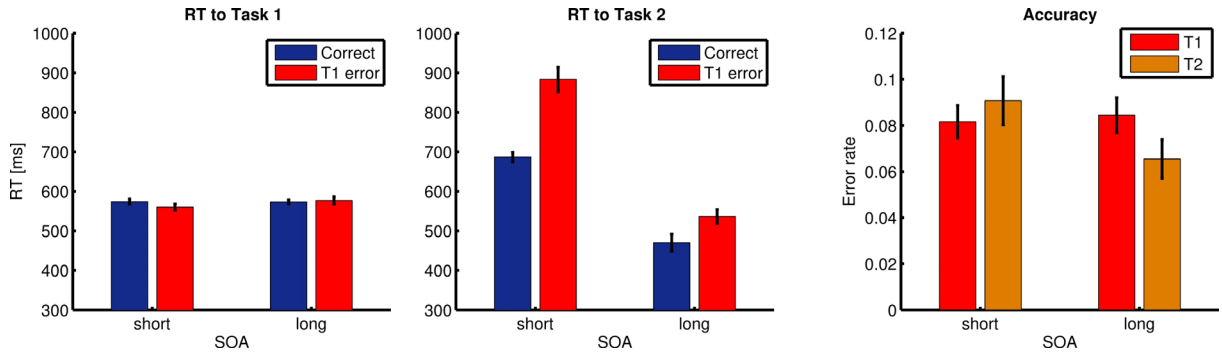

Figure 2. Behavioral results of Experiment 1. RTs to Task 1 and Task 2 are depicted in the left and middle panel, respectively. Error rates are presented in the right panel. Error bars indicate within-subject standard errors of the mean (Cousineau, 2005; Morey, 2008). RT = response time. SOA = stimulus onset asynchrony. T1 = Task 1. T2 = Task 2.

Ne/ERN and Pe. To investigate our rescheduling hypothesis, we analyzed error-related brain activity in short-SOA trials, which induce overlapping task execution, and long-SOA trials, which allow for serial task execution. Our central prediction was that, with a short SOA, the Pe associated with T1 errors should partially be rescheduled to the end of the dual-task. This would result in a reduced Pe following incorrect T1 responses but the emergence of

a Pe after correct T2 responses. We had no explicit hypothesis on the Ne/ERN but nevertheless analyzed this component to determine whether comparable effects can be found for this early form of error processing.

We initially analyzed T1-response-locked ERPs to find out about immediate neural correlates of error processing after T1 errors. A distinct parietal positivity, the Pe, was clearly observable for both SOA conditions (Fig 3A). For an analysis of the Pe, we subjected mean amplitudes at electrode POz to a repeated measures ANOVA on the variables SOA (short vs. long) and T1 Correctness (correct vs. error). A Pe was evident across both SOA conditions, as indicated by a main effect of T1 Correctness, $F(1, 23) = 28.6$, $p < .001$, $\eta_p^2 = .55$, but a significant interaction revealed that this Pe was far less pronounced in trials with short SOA than in trials with long SOA, $F(1, 23) = 11.0$, $p = .003$, $\eta_p^2 = .32$. Raw ERP waves in Figure 3 suggest that this interaction is mainly driven by a difference in the amplitudes of correct trials. This is likely caused by the fact that in most trials of the short SOA condition, the stimulus of T2 is presented and processed sometime within the time window observed here, eliciting stimulus-locked ERPs that are superimposed on the raw ERP waveforms of both correct and error trials in the short SOA condition (see Sigman & Dehaene, 2008). For this reason, we quantify the Pe not from error trials alone but as the difference between correct and error trials. To additionally rule out that the observed interaction may truly be rooted in differences in correct trials, we correlated the difference wave of conditions correct long and correct short with the difference wave of the interaction term of the above analysis. We found no significant correlation, $r = 0.14$, $p = .50$, indicating that the Pe reduction in trials with short SOA is not a mere consequence of the amplitude difference in correct trials.

Figure 3B shows that also a clear frontocentral Ne/ERN was observable in both SOA conditions. Indeed, subjecting mean amplitudes at electrode FCz to a repeated measures ANOVA of the same variables as above revealed that a significant Ne/ERN across both SOAs

was obtained, $F(1, 23) = 22.8$, $p < .001$, $\eta_p^2 = .50$. In addition, a main effect of SOA demonstrated that mean amplitudes of correct as well as error trials were more negative in trials with short SOA, $F(1, 23) = 13.5$, $p = .001$, $\eta_p^2 = .37$. However, the interaction between both variables did not reach significance, $F(1, 23) = 2.63$, $p = .12$, $\eta_p^2 = .10$.

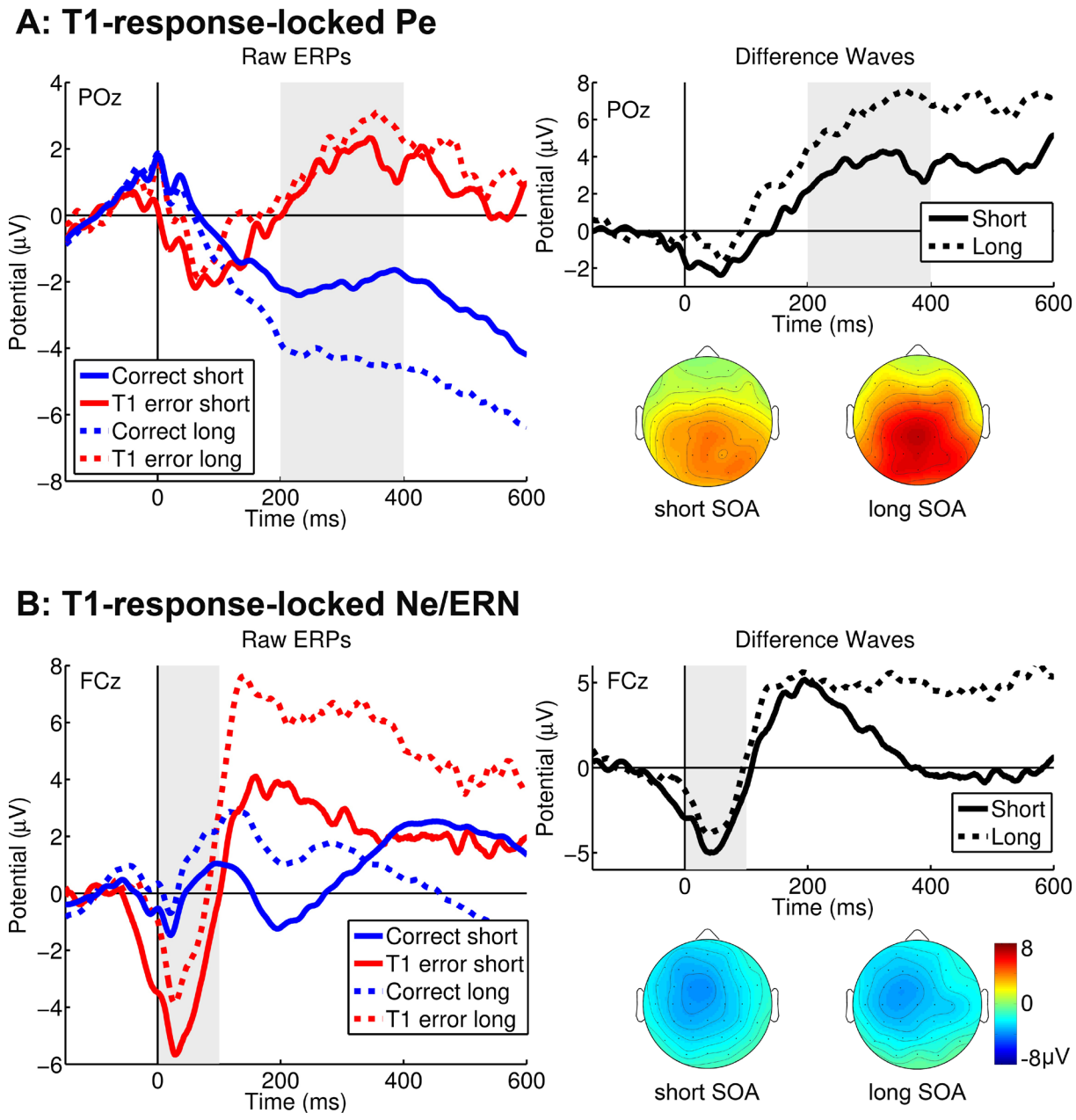

Figure 3. ERPs locked to the T1-response at posterior (A) and frontocentral (B) electrodes. Difference waves are computed from the respective T1 error minus T1 correct raw ERPs. Scalp topographies represent these difference waves. Gray areas indicate the time

intervals for statistical testing for the Pe (A) and the Ne/ERN (B). R1 = Response 1. T1 = Task 1. SOA = stimulus onset asynchrony.

Having shown that conscious error detection reflected by the Pe was largely impaired in the short-SOA condition, we subsequently examined neural correlates of T1 error processing after completion of the whole dual-task, i.e., in T2-response-locked data (Fig. 4). We investigated the possible emergence of such a deferred Pe (Fig. 4A) by means of ANOVAs on the variables SOA (short vs. long) and T1 Correctness (T1 correct vs. T1 error). It must be noted, though, that both conditions of the variable T1 Correctness here represent trials whose T2 was answered correctly, i.e., the "error-related" brain activity reported here was locked to a correct T2-response. Nonetheless, we obtained a significant interaction between SOA and T1 Correctness on Pe amplitudes, $F(1, 23) = 5.55$, $p = .027$, $\eta_p^2 = .19$, reflecting that correct T2-responses elicited a significant Pe on short-SOA trials, $t(23) = 2.00$, $p = .029$, $d = .27$, but not on long-SOA trials.

The same analysis on the Ne/ERN amplitudes revealed only a marginally significant main effect of T1 Correctness, $F(1, 23) = 3.48$, $p = .075$, $\eta^2_{\text{part.}} = .13$, indicating a negativity after T1 errors compared to T1 correct trials, but no significant interaction, $F(1, 23) = 1.73$, $p = .20$. Visual inspection of the topographies (Fig. 4B) revealed that this effect peaked earlier than the typical Ne/ERN and had a more frontal distribution than the Ne/ERN in T1-response-locked potentials (however, the same analysis at electrode Fz showed similar results). This suggests that, while there appears to be some frontocentral activity related to the T1 error in T2-response-locked data, this effect lacks robustness and differs from the usually observed Ne/ERN.

To sum up, we could provide evidence that overlapping task execution goes along with a reduction of the Pe to the incorrect T1 response whereas a sizeable Pe emerges after

the correct T2-response. We interpret this *deferred Pe* as a rescheduling of conscious error processing to the end of the dual-task trial.

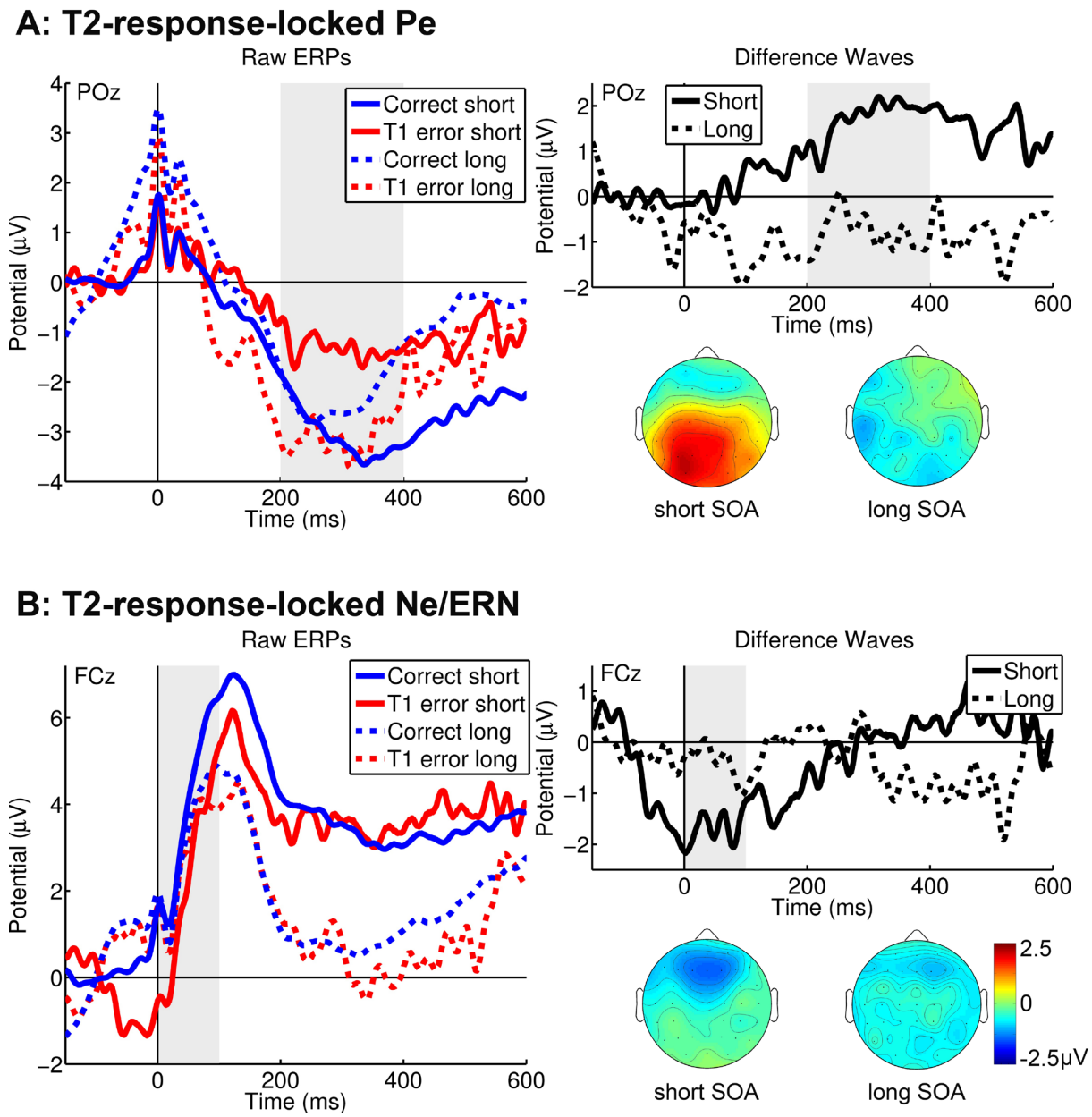

Figure 4. ERPs locked to the (correct) T2-response at posterior (A) and frontocentral (B) electrodes. Difference waves are computed from the respective T1 error minus T1 correct raw ERPs. Scalp topographies represent these difference waves. Gray areas indicate the time intervals for statistical testing for the Pe (A) and the Ne/ERN (B). R2 = Response 1. T1 = Task 1. SOA = stimulus onset asynchrony.

Relationship between immediate and deferred error processing. Although the immediate Pe that followed the T1 response was strongly reduced in the short-SOA condition, it was observable there nonetheless. This suggests two possible ways of how error monitoring is carried out under conditions of overlapping task execution. On the one hand, the accumulation of evidence that a T1 error has occurred could be evenly distributed between immediate (following T1 response) and deferred (following T2 response) error monitoring. On the other hand, error monitoring could occur for some trials immediately after the erroneous T1 response, whereas it is deferred until after the T2 response in other trials. To distinguish between these two accounts, we examined the trial-wise relationship between immediate and deferred Pe in two ways.

First, we investigated the inverse relationship of immediate and deferred error processing by deriving single-trial estimates for the R1-locked and the R2-locked Pe from mean amplitudes in raw data and comparing them in a regression-based analysis. This analysis had to be limited to trials with IRI > 400 ms, however, because raw EEG data within the same epoch are highly susceptible to autocorrelations in overlapping time periods. Importantly, this does not affect the remaining analyses of the present study, because EEG data from different, individually baseline-corrected epochs are compared. The restriction to IRIs above 400 ms resulted in M = 31.04 trials per participant. Figure 5 depicts strong negative beta weights in a significant cluster from 290 ms onwards in R1-locked data and 170 ms in R2-locked data, that is, around the time of the R1-locked and R2-locked Pe. This is evidence for a trade-off of the immediate and the delayed Pe, supporting the account that error processing happens immediately after some erroneous responses, whereas it is deferred on other trials.

Second, we used multivariate pattern analysis (MVPA) to create a set of classifiers over consecutive time windows that optimally distinguished between correct trials and T1

error trials based on T1-response-locked brain activity, which allowed us to examine all trials irrespective of their IRI. Mean classifier accuracy peaked around 190 ms with an Az of .61 but inspection of participants' individual classifier accuracies revealed that one participant (Subject 13) exhibited a remarkably low classification accuracy of .34, which is more than 2.5 standard deviations below the mean classification accuracy across all participants. This apparently failed attempt to compute a successful MVPA for this participant is likely due to the small number of 20 trials that entered the training set (the average training set size of all participants was 120.63 trials). For this reason, data from Subject 13 was excluded from subsequent MVPA-based analyses. Nonetheless, all effects remain significant also when including that participant (all $ps < .05$). Based on the peak classifier window, we computed the prediction value for each trial. These prediction values represent the degree to which each trial elicits error-related brain activity in this time window (Boldt & Yeung, 2015; Steinhauser & Yeung, 2010), and thus represent a single-trial estimate of the T1-response-locked Pe. We utilized these prediction values to split the error trials of each participant into three equally-sized bins that represent a small, medium, and large T1-response-locked Pe (Fig. 5A). A one-way repeated measures ANOVA on mean Pe amplitudes with the variable Bin (small Pe vs. medium Pe vs. large Pe) showed that this MVPA-based separation process was in fact able to separate trials according to Pe size, $F(2, 44) = 4.72$, $p = .015$, $\eta_p^2 = .18$. Contrasts revealed that this effect was mainly driven by trials in the large Pe bin having a larger T1-response-locked Pe than trials in the small Pe bin, $t(22) = 2.90$, $p = .008$, $d = .52$. Crucially, an equivalent analysis of T2-response-locked Pe values based on this T1-response-based bin separation (Fig. 5B) shows the opposite pattern, $F(2, 44) = 6.75$, $p = .004$, $\eta_p^2 = .23$. Trials with a large immediate Pe following the T1-response showed a particularly small deferred Pe following the T2-response, $t(22) = 4.10$, $p < .001$, $d = .63$. This conceptually replicates the above regression-based findings on a trade-off of the immediate and deferred Pe and thus

provides additional support for the idea that the rescheduling of conscious error processing is all-or-nothing. Apparently, T1-error processing is deferred in some trials to the end of the dual-task, whereas in other trials it occurs immediately after the incorrect T1 response.

The MVPA-based differentiation of trials with small, medium and large immediate Pe eventually also allowed us to investigate how immediate error processing affects selecting and executing the response to Task 2 (Fig. 5C). We compared IRIs in a one-way repeated measures ANOVA (small vs. medium vs. large immediate Pe) and found a significant difference between the conditions, $F(2,44) = 3.42$, $p = .045$, $\eta_p^2 = .13$. Contrasts confirmed that trials with a large immediate Pe were followed by slower responses to T2 than trials with a small immediate Pe, $t(22) = 2.94$, $p = .001$, $d = .40$, which in turn supports the idea that immediate error processing negatively affects efficient execution of the second task.

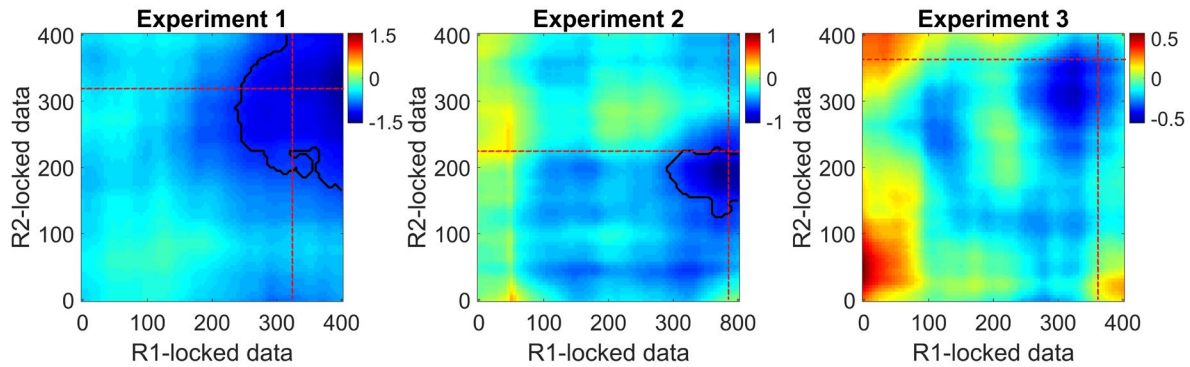

Figure 5 Regression weights of the regression-based single-trial analyses for Experiments 1-3, in which amplitudes of R1-locked data predict amplitudes of R2-locked data. Black lines indicate the borders of significant clusters as revealed by a cluster-based permutation test. Red dotted lines indicate the peak amplitude of the Pe in the respective difference waves. While on the numerical level, a negative association of the R1-locked Pe and R2-locked Pe becomes evident in all three experiments, statistical significance is reached only in Experiments 1 and 2.

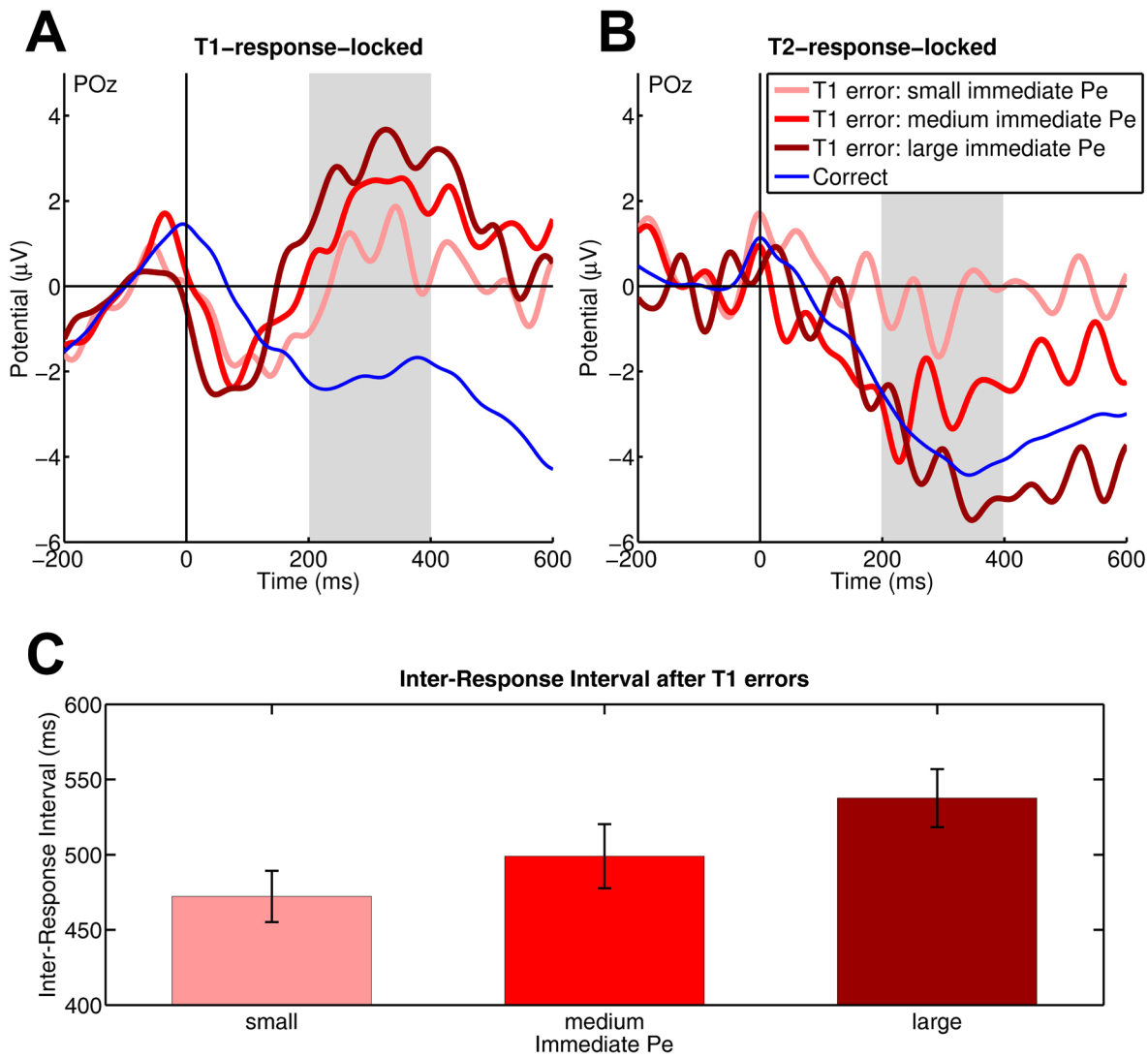

Figure 6. ERPs at a posterior electrode locked to the T1-response (A) and T2-response (B) and inter-response intervals after T1 errors (C). T1 error trials are divided into three separate conditions – small, medium, and large immediate Pe – based on a single-trial estimate of the Pe. Correct trials are presented for comparison (thin line). A 20 Hz lowpass filter was applied for better visibility. Gray areas indicate the time intervals for statistical testing for the Pe. R1 = Response 1. R2 = Response 2. T1 = Task 1. T2 = Task 2.

Experiment 2

Individual trials that elicit a deferred Pe with simultaneous omission of an immediate Pe could also be rooted in the problem of credit assignment (Fu & Anderson, 2008; Sutton & Barto, 1998; Walsh & Anderson, 2011). It is conceivable that trials with a deferred Pe actually reflect that error monitoring occasionally misattributed internal error signals to T2, and therefore falsely detected a T2 error. Rather than a rescheduling of T1-error processing, such a Pe would represent immediate T2-error processing. In a follow-up experiment with 24 different participants (out of which one participant had to be excluded due to technical difficulties during data acquisition), we did not only want to replicate our initial findings but also investigated whether this alternative explanation could account for our findings.

In Experiment 2, participants worked on the same PRP paradigm but reported after each trial by key press, whether they had committed an error in T1, T2, or both. This allowed us to distinguish between correctly assigned T1 errors (T1 hits) and T1 errors that were mistakenly reported as T2 errors (T2 false alarms). Rates of averaged error reports are depicted in Figure 7A. In fact, only 4.9% of all T1 errors were T2 false alarms (on average 0.7 trials per participant) and thus occurred very rarely. As only 17.0% of T1 errors were reported to be correct (T1 misses; on average 6.78 trials per participant), we could base our further ERP analyses on a solid subset of 73.8% of T1 hits (on average 31.0 trials per participant).

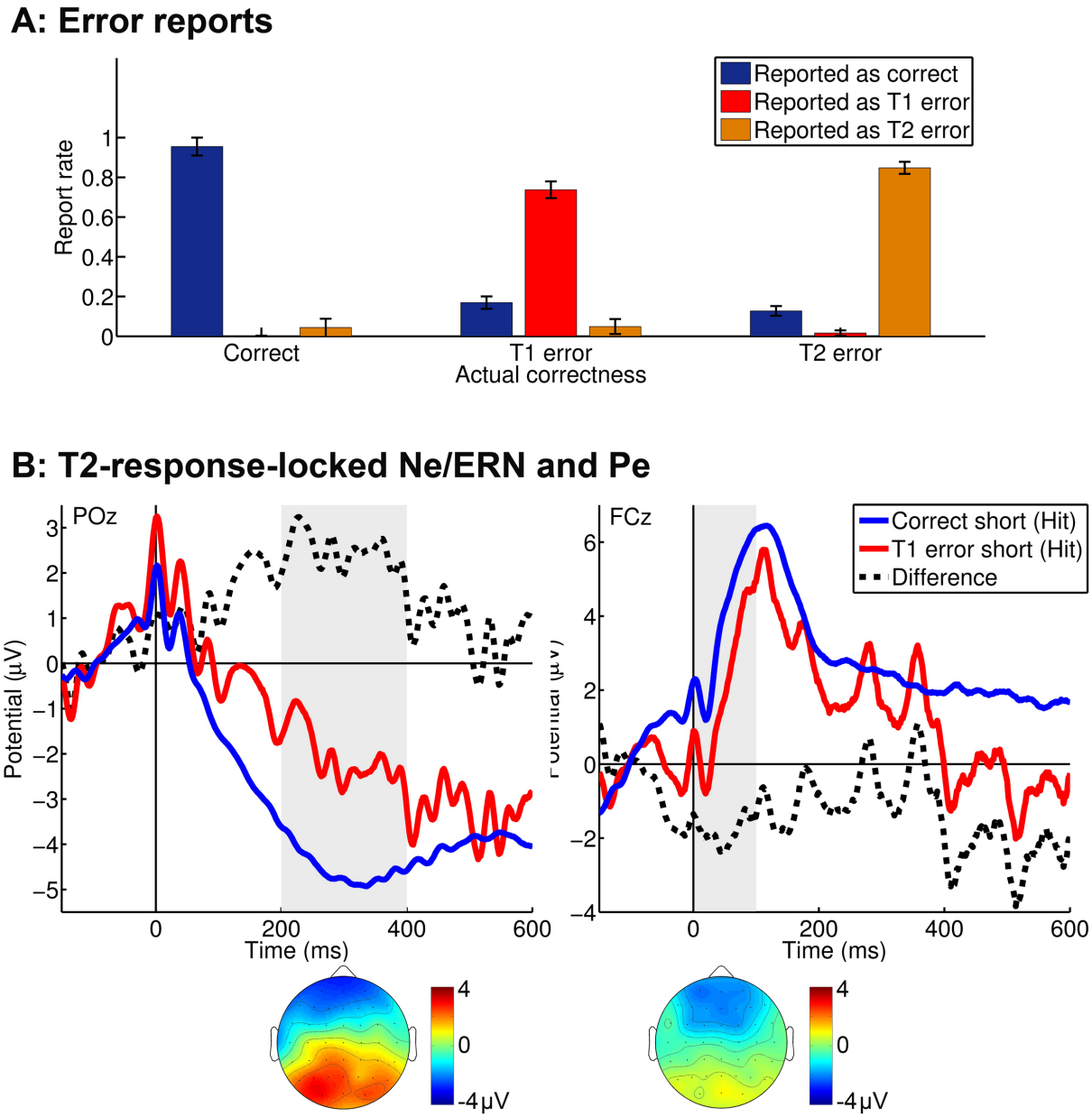

Figure 7. Error reports (A) and ERP results (B) of Experiment 2. Error bars in Panel A indicate within-subject standard errors of the mean (Cousineau, 2005; Morey, 2008). Panel B represents ERPs locked to the T2-response at posterior (left) and frontocentral (right) electrodes. Difference waves (dotted lines) are computed from T1 error minus correct raw ERPs. Scalp topographies represent these difference waves. Gray areas indicate the time intervals for statistical testing for the Pe (left) and the Ne/ERN (right). R1 = Response 1. R2 = Response 2. T1 = Task 1. T2 = Task 2.

To compensate for the increased trial duration due to trial-wise error reporting, Experiment 2 was restricted to the short SOA condition. Only T2-response-locked data are reported here as no immediate Ne/ERN and Pe could be compared between SOA conditions. Although only T1 hits were included in the analysis, a clear deferred Pe could be observed after T2-response-locked data (Fig. 6B), $t(22) = 2.86$, $p = .009$, $d = .57$.

As in Experiment 1, the analysis of the Ne/ERN in T2-response-locked data revealed only a marginally significant negativity for T1 errors relative to correct trials at electrode FCz, $t(22) = 1.87$, $p = .075$, $d = .51$. However, inspection of the topographies (Fig. 7B) showed that the peak of the Ne/ERN was again at more frontal electrodes. Correspondingly, additional testing was conducted at electrode Fz and now revealed a significant deferred Ne/ERN, $t(22) = 2.26$, $p = .034$, $d = .62$.

The regression-based approach on raw data with IRI > 400ms yielded a pattern with close similarity to that in Experiment 1, albeit somewhat earlier in R2-locked data (Figure 6). Again, strong negative beta weights form a significant cluster from 290 ms onwards in R1-locked data and between 125 ms and 230 ms in R2-locked data. We tentatively suggest that this negative correlation between the R1-locked Pe and the R2-locked Pe is reduced in the R2-locked time window beyond 250 ms due to the strong influence of a pronounced overlaying Contingent Negative Variation in anticipation of the error report (see Figure 7B). However, a robust MVPA-based bin separation of trials with large vs. medium vs. small immediate Pe as in Experiment 1 was not possible for Experiment 2 because classifier training did not yield a classification accuracy above the 5% significance threshold as established by a permutation test, likely due to the smaller size of the training data set.

Taken together, we could replicate the deferred Pe in Experiment 2. As this analysis was restricted to those error trials that were correctly reported as T1 errors, we can rule out that this deferred Pe originates in occasional T1-error trials that are misattributed as T2-errors.

Experiment 3

While a link between the Pe and conscious error awareness has been established in previous studies (Boldt & Yeung, 2015; Maier, Di Gregorio, Muricchio, & Di Pellegrino, 2015b; Nieuwenhuis et al., 2001; Overbeek et al., 2005; M. Steinhauser & Yeung, 2010), direct empirical evidence for this assumption within the scope of the present paradigm is still missing. Although Experiment 2 included behavioral error reports, a comparison of consciously aware and unaware errors was impossible due to the small number of T1 error trials that were reported as correct. We consequently conducted a variant of Experiment 2 as an additional experiment with a larger number of 40 new subjects (of which 3 subjects had to be excluded due to technical problems during data acquisition), which allowed us to conduct analyses on sub-pools of subjects with a minimal number of 6 trials in all relevant conditions. This was previously suggested as a minimum threshold to obtain reliable measures of error processing in EEG data (Olvet & Hajcak, 2009). To additionally address the question if the objective delay of the Pe in multitasking directly relates to a subjective feeling of later error detection, we added a second question after every trial. Here, participants were asked if they had the subjective feeling of error awareness at an early or late point of time (see also Di Gregorio, Maier, & Steinhauser, 2020). As a result, Experiment 3 provided two variants of behavioral measures for conscious error awareness.

Error reports in Experiment 3 closely resemble those in Experiment 2. 72.34 % of the T1 errors were correctly reported, whereas 24.77 % were not reported at all and only 2.9 % were falsely reported as T2 errors. Of the correctly reported T1 errors, 61.94 % were reported

with a subjective feeling of immediate error awareness and 35.85 % with a feeling of delayed error awareness, whereas only 2.22% were not classified in this regard at all, replicating the pattern found by Di Gregorio et al. (2020) in a single-task scenario. The larger number of participants in Experiment 3 allowed us to create a sub-pool of 21/22 participants for an ERP analysis of reported vs. missed T1 errors (1 subject fell below the threshold of 6 trials only in R1-locked analysis as artefact rejection was conducted separately on the two datasets; see Methods section), and a sub-pool of 25 participants for an ERP analysis of reported T1 errors with a feeling of immediate vs. delayed error awareness.

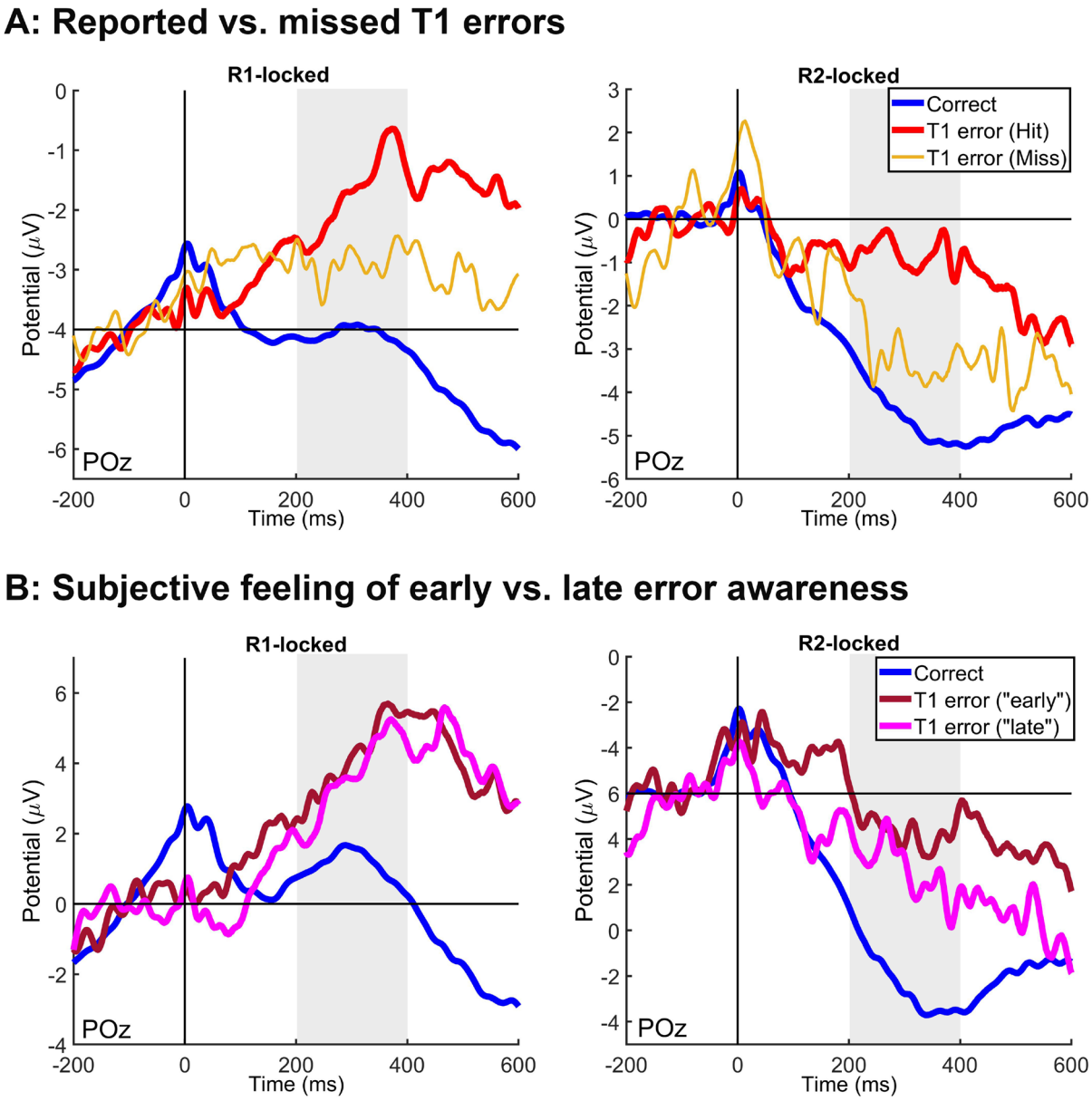

Figure 8. ERP results of the two main ERP analyses of Experiment 3. Error reports (A) and ERP results (B) of Experiment 2. Gray areas indicate the time intervals for statistical testing for the Pe in R1-locked data (left) and R2-locked data (right). R1 = Response 1. R2 = Response 2. T1 = Task 1. T2 = Task 2.

Figure 8 displays a pronounced Pe both in R1-locked data, $F(2,40) = 7.03$, $p = .004$, $\eta_p^2 = .26$, and in R2-locked data, $F(2,42) = 6.86$, $p = .004$, $\eta_p^2 = .25$, once again replicating our initial finding of a deferred Pe after execution of the correct second task. Also a considerable reduction of the Pe in misses vs hits can be seen in both analyses. While in R1-locked data, this difference only crosses the significance threshold in a somewhat later time window of 400-600 ms, $t(20) = 2.64$, $p = .016$, $d = .51$, in R2-locked data, misses and hits exhibit a significant difference also in the a-priori time window of 200-400 ms, $t(21) = 2.23$, $p = .037$, $d = .48$. These findings suggest a direct link between conscious error reports and both the immediate and the deferred Pe.

In contrast to this previous analysis, differentiating T1 error trials with the subjective feeling of immediate vs. deferred error awareness, however, did not yield any significant differences in Pe amplitudes, all $ts(24) < .72$, all $ps(24) > .47$.

And equivalently to Experiment 2, an MVPA-based bin separation was not possible here, either, because classifier training did not yield a classification accuracy beyond the 5.0 % significance threshold. The regression-based approach on trials with IRI > 400ms, however, though not yielding a significant cluster, exhibits a descriptive pattern similar to that in Experiment 1, with a peak in negative beta weights around 325 ms in R1-locked data and 315 ms in R2-locked data, temporally close to the peak amplitudes of the R1-locked and R2-locked Pe.

Discussion

In three ERP experiments, we investigated neural correlates of error monitoring in dual-tasking. Previous studies suggested that increased interference due to temporally overlapping tasks is countered by an active deferment of task components (Hesselmann et al., 2011; Meyer & Kieras, 1997; Sigman & Dehaene, 2008). Based on this idea, we hypothesized that because conscious processing of T1 errors can interfere with the execution of T2 (M. Steinhauser et al. 2017), also performance monitoring, more precisely the conscious evaluation of T1 errors, is adaptively rescheduled to after the completion of the whole dual-task. To this end, we analyzed the Pe, an ERP component that can be linked to conscious error awareness based on previous studies (Boldt & Yeung, 2015; Di Gregorio et al., 2018, 2020; Maier et al., 2015a; Nieuwenhuis et al., 2001; Overbeek et al., 2005; M. Steinhauser & Yeung, 2010) as well as based on an empirical confirmation from behavioral error reports in Experiment 3. In all three experiments, we found a result pattern that confirms our prediction. When the execution of two tasks overlaps, errors in the first task lead to a reduced immediate Pe whereas a deferred Pe appears after the response to the second task.

Our results are in line with the idea that conscious error processing as reflected by the Pe interferes with the execution of a subsequent task (Buzzell et al., 2017; Weißbecker-Klaus et al., 2016). In our study, only the Pe but not the Ne/ERN after T1 errors was reduced when task execution was overlapping. This conforms with recent findings by Weißbecker-Klaus et al. (2016), who found a similar reduction of the Pe when a flanker task was executed concurrently with a semantic task. In both studies, this impairment to immediate conscious error detection likely originates from interference by T2, which requires cognitive resources for T2-response selection at the same time as T1-error processing would occur. The decision process for selecting the correct response to T2 was previously suggested to rely on similar resources as conscious error processing (Hochman & Meiran, 2005). On the neural level, this

can well be explained by recent electrophysiological studies that view also conscious error processing as a decision process (Wessel, 2012; Wessel, Danielmeier, & Ullsperger, 2011) and studies that highlight the physiological similarities of the P300 and the Pe (Leuthold & Sommer, 1999; Overbeek et al., 2005). Given that interference between central decision processes is seen as the main origin of dual-task costs (Pashler 1994; Tombu and Jolicœur, 2003), this explains why particularly the Pe as a neural correlate of conscious error detection (Overbeek et al., 2005) is impaired in dual-tasking.

In addition to a reduced immediate Pe after the erroneous T1 response, we found a deferred Pe after completion of the dual-task even though the second response itself was correct. MVPA-based single-trial estimates of the Pe as well as a regression-based approach on raw data suggest that the deferred Pe occurred predominantly on trials with little immediate conscious error processing. This temporal detachment of conscious error processing from the early error signals that are represented by the Ne/ERN – on average by more than 580 ms – goes beyond a mere serialization as it was previously found for the P300 in response selection (Hesselmann et al., 2011; Marti et al., 2015; Sigman & Dehaene, 2008). In the present study, conscious error detection appears to be suspended until a whole additional task has been executed, which is potentially implemented on the neural level as the result of error evidence accumulation being stored in a buffer because the fronto-parietal network that represents conscious access is still engaged in stimulus processing of Task 2 (Meyer & Kieras, 1997). Analysis of IRIs furthermore confirmed that the deferment of error processing is associated with faster responses to T2, suggesting reduced dual-tasking interference on such trials, in line with Buzzell et al.'s (2017) finding that a smaller Pe is also linked to improved sensory processing in the subsequent trial. This is strong support for the adaptive rescheduling account of error detection in dual-tasking, which suggests such a

mechanism to reduce interference between the subtask representations and hence to maintain the ability to detect and evaluate errors.

Analyzing brain activity in a PRP paradigm bears the risk that activity related to T1 is accidentally attributed to T2 (Sigman & Dehaene, 2008). For instance, what appears to be a deferred Pe after the T2-response could simply be a carry-over of the immediate Pe after the T1-response. However, there are several reasons why this cannot account for the present results. First, the deferred Pe shows a distinct time course, emerging about 200 ms after the T2 response, which is comparable to the latency of the immediate Pe as well as of the Pe in other studies (Overbeek et al., 2005). Second, even with a short SOA of 300 ms, the average interval between the T1-response and T2-response on error trials was 587 ms. That is, the onset of the immediate Pe is clearly prior to the T2-response, and any sustained differences between T1 errors and correct trials emerging prior to the T2-response are controlled by the pre-response baseline. Finally, and most importantly, our results indicate an inverse correlation between immediate Pe and deferred Pe. If the deferred Pe reflected a carry-over of activity from the immediate Pe, both effects should be positively correlated.

In Experiments 2 and 3, we could also rule out that the deferred Pe results from internal misattribution of occasional T1 errors as T2 errors, which would imply a credit assignment problem (Fu & Anderson, 2008; Sutton & Barto, 1998; Walsh & Anderson, 2011) as the origin of our findings. Trial-wise error-reports in Experiments 2 and 3 revealed that credit-assignment errors occurred extremely rarely in our tasks. Moreover, we obtained a deferred Pe even on the subset of correctly assigned T1 errors. This suggests that even when conscious processing of the T1 error is postponed, these errors are still detected and correctly assigned to their corresponding task. One could speculate that the proposed scheduling mechanism even contributes to solving the credit assignment problem. On the one hand, rescheduling counteracts interference between T1 error processing and T2 response selection

which otherwise would increase the risk of confusing error signals from T1 and T2 responses. On the other hand, when conscious error evaluation occurs after both responses were executed, the involved decision process could utilize cues from both tasks to correctly assign error signals (like post-response conflict) to their corresponding tasks.

Experiment 3 was an important addition to this study, because only by being able to analyze a considerably large number of participants with regard to reported vs. missed T1 errors, we could empirically confirm that the Pe represents conscious error awareness also in the present dual-tasking paradigm. Replicating a core finding of previous single-task studies (Boldt & Yeung, 2015; Di Gregorio, Steinhauser, & Maier, 2016; Maier et al., 2015a; Nieuwenhuis et al., 2001; Overbeek et al., 2005; M. Steinhauser & Yeung, 2010), the Pe was reduced immediately after an erroneous T1 response when participants did not report that error. Interestingly, an equivalent reduction was also found on missed T1 errors for the deferred Pe following the T2 response. This substantiates that also this later ERP component is directly linked to conscious error awareness. The analysis with regard to a subjective feeling of early vs. late awareness, on the other hand, did not yield any significant results. The immediate and the deferred Pe showed similar amplitudes in both conditions. This is, however, in line with Di Gregorio et al.'s (2020) argumentation that the subjective feeling of early vs. late error awareness may well be a metacognitive illusion that is created at a later point in time. It is hence plausible that the trial-wise objective delay of the Pe is independent from subjective deferments of error awareness.

Though far less robust than the deferred Pe, we also obtained frontocentral activity related to T1 errors following the T2 response. This could reflect a deferred Ne/ERN, although this activity had a slightly different time course and spatial distribution as the immediate Ne/ERN after T1-responses. However, this deferred Ne/ERN differed in a crucial aspect from the deferred Pe as it was not associated with a reduced immediate Ne/ERN.

Without this trade-off between immediate and deferred error processing, this effect cannot reflect a scheduling mechanism. It is also implausible that the deferred Ne/ERN reflects post-response conflict induced by a corrective response tendency (Yeung et al., 2004) or a reward prediction error (Holroyd & Coles, 2002) as both these mechanisms are temporally linked to the erroneous response. Instead, the deferred Ne/ERN could represent a negative affective signal to the T2-response. Several studies could show that the Ne/ERN involves affective processing of an error (Aarts, Houwer, & Pourtois, 2013; Maier, Scarpazza, Starita, Filogamo, & Ladavas, 2016). The deferred Ne/ERN hence could reflect that committing an error in one task leads to a devaluation also of the second task, indicating that the whole dual-task trial is processed as erroneous.

Taken together, our results suggest that reliable error detection in dual-tasking is maintained by a mechanism that adaptively reschedules the accumulation of evidence underlying conscious error processing to after the completion of the whole dual-task. Our findings indicate that the flexible reorganization of componential processes under dual-tasking is not only a viable strategy to prevent interference between decision processes involved in task execution (Hesselmann et al., 2011; Marti et al., 2015; Meyer & Kieras, 1997; Sigman & Dehaene, 2008), but also serves to reduce interference between task execution and performance monitoring.

Materials and methods

Participants

We conducted three separate, consecutive experiments with different groups of participants. 24 healthy students (2 male; age: M = 22.3 years; SD = 4.5 years; 2 left-handed)

participated in Experiment 1, 24 healthy students (2 male; age: M = 24.0 years; SD = 3.9; 2 left-handed) participated in Experiment 2 and 40 healthy students participated in Experiment 3 (4 male; age: M = 21.7 years; SD = 2.5; 2 left-handed). With an assumed correlation of .5 among cell means, a sample size of N = 24 should allow for detecting medium-sized effects (f = .25) in the present repeated-measures design with a statistical power of .82. All participants were recruited from the Catholic University of Eichstätt-Ingolstadt and received course credit or payment (8€ per hour). Informed consent was provided by all participants and the study was approved by the ethics committee of the Catholic University of Eichstätt-Ingolstadt. One subject in Experiment 2 and 3 subjects in Experiment 3 had to be excluded from further analysis due to technical problems during data acquisition.

Task and procedure

Experiment 1. Adopting the PRP paradigm from Steinhauser, Ernst, and Ibald (2017), we combined an error-prone three-choice color flanker task and a two-choice pitch discrimination task. The flanker task stimulus (see Fig. 1) consisted of three horizontally arranged squares of .82° edge length. The central target square and the two flanker squares were either red, yellow or blue. While both flanker squares had the same color, this color was always different from that of the target. As three colors were used, it was possible to exclusively present these more error-prone incongruent stimuli without enabling participants to infer the target color from the flankers. Stimuli for the secondary pitch task were low (400 Hz) and high (900 Hz) sine tones.

The trial procedure is depicted in Figure 1. Participants first had to respond to the flanker task (T1) and then to the pitch discrimination task (T2). Each trial started with the presentation of a fixation cross for 500 ms. Following this, the flanker task was presented by initially displaying the two flankers for 60 ms and then the target together with the flankers

for 200 ms. The pitch discrimination task started after an SOA of either 300 ms (short) or 1200 ms (long), which was chosen randomly within the blocks. The stimulus for the pitch task was presented for 150 ms. After responses to both tasks were given, the fixation cross of the next trial was presented 500 ms after the response. The short SOA was supposed to create a dual-tasking situation with overlapping task execution while at the same time preventing excessive response grouping, which was observed in a piloting phase with even shorter SOAs. In contrast, the long SOA was intended to serve as a baseline that allows for serial task execution.

For the flanker task, participants had to indicate the color of the central target square by pressing the “Y”, “X”, or “C” button on a German QWERTZ keyboard (in which Y and Z are exchanged) with their left hand (the T1-response). For the pitch discrimination task, participants entered their response with “arrow down” for a low pitch and “arrow up” for a high pitch (the T2-response). The mapping of colors to keys was counterbalanced across participants, and for one half of the participants, the task-to-hand assignment was reversed (resulting in “,”, “.”, “-” for the flanker task and “A” and “Y” for the pitch task). If the T2-response was given before the onset of the pitch stimulus, written feedback (“AUFGABE 2 ZU FRÜH!”, *engl*. task 2 too early) was given immediately and the trial was excluded from data analysis.

The experiment started with a series of practice blocks that ensured that participants had thoroughly learned the experimental paradigm and the mappings of colors/pitches to keys. First, in three blocks of 24 trials the flanker task was presented alone. Then, one block of 16 trials served to practice the pitch task alone, and in one subsequent block of 36 trials, both tasks were responded to together. Actual testing consisted of 10 blocks of 108 trials each, resulting in a total number of 1080 trials. Oral feedback to respond faster was given after blocks, if the error rate for any of the tasks was below 10%.

Experiment 2. Experiment 2 adopted the experimental method of Experiment 1 but added a trial-wise report of errors. To this end, response keys to the two tasks were changed ("A", "S", "D" and "L", "P" for one half of participants, "L", "Ö", "Ä" and "A", "Q" for the other half) and two additional keys, "alt" and "alt gr", served to report errors. In each trial, 600 ms after the response to Task 2, an additional instruction ("FEHLER?", *engl*. error?) appeared for 1000 ms in the center of the screen. During this time, participants should indicate if they had just committed an error in the task of their left hand ("ALT") or their right hand ("ALT GR"), or both. No response was required if the participants considered both their responses correct. To account for the increased trial duration, Experiment 2 was restricted to only the short SOA of 300 ms. With 10 blocks of 84 trials each, this resulted in 840 trials for the condition with overlapping task execution.

Experiment 3. Experiment 3 was a replication of Experiment 2 with the following two changes. First, a third response possibility was included in the error report. By pressing the "SPACE" key with one of their thumbs, participants should also indicate when no error had occurred. This was added to reduce the effect of the frontal Readiness Potential that we observed in Experiment 2 after the response only on error trials because only there, a key press would soon have to be made to report the error (Brunia, van Boxtel, & Böcker, 2012). 500 ms after the error report, a second instruction ("FRÜH ODER SPÄT?", *engl.* early or late?) required participants to decide whether they had a subjective feeling of "early" or "late" conscious error awareness. Following the basic rationale of Di Gregorio et al. (2020), participants were told in the initial instruction of the experiment that some errors are detected already at or before the time of the response button press, whereas others are detected later. Participants hence should distinguish between such early and late detected errors. Responses to this second question were given with the same keys as for the first one (left hand: "early"; right hand: "late"; space key: "I don't know").

Data acquisition

The EEG was recorded from 64 electrodes at a sample rate of 512 Hz 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 as well as the left and right mastoid). The CMS (Common Mode Sense) and DRL (Driven Right Leg) electrodes were used as reference and ground electrodes. Vertical and horizontal electrooculogram (EOG) was recorded from electrodes above and below the right eye and on the outer canthi of both eyes. All electrodes were off-line re-referenced to averaged mastoids.

Experimental design and statistical analysis

Design. Each trial was assigned to a condition based on the SOA (short, long) and T1 Correctness (correct, error). T1 Correctness relied on the post-hoc classification of whether the response in the flanker task (T1) was correct or not. Trials on which an error in the pitch discrimination task (T2) occurred were removed from all analyses of response times (RTs) and ERP data.

Data analysis. For the analysis of RTs, trials were excluded whose RT deviated more than three standard deviations from the RT mean of each condition and participant. Error rates were arcsine-transformed prior to statistical testing (Winer, Brown, & Michels, 1991). All analyses were performed using custom MATLAB v8.2 (The Mathworks, Natick, MA) scripts and EEGLAB v12.0 (Delorme & Makeig, 2004) functions. All data and analysis scripts are publicly available in an online repository hosted by the Open Science Framework (https://osf.io/5ub8z/). For Experiment 1, continuous EEG data was initially band-pass

filtered to exclude frequencies below 0.1 Hz and above 40 Hz. Then, epochs were created for two separate analyses, first from -500 ms before to 1000 ms after the T1-response (T1-response-locked dataset) and, second, from -500 ms before to 1000 ms after the T2-response (T2-response-locked dataset). Epochs in both analyses were baseline-corrected by subtracting mean activity between -150 ms and -50 ms before the response, as neural correlates of performance monitoring were previously found to emerge slightly before the response button press. Separately for each dataset, electrodes were interpolated using spherical spline interpolation if they met the joint probability criterion (threshold 5) as well as the kurtosis criterion (threshold 5) in EEGLAB's channel rejection routine (pop_rejchan.m). Epochs were removed that contained activity exceeding +/-300 µV in any channel except AF1, Fp1, Fpz, Fp2, AF8 (to prevent exclusion of blink artifacts, which were corrected at a later stage) and whose joint probability deviated more than 5 standard deviations from the epoch mean. To correct for eye blinks and muscular artefacts, an infomax-based ICA (Bell & Sejnowski, 1995) was computed and components with time courses and topographies typical for these artefacts were removed after visual inspection.

In both experiments and for each dataset (T1-response-locked and T2-response-locked), following a considerable number of previous studies that feature a Pe peaking over parieto-occipital electrodes (Beatty et al., 2018; Endrass, Reuter, & Kathmann, 2007; Shalgi, Barkan, & Deouell, 2009; M. Steinhauser & Yeung, 2010, 2012), the Pe was quantified by comparing mean amplitudes from 200 ms to 400 ms after the respective response at electrode POz between T1 error trials and correct trials. The Ne/ERN was quantified by comparing mean amplitudes from 0 ms to 100 ms after the respective response at electrode FCz between T1-error trials and correct trials. In Experiment 2, these analyses were restricted to T1-error trials that were reported as T1 errors, and correct trials that were reported as correct trials.

As we found that T1 errors elicited both an immediate Pe in T1-response-locked data and a deferred Pe in T2-response-locked data, we aimed to investigate the relationship between both types of Pe on a single-trial level. Our approach was to categorize trials according to the size of the immediate Pe and then analyze differences between these trials with respect to the deferred Pe. To acquire a robust single-trial estimate of the T1-response-locked Pe, we used a multivariate pattern analysis (MVPA) based on the linear integration method introduced by Parra et al. (2002, 2005). This method has previously been used to quantify error-related brain activity on a single-trial level (Boldt & Yeung, 2015; M. Steinhauser & Yeung, 2010). Here, we provide only a brief description of this method while details can be found elsewhere (e.g., Steinhauser & Yeung 2010). In a first step, we computed a set of classifiers on T1-response-locked data that discriminated optimally between correct trials and T1 errors. Classifiers were constructed for consecutive, partially overlapping time windows from 0 ms to 400 ms after the T1-response (width 50 ms, step size 10 ms). All classifiers were trained, separately for each participant, on T1 error trials and the same number of randomly drawn correct trials. Then, the classifier was selected that featured the highest discrimination sensitivity, as indicated by the Az score. To prevent overfitting, Az was computed using leave-one-out cross-validation. Using this classifier, we calculated prediction values for each error trial. These prediction values represent single-trial estimates of error-related brain activity in the respective classifier window. Based on these estimates, we assigned each error trial to one of three equally sized bins: small Pe, medium Pe, and large Pe in T1-response-locked data. By analyzing Pe amplitudes in T2-response-locked data according to these bins, we were able to investigate the relationship between immediate and deferred Pe elicited by T1 errors.

As a second way of investigating a possible trade-off between the R1-locked Pe and the R2-locked Pe, we computed an analysis based on linear regression on raw data. To get a

sufficiently robust measure of the Pe in single trials, we computed mean amplitudes of trial-wise EEG data in a 100 ms window on a cluster of posterior electrodes (Pz, P1, P2, POz, P03, P04). This was done in steps of 10 ms on time windows centering from 0 ms to 400 ms after both responses. For each such combination of single-trial Pe values in R1-locked data and R2-locked data, the following linear regression model was fitted on the data of every participant:

$$EEG_{R2} = \beta_0 + \beta_{EEG_{R1}} * EEG_{R1} + \beta_{IRI} * IRI$$

where EEG_{R2} represents z-scored posterior response 2-locked EEG activity, EEG_{R1} represents z-scored posterior response 1-locked EEG activity, IRI represents the z-scored inter-response interval of the current trial. All z-scores were computed within-subject (i.e., using subject-specific across-trial mean and standard deviations) and β_i represent within-subject regression coefficients. Participants' regression coefficients were subsequently standardized by their SDs and tested against zero by means of Student's t-tests. Due to the large number of the resulting t-tests, we corrected for multiple comparisons by means of a cluster-based permutation test with 100.000 permutations, a cluster inclusion threshold of $p = .01$ and an output threshold of $p = .05$, utilizing the Mass Univariate ERP Toolbox (Groppe, Urbach, & Kutas, 2011).

Acknowledgements

This work was supported by a grant within the Priority Program, SPP 1772 from the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG, grant number STE 1708/4-1). The open access publication of this article was supported by the Open Access Fund of the Catholic University Eichstätt-Ingolstadt.

References

Aarts, K., Houwer, J. De, & Pourtois, G. (2013). Erroneous and correct actions have a different affective valence: Evidence from ERPs. *Emotion*, *13*, 960–973.

Beatty, P. J., Buzzell, G. A., Roberts, D. M., & McDonald, C. G. (2018). Speeded response errors and the error-related negativity modulate early sensory processing. *NeuroImage*, *183*, 112–120.

Bell, A. J., & Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, *7*, 1129–1159.

Boldt, A., & Yeung, N. (2015). Shared neural markers of decision confidence and error detection. *Journal of Neuroscience*, *35*, 3478–3484.

Brunia, C. H. M., van Boxtel, G. J. M., & Böcker, K. B. E. (2012). *Negative slow waves as indices of anticipation: the Bereitschaftspotential, the Contingent Negative Variation, and the Stimulus-Preceding Negativity*. *The Oxford Handbook of Event-Related Potential Components*.

Buzzell, G. A., Beatty, P., Paquette, N., Roberts, D., & McDonald, C. (2017). Error-induced blindness: Error detection leads to impaired sensory processing and lower accuracy at short response-stimulus intervals. *Journal of Neuroscience*.

Cousineau, D. (2005). Confidence intervals in within-subject designs: a simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, *1*, 42–45.

Dehaene, S., Kerszberg, M., & Changeux, J.-P. (1998). A neuronal model of a global workspace in effortful cognitive tasks. *Proceedings of the National Academy of Sciences*, *95*, 14529–14534.

Del Cul, A., Baillet, S., & Dehaene, S. (2007). Brain dynamics underlying the nonlinear threshold for access to consciousness. *PLoS Biology*, *5*, 2408–2423.

Dell'Acqua, R., Jolicoeur, P., Vespignani, F., & Toffanin, P. (2005). Central processing overlap modulates P3 latency. *Experimental Brain Research*, *165*, 54–68.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*, 9–21.

Di Gregorio, F., Maier, M. E., & Steinhauser, M. (2018). Errors can elicit an error positivity in the absence of an error negativity: Evidence for independent systems of human error monitoring. *NeuroImage*.

Di Gregorio, F., Maier, M. E., & Steinhauser, M. (2020). Are errors detected before they occur? Early error sensations revealed by metacognitive judgments on the timing of error awareness. *Consciousness and Cognition*, *77*, 102857.

Di Gregorio, F., Steinhauser, M., & Maier, M. E. (2016). Error-related brain activity and error awareness in an error classification paradigm. *NeuroImage*, *139*, 202–210.

Endrass, T., Reuter, B., & Kathmann, N. (2007). ERP correlates of conscious error recognition: Aware and unaware errors in an antisaccade task. *European Journal of Neuroscience*, *26*, 1714–1720.

Falkenstein, M., Hohnsbein, J., Hoormann, J., & Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. *Electroencephalography and Clinical Neurophysiology*, *78*, 447–455.

Fu, W. T., & Anderson, J. R. (2008). Solving the credit assignment problem: Explicit and implicit learning of action sequences with probabilistic outcomes. *Psychological Research*, *72*, 321–330.

Gehring, W. J., Goss, B., & Coles, M. G. H. (1993). A neural system for error detection and compensation. *Psychological Science*, *4*, 385–390.

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass Univariate Analysis of Event-

Related Brain Potentials/ Fields II: Simulation Studies. *Psychophysiology*, *48*, 1726–1737.

Gross, J., Schmitz, F., Schnitzler, I., Kessler, K., Shapiro, K., Hommel, B., & Schnitzler, A. (2004). Modulation of long-range neural synchrony reflects temporal limitations of visual attention in humans. *Proceedings of the National Academy of Sciences*, *101*, 13050–13055. Retrieved from papers3://publication/uuid/D192C02A-B83A-4F02-AEA5-BC0E51C10D82

Hesselmann, G., Flandin, G., & Dehaene, S. (2011). NeuroImage Probing the cortical network underlying the psychological refractory period : A combined EEG – fMRI study. *NeuroImage*, *56*, 1608–1621.

Hester, R., Foxe, J. J., Molholm, S., Shpaner, M., & Garavan, H. (2005). Neural mechanisms involved in error processing: A comparison of errors made with and without awareness. *NeuroImage*, *27*, 602–608.

Hochman, E. Y., & Meiran, N. (2005). Central interference in error processing. *Memory & Cognition*, *33*, 635–643.

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*, 679–709.

Hommel, B. (1998). Automatic Stimulus-Response Translation in Dual-Task Performance. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 1368–1384.

Jentzsch, I., & Dudschig, C. (2009). Why do we slow down after an error? Mechanisms underlying the effects of posterror slowing. *The Quarterly Journal of Experimental Psychology*, *62*, 209–218.

Jersild, A. T. (1927). Mental set and shift. *Archives of Psychology*, *14*, 81.

Klawohn, J., Endrass, T., Preuss, J., Riesel, A., & Kathmann, N. (2016). Modulation of hyperactive error signals in obsessive–compulsive disorder by dual-task demands. *Journal of Abnormal Psychology*, *125*, 292–298.

Leuthold, H., & Sommer, W. (1999). ERP correlates of error processing in spatial S-R compatibility tasks. *Clinical Neurophysiology*, *110*, 342–357.

Maier, M. E., Di Gregorio, F., Muricchio, T., & Di Pellegrino, G. (2015a). Impaired rapid error monitoring but intact error signaling following rostral anterior cingulate cortex lesions in humans. *Frontiers in Human Neuroscience*, *9*, 1–15.

Maier, M. E., Di Gregorio, F., Muricchio, T., & Di Pellegrino, G. (2015b). Impaired rapid error monitoring but intact error signaling following rostral anterior cingulate cortex lesions in humans. *Frontiers in Human Neuroscience*, *9*, 339.

Maier, M. E., Scarpazza, C., Starita, F., Filogamo, R., & Ladavas, E. (2016). Error monitoring is related to processing internal affective states. *Cognitive, Affective and Behavioral Neuroscience*, *16*, 1050–1062.

Maier, M. E., & Steinhauser, M. (2017). Working memory load impairs the evaluation of behavioral errors in the medial frontal cortex. *Psychophysiology*, *54*, 1472–1482.

Marti, S., King, J.-R., & Dehaene, S. (2015). Time-Resolved Decoding of Two Processing Chains during Dual-Task Interference. *Neuron*.

Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part 1. Basic mechanisms. *Psychological Review*, *104*, 3.

Morey, R. D. (2008). Confidence Intervals from Normalized Data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, *4*, 61–64.

Moser, J. S., Moran, T. P., Schroder, H. S., Donnellan, M. B., & Yeung, N. (2013). On the relationship between anxiety and error monitoring: A meta-analysis and conceptual

framework. *Frontiers in Human Neuroscience*, *7*, 466.

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

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-à-vis the ERN/Ne. *Journal of Psychophysiology*, *19*, 319–329.

Parra, L. C., Alvino, C., Tang, A., Pearlmutter, B., Yeung, N., Osman, A., & Sajda, P. (2002). Linear Spatial Integration for Single-Trial Detection in Encephalography. *NeuroImage*, *17*, 223–230.

Parra, L. C., Spence, C. D., Gerson, A. D., & Sajda, P. (2005). Recipes for the linear analysis of EEG. *NeuroImage*, *28*, 326–341.

Pashler, H. (1994). Dual-task interference in simple tasks: data and theory. *Psychological Bulletin*, *116*, 220–244.

Pashler, H., & Johnston, J. C. (1989). Chronometric evidence for central postponement in temporally overlapping tasks. *The Quarterly Journal of Experimental Psychology Section A*, *41*, 19–45.

Rabbitt, P. (2002). Consciousness is slower than you think Consciousness is slower than you think. *The Quaterly Journal of Experimental Psychology: Section A*, *55*, 1081–1092.

Sergent, C., Baillet, S., & Dehaene, S. (2005). Timing of the brain events underlying access to consciousness during the attentional blink. *Nature Neuroscience*, *8*, 1391–1400.

Sergent, C., & Dehaene, S. (2004). Neural processes underlying conscious perception: Experimental findings and a global neuronal workspace framework. *Journal of*

Physiology Paris, *98*, 374–384.

Shalgi, S., Barkan, I., & Deouell, L. Y. (2009). On the positive side of error processing: Error-awareness positivity revisited. *European Journal of Neuroscience*, *29*, 1522–1532.

Sigman, M., & Dehaene, S. (2008). Brain mechanisms of serial and parallel processing during dual-task performance. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *28*, 7585–7598.

Steinhauser, M., Ernst, B., & Ibald, K. W. (2017). Isolating component processes of posterror slowing with the psychological refractory period paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*, 653–659.

Steinhauser, M., & Yeung, N. Decision processes in human performance monitoring, 30 Journal of Neuroscience § (2010).

Steinhauser, M., & Yeung, N. (2012). Error awareness as evidence accumulation: effects of speed-accuracy trade-off on error signaling. *Frontiers in Human Neuroscience*, *6*, 1–12.

Steinhauser, R., & Steinhauser, M. (2018). Preparatory brain activity in dual-tasking. *Neuropsychologia*, *114*, 32–40.

Sutton, R. S., & Barto, a G. (1998). Reinforcement learning: an introduction. *IEEE Transactions on Neural Networks / a Publication of the IEEE Neural Networks Council*, *9*, 1054.

Telford, C. W. (1931). The refractory phase of voluntary and associative responses. *Journal of Experimental Psychology*, *14*, 1–36.

Tombu, M., & Jolicœur, P. (2003). A central capacity sharing model of dual-task performance. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 3–18.

Ullsperger, M., Danielmeier, C., & Jocham, G. (2014). NEUROPHYSIOLOGY OF PERFORMANCE, 35–79.

Ullsperger, M., Fischer, A. G., Nigbur, R., & Endrass, T. (2014). Neural mechanisms and temporal dynamics of performance monitoring. *Trends in Cognitive Sciences*, *18*, 259–267.

Walsh, M. M., & Anderson, J. R. (2011). Modulation of the feedback-related negativity by instruction and experience. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 19048–19053.

Weißbecker-Klaus, X., Ullsperger, P., Freude, G., & Schapkin, S. A. (2016). Impaired Error Processing and Semantic Processing During Multitasking. *Journal of Psychophysiology*, 1–12.

Welford, A. T. (1952). The "psychological refractory period" and the timing of high-speed performance - a review and a theory. *The British Journal of Psychology General Section*, *43*, 2–19.

Wessel, J. R. (2012). Error awareness and the error-related negativity: evaluating the first decade of evidence, 1–16.

Wessel, J. R., Danielmeier, C., & Ullsperger, M. (2011). Error awareness revisited: accumulation of multimodal evidence from central and autonomic nervous systems. *Journal of Cognitive Neuroscience*, *23*, 3021–3036.

Winer, B. J., Brown, D. R., & Michels, K. M. (1991). *Statistical Principles in Experimental Design* (3rd ed.). McGraw-Hill.

Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The Neural Basis of Error Detection: Conflict Monitoring and The Error Related Nagativity. *Psychological Review*, *11*, 931–959.