Are errors detected before they occur? Early error sensations revealed by metacognitive judgments on the timing of error awareness

Published Version:
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, 1-14 [10.1016/j.concog.2019.102857].

Availability:
This version is available at: https://hdl.handle.net/11585/950421 since: 2023-12-11

Published:
DOI: http://doi.org/10.1016/j.concog.2019.102857

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 *Conscious. Cogn., 77,* Article 102857, https://doi.org/10.1016/j.concog.2019.102857

Running head: Early error sensations

Are errors detected before they occur? Early error sensations revealed by metacognitive judgments on the timing of error awareness

Francesco Di Gregorio[1,2], Martin E. Maier[1] & Marco Steinhauser[1]

[1]Catholic University of Eichstätt-Ingolstadt, Germany

[2]Casa Dei Risvegli Luca De Nigris - Centro Studi per la Ricerca sul Coma, Bologna, Italy

Correspondence should be addressed to:

Francesco Di Gregorio

Catholic University of Eichstätt-Ingolstadt

Ostenstraße 27

85072 Eichstätt

Germany

E-mail: francesco.digregorio@ku.de

Phone: +int (0)8421-93-21364

Keywords: error awareness; error detection; metacognition.

Introduction

Research on human performance monitoring has shown that errors in choice tasks are detected fast and reliably. For instance, if participants have to classify stimuli under high time pressure, occasional errors are almost always recognized and can be corrected within a few hundreds of milliseconds (Rabbitt, 1968, 2002). Frequently, participants in these experiments report that they sometimes became aware of their errors even slightly before the erroneous response was executed. Similar observations can be made in everyday behavior. For example, when writing an email in a hurry, we sometimes have the feeling that an error is about to occur even before an actual typo is made. These anecdotal reports of early error sensations, that is, subjective feelings of early error detection, stand in stark contrast to findings from research on the neural basis of error detection. Here, the predominant view is that correlates of conscious error detection emerge not until several hundreds of milliseconds after the erroneous response (Nieuwenhuis, Ridderinkhof, Blom, Band, & Kok, 2001; Steinhauser & Yeung, 2010). The present study aimed to test whether anecdotal evidence of early error sensations can be corroborated in a controlled study in which participants were asked to indicate whenever this early error detection has occurred. Robust evidence for early error sensations would pose strong constraints on theoretical accounts of the emergence of conscious error awareness.

The time course of error detection has frequently been investigated using event-related potentials (Gehring, Goss, Coles, Meyer, & Donchin, 2018; Ullsperger, Fischer, Nigbur, & Endrass, 2014). Corresponding studies revealed that errors in choice tasks elicit a cascade of error-related brain activity that starts almost immediately after the erroneous response. Within 50 to 100 ms after the error, a so-called error-related negativity (Ne or ERN; Falkenstein, Hohnsbein, Hoormann, & Blanke, 1991; Gehring, Goss, Coles, Meyer, & Donchin, 1993) is observed, which is supposed to reflect the rapid detection of a mismatch, conflict, or prediction error indicating a discrepancy between the correct and the executed response (Falkenstein, Hoormann, Christ, & Hohnsbein, 2000; Holroyd & Coles, 2002; Yeung, Botvinick, & Cohen, 2004). This Ne/ERN is followed by an error positivity (Pe; Falkenstein et al., 1991; Overbeek, Nieuwenhuis, & Ridderinkhof, 2005) which emerges about 300 ms after the error.

Several studies investigated how the Ne/ERN and Pe are involved in the emergence of conscious error detection by asking participants to press a key whenever an error has occurred or to rate the subjective confidence that they have made an error on each trial. Most frequently, it has been found that the late Pe but not the early Ne/ERN is predictive of conscious error awareness or confidence (Boldt & Yeung, 2015; Nieuwenhuis et al., 2001; Overbeek et al., 2005; Steinhauser & Yeung, 2010). Although the Ne/ERN was larger for detected errors than for undetected errors in some studies (for a review, see Wessel, 2012), it has been argued that this reflects differences in intrinsic features of task performance, such as response conflict, between detected and undetected errors (Steinhauser & Yeung, 2010). This receives support from the finding that the Ne/ERN can be larger for errors associated with a lower degree of conscious awareness (Di Gregorio, Steinhauser, & Maier, 2016), and that the Pe and error awareness can still occur under conditions where the Ne/ERN is impaired or even absent (Di Gregorio, Maier, & Steinhauser, 2018; Maier, Di Gregorio, Muricchio, & di Pellegrino, 2015). These results strongly suggest that the later Pe rather than the earlier Ne/ERN is the neurophysiological correlate of conscious error detection, and that conscious error detection emerges not until several hundreds of milliseconds after an error. This conclusion receives further support from the observation that the ability to indicate own errors is already impaired when the interval between a response and a subsequent stimulus is shorter than 800 ms (Rabbitt, 2002).

Based on these considerations, it becomes clear that there is a discrepancy between current findings on the time course of error awareness and the anecdotal reports of early error sensations in experiments. Although conscious error detection emerges several hundred milliseconds after the error, participants in experimental studies frequently report that they already knew that a response would be an error before they actually executed it. This apparent contradiction can be explained in at least two ways: First, it is possible that the Pe is not the neural correlate of error awareness as frequently assumed, but that error awareness emerges much earlier. Second, early error sensations could be a metacognitive illusion that serves to temporally synchronize metacognitive content (error awareness) with objective events (the erroneous response). Indeed, similar mechanisms have been proposed in the field of visual awareness. The conscious perception of a visual stimulus has been suggested to emerge at around 300 ms after presentation of this stimulus (Dehaene & Naccache, 2001; Sergent, Baillet, & Dehaene, 2005). The fact that we subjectively attribute the emergence of visual awareness to the onset of the stimulus has been explained by a backward referral process that aims to create a coherent perception of the objective world in our stream of consciousness (Libet, Gleason, Wright, & Pearl, 1983; Libet, Wright, Feinstein, & Pearl, 1979). The present study did not aim to distinguish between these alternatives. Rather, our goal was to set the stage for further research by providing first experimental evidence for the existence of early error sensations.

The present study

In this study, we investigated whether early error sensations exist and whether participants are able to reliably report it. To this end, we conducted four experiments using two different experimental approaches. In each experiment, participants had to perform a *primary task* in which errors could occur. Then, a metacognitive *secondary task* was applied, in which a) participants had to indicate whether an error in the primary task has occurred and whether it was detected before or after execution of the erroneous response (Experiments 1a and 1b), or b) participants had to rate the subjective confidence that such an early error sensation has occurred (Experiment 2a and 2b). The secondary task essentially required participants to solve a signal detection task. That is, they had to detect a signal (the early error sensation) among noise. The standard framework for analyzing tasks like this is signal detection theory (Green & Swets, 1966) or its applications to metacognitive judgments (Maniscalco & Lau, 2012). Within all these frameworks, the ability to detect a signal is quantified based on the rates of correctly and falsely detected signals (i.e., hits and false alarms). However, the particular challenge in the present case is that we cannot discriminate between hits and false alarms, as we do not know for which errors early error sensations are present or whether early error sensations exist at all. It is possible that participants simply report early error sensations because they are instructed to do so, thus forming expectations about its existence or its frequency. In other words, the data could reflect an instruction bias or an expectation bias rather than real signal detection.

To deal with this problem, we adopted several experimental strategies: First, we measured early error sensations using different tasks and methods. In Experiments 1a and 1b, the primary task was a flanker task (Eriksen & Eriksen, 1974) and the secondary task required participants to simply indicate whether an early error sensation had occurred or not. In Experiment 2a and 2b, the primary task was a letter/number discrimination task and the secondary task employed post-decision wagering (Persaud, McLeod, & Cowey, 2007), a method that has been proposed as an effective measure of visual awareness (Persaud et al., 2007) or metacognitive content (Seth, 2008). Finding similar rates of early error sensations across different primary and secondary tasks would speak for the robustness of this phenomenon. Second, we aimed to improve metacognitive judgments on early error sensations by introducing a second task that served as a reference point for judgments on early error sensations. Although post-decision wagering has been assumed to generally improve judgments about the contents of consciousness (Persaud et al., 2007; Seth, 2008), participants may not have an objective criterion to judge whether early error sensations have actually occurred. Thus, we sought to provide a reference point to guide participants' metacognitive judgments. Participants initially performed a task involving metacognitive judgments on their response accuracy in a Visual Awareness task. In the subsequent main task, they were instructed that their reports of early error sensations should be based on the same level of confidence as their previous judgments in the Visual Awareness task. Moreover, we directly investigated the impact of expectations on reports of early error sensations. We varied the difficulty of the initial Visual Awareness task across Experiments 2a and 2b, and asked whether the pattern of metacognitive judgments in this task is carried over to the pattern of reports of early error detection in the main task, thus indicating an expectation bias.

Experiment 1

The goal of the first experiment was to study whether and how frequently participants report early error sensations in a flanker task – one of the most frequently used tasks in studies on error detection. On each trial, participants first had to perform a flanker task in which they responded to the direction of a central target arrow (left/right) while ignoring congruent or incongruent distractor arrows. Here, errors mainly occur on incongruent trials. Then, participants had to classify their responses according to whether they were correct, or whether an early error (detected before response execution) or a late error (detected after response execution) occurred. Because participants in Experiment 1a frequently reported that they were unsure about when an error was detected, we conducted Experiment 1b in which they could additionally indicate that they did not know whether error detection occurred early or late.

We were primarily interested in whether a substantial amount of error trials show early error sensations. If early error sensations reflected a metacognitive illusion resulting from the synchronization of metacognitive content with objective motor events, the frequency of this phenomenon should be rather high. In addition, we also compared early and late errors with respect to their response time (RT) as this could provide valuable information about the source of early error sensations. If early error sensations reflected the described metacognitive illusion, there is no reason to assume that early and late errors differ with respect to RT. If, however, early error sensations represented the true time course of error detection relative to the response, we might observe a relation between RT and the subjective timing of error awareness, which could emerge for different reasons: First, early errors could occur mainly on trials with slow RTs, i.e., when the motor command was delayed and sent after error detection has already occurred (e.g., based on the preceding decision process). Second, early errors could occur mainly on trials with fast RTs, as studies on the latency of error detection have found that lower response criteria (i.e., less cautious responding) lead to both shorter RTs and shorter latencies of error detection (Steinhauser, Maier, & Hübner, 2008).

Method

Participants. 23 participants (4 male) between 19 and 29 years of age (M = 21.8, SE = 0.6) participated in Experiment 1a. A new group of 20 participants (6 male) between 19 and 28 years of age (M = 22.1, SE = 0.58) participated in Experiment 1b. All participants had normal or corrected to normal vision, were recruited from the student population at the Catholic University of Eichstätt-Ingolstadt, and received course credit or 8 Euro per hour for participation. The study was approved by the ethical committee of the Catholic University of Eichstätt-Ingolstadt, and informed consent was obtained from all participants.

Apparatus. A PC running presentation software (Neurobehavioral Systems, Albany, CA) controlled stimulus presentation and response registration. Stimuli were presented on a 21-inch color monitor (60 Hz refresh rate) at a viewing distance of 70 cm.

Task and procedure. Both experiments consisted of a primary task in which participants performed a flanker task and a secondary task in which participants classified their response on the primary task. Stimuli of the flanker task were strings of five arrow heads (e.g., < < > < <) in Arial font, subtending a visual angle of 4.1° horizontally and 1.4° vertically. The central arrow head in each string was designated as the target and the lateral arrows were designated as the flankers. Flankers could have either the same direction as the target (*congruent* condition) or the opposite direction (*incongruent* condition).

Each trial started with the presentation of a fixation cross for 350 ms. Then, the stimulus of the flanker task was presented for 200 ms followed by a black screen until a response was given. Participants had to identify the direction of the target by pressing the “left arrow” key or the “right arrow” key on a standard keyboard with the index or the middle finger of the right hand (*primary task response*). After the response, another black screen was presented for 1000 ms. Then, a question mark appeared in the screen center to prompt participants to classify their response in the primary flanker task. To this end, participants had to execute one of three classification responses on the same keyboard with the left hand (*secondary task response*). Participants indicated whether a) they felt they had responded correctly (“A” key with ring finger), b) they felt they had committed an error (i.e., they pressed the wrong button on the primary task), and this feeling has emerged already before actual response execution (early error; “S” key with middle finger), or c) they had committed an error without this early error sensation (late error; “D” key with index finger). In Experiment 1b, a fourth response alternative could be provided to indicate if they detected an error but did not know whether this was detected before or after the error was committed (“F” key, also with index finger).

Both experiments consisted of eight test blocks with 64 trials per block. Each block contained 16 instances of each of the four possible flanker stimuli in randomized order. Prior to the test blocks, participants performed two practice blocks (32 trials each) without a secondary response to practice the primary task, and two further practice blocks (32 trials each) in which the secondary task response was introduced. In all practice blocks, whenever the error rate in the preceding block was below 25%, participants were instructed to respond faster on the primary task prior to the next block. Before the secondary task was introduced, participants were instructed about early error sensations. We told participants that errors are sometimes accompanied by the sensation that they already knew that they commit an error before the incorrect button was pressed, and that this is called “early error sensation”. In other words, participants were instructed not only to report errors, but also to focus on the timing of error detection. Similar instructions were used to introduce early error sensation in all experiments.

Data analysis. Trials were classified according to stimulus congruency (congruent, incongruent), primary task response (correct, error), and secondary task response (correct, early error, late error – in both experiments - and I don’t know error in Exp. 1b). RT in the primary task was defined as the time interval between the onset of the stimulus and the subsequent button press. To control for outliers, trials were excluded whenever the response time of the primary task response was 3 standard deviations above or below the condition mean (<1%). All frequency data were arcsine transformed before statistical analyses (Winer, 1971).

Data were analyzed using analyses of variance (ANOVAs) with repeated measurement and planned comparisons using two-tailed t-tests for dependent samples. To compensate for violations of sphericity, Greenhouse–Geisser corrections were applied whenever appropriate (Greenhouse & Geisser, 1959), and corrected p values (but uncorrected degrees of freedom) are reported.

Table 1: Experiments 1a and 1b. Relative frequencies (in %) of secondary task responses for each stimulus condition and primary task response. Please note that the error rate cannot be derived from this table as these frequencies reflect secondary task responses only.

		Primary Task			
		Congruent		Incongruent	
		Correct	Error	Correct	Error
Secondary Task (1a)	Correct	99.66 (0.1)	0.96 (0.53)	98.4 (0.5)	3.26 (0.64)
	Early Error	0.25 (0.1)	66.7 (7.9)	1.01 (0.4)	69.1 (3.5)
	Late Error	0.09 (0.04)	32.3 (7.9)	0.61 (0.13)	27.7 (3.5)
Secondary Task (1b)	Correct	98.8 (1.7)	2.91 (1.24)	97.9 (0.5)	3.75 (0.74)
	Early Error	0.76 (0.43)	61.45 (7.04)	0.89 (0.45)	57.2 (6.1)
	Late Error	0.12 (0.07)	28.88 (7.51)	0.17 (0.09)	28.4 (5.4)
	Don’t know	0.29 (0.11)	6.76 (2.88)	1.02 (0.25)	10.7 (3.2)

Notes: Brackets contain standard errors of the mean.

Results

Experiment 1a. In a first step, we analyzed the overall performance in the primary flanker task. The mean error rate was 1.68% (SE = 0.49%) for congruent errors and 13.8% (SE = 2%) for incongruent errors (corresponding to 6.6 errors for congruent trials and 54.5 errors for incongruent trials). A t-test on the frequency data showed a congruency effect, $t(22) = 9.89$, $p < .001$, $d = 2.11$. We also investigated the congruency effect in RTs by comparing correct trials in the congruent and incongruent conditions. Statistical analyses showed a standard congruency effect with faster RTs for congruent corrects (M = 378 ms, SE = 8 ms) than incongruent corrects (M = 478 ms, SE = 14 ms), $t(22) = 11.44$, $p < .001$, $d = 2.38$.

We then considered the frequencies of secondary task responses for each primary task response and stimulus condition (see Tab. 1). As observed in previous studies (e.g., Steinhauser, Maier, & Hübner, 2008), the mere detection of errors (independent of error type) was very reliable. 97.8% (SE = 0.6%) of objective errors were categorized as either early errors or late errors, and this rate was higher for congruent stimuli (M = 99.1%, SE = 0.52%) than for incongruent stimuli (M = 96.7%, SE = 0.64%), $t(22) = 2.71$, $p = .012$, $d = 0.57$. Only 0.9% (SE = 0.01%) of correct responses were categorized as errors, which was also higher for incongruent (M = 1.63%, SE = 0.5%) than for congruent stimuli (M = 0.34%, SE = 0.12%), $t(22) = 3.08$, $p = .005$, $d = 0.64$.

Crucially, a considerable number of errors was categorized as early errors, that is, as errors accompanied by an early error sensation. Figure 1A shows the proportion of objective errors categorized as early and late errors among all detected objective errors. The proportion of early errors was similar for congruent (M = 76.1%, SE = 5.1%) and incongruent trials (M = 71.4%, SE = 4%), $t(22) = 0.96$, $p = .347$, $d = 0.21$. Additionally, we calculated confidence intervals (CI) to show the range of frequencies of early errors. The 95%-CIs ranged from 65.4% to 86.9% for congruent trials and between 63.9% and 78.9% for incongruent trials.

Additionally, we investigated the RTs of the different error types. Only incongruent trials were considered for this analysis because 12 participants had no errors in at least one condition of the congruent trials. Moreover, only correct trials classified as correct and error trials classified as errors were entered into this analysis. The mean number of trials included was 143.3 (SE = 5.0) for correct trials, 41.1 (SE = 6.9) for early errors, and 16.0 (SE = 3.2) for late errors. The results of the one-way ANOVA (corrects, early error, late error) showed a significant effect, $F(2,44) = 5.86$, $p = .022$, $\eta_p^2 = .21$. Planned contrasts revealed larger RTs for correct trials (M = 474 ms, SE = 13 ms) than for early errors (M = 367 ms, SE = 15 ms), $t(22) = 12.1$, $p < .001$, $d = 2.52$, but no significant difference between correct trials and late errors (M = 448 ms, SE = 46 ms), $t(22) = 0.68$, $p = .51$, $d = 0.14$. Furthermore, the difference between early and late errors was marginally significant, $t(22) = 2.07$, $p = .051$, $d = 0.42$.

Experiment 1b. The data were analyzed in the same way as in Experiment 1a. In the primary task, the mean error rate was smaller for congruent errors (M = 5.56%, SE = 1.58%) than for incongruent errors (M = 27.8%, SE = 2.4%), $t(19) = 8.49$, $p < .001$, $d = 1.89$. RTs were faster for congruent correct trials (M = 378 ms, SE = 8 ms) than for incongruent correct trials (M = 463 ms, SE = 12 ms), $t(19) = 10.2$, $p < .001$, $d = 2.34$.

Frequencies of secondary task responses (Tab. 1) again revealed that the detection of errors (independent of error type) was very reliable. 96.6% (SE = 0.93%) of objective errors were categorized as either early errors, late errors or I don't know errors, and this rate was comparable between congruent stimuli (M = 97.1%, SE = 1.23%) and incongruent stimuli (M = 96.3%, SE = 0.73%), $t(19) = 0.61$, $p = .55$, $d = 0.13$. Only 1.63% (SE = 0.5%) of correct responses were categorized as errors, which was higher for incongruent (M = 2.08%, SE = 0.51%) than for congruent stimuli (M = 1.17%, SE = 0.44%), $t(19) = 3.49$, $p = .002$, $d = 0.78$.

Crucially, the results again showed that participants consistently reported errors accompanied by early error sensations (Fig. 1B). In this experiment, participants could classify their errors as I don't know errors if they felt unable to classify them as early or late errors. However, the proportion of these errors was low (M = 9.03%, SE = 3.15%; 95%-CI = 7.55% - 10.5%), and comparable for congruent and incongruent stimuli, $t(19) = 1.05$, $p = .31$, $d = 0.23$. Due to this third category, we could now analyze the proportion of early and late errors in an ANOVA with the variables congruency (congruent, incongruent) and error type (early errors, late errors). This analysis revealed a significant main effect of error type, $F(1,19) = 6.62$, $p = .019$, $\eta_p^2 = .25$, indicating that early errors (M = 59.1%, SE = 7.08%; 95%-CI = 55.8% - 62.4%) were more frequent than late errors (M = 29.4%, SE = 6.69; 95%-

CI = 26.3% - 32.5%). However, neither a significant main effect of congruency, $F(1,19) = 0.321$, $p = .861$, $\eta_p^2 = .02$, nor a significant interaction, $F(1,19) = 0.321$, $p = .861$, $\eta_p^2 = .02$, was obtained.

Finally, we investigated RTs for correct responses, early and late errors. Again, only incongruent trials were considered for this analysis because 13 participants had no error trials in at least one condition of the congruent trials. Moreover, in line with Experiment 1a, only correct trials classified as correct and error trials classified as errors were entered into this analysis. The mean number of trials included was 209.0 (SE = 7.5) for correct trials, 47.5 (SE = 6.6) for early errors, and 23.1 (SE = 4.0) for late errors. The results of the one-way ANOVA (corrects, early error, late error) showed a significant effect, $F(2,38) = 159$, $p < .001$, $\eta_p^2 = .89$. RTs were larger for correct trials (M = 463 ms, SE = 12 ms) than for the other two error types, $ts > 13.2$, $ps < .001$, $ds > 2.95$. Nevertheless, there was no significant difference between early (M = 358 ms, SE = 10 ms) and late errors (M = 353 ms, SE = 13 ms), $t(19) = 0.71$, $p = .486$, $d = 0.16$.

Comparison of Experiments 1a and 1b. In a last step, we compared frequencies of early errors (among all errors) in the congruent and incongruent conditions between Experiments 1a and 1b. The resulting mixed-model ANOVA with the within-participant variable congruency and the between-participant variable Experiment showed a significant effect of Experiment indicating a lower overall frequency of early errors in Experiment 1b (M = 59.1%, SE = 7.1%) compared to Experiment 1a (M = 73.7%, SE = 4.5%), $F(1,41) = 4.31$, $p = .044$, $\eta_p^2 = .095$. Notably, this analysis revealed neither a significant main effect of congruency, $F(1,41) = 0.26$, $p = .607$, $\eta_p^2 = .006$, nor a significant interaction, $F(1,41) = 0.414$, $p = .524$, $\eta_p^2 = .01$.

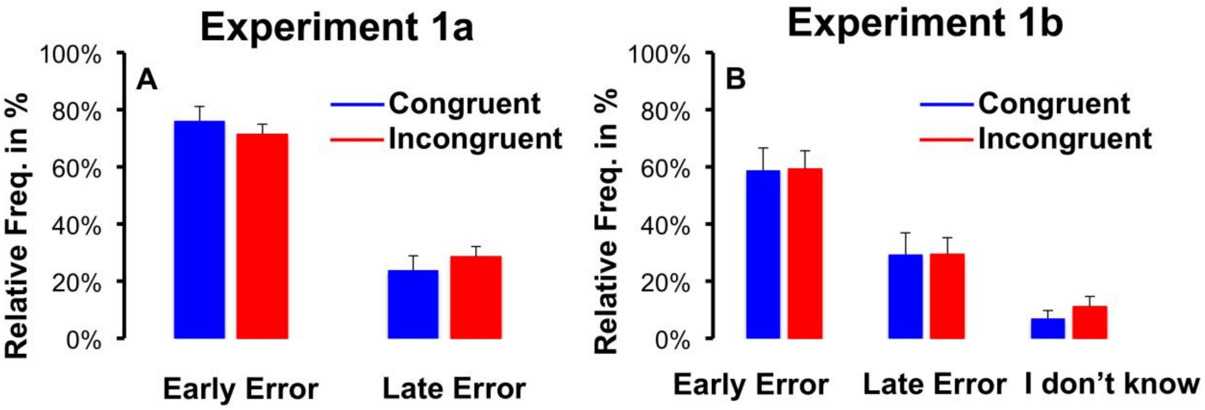

Figure 1: Data from Experiments 1a and 1b. A: Proportion of errors classified as early or late in the congruent and incongruent conditions for Experiment 1a. B: Proportion of errors classified as early, late or I don’t know in the congruent and incongruent conditions for Experiment 1b. Error bars represent standard errors of the mean.

Discussion

Data from Experiment 1 showed that our task successfully evoked reports of early error sensations and participants reported this feeling on a substantial proportion of error trials. Subjective reports on early error sensations were independent from other task-related features like stimulus congruency and RT. Indeed, participants reported early error sensations in a similar proportion for congruent and incongruent errors in both Experiments 1a and 1b.

In Experiment 1a, participants had to guess whenever they did not know when they detected the error. By adding a fourth classification response (I don’t know), this effect of guessing was controlled in Experiment 1b. As a result, the proportion of early detected errors decreased from 76% in Experiment 1a to 59% in Experiment 1b, which presumably reflects that participants frequently classified errors as early detected errors when they were actually unsure about when the error was detected. Introducing the “I don’t know” category in Experiment 1b eliminated this effect suggesting that the 59% early detected errors in Experiment 1b provides a more valid estimate of the true proportion of trials with early error sensations. Furthermore, the marginally significant difference in RTs between early and late errors in Experiment 1a was absent in Experiment 1b, which suggests that this difference reflects that many early errors were indeed guesses in the secondary task, which seem to be associated with shorter RTs in the primary task.

In general, participants reported early errors with a high frequency and we did not find robust evidence of differences in task-related features (i.e., stimulus congruency or RT) between early and late errors. Both results could suggest that early error sensations result from a functionally relevant process that serves to temporally synchronize the execution of the erroneous action and the later occurring error awareness.

Experiment 2

Participants in our first experiments reported that they had detected an error already before response execution on the majority of trials. This was the case even when they also could have classified the error as “I don’t know” in Experiment 1b, suggesting that they classified the remaining errors as early or late with considerable confidence. However, because the reported early error sensations are subjective reports on the participants' own metacognitive contents, there is no objective information about whether an early error sensation has really occurred. Therefore, it may still be argued that results reflect an expectation bias. Namely, participants could simply have expected that some errors must have been early errors, because we instructed them about their existence. In Experiment 2, we aimed to directly test how confident participants are about early error sensations and if their reports of early error sensations are subject to an expectation bias.

The procedure we used was post-decision wagering. In typical studies using this method, participants are instructed to perform a challenging perceptual discrimination task and then are asked how much money they would wager on the correctness of their response (Persaud et al., 2007). Post-decision wagering is particularly suitable to assess metacognitive contents as monetary incentives can improve the accuracy of these judgments and wagering provides a direct and intuitive measure to rate subjective confidence associated with a decision (Persaud et al., 2007; Windey & Cleeremans, 2015). In our Error Awareness task, we modified this procedure by having participants wager on their early error sensations. More specifically, after each response, participants had to place one of three bets on whether they have experienced an early error sensation: No bet, a low bet, or a high bet. Even though we could not objectively verify participants' bets, we instructed them that their wagering should correspond to their subjective confidence of having experienced an early error sensation. While we informed participants that they would not earn any money in this task, they were not explicitly told that their bets could not be verified.

To further improve the participants' metacognitive assessments, we provided a more objective reference point that should guide their wagering on early error sensations. Prior to the Error Awareness task, participants performed a Visual Awareness task similar to those typically used with the post-decision wagering method (Persaud et al., 2007). In this task, participants first performed a difficult perceptual letter/number discrimination with masked stimuli and after their response, they wagered on whether they had seen the stimulus, and hence, on their response accuracy. In the subsequent Error Awareness task, which used the same stimuli but a longer masking interval, participants were instructed to apply similar levels of confidence when placing their bets as in the Visual Awareness task. That is, they should place a high bet only if they were similarly confident of having experienced an early error in the Error Awareness task as of having executed a correct response in the Visual Awareness task. In this way, we induced a common metric for the metacognitive judgments in the two tasks. High bets in the Error Awareness task should be associated with the same confidence as high bets in the Visual Awareness task, with the advantage that we have objective data on the accuracy of the latter bets.

Finally, this two-stage design allowed us to evaluate the possible impact of expectation biases on the reports of early error sensations. If participants reported early errors only because they expected that early error sensations exist, their judgments should be strongly influenced by the proportion of low and high bets in the preceding Visual Awareness task. Performance in the Visual Awareness task should serve as an anchor or reference point for how many low and high bets should be expected, if no objective signal about the timing of error detection is available. We therefore conducted two experiments (Experiment 2a and 2b) in which we varied the difficulty of the perceptual discrimination in the Visual Awareness task. In particular, we calibrated stimulus-masking intervals to set individual response accuracy to around 75% for Experiment 2a and to around 50% for Experiment 2b. In this way, we varied perceptual difficulty and thus the number of low and high bets in the Visual Awareness task, while holding the difficulty of the Error Awareness task constant. If the reported rate of early error sensations is influenced by the reference point set by the Visual Awareness task, we should obtain more high bets on early error sensations when the Visual Awareness task was easier (Exp. 2a) than when it was more difficult (Exp. 2b).

Method

Participants. 20 participants (3 male), which did not participate in Experiment 1, between 19 and 26 years of age (M = 22.6, SE = 0.7) with normal or corrected to normal vision participated in Experiment 2a. A new group of 20 participants (5 male) between 19 and 31 years of age (M = 23.4, SE = 0.9) participated in Experiment 2b.

Overview. Both experiments consisted of two separate sub-tasks, which always occurred in the same order: Participants first performed a Visual Awareness task and then an Error Awareness task. In the Visual Awareness task, participants performed a difficult perceptual discrimination task (primary task), and then wagered 0 cents (no bet), 1 cent (low bet) or 10 cents (high bet) on the accuracy of the primary task response (secondary task). In the Error Awareness task, participants performed a speeded choice task (primary task), and then wagered on whether they had the feeling of early error detection in case of an error (secondary task).

Visual Awareness task. Each target stimulus was a letter (W, B, G, T, S, C, P, N, or R) or a number (1, 2, 3, 4, 5, 6, 7, 8, or 9) subtending a visual angle of 0.5° horizontally and 1.4° vertically and presented in the center of the screen. The procedure of a trial is depicted in Figure 2. The primary task required participants to classify the target as a letter or a number by pressing the "left arrow" key or the "right arrow" key with the index or middle finger, respectively, of the right hand. The category-response mapping was counterbalanced across participants. Each trial started with the presentation of a fixation cross for 350 ms, which was followed by the target. After a stimulus-mask interval (SMI, see below), a mask appeared for 200 ms that consisted of random feature patterns created by randomly rearranging features of the letter and number stimuli. Then, a black screen was presented until a response was provided. After the response, the German equivalent of the text "Bet that you have seen the number or the letter: 0 cents, 1 cent, 10 cents" was presented. As secondary task, participants had to provide one of the three wagering responses (0, 1, 10) on the keyboard with index (D key), middle (S key) or ring finger (A key), respectively, of the left hand. Participants were instructed that whenever they responded correctly, they would earn the corresponding wagered money, and that in case of an error, they would lose the wagered money. After the response, another black screen was presented for 1500 ms followed by the next trial.

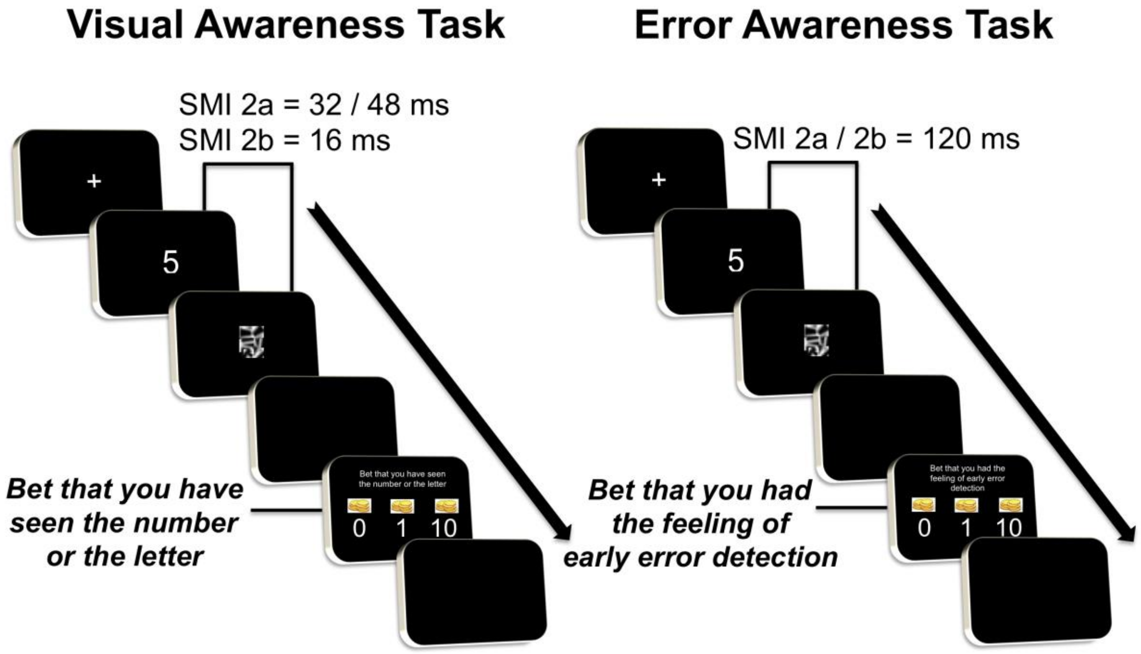

Figure 2: Procedure in Experiments 2a and 2b. Procedure of a trial in the Visual Awareness task (left side) and the Error Awareness task (right side). SMI 2a and 2b refer to the intervals used in Experiment 2a and 2b, respectively. SMI = stimulus-masking interval.

The SMI for Experiment 2a was determined in a pilot study in which 6 participants worked only on the Visual Awareness task. The SMI was calibrated using a staircase procedure to obtain a mean error rate of 25%. The experiment started with an SMI of 32 ms. Whenever participants committed less than 15% of errors in the last 20 trials, the SMI was decreased by one step (-16 ms). Whenever participants committed more than 35% of errors in the last 20 trials, the SMI increased by one step (+16 ms). The task terminated when a constant error rate of 25% was observed in the last 60 trials. The resulting SMI was 32 ms in 5 participants and 48 ms in 1 participant. Therefore, an SMI of 32 ms was used in Experiment 2a.

Experiment 2a started with two practice blocks (72 trials each) in which the Visual Awareness task was practiced and the SMI was re-calibrated for each participant (using the same method as in the pilot study and starting with an SMI of 32 ms). The resulting SMIs used in the test blocks were 32 ms for 15 participants and 48 ms for 5 participants. Only in these practice blocks, participants received feedback about whether they wagered advantageously (i.e., correct response followed by 10 cents bet and error followed by 0 cents bet) or not (correct response followed by 1 or 0 cents bet and error followed by 10 or 1 cents bet). No feedback was provided in the test blocks to make the Visual Awareness task as similar as possible to the subsequent Error Awareness task in which no feedback was possible. After these practice blocks, participants completed 6 test blocks (72 trials each). At the end of all blocks in visual awareness task, participants were informed about the amount of money they won.

In Experiment 2b, we used an SMI of 16 ms for all participants. This was done to further limit stimulus processing relative to Experiment 2a, and thus, to set a different reference point for the later Error Awareness task. Manipulating the SMI across Experiments 2a and 2b would allow for studying the impact of an expectation bias on the report of early error sensations in the Error Awareness Task.

Error Awareness task. The Error Awareness task in both experiments employed a similar trial sequence and the same stimuli as the Visual Awareness task (see Fig. 2) but differed in the SMI, the instruction, and the secondary task. We now used a fixed SMI of 120 ms, which should prevent errors due to data limitation and thus should guarantee that error detectability is high (Del Cul, Dehaene, Reyes, Bravo, & Slachevsky, 2009; Scheffers & Coles, 2000). To obtain a mean error rate of 25%, participants were instructed to respond as quickly as possible. Moreover, whenever the average error rate in a block fell below 20%, participants were prompted to respond more quickly prior to the beginning of the next block. The secondary task was to bet on the feeling of early error detection. The equivalent of the German text was: “Bet that you had the feeling of early error detection: 0 cents, 1 cent, 10 cents”, and participants had to execute one of the three wagering responses (0, 1, 10).

Participants were instructed to wager 0 cents in case of correct responses. The Error Awareness task was always conducted after the last block of the Visual Awareness task. It consisted of 2 practice blocks and 12 test blocks with 72 trials each. In contrast to the Visual Awareness task, participants did not receive any money in the Error Awareness task, as there is no possibility to determine whether a wager was successful or not. Prior to the task, participants were explicitly instructed to apply similar levels of confidence as in the Visual Awareness task when placing their bets in the Error Awareness task. Specifically, they were instructed to place a high bet in the Error Awareness task only if they were similarly confident of having experienced an early error sensation as of having seen the stimulus (on high-bet trials) in the Visual Awareness task.

Data analysis. RTs and frequencies were analyzed in the same way as in Experiment 1. For both the Visual Awareness task and the Error Awareness task, the correctness of the primary task response and the type of wagering response were used to distinguish the following trial~~s~~ types: (1) *correct/high bet* on which the response in the primary task was correct and the wagered amount of money was 10 cents; (2) *correct/low bet* on which the response in the primary task was correct and the wagered money was 1 cent; (3) *correct/no bet* on which the response in the primary task was correct and the wagered money was 0 cents; (4) *error/high bet* on which the response in the primary task was an error and the wagered money was 10 cents; (5) *error/low bet* on which the response in the primary task was an error and the wagered money was 1 cent; (6) *error/no bet* on which the response in the primary task was an error and the wagered money was 0 cents.

Results

Experiment 2a: Visual Awareness Task. Analysis of the primary task response in the Visual Awareness task indicated a mean error rate of 21.8% (SE = 3.5%), which corresponded to a mean number of errors of 94.2 (SE = 15.12). Moreover, participants were faster on correct trials (M = 904 ms, SE = 77 ms) than on error trials (M = 1302 ms, SE = 122 ms), *t*(19) = 3.34, *p* = .003, *d* = 0.79. Table 2 provides the frequencies for each secondary task response among correct and error trials. On correct trials, participants had a larger frequency of advantageous wagering (correct/high bet; M = 66.1%; SE = 6.7%; 95%-CI = 62.9% - 69.2%) compared to correct/low bet (M = 25.5%; SE = 5.2%; 95%-CI = 23.1% - 27.9%), *t*(19) = 3.51, *p* = .002, *d* = 0.78, and a larger frequency of correct/low bet compared to correct/no bet (M = 8.46%; SE = 3.4%; 95%-CI = 6.9% - 10.1%), *t*(19) = 3.01, *p* = .007, *d* = 0.67 (see also Figure 3). On error trials, error/low bets (M = 40%; SE = 4.7%; 95%-CI = 37.8% - 42.2%) were more frequent than error/high bets (M = 24.7%; SE = 4.1%; 95%-CI = 22.8% - 26.6%), *t*(19) = 2.53, *p* = .021, *d* = 0.56, while error/no bets (M = 35.3%; SE = 1.3%; 95%-CI = 34.7% - 35.9%) were comparable with error/low bets, *t*(19) = 0.45, *p* = .657, *d* = 0.11. The considerable number of low bets, particularly on error trials, suggests that data limitation led to a high response uncertainty. Because the Visual Awareness Task is not a speeded choice task, and because several participants had only few responses for single conditions, we did not further analyze RTs as a function of the secondary task response.

Table 2: Experiment 2a and 2b. Relative frequencies (in %) of secondary task responses for each wagering condition and primary task response. Please note that the error rate cannot be derived from this table as these frequencies reflect secondary task responses only.

		Primary Task			
		Visual Awareness		Error Awareness	
		Correct	Error	Correct	Error
Secondary Task (2a)	High Bet	66.1 (6.7)	24.7 (4.1)	12.3 (6.3)	62.4 (5.8)
	Low Bet	25.5 (5.2)	40 (4.7)	2.44 (0.86)	19.8 (4.2)
	No Bet	8.46 (3.4)	35.3 (1.3)	85.3 (6.5)	17.8 (4.7)
Secondary Task (2b)	High Bet	36.5 (5.1)	18.1 (3.2)	5.67 (2.79)	70.9 (5.2)
	Low Bet	37.4 (3.7)	43.7 (4.1)	8.91 (4.45)	12.8 (3)
	No Bet	26.1 (4.7)	38.2 (5.1)	85.4 (5.7)	16.3 (3.5)

Notes: Brackets contain standard errors of the mean.

Experiment 2a: Error Awareness Task. In the Error Awareness task, we reduced errors due to data limitation using a longer SMI but instead induced a high time pressure. The mean error rate was 20.1% (SE = 2.5%), which corresponded to a mean number of errors of 172.8 (SE = 21.6). Participants were slower for correct trials (M = 296 ms, SE = 13 ms) than for error trials (M = 234 ms, SE = 14 ms), $t(19) = 7.33$, $p < .001$, $d = 1.82$. Table 2 provides the frequencies of secondary task responses among correct trials and error trials. Now, participants had to bet on having experienced an early error sensation. As merely detecting an error is rather easy in this task, it is not surprising that, on correct trials, the frequency of correct/no bets (M = 85.3%; SE = 6.5%; 95%-CI = 82.1% - 88.2%) was much larger than the frequencies of correct/low bets (M = 2.4%; SE = 0.9%; 95%-CI = 1.97% - 2.82%), $t(19) = 12.2$, $p < .001$, $d = 2.75$, and correct/high bets (M = 12.3%; SE = 6.3%; 95%-CI = 9.35% -

15.2%), *t*(19) = 5.71, $p < .001$, $d = 1.28$. Crucially, however, on error trials, the frequency of error/high bets (M = 62.4%; SE = 5.8%; 95%-CI = 59.7% - 65.1%) was larger than the frequency of error/low bets (M = 19.8%; SE = 4.2%; 95%-CI = 17.8% - 21.8%), *t*(19) = 4.72, $p < .001$, $d = 1.05$, and the frequency of error/no bets (M = 17.8%; SE = 4.7%; 95%-CI = 15.6% - 19.9%), *t*(19) = 4.58, $p < .001$, $d = 1.03$ (see also Figure 3). This implies that participants were highly confident in experiencing early error detection on a large number of trials.

For analysis of primary task RTs, 3 participants who had no trial in one of the error conditions were excluded. However, the one-way ANOVA did not reveal any significant effects (error/high bets: M = 241 ms, SE = 16 ms; 113 trials; SE = 21.3 trials; error/low bets: M = 282 ms, SE = 25 ms; 33.9 trials; SE = 7.1 trials; error/no bets: M = 361 ms, SE = 61 ms; 26.1 trials; SE = 6.1 trials), *F*(2,32) = 1.93, $p = .18$, $\eta_p^2 = .11$. When 7 further participants with fewer than 10 trials in one of the conditions were excluded, again no significant difference was revealed (error/high bets: M = 228 ms, SE = 17 ms; 122 trials; SE = 22.9 trials; error/low bets: M = 250 ms, SE = 26 ms; 38.1 trials; SE = 7.3 trials; error/no bets: M = 348 ms, SE = 77 ms; 30.9 trials; SE = 6.2 trials), *F*(2,22) = 1.59, $p = .22$, $\eta_p^2 = .13$.

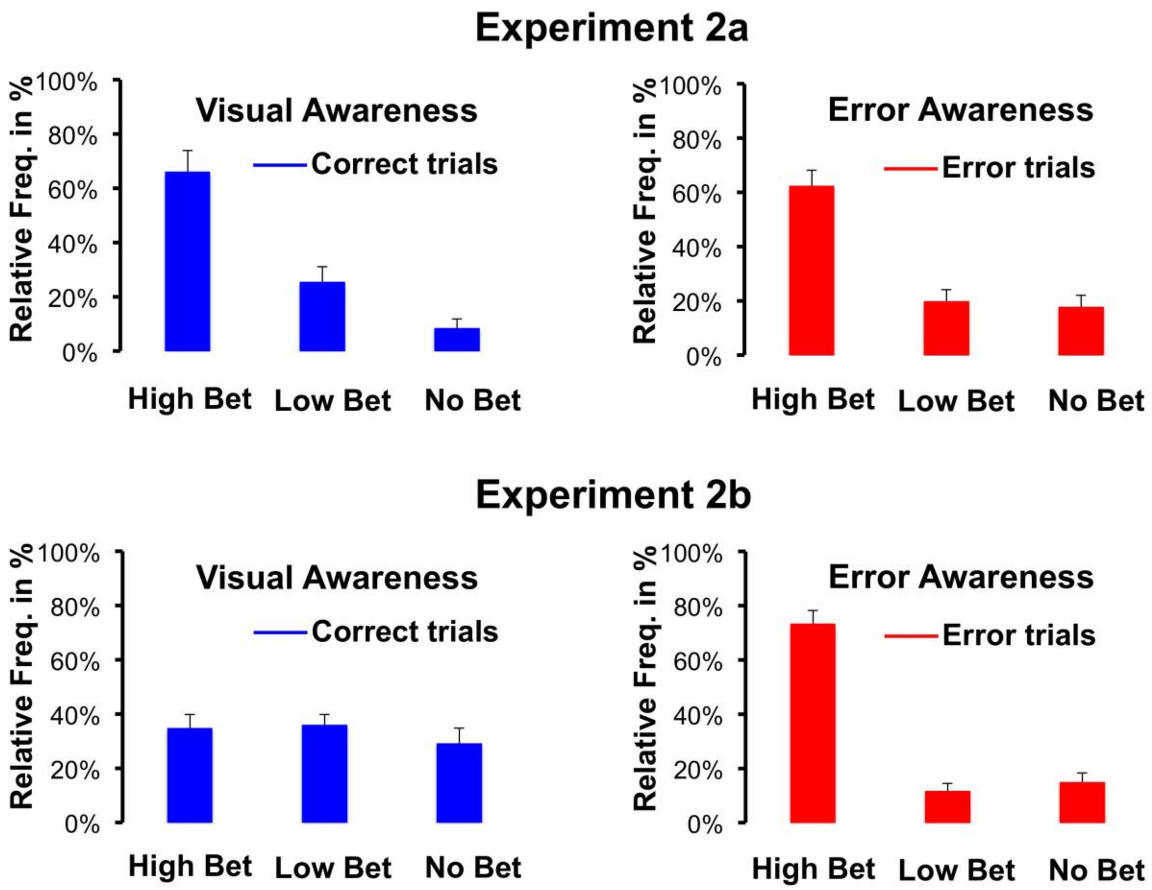

Figure 3: Results of Experiments 2a and 2b. For the Visual Awareness task (left column), the proportions of correct responses followed by high, low or no bets are reported for Experiments 2a and 2b. For the Error Awareness task (right column), the proportions of errors followed by high, low or no bets are reported for Experiments 2a and 2b. Error bars represent standard errors of the mean.

Experiment 2b: Visual Awareness Task. The mean error rate in the Visual Awareness task was 38.4% (SE = 2.3%), which corresponded to a mean number of errors of 165.9 (SE = 9.94). The RT difference between correct trials (M = 1176 ms, SE = 126 ms) and error trials (M = 1264 ms, SE = 119 ms) was marginally significant, $t(19) = 1.77$, $p = .093$, $d = 0.39$. Frequencies of secondary task responses are provided in Table 2. Notably, the reduced SMI in Experiment 2b led to a drastic change of the pattern of wagering. On correct trials, neither the difference between the frequencies of correct/high bets (M = 36.5%; SE = 5.1%; 95%-CI = 34.1% - 38.8%) and correct/low bets (M = 37.4.7%; SE = 3.7%; 95%-CI = 35.7% - 39.1%), $t(19) = 0.12$, $p = .907$, $d = 0.03$ nor the difference between the frequencies of correct/low bets and correct/no bets (M = 26.1%; SE = 4.8%; 95%-CI = 23.9% - 28.3%) was significant, *t*(19) = 1.29, *p* = .212, *d* = 0.28 (see also Figure 3). On error trials, the frequency of high bets (M = 18.1%; SE = 3.3%; 95%-CI = 16.6% - 19.6%) was significantly lower than the frequency of error/low bets (M = 43.7%; SE = 4%; 95%-CI = 41.8% - 45.6%), *t*(19) = 4.61, *p* < .001, *d* = 1.03, while the frequencies of error/low bets and error/no bets (M = 38.2%; SE = 5.1%; 95%-CI = 35.8% - 40.6%) did not differ significantly, *t*(19) = 0.55, *p* = .589, *d* = 0.12. These data show that the small SMI led to considerable response uncertainty in this task.

Experiment 2b: Error Awareness Task. The mean error rate was 19.1% (SE = 1.9%), which corresponded to a mean number of errors of 165 (SE = 17.28). Participants were slower on correct trials (M = 331 ms, SE = 10 ms) than on error trials (M = 275 ms, SE = 15 ms), *t*(19) = 5.09, *p* < .001, *d* = 1.13. Table 2 provides the frequencies of secondary task responses. Among correct trials, the frequency of correct/no bets (M = 85.4%; SE = 5.7%; 95%-CI = 82.7% - 88.1%) was again larger than the frequencies of correct/low bets (M = 8.91%; SE = 4.45%; 95%-CI = 6.8% - 10.9%), *t*(19) = 7.76, *p* < .001, *d* = 1.73, and correct/high bets (M = 5.67%; SE = 2.79%; 95%-CI = 4.4% - 6.97%), *t*(19) = 10.2, *p* < .001, *d* = 2.28. On error trials, the frequency of error/high bets (M = 70.9%; SE = 5.2%; 95%-CI = 68.5% - 73.3%) was larger than the frequency of error/low bets (M = 12.8%; SE = 3%; 95%-CI = 11.5% - 14.3%), *t*(19) = 7.53, *p* < .001, *d* = 1.68, and the frequency of error/no bets (M = 16.3%; SE = 3.5%; 95%-CI = 14.7% - 17.9%), *t*(19) = 6.55, *p* < .001, *d* = 1.47 (see also Figure 3). This again suggests that participants rather consistently experienced early error detection in this task.

The one-way ANOVA on the primary task RTs did not reveal any significant effect (error/high bets: M = 271 ms, SE = 21 ms; 115 trials; SE = 14.9 trials; error/low bets: M = 437 ms, SE = 130 ms; 18.8 trials; SE = 5.5 trials; error/no bets: M = 347 ms, SE = 44 ms; 24.8 trials; SE = 4.9 trials;), *F*(2,38) = 2.73, *p* = .11, η_p^2 = .13. When 9 participants were excluded that had fewer than 10 trials in one of the conditions, again no significant difference was obtained (error/high bets: M = 283 ms, SE = 23 ms; 108 trials; SE = 14.7 trials; error/low bets: M = 301 ms, SE = 22 ms; 31.6 trials; SE = 6.1 trials; error/no bets: M = 303 ms, SE = 23 ms; 32.8 trials; SE = 5.9 trials), $F(2,20) = 0.51$, $p = .61$, $\eta_p^2 = .05$.

Comparison across tasks and Experiments 2a and 2b. In a further analysis, we compared the frequencies of high bets in the Visual Awareness task and Error Awareness task across Experiments 2a and 2b to investigate whether the reference point set in the Visual Awareness task influenced wagering in the Error Awareness task. To this end, we included data from correct trials in the Visual Awareness task (because participants placed bets on being correct) and from errors in the Error Awareness task (because participants placed bets on experiencing early error detection) in the analysis, as summarized in Figure 3. The resulting mixed-model ANOVA with the within-participants variable task (Visual Awareness, Error Awareness) and the between-participants variable experiment (Exp. 2a, Exp. 2b) revealed a significant main effect of task, $F(1,38) = 7.79$ $p = .008$, $\eta_p^2 = .173$, and a significant interaction between both variables, $F(1,38) = 9.31$ $p = .004$, $\eta_p^2 = .197$. Independent samples t-tests showed that, in the Visual Awareness task, the frequency of high bets was larger for correct trials in Experiment 2a than in Experiment 2b, $t(38) = 3.39$, $p = .002$, $d = 0.76$, thus reflecting the increased difficulty of the Visual Awareness task in Experiment 2b. In contrast, in the Error Awareness task, the frequency of high bets on errors did not differ between experiments, $t(38) = 1.09$, $p = .282$, $d = 0.21$, suggesting that the different reference points in the Visual Awareness task did not influence wagering in the Error Awareness task. Indeed, the frequency of high bets across tasks differed significantly only in Experiment 2b, $t(19) = 4.26$, $p < .001$, $d = 0.95$, but not in Experiment 2a, $t(19) = 0.39$, $p = .71$, $d = 0.09$.

Early error detection across Experiments 1 and 2. In a final analysis step, we compared the frequencies of early error sensations across Experiment 1 and 2 to investigate whether early detected errors occur similarly across experimental methods. We assume that error/high bet trials from Experiment 2 correspond to trials with truly experienced early error sensations. As Experiment 1a presumably overestimated the rate of early error sensations because classified early errors also included guesses, we compared the rate of early error sensations among errors from Experiment 1b with the rates of high bet trials among errors from Experiments 2a and 2b. The one-way ANOVA between groups on the frequency of early detected errors showed no significant difference, $F(2,57) = 2.02$ $p = .142$, $\eta_p^2 = .033$. Moreover, we directly compared the two methodological approaches contrasting participants' rates of early error sensations in Experiment 1b (20 participants) and 2a/2b (40 participants). The result of the independent samples t-test also showed no significant difference, $t(58) = 1.35$, $p = .18$, $d = 0.17$. The estimated mean of early detected errors across the two experimental methods was 63.5% (SE = 3.71).

Discussion

Experiment 2 employed post-decision wagering to measure the confidence associated with reports of early error sensations. Moreover, a Visual Awareness task was used to provide a reference point that should guide participants' wagering on early error sensations. They were instructed to place high bets on early error sensations only if their confidence was comparable to high bets in the Visual Awareness task. The results indicate that, averaged across Experiments 2a and 2b, participants perceived early error detection with high confidence on about 67% of trials. Impressively, this is very similar to (and does not differ significantly from) the rate of early detected errors in Experiment 1b (59%).

A further goal of this experiment was to investigate whether reports on early error detection are influenced by an expectation bias. To this end, the Error Awareness task was preceded by either a less (Exp. 2a) or more (Exp. 2b) difficult Visual Awareness task.

Manipulating the difficulty of this task successfully led to strong shift in wagering. Whereas 66% of correct responses led to a high bet in Experiment 2a, only 37% of correct responses led to a high bet in Experiment 2b. However, wagering in the Error Awareness task was fully independent of this manipulation. This clearly shows that expectations about how many high bets might be expected based on the Visual Awareness task had no effect at all on wagering in the Error Awareness task.

General Discussion

Participants in experiments on error detection frequently report that they already knew that an error has occurred before the response was executed, a phenomenon we term early error sensation. The goal of the present study was to investigate whether these anecdotally reported early error sensations exist and whether they can be reliably reported. In four experiments using two experimental approaches, we provided evidence that early error sensations indeed exist, and that they occur on the majority of error trials. When participants were asked to classify responses in a flanker task either as being correct, as early detected errors, or as late detected errors in Experiment 1a, they reported early errors in 73.7% of errors. When an additional category for detected errors with unclear timing was introduced in Experiment 1b, early errors were reported in 59.1% of trials. When participants had to wager on the feeling of early error detection, they placed high bets on 62.4% (Exp. 2a) and 70.9% (Exp. 2b). These data demonstrate that early error sensations are reported very consistently across different primary tasks (flanker task vs. number/let discrimination) and secondary tasks (error classification vs. post-decision wagering).

Crucial, however, is the question whether these introspective reports indeed reflect that errors were detected before the response, or whether participants were unable to discriminate between early and late errors and simply guessed that early errors must occasionally occur. A challenging problem for measuring early error sensations is that we cannot objectively determine whether a given error was detected early or late. To deal with this problem, we introduced a reference for the metacognitive reports of early error sensations. In Experiment 2, we used a Visual Awareness task in which participants had to wager on the accuracy of their responses. In the subsequent Error Awareness task, we instructed participants to place high bets on early error sensations only if they were similarly confident as for the high bets in the Visual Awareness task. We argued that this induces a common metric for judging confidence of the two tasks, which allowed us to interpret the metacognitive reports of early error detection with respect to the metacognitive judgments of visual awareness. This reasoning receives support from previous findings showing that humans represent confidence in a task-unspecific format which allows them to compare confidence across tasks with a similarly high precision as confidence within tasks (de Gardelle & Mamassian, 2014). Moreover, it has recently been suggested that integrating information from different sources into a common metric might even be the major purpose of metacognition (Shea & Frith, 2019). In Experiment 2a, the frequencies of high bets were coincidentally similar in both tasks. We can thus infer that the average confidence by which participants reported early error sensations in this experiment corresponded to the average confidence by which they were aware of the visual stimuli in the Visual Awareness task. This confidence level ought to be rather high given that the objective performance in the Visual Awareness task was far above chance level.

We found no evidence that metacognitive reports of early error sensations were subject to an expectation bias. If participants simply guessed that early error sensations must occasionally occur, these guesses should be influenced by expectations about the frequency of early error sensations. To investigate whether such an expectation bias exists, we manipulated the difficulty of the Visual Awareness task, and thus the frequency of high bets in this task. However, whereas the frequency of high bets in the Visual Awareness task varied between

Experiments 2a and 2b, the frequency of high bets in the Error Awareness task remained constant across the two experiments. This suggests that metacognitive judgments about early error sensations are not influenced by a specific expectation bias induced by the frequency of high bets in the Visual Awareness task. While we cannot fully exclude a general bias towards instruction-driven expectations about early error sensations, our results strongly suggest that metacognitive judgments on early error sensations are very consistent and reliable across experimental procedures.

We found no evidence that early and late detected errors differ with respect to any objective features. It has been reported that uncertainty or conflict during response selection can influence post-response decision process and metacognitive judgments about errors (Steinhauser et al., 2008; Yeung & Summerfield, 2012). As a consequence, variables like stimulus congruency or RT could potentially influence subjective judgments about early error sensations. However, we found no robust evidence that this was the case in the present study. Participants reported early error sensations in a similar proportion for congruent and incongruent errors in Experiment 1. Moreover, RTs were similar across all error types. A small RT difference between early and late detected errors in Experiment 1a disappeared when we controlled for errors with unclear timing in Experiment 1b. This suggests that the emergence of early error sensations is not related to specific features of task processing like stimulus congruency or RTs. Thus, our data provide little evidence that early error sensations reflect the objective latency of error detection, which has been found to correlate with RT when response speed was directly manipulated (Steinhauser et al., 2008).

An important question is why early error sensations occurred on the majority of trials whereas the neural correlates of error awareness emerge not until 300 ms after an error (e.g., Steinhauser & Yeung, 2010). There are at least two possible explanations. A first explanation is that conclusions about the timing of error awareness from EEG measures like the Pe are incorrect. The Pe is often considered the earliest neural correlate of error awareness and the role of the Pe for the emergence of error awareness has been described within an evidence accumulation account (Steinhauser & Yeung, 2010; Ullsperger, Harsay, Wessel, & Ridderinkhof, 2010). It is assumed that the Pe reflects the accumulated evidence that an error has occurred, and that error awareness emerges when this evidence exceeds a threshold. The evidence is provided by cognitive, autonomous, motor and sensory processing (Bode & Stahl, 2014; Wessel, Danielmeier, Morton, & Ullsperger, 2012; Wessel, Danielmeier, & Ullsperger, 2011), but does not necessarily rely on early error processing represented by the Ne/ERN (Di Gregorio, Maier, & Steinhauser, 2018). One possibility is that the feeling of error awareness emerges already before the Pe, for instance, at the time point of the Ne/ERN or even earlier (Bode & Stahl, 2014). The Pe could represent a later stage of metacognitive processing, perhaps related to the emergence of confidence about response accuracy (Boldt & Yeung, 2015).

A second explanation is that early error sensations are a metacognitive illusion. Error awareness could emerge at the time of the Pe but the illusion is created that the error has been detected already before the response. This mechanism could serve to subjectively synchronize error awareness with the timing of the objective error in the same way as visual awareness is subjectively aligned with the onset of a visual stimulus. In the context of visual awareness, expectations and other top-down variables can influence the accumulation of sensory evidence and consequentially metacognitive judgments about stimulus awareness (de Lange, Jensen, & Dehaene, 2010; Kouider, de Gardelle, Sackur, & Dupoux, 2010) . Moreover, a backward referral process has been assumed to synchronize the subjective time point of visual awareness with the objective stimulus to create a coherent perception in the stream of consciousness (Libet et al., 1979, 1983). A similar process could align the subjective time point of error awareness with the emergence of the objective error. This temporal alignment of actions (i.e., a response) and their effects (i.e., the feeling of being incorrect) could further serve to evoke a sense of agency, i.e., the feeling of having caused an effect. Indeed, previous studies have shown that action-effect contingencies are influenced by their temporal contiguity and vice versa. Humans tend to perceive two events more causally related the closer they occur in time (Greville & Buehner, 2010), and causality judgments correlate with the perceived temporal contiguity between actions and their sensory effects (Haering & Kiesel, 2016). In other words, these metacognitive illusions on early error sensations could serve to reconstruct temporal contiguity between perception, action and metacognitive contents (Kouider, de Gardelle, Sackur, & Dupoux, 2010).

While we obtained clear and robust results across several experiments, the present method has also some limitations. A first limitation is that using a categorical measure for the timing of error detection implies a loss of information as time is a continuous phenomenon. However, differentiating only between errors detected before and after the response has the advantage of imposing considerably lower cognitive load than using a continuous measure. For instance, in the classical Libet studies (Libet et al., 1983), participants had to indicate the time of voluntary action initiation on a visual clock. However, in addition to considerable methodological weaknesses (Trevena & Miller, 2002), monitoring a clock represents a difficult secondary task that presumably interferes with both, the primary task and the task to detect errors. In contrast, our categorical measure uses the response as a reference rather than a continuous timer. As error detection already involves response monitoring (Steinhauser et al., 2008), only minimal additional load should be imposed.

As already discussed, a second limitation is that we have no objective measure that verifies the existence of early error sensations. Future studies could solve this problem by measuring neural correlates of early error sensations. Strong evidence for the existence of early error sensations would be provided if not only the Pe but also the earlier Ne/ERN would correlate with early error sensations. If only the Pe differed between early and late detected errors, this would suggest that early error sensations emerge during the later stage of conscious error processing. However, if such a difference was found also for the Ne/ERN, this would point to early error signals such as response conflict (Yeung et al., 2004) or prediction errors (Holroyd & Coles, 2002) as the origin of early error sensations. It is even possible that brain activity preceding the response can affect metacognitive judgments on early error sensations. ERP differences between errors and correct responses have been found prior to the response (Bode & Stahl, 2014) or even on the previous trial of simple tasks (Hajcak, Nieuwenhuis, Ridderinkhof, & Simons, 2005; Hoonakker, Doignon-Camus, & Bonnefond, 2016; Ridderinkhof, Nieuwenhuis, & Bashore, 2003), as well as in tasks involving complex sequences of motor programs such as piano playing (Maidhof, Rieger, Prinz, & Koelsch, 2009). In a similar vein, a study using self-report measures has revealed that internal error prediction occurs before responses in skilled typing (Rieger & Bart, 2016). Here, the question arises whether this activity serves as a cue for metacognitive judgments, or whether metacognition relies on direct access to the timing of these neural events.

A further question is whether early error sensations are related to early incorrect response activation. On correct trials, early incorrect response activation leads to a phenomenon called partial errors (Burle, Possamaï, Vidal, Bonnet, & Hasbroucq, 2002; Coles, Scheffers, & Fournier, 1995; Endrass, Klawohn, Schuster, & Kathmann, 2008), which can be consciously reported by participants (Rochet, Spieser, Casini, Hasbroucq, & Burle, 2014). Future studies could investigate whether such early incorrect response activation on error trials is responsible for early error sensations. Indeed, lower response force for errors than correct responses has been shown in skilled typing (Rabbitt, 1978). As this phenomenon has been interpreted as resulting from inhibition of the error response before actual response execution, it could be taken as indirect evidence for early error sensations. Future studies could examine whether errors accompanied by early error sensations are executed with lower response force than late errors.

The present study provides first evidence that participants have the subjective feeling of detecting errors already before they occurred. We show that these early error sensations can be robustly measured across different tasks and metacognitive judgments. Our results add to the broad body of evidence that humans have metacognitive access to a multitude of performance parameters. Previous studies could show that participants are able to report whether an error has occurred or not (Rabbitt, 1968, 2002), to provide graded confidence judgments on the accuracy of their response (Boldt & Yeung, 2015), to classify the type of error they committed (i.e., to which distractor stimulus they responded; Di Gregorio, Steinhauser, & Maier, 2016), and to estimate their RTs in choice tasks (Bryce & Bratzke, 2014). These metacognitive contents are used for optimizing decision processes (Desender, Boldt, & Yeung, 2018; Desender, Van Opstal, & Van den Bussche, 2014). Metacognitive representations on the timing of error detection could form another piece of information to support this optimization.

Acknowledgements: This work was supported by the Deutsche Forschungsgemeinschaft (DFG; grant numbers: MA 4864/3-1) granted to MEM.

References

Bode, S., & Stahl, J. (2014). Predicting errors from patterns of event-related potentials preceding an overt response. *Biological Psychology*, *103*, 357–369. http://doi.org/10.1016/j.biopsycho.2014.10.002

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

Bryce, D., & Bratzke, D. (2014). Introspective reports of reaction times in dual-tasks reflect experienced difficulty rather than timing of cognitive processes. *Consciousness and Cognition*, *27*(1), 254–267. http://doi.org/10.1016/j.concog.2014.05.011

Burle, B., Possamaï, C. A., Vidal, F., Bonnet, M., & Hasbroucq, T. (2002). Executive control in the Simon effect: An electromyographic and distributional analysis. *Psychological Research, 66*, 324–336. http://doi.org/10.1007/s00426-002-0105-6

Coles, M. G., Scheffers, M. K., & Fournier, L. (1995). Where did you go wrong? Errors, partial errors, and the nature of human information processing. *Acta Psychologica*, *90*(95), 129–144. http://doi.org/10.1016/0001-6918(95)00020-U

de Gardelle, V., & Mamassian, P. (2014). Does confidence use a common currency across two visual tasks? *Psychological Science*, *25*(6), 1286–1288. http://doi.org/10.1177/0956797614528956

de Lange, F. P., Jensen, O., & Dehaene, S. (2010). Accumulation of evidence during sequential decision making: the importance of top-down factors. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*, 731–738. http://doi.org/10.1523/JNEUROSCI.4080-09.2010

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition*, *79*, 1–37.

http://doi.org/10.1016/S0010-0277(00)00123-2

Del Cul, A., Dehaene, S., Reyes, P., Bravo, E., & Slachevsky, A. (2009). Causal role of prefrontal cortex in the threshold for access to consciousness. *Brain*, *132*(9), 2531–2540. http://doi.org/10.1093/brain/awp111

Desender, K., Boldt, A., & Yeung, N. (2018). Subjective Confidence Predicts Information Seeking in Decision Making. *Psychological Science*, *29*(5), 761–778. http://doi.org/10.1177/0956797617744771

Desender, K., Van Opstal, F., & Van den Bussche, E. (2014). Feeling the Conflict: The Crucial Role of Conflict Experience in Adaptation. *Psychological Science*, *25*(3), 675–683. http://doi.org/10.1177/0956797613511468

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

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. http://doi.org/10.1016/j.neuroimage.2016.05.074

Endrass, T., Klawohn, J., Schuster, F., & Kathmann, N. (2008). Overactive performance monitoring in obsessive-compulsive disorder: ERP evidence from correct and erroneous reactions. *Neuropsychologia*, *46*, 1877–1887. http://doi.org/10.1016/j.neuropsychologia.2007.12.001

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*. http://doi.org/10.3758/BF03203267

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

Falkenstein, M., Hoormann, J., Christ, S., & Hohnsbein, J. (2000). ERP components on reaction errors and their functional significance: a tutorial. *Biological Psychology*, *51*(2–3), 87–107.

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

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

Greenhouse, S., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*(2), 95–112. http://doi.org/10.1007/BF02289823

Greville, W. J., & Buehner, M. J. (2010). Temporal Predictability Facilitates Causal Learning. *Journal of Experimental Psychology*, *139*, 756–771. http://doi.org/10.1037/a0020976

Haering, C., & Kiesel, A. (2016). Time perception and the experience of agency. *Psychological Research*, *80*(2), 286–297. http://doi.org/10.1007/s00426-015-0654-0

Hajcak, G., Nieuwenhuis, S., Ridderinkhof, K. R., & Simons, R. F. (2005). Error-preceding brain activity: Robustness, temporal dynamics, and boundary conditions. *Biological Psychology*, *70*, 67–78. http://doi.org/10.1016/j.biopsycho.2004.12.001

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

Hoonakker, M., Doignon-Camus, N., & Bonnefond, A. (2016). Performance monitoring mechanisms activated before and after a response: A comparison of aware and unaware errors. *Biological Psychology*, *120*, 53–60.

http://doi.org/10.1016/j.biopsycho.2016.08.009

Kouider, S., de Gardelle, V., Sackur, J., & Dupoux, E. (2010). How rich is consciousness? The partial awareness hypothesis. *Trends in Cognitive Sciences*, *14*(7), 301–307. http://doi.org/10.1016/j.tics.2010.04.006

Libet, B., Gleason, A. C., Wright, E. W., & Pearl, D. K. (1983). Time of Conscious Intention To Act in Relation To Onset of Cerebral Activity (Readiness-Potential). *Brain*, *106*(3), 623–642. http://doi.org/10.1093/brain/106.3.623

Libet, B., Wright, E. W., Feinstein, B., & Pearl, D. K. (1979). Subjective referral of the timing for a conscious sensory experience: A functional role for the somatosensory specific projection system in-man. *Brain*, *102*(1), 193–224. http://doi.org/10.1093/brain/102.1.193

Maidhof, C., Rieger, M., Prinz, W., & Koelsch, S. (2009). Nobody is perfect: ERP effects prior performance errors in musicians indicate fast monitoring processes. *PLoS ONE*, *4*(4). http://doi.org/10.1371/journal.pone.0005032

Maier, M. E., Di Gregorio, F., Muricchio, T., & di Pellegrino, G. (2015). Impaired rapid error monitoring but intact error signaling following rostral anterior cingulate cortex lesions in humans. *Frontiers in Human Neuroscience*, *9*(June), 1–15. http://doi.org/10.3389/fnhum.2015.00339

Maniscalco, B., & Lau, H. (2012). A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness and Cognition*, *21*(1), 422–430. http://doi.org/10.1016/j.concog.2011.09.021

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

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. http://doi.org/10.1027/0269-8803.19.4.319

Persaud, N., McLeod, P., & Cowey, A. (2007). Post-decision wagering objectively measures awareness. *Nature Neuroscience*, *10*(2), 257–61. http://doi.org/10.1038/nn1840

Rabbitt, P. (1968). Repetition effects and signal classification strategies in serial choice-response tasks. *The Quarterly Journal of Experimental Psychology*, *20*(March 2014), 232–240. http://doi.org/10.1080/14640746808400157

Rabbitt, P. (1978). Detection of Errors by Skilled Typists. *Ergonomics*, *21*(11), 945–958. http://doi.org/10.1080/00140137808931800

Rabbitt, P. (2002). Consciousness is slower than you think. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *55*(4), 1081–1092. http://doi.org/10.1080/02724980244000080

Ridderinkhof, K. R., Nieuwenhuis, S., & Bashore, T. R. (2003). Errors are foreshadowed in brain potentials associated with action monitoring in cingulate cortex in humans. *Neuroscience Letters*, *348*, 1–4. http://doi.org/10.1016/S0304-3940(03)00566-4

Rieger, M., & Bart, V. K. E. (2016). Typing style and the use of different sources of information during typing: An investigation using self-reports. *Frontiers in Psychology*, *7*(DEC). http://doi.org/10.3389/fpsyg.2016.01908

Rochet, N., Spieser, L., Casini, L., Hasbroucq, T., & Burle, B. (2014). Detecting and correcting partial errors: Evidence for efficient control without conscious access. *Cognitive, Affective and Behavioral Neuroscience*, *14*(3), 970–982. http://doi.org/10.3758/s13415-013-0232-0

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

Sergent, C., Baillet, S., & Dehaene, S. (2005). Timing of the brain events underlying access to consciousness during the attentional blink. *Nature Neuroscience*, *8*(10), 1391–1400. http://doi.org/10.1038/nn1549

Seth, A. K. (2008). Post-decision wagering measures metacognitive content, not sensory consciousness. *Consciousness and Cognition*, *17*(3), 981–983. http://doi.org/10.1016/j.concog.2007.05.008

Shea, N., & Frith, C. D. (2019). The Global Workspace Needs Metacognition. *Trends in Cognitive Sciences*, *23*(7), 560–571. http://doi.org/10.1016/j.tics.2019.04.007

Steinhauser, M., Maier, M., & Hübner, R. (2008). Modeling behavioral measures of error detection in choice tasks: response monitoring versus conflict monitoring. *Journal of Experimental Psychology. Human Perception and Performance*, *34*(1), 158–176. http://doi.org/10.1037/0096-1523.34.1.158

Steinhauser, M., & Yeung, N. (2010). Decision processes in human performance monitoring. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(46), 15643–53. http://doi.org/10.1523/JNEUROSCI.1899-10.2010

Trevena, J. A., & Miller, J. (2002). Cortical movement preparation before and after a conscious decision to move. *Consciousness and Cognition*, *11*(2), 162-190-325. http://doi.org/10.1006/ccog.2002.0548

Ullsperger, M., Fischer, A. G., Nigbur, R., & Endrass, T. (2014). Neural mechanisms and temporal dynamics of performance monitoring. *Trends in Cognitive Sciences*, *18*(5), 259–67. http://doi.org/10.1016/j.tics.2014.02.009

Ullsperger, M., Harsay, H. A., Wessel, J. R., & Ridderinkhof, K. R. (2010). Conscious perception of errors and its relation to the anterior insula. *Brain Structure & Function*, *214*(5–6), 629–43. http://doi.org/10.1007/s00429-010-0261-1

Wessel, J. R. (2012). Error awareness and the error-related negativity: evaluating the first decade of evidence. *Frontiers in Human Neuroscience, 6*(April), 1–16.

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

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*(10), 3021–36. http://doi.org/10.1162/jocn.2011.21635

Windey, B., & Cleeremans, A. (2015). Consciousness as a graded and an all-or-none phenomenon : A conceptual analysis. *Consciousness and Cognition*, *35*, 185–191. http://doi.org/10.1016/j.concog.2015.03.002

Winer, B. J. (1971). Use of analysis of variance to estimate reliability of measurements. In B. J. Winer (Ed.), *Statistical Principles in Experimental Design* (2nd ed., pp. 283–295). New York: McGraw-Hill.

Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychological Review*, *111*(4), 931–959. http://doi.org/10.1037/0033-295X.111.4.931

Yeung, N., & Summerfield, C. (2012). Metacognition in human decision-making: confidence and error monitoring. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*, 1310–1321. http://doi.org/10.1098/rstb.2011.0416