Error awareness and the error-related negativity: evaluating the first decade of evidence

Jan R. Wessel[1,2]*

[1] Max Planck Institute for Neurological Research, Cologne, Germany
[2] Psychology Department, University of California San Diego, San Diego, USA

From its discovery in the early 1990s until this day, the error-related negativity (ERN) remains the most widely investigated electrophysiological index of cortical error processing. When researchers began addressing the electrophysiology of subjective error awareness more than a decade ago, the role of the ERN, alongside the subsequently occurring error positivity (Pe), was an obvious locus of attention. However, the first two studies explicitly addressing the role of error-related event-related brain potentials (ERPs) would already set the tone for what still remains a controversy today: in contrast to the clear-cut findings that link the amplitude of the Pe to error awareness, the association between ERN amplitude and error awareness is vastly unclear. An initial study reported significant differences in ERN amplitude with respect to subjective error awareness, whereas the second failed to report this result, leading to a myriad of follow-up studies that seemed to back up or contradict either view. Here, I review those studies that explicitly dealt with the role of the error-related ERPs in subjective error awareness, and try to explain the differences in reported effects of error awareness on ERN amplitude. From the point of view presented here, different findings between studies can be explained by disparities in experimental design and data analysis, specifically with respect to the quantification of subjective error awareness. Based on the review of these results, I will then try to embed the error-related negativity into a widely known model of the implementation of access consciousness in the brain, the global neuronal workspace (GNW) model, and speculate as the ERN's potential role in such a framework. At last, I will outline future challenges in the investigation of the cortical electrophysiology of error awareness, and offer some suggestions on how they could potentially be addressed.

Keywords: consciousness, ERN, error awareness, event-related potentials, performance monitoring, cognitive control

INTRODUCTION: THE ERROR-RELATED NEGATIVITY

In the cognitive neuroscience of error processing, the discovery of an event-related brain potential (ERP) whose amplitude is different depending on the success or failure of an action was a groundbreaking step. Before Falkenstein and colleagues published the first peer-reviewed article about said potential in the human scalp EEG and termed it "Error Negativity" (Ne; Falkenstein et al., 1991 alternatively, and somewhat more commonly today called the "error-related negativity"; ERN, Gehring et al., 1993), the neuroscientific community was largely ignorant toward error processing, even though much of the experimental groundwork had been laid in the 1960s, prominently by Rabbitt and colleagues (Rabbitt, 1966, 1967). The discovery of this first measurable index of performance monitoring-related brain activity coincides with a continuously growing interest in the neuroscience of the more general area of cognitive control, signified by an exponential increase of publications in the field.

Since the 1990s, during which most of the studies about the ERN were published in journals focusing on behavioral rather than neuroscientific research, the differential properties of the ERN had been probed in a number of early studies. This early

empirical work culminated in the emergence of (at least) four main branches of theories of what exactly drives the ERN amplitude: the error detection or "mismatch"-theories (Falkenstein et al., 1991; Coles et al., 2001) postulate the amount of difference between an intended and the actually performed action as the main influence on the amplitude of the ERN, with the latter represented as early as in the motor efference copy. According to the reinforcement learning theories of the ERN on the other hand (Holroyd and Coles, 2002), this comparison is carried out on the subcortical level of the basal ganglia instead, whereas the amplitude of the ERN amplitude is influenced by a learning signal carried forward into the cortical generators of the ERN by the mesencephalic dopamine system. A third perspective of ERN functionality is offered by the conflict monitoring accounts (Botvinick et al., 2001; Yeung et al., 2004), which move away from the accuracy of the action per se as the main determinant of ERN amplitude. Instead they postulate the degree of motor response-conflict, i.e., the arithmetic product of the activation of the erroneous and correct response tendencies at the time of the response as the decisive factor in ERN amplitude. A last branch of theories implicate the perceived probability of the occurrence of

an error in a given experimental trial as the main determinant of ERN amplitude on that trial (Brown and Braver, 2005).

On the descriptive level, the ERN has a prominent fronto-central radial voltage distribution on the scalp and is consequently mostly quantified at electrode FCz in the extended 10–20 system of the EEG. Its neuronal generator has been located to the medial wall of the posterior medial frontal cortex (pMFC, Dehaene et al., 1994; Holroyd et al., 1998; Ullsperger and von Cramon, 2001; Gehring and Willoughby, 2002; Van Veen and Carter, 2002; Debener et al., 2005), the human homologue of the monkey rostral cingulate zone (RCZ, Ridderinkhof et al., 2004), a region also referred to as dorsal anterior cingulate cortex (dACC). It is followed by a complex of positive voltage deflections, commonly referred to as the error positivity (Pe, Falkenstein et al., 2000), which itself consists of at least two distinct components (late and early Pe, respectively) with partially dissociable features (Overbeek et al., 2005; Ridderinkhof et al., 2009).

The role of the ERN in subjective error awareness, i.e., the question of whether or not the ERN is related to humans' conscious awareness of the accuracy of their own action, had not been studied until 10 years after the initial discovery of the ERN. The relation between a neuronal correlate of error processing on the one hand, and the emergence of explicit awareness of one's own errors on the other hand is of pressing interest for the cognitive neurosciences of cognitive control, as the subjective perception of errors has obvious implications for remedial actions following errors (e.g., with respect to immediate corrective behaviors, learning from errors, or other behavioral adaptations, particular such that are in any sense intentional). Ultimately, one would want to be able to exploit the neuronal correlates of error processing for everyday life, e.g., in the context of brain-computer interfaces that inform a person of whether an error was made or not, which is why it is very important to identify which neuronal correlates influence the emergence of the subjective, spontaneous realization of having committed an error. The ERN is a prime candidate for this as it is (a) chronologically the first physiological manifestation of error-related processing following the response, peaking in the first 50–100 ms after an errors, (b) unlike the Pe, for which there exist many source localization attempts with quite variable results, it is reliably located to a specific, very circumscribed part of cortex, and (c) there is a huge body of literature about which factors influence the ERN per se, making it interesting if and how these factors are related to subjective error awareness.

The first study that explicitly probed the ERN's sensitivity toward the degree of subjectively perceived accuracy was published in 2000 (Scheffers and Coles, 2000). It was followed by the emergence of a complex and ambivalent picture in subsequent studies of subjective "error awareness," which either backed up the general finding of that initial study, which was that the processes underlying the ERN influence the subjective certainty of error perception, or seemingly contradicted it. As a matter of fact, just a year later, an influential study (Nieuwenhuis et al., 2001) failed to find a difference in ERN amplitude with respect to subjective error awareness. In the following, I will review the first decade of studies that dealt with the ERN and subjective error awareness, and try to find underlying factors that might contribute to either view. First, however, I will try to characterize and define what is meant by "error awareness" in a philosophical and empirical sense.

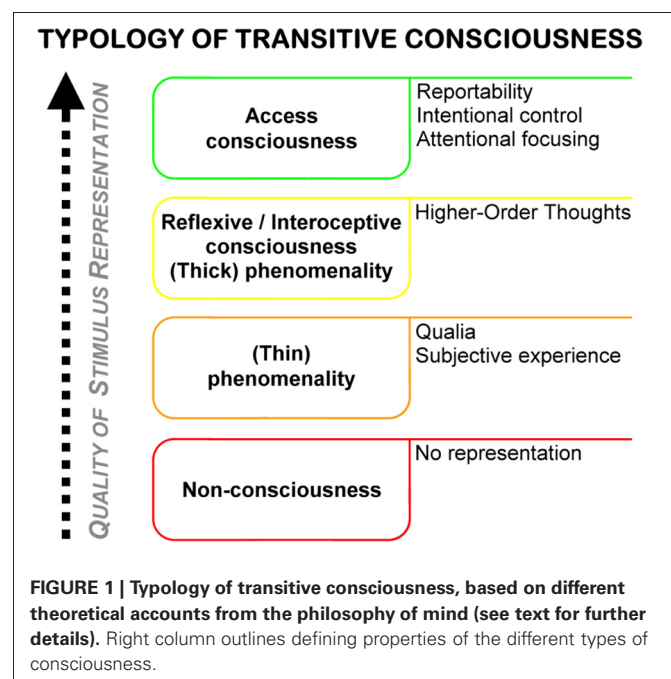## AWARENESS AND CONSCIOUSNESS: SOME DEFINITIONS

In order to be able examine error awareness and its influence on the brain processes that underlie performance monitoring (or any brain process that could potentially be influenced by awareness and vice versa) one must first define what exactly is meant by (error) "awareness."

Consciousness and subjective awareness lie at the core of the discipline of philosophy of mind. As will be seen later on, what researchers mostly meant by "awareness" in the context of subjective error perception is called "access consciousness" in that branch of philosophy (Block, 2007). Access consciousness is defined as follows

> "A mental state is *access conscious* when a subject has a certain sort of access to the content of the state. More precisely, a state is access conscious if by virtue of having the state, the content of the state is available for verbal report, for rational inference, and for the deliberate control of behavior."
>
> (Bayne and Chalmers, 2003, p. 6)

Access consciousness is characterized as the highest quality of representation in transitive (object-related) consciousness. The concept of accessibility, which is at the center of what characterizes an access conscious state, is in practice mainly operationalized by reportability, i.e., the availability of the presence of a stimulus for spontaneous verbalization by the (cognitive) system. Access consciousness and other types of transitive consciousness can be distinguished on the basis of the strength and quality of the subjective representation of a either a stimulus in a system or an internal state of a system in that system itself (see **Figure 1**).

FIGURE 1 | Typology of transitive consciousness, based on different theoretical accounts from the philosophy of mind (see text for further details). Right column outlines defining properties of the different types of consciousness.

The degree of awareness of the presence of a certain stimulus is a good example for illustration: a (cognitive) system can be completely ignorant with respect to the presence of a stimulus, with no evidence of processing being present at any stage of the system. In such a case, the stimulus would consequently be classified as being "unperceived" in the narrow sense; the system would be non-conscious of it. The minimum of representation that must be evident in a system to indicate a type of consciousness is what constitutes phenomenological consciousness (Block, 1995), or phenomenality (Rosenthal, 2002). Quantifying this representation is called the "hard problem" of consciousness (Chalmers, 1995), or the problem of "qualia" (i.e., the "redness of red," Crick and Koch, 2003), as opposed to the "easy problem" of consciousness, which is the problem of access consciousness ("easy" problem presumably because access consciousness is relatively easily quantifiable on the basis of overt behavior/verbal reporting). A fourth kind of conscious state is called reflexive consciousness by Block (synonyms: monitoring/interospective consciousness Block, 2001), and is characterized by the presence of Higher-Order Thoughts (Rosenthal, 2002), i.e., "thoughts about thoughts." This ipsoreflexive quality distinguishes reflexive consciousness from mere phenomenality (or "thick" from "thin" phenomenality in Rosenthal's terminology, where thick phenomenality is a synonym for what Block calls reflexive consciousness, and thin phenomenality is phenomenality in Block's original sense). Importantly, (thin) phenomenanilty is indistinguishable from non-consciousness both empirically and for the system itself[1].

Beyond being able to formulate a clear working definition of what one is researching on, what is interesting about these formal and theoretical classifications for empirical performance monitoring research, is the question of what is potentially examinable using the battery of methods available to psychological and neuroscientific research. Research in the area of error awareness usually employs behavioral procedures aimed at an operationalization of access consciousness (in a sense that subjects are mostly presented with the computerized version of a verbal report, i.e., the pressing of a button to indicate conscious availability). However, reflexive consciousness ("gut feelings") is also potentially examinable using standard experimental psychological methods. The methodological repertoire of research on so called "meta-cognitive feelings" (Koriat, 2007), i.e., feelings of the presence of a certain state in absence of the ability to explicitly fully characterize its nature, can potentially be utilized in error awareness research as well, e.g., by using wagering procedures (Persaud et al., 2007, see "Future directions" for more details). Also, a big virtue of neuroscientific compared to behavioral methods is that it is theoretically possible to detect the representation of a stimulus in the absence of any higher-order thought or access consciousness. For example, stimulus-evoked activity in primary sensory areas like V1 or the primary auditory cortex might well be a physiological manifestation of "thin" phenomenality, which is per definition unexaminable using behavioral methods.

---

[1]This begs the question if it is a valid state of what would commonly be called "consciousness" to begin with, as it appears to be more of a theoretical construct (Rosenthal, 2002).

For the purposes of this review, unless otherwise declared, I will talk about access consciousness when referring to (error) awareness. What distinguishes "consciously perceived/aware errors" from "non-consciously perceived/unaware errors" is reportability: is the subject able to report the inaccuracy of its action or not? Since there is also an ambiguity in the literature concerning the naming of error types depending on the presence or absence of access consciousness, I will refer to errors with access consciousness as "reported errors" (REs) and to errors in the absence of access consciousness as "non-reported errors" (NREs), unless otherwise specified.

ERROR AWARENESS AND THE ERN: A CHRONOLOGY

In this paragraph, I will introduce and discuss the studies that reported findings with respect to the influence of ERN amplitude on subjective error awareness (or vice versa). This paragraph should give a comprehensive overview that outlines the respective details and findings of these studies. A summary of these details can be found in **Table 1**.

The first study that explicitly addressed the sensitivity of the ERN amplitude to subjective error awareness was published in 2000 by Scheffers and Coles (2000). The authors presented subjects with a letter version of the classic flanker paradigm (Eriksen and Eriksen, 1974). After each trial, they prompted subjects to rate their confidence in their response on a five-point scaling ranging from "sure correct" to "sure incorrect," with a neutral "don't know" rating in between. They carried out two main analyses to address the question of the influence of error awareness on the ERN. The first analysis compared ERN amplitudes between all five confidence ratings, showing that ERN amplitude increased with growing error awareness. This result was confirmed in a second analysis which focused only on the three rating bins "don't know", "not sure incorrect", and "sure incorrect," as only eight participants had sufficient error numbers to warrant inclusion in the full analysis. Even more so: the same pattern appeared to be true for the negativity on correct trials that were examined in the full analysis (correct-related negativity, CRN, Vidal et al., 2000; Roger et al., 2010): the larger the ERN/CRN, the more the subjects consciously felt that they had made an error, even on correct trials. It has to be said that the CRN and ERN represent the activity of the same underlying neuronal network (Roger et al., 2010), and therefore, ostensibly reflect the same process.

This seemingly clear cut pattern of results was subsequently contradicted just 1 year later, though: Nieuwenhuis et al. (2001) published results from an eye-movement experiment, an anti-saccade task (AST), which demonstrated a null effect of error awareness on the ERN. In the anti-saccade paradigm, subjects must inhibit a prosaccade to a target stimulus appearing on one side of the screen and initiate an "anti"-saccade to the opposite site. Similarly (but not identically) to Scheffers and Coles (2000), Nieuwenhuis and colleagues prompted their subjects to assess the accuracy of their action after each trial: subjects had a limited time following the onset of the display of a cross on the correct side of the screen in order to press a button when they thought they had committed an erroneous prosaccade to the wrong side of the screen. Whereas the error positivity was significantly enlarged for

Table 1 | Details of the studies that report testing of ERN amplitude differences for reported vs. non-reported errors, either as part of their main hypothesis or as auxiliary analyses.

Study	Year	Task	Awareness signaling	N	p(RE = NRE)	Statistical test	Additional information
Scheffers and Coles (all)	2000	Flanker task (letter version)	Five-point scale ranging from "surely incorrect" to "surely correct"	8	**0.005**	ANOVA (two-sided)	
Scheffers and Coles (partial)	2000			15	**0.002**		"Don't know" to "surely incorrect"
Nieuwenhuis et al.	2001	Anti-saccade task	Awareness button (1250 ms time)	15	0.28	ANOVA (two-sided)	
Endrass et al.	2005	Oculomotor stop-signal task	Binary rating (1300 ms time)	20	N.A.	ANOVA (two-sided)	Trials without a rating were potentially discarded
Endrass et al.	2007	Anti-saccade task	Binary rating with an "unsure" option (press both buttons)	19	0.55	t-test (two-sided)	
O'Connell et al.	2007	Manual Go-NoGo Task, visual stimuli	Awareness button on next trial, abolish Go response	12	0.872	ANOVA (two-sided)	Minimum for inclusion: 20 errors of both types (initial N = 19)
Maier et al.	2008	Flanker task (letter version) with additional neutral stimuli	Awareness button (1200 ms time, including RT on primary task)	14	**<0.001**	ANOVA (two-sided)	
Shalgi et al.	2009	Manual Go-NoGo Task, auditory stimuli	Awareness button on next trial, abolish Go response	16	0.187	t-test (two-sided)	
Woodman	2010	Visual search with non-masked and masked stimuli	N2pc, binary rating	7	**<0.01**	ANOVA (two-sided)	
Steinhauser and Yeung	2010	Visual pattern discrimination	Awareness button (1000 ms time)	16	**0.046**	t-test (two-sided)	
Hughes and Yeung	2011	Flanker task (arrow version) with additional masked stimuli	Awareness button (1000 ms time)	8	**0.086**	t-test (two-sided)	Minimum for inclusion: 6 errors of both types (initial N = 20)
Wessel et al. (Exp. 1)	2011	Anti-saccade task	Binary rating	17	**0.027**	ANOVA, planned contrast	
Wessel et al. (Exp. 2)	2011	Anti-saccade task	Binary rating (with post-hoc "sureness"-quantification based on rating times)	17	**0.018**	ANOVA, planned contrast	
Hewig et al.	2011	Semi-blind digit-entering	Three-point scale ranging from "surely incorrect" to "surely correct"	16	**0.003**	ANOVA, *post-hoc*	
Dhar et al.	2011	Manual Go-NoGo Task, visual stimuli	Awareness button (1500 ms time)	14	0.467	t-test (two-sided)	No significant ERN-CRN difference for either error type. ERN source (pCMA) has RE > NRE effect, **p = 0.004**

pCMA, posterior cingulate motor area; p(RE = NRE), probability of the null hypothesis of equal ERN amplitudes between reported and non-reported errors; ANOVA, analysis of variance.

reported as compared to non-reported errors, the ERN, contrary to Scheffers and Coles findings, was not.

Surprisingly, in the 4 years after these two initial studies, there were no further publications that tried to explain the disparity between them. Following a 2003 study by Dehaene et al. (2003), which found conflict-related effects in the dorsal ACC/RCZ, the neuronal generator of the ERN, only for unmasked conflicting primes as compared to fully masked primes, Mayr (2004) concluded

> "There is some convergence across studies in that awareness seems crucial [. . .] for indications of ACC-related activity. At the same time, enough inconsistencies remain to preclude any firm conclusion in this regard."
> (Mayr, 2004, p. 147, references removed from original text)

Mayr cites Scheffers and Coles (2000) study, alongside Dehaene et al. (2003) and another fMRI study (Stephan et al., 2002) as evidence for the first part of this statement, whereas Nieuwenhuis et al. (2001) study serves as reference for the second part.

It took until 2005 until the issue was addressed again, when Endrass et al. (2005) published data from a third type of paradigm, a stop-signal task in the oculomotor domain, which also introduced another slightly different scoring method for error awareness: similar to Scheffers and Coles (2000), people had to indicate their perceived response accuracy in both cases (errors and correct trials), but as in Nieuwenhuis et al. (2001), the rating was binary (error or correct, as compared to the five-point scale employed by Scheffers and Coles) and people had only limited time to make their assessment. In this stop-signal experiment, Endrass and colleagues again reported a null-finding with respect to the ERN and error awareness.

Comparable results were obtained in the two next studies dating from 2007 (Endrass et al., 2007; O'Connell et al., 2007). The 2007 study by Endrass and colleagues employed a similar AST as Nieuwenhuis et al. (2001), but the rating procedure was identical to their previous study (Endrass et al., 2005), with the exception that this time, the response to the accuracy-prompt was not under time pressure. O'Connell et al. (2007) combined EEG with concurrent measurements of autonomic nervous system (ANS) activity, as measured by the skin-conductance response (SCR). They also employed a novel paradigm into the study of the effects of error awareness on the ERN, that has been previously used in the fMRI domain by Hester et al. (2005) to probe the activity of the RCZ on reported and non-reported error trials (see below). They employed a Go-Nogo paradigm with Stroop-like stimuli (color-words in different ink color, Stroop, 1935) that they called "error awareness task" (EAT). In the EAT subjects have to perform a Go-response (button-press) unless one of two NoGo-situations is encountered: (1) a mismatch between word-ink and meaning of the word (Stroop NoGo); (2) a repetition of the previous word (Repeat NoGo). With those two complex rules, one engaging the psychological processes associated with the Stroop effect and the other engaging working memory effects similar to a one-back task, a sufficiently high number of non-reported errors can be achieved (a methodological problem in all error awareness studies) to warrant statistical comparison. The rating procedure to assess subjectively perceived accuracy was also arguably more

complex than in previous paradigms: in case subjects thought they made an error (i.e., a Go-response in one of the two NoGo-situations), they had to abolish the Go-Response on the next trial and press an error-awareness button instead. Both these studies (Endrass et al., 2007; O'Connell et al., 2007) failed to find an error awareness effect on ERN amplitude, speaking in favor of the ERN being unrelated to subjective error awareness, and contradicting the initial findings of Scheffers and Coles (2000). Also, the findings of O'Connell et al. (2007)[2] were later replicated in a slightly larger sample using auditory cues by Shalgi et al. (2009).

To add to the apparent confusion, however, in the last 4 years, seven more studies were published which all, to different extents, apparently backed up the findings of Scheffers and Coles (2000), reporting differences in ERN amplitude or source level RCZ activity between reported and non-reported errors. The closest replication of Scheffers and Coles' findings with respect to experimental conditions was done by Maier et al. (2008), who also used a letter version of the flanker task. However, they employed the rating procedure from Nieuwenhuis et al. (2001), having people press an "error awareness button" in case of a reported error. They found highly significant differences in ERN amplitude with respect to subjective error awareness.

In 2010, Steinhauser and Yeung (2010) manipulated subjects' incentives to either signal or not signal an error, effectively introducing two different response-bias conditions. They could show that it is primarily the error positivity that represents the input variables of the decision process that leads to signaling or not signaling an error, but they also found differences between reported and non-reported errors in the overall ERN in their perceptual discrimination task, with ERN amplitude being significantly increased for reported errors. That same year, Woodman (2010) published a study that differed from all previous studies to certain extent. Not only did he introduce a previously unseen paradigm into the error awareness literature (a visual search paradigm with masked or non-masked stimuli), but he also introduced a special quantification of awareness. The main task was to detect the presence of a stimulus in a visual search array by pressing a button when it was perceived as present in the array and another when it was supposedly absent. The stimulus was either masked by simultaneous-offset mask, or by delay-offset mask, with the latter reducing overt stimulus detection to chance level, whereas the simultaneous-offset mask left aware stimulus perception intact. It could be shown that an ERN was only elicited in the condition in which the mask did not disturb conscious stimulus perception (simultaneous-offset mask), whereas it was absent in the delayed-masking, pre-conscious condition. Furthermore, and most interestingly, an N2pc wave could be seen on target trials in either condition, irrespective of masking condition. The N2pc is an index of a shift in visuo-spatial attention following the presence of target stimuli (Luck and Hillyard, 1994). In essence, this shows a dissociation between intact target-stimulus representation (as indexed by the N2pc) and performance monitoring (as indexed by the ERN), possibly also dissociating a neural

---

[2]O'Connell et al. also reported another null-finding with respect to error awareness and ERN amplitude in O'Connell et al. (2009), yet the sample in that study was overlapping with the sample used in O'Connell et al. (2007).

correlate of classic access conscious "awareness" of an error and phenomenologically conscious representations of a stimulus (see above). It also provides evidence that the ERN is related to the quality of awareness of an error.

The year 2011 brought four more studies that measured ERN amplitude in error awareness experiments. Hughes and Yeung (2011) tried to dissociate response-conflict from error awareness using a flanker task with additional masked stimuli. They reported a null-finding with respect to error awareness and ERN-amplitude in a limited sample[3]. They did, however, find an association between ERN amplitude and error awareness on a single-trial level, which larger ERN amplitudes being beneficial for error awareness. In yet another recent study that investigated concurrent EEG and ANS measurements (heart rate and pupil diameter) during error awareness, our group (Wessel et al., 2011) reported a significantly enlarged ERN amplitude for reported compared to non-reported errors in the anti-saccade experiment, alongside differential effects of error awareness on both heart-rate and pupil diameter. In the first experiment, we used a binary rating for the assessment of error awareness, similarly to Endrass et al. (2007). In a second experiment, we tried to replicate the findings using the exact same stimulus layout and timing as in the first study of error awareness in the AST (Nieuwenhuis et al., 2001). Instead of the awareness button used in their study, however, we used a twofold procedure to get a more detailed picture of the degree of error awareness in this experiment. To that end, we used the same binary rating as in the first experiment, i.e., subjects had to push a button when they thought they made an error and a different button when they thought they did not. Then, we subsequently split the experimental trials for each subject and error type in half, based on the time it took for the subject to make the assessment of their own accuracy. This was done with the rationale that ratings that were made very fast were made with a higher degree of certainty than those which took the subjects longer to make. Not only did we again find a significantly enlarged ERN for reported compared to non-reported errors, but we also found that almost all of this difference was explained by the subsample of aware errors that was signaled very quickly, i.e., with high certainty, again providing evidence that ERN and error awareness are directly or indirectly related. Another recent study backed up this finding (and earlier ones that found an enlarged ERN for reported errors), this time using another novel task: Hewig et al. (2011) used a semi-blind digit-entering task and a three-point rating scale ("correct", "unsure", "incorrect") after each trial and found significant ERN-CRN differences exclusively for incorrect trials judged "incorrect," i.e., reported errors. "Unsure" and "correct"-rated error trials did not differ from their respective correct counterparts, confirming the results from the rating-reaction-time split in Experiment 2 in Wessel et al. (2011).

---

[3]However, as noted by the authors in the discussion, the low number of samples hampers the acceptance of a null-finding in this study. This is especially true since, even despite the low sample size, the significant tendency ($p = 0.086$, two-sided) would turn into a positive finding if tested in a one-sided fashion [which would be justified in principle, in light of the previous results from flanker studies of error awareness, i.e., Scheffers and Coles (2000) and Maier et al. (2008)].

To this day, the latest study regarding the cortical electrophysiology of error awareness (Dhar et al., 2011) did not explicitly focus on ERPs, but rather on EEG source imaging. Dhar and colleagues had subjects perform a visual Go-NoGo task with the option of pressing an awareness button whenever subjects felt they made an error. Even though they did not find a significantly enlarged ERN for reported errors compared to non-reported errors at FCz (in fact, there was no difference between either error trial and correct trials at FCz, i.e. no ERN), they did find significant differences in that direction at more left-lateralized frontal electrode sites, which is in line with their left-lateralized source-solution for the ERN in the left posterior cingulate motor area (lPCMA, MNI coordinates: $x = -5\ y = -15\ z = 55$) and also with the voltage distribution of the ERN in their study (see **Figure 2** in their manuscript). Consequently, the activity in the lPCMA source was significantly enlarged on reported errors as compared to non-reported errors in their study.
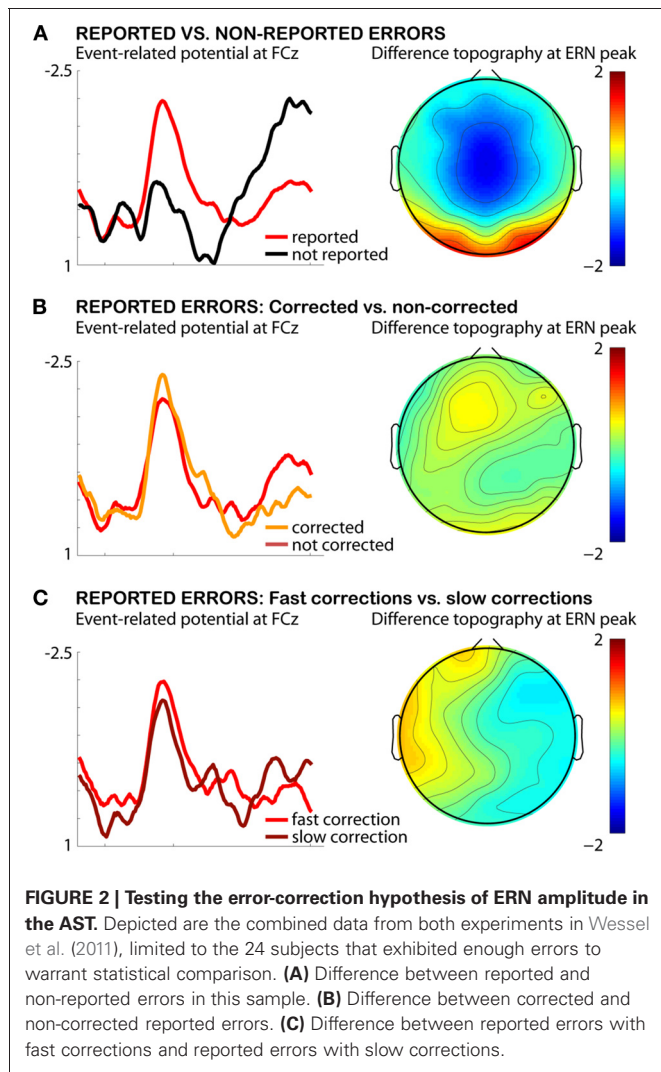
As is evident, there is considerable disparity between studies as to whether error awareness is unrelated to the ERN (or vice versa) or not. Whereas there are several findings that strongly point to the fact that the ERN does coincide with higher degrees of error awareness (Scheffers and Coles, 2000; Maier et al., 2008; Steinhauser and Yeung, 2010; Woodman, 2010; Dhar et al., 2011; Hewig et al., 2011; Wessel et al., 2011), there are enough null-findings to shy away from too optimistic inferences (Nieuwenhuis et al., 2001; Endrass et al., 2005, 2007; O'Connell et al., 2007; Shalgi et al., 2009).

STUDIES OF THE ERN IN ERROR AWARENESS: COMMONALITIES AND DIFFERENCES

Because of the discrepancies in findings between studies, it is essential to review the commonalities and differences in these studies (the details of each study are listed in **Table 1**), and look for common patterns that might explain either finding, which I will do in the following.

FACTORS OF THE TASK: DIFFERENT PARADIGMS, DIFFERENT FINDINGS?

The paradigms used to investigate error awareness in relation to the ERN and Pe span many of the central paradigms of performance monitoring or cognitive control research in general. Of the abovementioned 13 studies addressing the topic, three utilize variants of the classic flanker task (Scheffers and Coles, 2000; Maier et al., 2008; Hughes and Yeung, 2011), four use Go-NoGo or stop signal paradigms (Endrass et al., 2005; O'Connell et al., 2007; Shalgi et al., 2009; Dhar et al., 2011), and three use the anti-saccade task (AST, Nieuwenhuis et al., 2001; Endrass et al., 2007; Wessel et al., 2011), which is essentially a combination of a Go-NoGo like paradigm and a forced choice reaction time task like the flanker task (in that one has to countermand an automatic response tendency and subsequently initiate another response). The three remaining studies used a visual discrimination task (Steinhauser and Yeung, 2010), a digit-entering task (Hewig et al., 2011), and a masked visual search paradigm (Woodman, 2010). One apparent tendency is that stop-signal/Go-NoGo studies (with the exception of Dhar et al., 2011) generally tend to yield null-findings, whereas flanker

FIGURE 2 | Testing the error-correction hypothesis of ERN amplitude in the AST. Depicted are the combined data from both experiments in Wessel et al. (2011), limited to the 24 subjects that exhibited enough errors to warrant statistical comparison. **(A)** Difference between reported and non-reported errors in this sample. **(B)** Difference between corrected and non-corrected reported errors. **(C)** Difference between reported errors with fast corrections and reported errors with slow corrections.

findings yield enlarged ERN amplitudes for reported compared to non-reported errors. The picture is less clear for the AST: whereas Nieuwenhuis et al. (2001) and Endrass et al. (2007) demonstrated null-findings; both experiments in Wessel et al. (2011) showed the error awareness amplitude effect for the ERN. All studies using other paradigms show significantly enlarged ERN amplitudes on reported errors.

While there seems to be a pattern in that studies using a task with a Go-NoGo/stop-signal component tend to yield null-effects whereas other tasks show enlarged ERN amplitudes for reported errors, it is hard to find an explanation for this. One reason might lie in the quantification of error awareness itself, or in the low ERN amplitudes and general effect sizes in these paradigms, both of which will be reviewed later on in this section. First, I will review two hypotheses concerning primary task performance (stimulus perception and error correction) that have recently been put forward as potentially influential in producing the presence or absence of ERN amplitude effects in error awareness experiments.

STIMULUS DEGRADATION AS POTENTIAL DETERMINANT OF ERN AMPLITUDE DIFFERENCES

It has been argued that degraded stimulus perception might underlie the lower ERN amplitude on non-reported errors (Steinhauser and Yeung, 2010), based on the fact that some of the studies that reported null-findings used either masking procedures (Maier et al., 2008) or degraded the stimulus material in order to obtain enough non-reported errors to warrant statistical comparison (Scheffers and Coles, 2000; Steinhauser and Yeung, 2010). However, more recent studies do demonstrate these differences in the absence of degraded or masked stimulus material (Dhar et al., 2011; Hewig et al., 2011; Wessel et al., 2011). Also, the dissociation between stimulus perception on the neuronal level (as quantified by the N2pc) in such masking paradigms on the one hand and error awareness effects on the ERN on the other hand (Woodman, 2010) speaks against the fact that degraded stimulus perception is the only influence that causes ERN differences between error types in error awareness experiments. "Objective" evidence of neuronal stimulus representation was identical between error types in that study.

Unless subjective awareness of the stimulus material itself is a determinant of ERN amplitude, which would be assuming a direct connection between ERN and (error) awareness, differences in stimulus representation seem unlikely as the exclusive determinant of ERN amplitude in error awareness studies.

ERROR CORRECTION: DIFFERENCES BASED ON AWARENESS AND THEIR POTENTIAL INFLUENCE ON THE ERN

Another explanation for the discrepancies between studies has been put forward by Steinhauser and Yeung (2010). They argue that

> "Ne/ERN amplitude should be determined primarily by variations in primary task performance rather than variations in error signaling. [. . .] Thus, the ERN increase for detected errors may not reflect its direct role in error processing, but might instead be a by-product of the fact that detected errors tend to occur when fast guess responses are subsequently corrected (cf. Scheffers and Coles, 2000), resulting in high levels of conflict. This interpretation is consistent with evidence from the anti-saccade task that Ne/ERN amplitude is similar for detected and undetected errors that are always corrected (Nieuwenhuis et al., 2001), although in some studies this relationship is less clear (Endrass et al., 2007)."
>
> (Steinhauser and Yeung, 2010, p. 15651)

It is in line with the evidence from the error awareness experiments that primary task performance does influence ERN amplitude [see later section: errors in the global workspace: the accumulating evidence (AE) account]. However, even though there is evidence from ERN studies not focusing on error awareness that error correction influences ERN amplitude (Rodriguez-Fornells et al., 2002), there is evidence that the instruction to explicitly withhold or carry out error correction tampers with the expectation of error likelihood, error significance (Fiehler et al., 2005), or a reduced motor threshold that account for differences in ERN amplitude found in these studies (Ullsperger and von Cramon, 2006) and are not directly related to error awareness.

In addition, behavioral findings across studies contradict the proposition that the ERN amplitude reflects additional response-conflict that results from the presence or absence of a corrective response (it should, however, still be influenced by "classic" response-conflict at the time of the response, cf. Yeung et al., 2004; Danielmeier et al., 2009). Steinhauser and Yeung mention that evidence for the error-correction hypothesis from the AST in Nieuwenhuis et al. (2001), who found identical error rates for both types of errors and also identical ERN amplitudes, is contradicted by the AST results from Endrass et al. (2007). In the latter study, a dissociation between error correction rate and ERN amplitude was found: significantly fewer reported errors than non-reported errors were subsequently corrected, despite identical ERN amplitudes. This pattern of behavioral results was confirmed in both AST experiments in Wessel et al. (2011), further contradicting the influence of corrective saccades on ERN amplitudes in error awareness AST studies. Also, the same pattern of results might also be present in Nieuwenhuis et al. (2001) data[4], speaking against the error correction as lone determinant of the ERN amplitude differences found in error awareness experiments. Based on significant differences in corrective saccade latency relative to the response, which is shorter for non-reported errors in all three studies, it seems that in actuality, non-reported errors are the ones that are corrected in a quick and automatic fashion. Following a response-conflict based rationale, this pattern of results would actually lead to the prediction of enlarged ERN amplitudes for non-reported errors, if the presence or absence or timing of a potential error correction would be the primary influence on ERN amplitude.

In addition to these arguments, I will in the following present empirical evidence against the influence of error correction (both frequency and speed of correctional saccades) on the ERN amplitude result found in our study (Wessel et al., 2011). **Figure 2A** displays a re-analysis of the reported errors from both datasets used in Wessel et al. (2011, see manuscript for details on the AST and details on data processing), split by whether they were corrected or not. Only 24 out of 34 participants rendered enough aware errors in both conditions (corrected and not corrected, threshold at a minimum of five reported errors in each condition), but for the present purposes, this sample size is sufficient to warrant a sufficiently low beta-error probability to enable the testing of a null hypothesis. As can be seen from **Figure 2A**, there

is no difference in ERN amplitude based on error correction in reported errors: $t_{(23)} = -0.2815$, $p > 0.7$. Also, as can be seen from **Figure 2B**, there is no difference between fast and slow corrections in reported errors (median split of correction times): $t_{(23)} = 0.6739$, $p > 0.5$.

MEASURING ERROR AWARENESS: WHAT IS AN "AWARE" ERROR?

As seen above, performance on the primary task itself does not seem to be able to account for the differences in findings. One interesting possibility is that the measurement of awareness/access consciousness itself could be a decisive factor instead. There are several different quantifications of access consciousness in studies examining error awareness and the ERN, presumably all aimed at the same process. Procedures differs in certain core aspects: (a) difference in signaling between errors and correct trials, (b) the scaling of the quantification (binary vs. parametric), (c) the presence or absence of a neutral option, and (d) the presence or absence of a time-limit to rate one's accuracy.

There is an even split between studies using a forced-choice rating (i.e., a button has to be pressed for both errors and corrects) and an error-signaling only (i.e., a button has to be pressed for errors only; nothing has to be done on subjectively correct trials). Seven studies use the latter approach, whereas seven other experiments (counting Experiment 1 and 2 from Wessel et al., 2011, as two separate experiments) use a forced choice rating. Amongst the studies using an "awareness button" are all studies using Go-NoGo paradigms. All studies using the "awareness button" method naturally set a time-limit for the subjects to make their decision (ranging from 1000 to 1500 ms), whereas all but one (Endrass et al., 2005) studies using forced-choice rating give subjects unlimited time to come up with their decision (the tasks will not commence until a decision for a trial has been made).

Strikingly, these methods of quantification potentially lead to different classifications of certain errors in terms of whether they count as reported/perceived or not. In a forced choice rating situation, subjects can still fully evaluate their (uncertain) situation and might still signal the error, or judge it as a "don't know" trial, if that category is present. When using an error awareness button, however, after a certain amount of time, the next trial will start and the previous trial will be marked as "participant thought he/she was correct," i.e., as an non-reported error, even though there might have been some residual error awareness, which then effectively contaminates the measurement. A good demonstration for this fact comes from examining false alarm rates in the different studies. False alarms in this scenario are rare events when subjects signal their correct responses as erroneous. A direct comparison is possible in the AST experiments: Nieuwenhuis et al. (2001), who used an awareness button, yielded a false alarm rate of 1.5%. Experiment 2 in Wessel et al. (2011), which used the exact same primary stimulus layout and task timing as Nieuwenhuis et al. (2001), but exchanged the awareness button rating with a forced choice rating, yielded a false alarm rate of 9.8%. This demonstrates that the usage of an awareness button not only potentially contaminates the "non-reported" errors with errors with residual access consciousness, but it also introduces a response bias toward not signaling an error. This is not

---

[4]Nieuwenhuis et al. (2001) show a plot of size and speed of the corrective saccades in their manuscript (**Figure 1** therein), depicting corrective saccades in the latency-ranges from 0 to 1200 ms following the response. In the design of their version of the AST [unlike the AST variants employed in Endrass et al. (2007) and Wessel et al. (2011)], a white cross was displayed on the correct side of the screen (opposite of the imperative stimulus) 1000 ms after the onset of the imperative stimulus. Based on RTs of 194 ms and 200 ms for reported and non-reported errors, respectively, this means that on average, the white cross was displayed around the 800 ms mark in their corrective-saccades plot, rendering the saccades following that onset prosaccades to the now-present target rather than spontaneous, endogenous corrections of the error. Given that there are visibly more corrective saccades depicted in these latency ranges in the aware errors, even though there were significantly more non-reported errors on absolute, this speaks in favor of the fact that also in their study, just like in Endrass et al. (2007) and Wessel et al. (2011), there might have been more corrections on non-reported errors than on reported errors.

only so because of the fact that unsure situations, where deciding to signal an error might take more time than allowed would be rated as "participant thought he/she was correct," but also simply because signaling an error by pushing a button is more effortful than not signaling an error by doing nothing.

While the usage of an awareness button is probably a suboptimal procedure, it cannot alone explain the differences between studies. Not only do two out of the seven studies using forced choice ratings demonstrate null-findings with respect to ERN amplitude (Endrass et al., 2005, 2007), but also, significantly enlarged ERN amplitudes on aware errors can be observed in three out of the seven studies using the awareness button (four if counting Hughes and Yeung, 2011). Ultimately, when deciding which quantification of consciousness to choose, one is faced with the decision of whether (a) one wants to have a set of non-reported errors that are clear of any sort of residual (potentially reflexive/interoceptive) conscious representation (in which case a forced choice rating is the method of choice), or (b) one wants to have a set of reported errors that include only very "highly" (access-) conscious errors and in turn risk contaminating the "unaware" errors with potentially reflexively conscious errors. However, a solution to this problem might lie in using a finer scale than a parametric yes/no rating (which some studies have done, e.g., Scheffers and Coles, 2000; Hewig et al., 2011). Be aware, though, that if choosing between a forced choice rating and an "awareness button" procedure, a forced choice is probably the better option, because it does not introduce a response bias toward signaling or not signaling an error.

Since the method of quantification of an "aware" error cannot on its own account for the different findings (see above), another issue has to be taken into consideration, which is the question of type-2 error probability, i.e., the probability of accepting a null hypothesis, even though the alternative hypothesis is true.

FACTORS OF ANALYSIS: WHEN IS A NULL-FINDING A NULL FINDING?

The question of type-2 error probability is a classic topic in introductory statistics, but is often neglected in many studies, especially in the (cognitive) neurosciences. A high probability of committing a type-2 error stems from either low-power, low effect sizes, or a combination of the two. Low power mostly results from small sample sizes used to test a null hypothesis. This is a common problem in the neurosciences in particular, because data acquisition is an expensive, time-consuming procedure, which oftentimes limits sample sizes of such studies to fewer than 20 samples. The average sample size of the ERN-error awareness studies reviewed so far is 14.7. The sample size of the six studies officially demonstrating null effects is 14.1. A lot of studies do find marked numeric differences in neuronal activity that would replicate the early findings of Scheffers and Coles (2000), but fail to find significances presumably because of low sampling size. I have already mentioned the low sample size in the null-finding from one study (Hughes and Yeung, 2011) as an example. Since no major inferences in that study were based on this result, and the authors outline the limited sample size for that result in the discussion, it can be used for demonstration without depletion of their main findings. If all subjects involved in that study ($N = 20$) would have met the inclusion criterion (which was a

minimum number of six errors in both conditions), the two-sided p-value would have been 0.06 (vs. 0.086 in the eight included subjects), provided the effect sizes would have remained constant. Considering the fact that all 12 subjects in that study who were not included in the actual test were excluded because they were statistically better at either the primary task (resulting in fewer overall errors) and/or at consciously detecting their errors (resulting in a lower ratio of non-reported to reported errors), it is not possible to justify the acceptance of a null-hypothesis. Similar arguments can in principle be applied to other studies that find numerical differences but no significances between error types. This is not to say that these results are of low value, particularly because the null-findings in ERN amplitude are oftentimes only remote points in the respective papers that do not lie at the core of the hypotheses tested. It does mean, however, that in case of a very low sample size, particularly when reporting low p-values for reported vs. non-reported errors, the acceptance of the null-hypothesis is not warranted from a statistical point of view.

Support for the low-power hypothesis presented here comes from the fMRI domain. Missing differential error awareness effects in the dACC/RCZ (Hester et al., 2005; Klein et al., 2007), the neural generator of the ERN, is oftentimes cited as supporting evidence in studies reporting the absence of an effect of error awareness on ERN amplitude. This is despite findings that demonstrate that response-conflict, which is also registered in the dACC/RCZ (Botvinick et al., 2001; Yeung et al., 2004) does not evoke such a RCZ response when elicited subliminally (Dehaene et al., 2003), and also despite the finding that consciously rejecting trials with a high subjective error-likelihood is correlated with activity in the RCZ (Magno et al., 2006). The three studies that explicitly address error awareness related activity in the RCZ in fMRI experiments (Hester et al., 2005, 2009; Klein et al., 2007) are an excellent illustration of the potential pitfalls of low samples sizes: Klein et al. (2007) report numerical differences in RCZ BOLD-activity, with reported errors eliciting more activity than non-reported errors (visible in **Figure 2C** in their manuscript), which fails to reach significance in the 13 subjects reported ($p = 0.211$, two-sided), leaving the anterior part of the left insular cortex as the only part of cortex sensitive to subjective error awareness. Hester et al. (2005) initially reported null-findings in the error-awareness task (EAT) with respect to RCZ activity as well, also in 13 subjects ($p = 0.59$ for the RCZ ROI). In a later study (Hester et al., 2009) using the same experiment in 16 subjects, however, they did find significant differences in that exact region.

All of this is not to argue that there is a definitive effect of error awareness on the amplitude of the ERN/RCZ activity, and all studies not demonstrating these effects fail to do so. There are certainly many factors that contribute to error awareness, and even more factors that potentially contribute to ERN amplitude. Error correction and stimulus representation might be among them, but they are unlikely to account for the differences found across several error awareness studies. Differences in study design or operationalization of subjective error awareness (see above) could account for many differences in findings.

In any case, based on the argument made in this paragraph, it is not possible to uphold the statement that the amplitude of the ERN is unrelated to subjective awareness. On the contrary: while

there are many studies that demonstrate enlarged ERN amplitudes with respect to subjective error awareness with a low enough type-1 error probability to warrant rejection of the null-hypothesis (Scheffers and Coles, 2000; Maier et al., 2008; Steinhauser and Yeung, 2010; Woodman, 2010; Dhar et al., 2011; Hewig et al., 2011; Wessel et al., 2011), there are few, if any, studies that have sufficiently low type-2 error probability to warrant an acceptance of that null hypothesis. Future studies should make sure to contain large enough sample sizes in order to allow for strong inferences in case of a potential null finding.

A PUTATIVE ROLE OF THE ERN IN AN OVERARCHING MODEL OF ACCESS CONSCIOUSNESS

After one establishes the fact that the ERN and error awareness are not unrelated, the obvious question is: what is its exact role in the emergence of error awareness? Does the amplitude of the ERN influence the emergence of error awareness or vice versa? Furthermore: what's the role of the Pe? What's the role of the ANS, which has been found to react differently to reported and non-reported errors (O'Connell et al., 2007; Wessel et al., 2011)? Ullsperger et al. (2010) have recently proposed a unified account of a putative role of these potentials in the emergence of error awareness, in which multiple sources of evidence accumulate over time and eventually culminate in error awareness (or blindness). Steinhauser and Yeung (2010) have convincingly demonstrated that this accumulating evidence (AE) is indeed reflected in the amplitude of the error-related potential following the ERN, the error positivity. In the following, I will try to link these accounts with each other and embed them in a prominent theory of the emergence of access consciousness in the brain, the global neuronal workspace (GNW) theory (Baars, 1988; Dehaene and Naccache, 2001).

THE GLOBAL NEURONAL WORKSPACE THEORY

The GNW theory is a unified theory about the neural mechanisms underlying the emergence of access consciousness of any stimulus in the brain. It is based early formulations of a "global workspace" of consciousness from Baars (1988) and on Fodor's distinction of the brain into different "modular facilities" that are distinguishable from an "isotropic system" that integrates information across these modules (Fodor, 1985). Consequently, Dehaene and Naccache (2001) and Dehaene and Changeux (2004) pose the existence of two distinct networks in the human brain: the network of processors on the one hand, and the "global neuronal workspace" (GNW) on the other.

There are multiple different separate entities that comprise the network of processors, which consists of modules that code simple visual information (area V1), motion (area MT), faces (fusiform face area), or sounds (auditory cortex areas in the temporal lobe), amongst many others. Although the information coded in these processors differs in complexity and level of abstraction, all these areas have in common that they are located at relatively early stages of the stimulus processing chain, and can relay information in a specialized, automated, and fast feed-forward fashion.

The second network, the GNW, constitutes the neuronal basis of access consciousness according to the theory. It consists of long-range excitatory axons, which allow the exchange, or "broadcasting" of many different kinds of information across the areas that comprise the network of processors. It is the process of entering the GNW that effectively constitutes the emergence of awareness in the GNW model.

Attention plays a critical role in the GNW theory. Just as in classic models of attention, a stimulus can enter the GNW through one out of two mechanisms: (a) the specific module/processor is already the current locus of attention (top-down allocated attention) or (b) the stimulus is of sufficient strength to attract top-down attention itself (bottom-up driven attention).

The existence of a GNW has been formulated over a decade ago and predictions derived from it have been experimentally tested in several studies (e.g., Del Cul et al., 2007). It addresses the question of the generation of access consciousness in a neurobiologically plausible way, which is why I will try to implement our recent theory about the emergence of error awareness in the human brain (Ullsperger et al., 2010) into this framework, specifically focusing on the role of the ERN.

ERRORS IN THE GLOBAL WORKSPACE: THE ACCUMULATING EVIDENCE ACCOUNT

A putative model of the emergence of error awareness is outlined in **Figure 3**. It embeds ideas from the AE account of emerging error awareness (Ullsperger et al., 2010) into the more general framework of the GNW model (Dehaene and Naccache, 2001). The general idea of the AE model fits in well with the basic principle of the GNW model: in the AE model, consistent with experimental findings, evidence about the accuracy of an action is available from multiple different cortical processors that code different types of information. This information accumulates over time and contributes to the reportability of an error in a feed-forward fashion. This kind of parallel processing in multiple different areas corresponds to the "network of processors" in the GNW model. Reportability of an error is then defined as access of that accumulating information to the GNW.

THE NETWORK OF PROCESSORS: CODING OF MULTIPLE SOURCES OF ERROR-EVIDENCE

Differences between reported and non-reported errors have been described on multiple levels of early and late nervous system processing. Much of this information is available at very early latency ranges, making it chronologically and logically unlikely to be a consequence of error awareness, and rather implicate it in feed-forward processing that contributes to emerging error awareness.

Sensory systems

It has been shown that errors that are subsequently reported differ from non-reported errors with respect to quantity and quality of the sensory information at hand. It is evident from correction rates in the AST studies (Endrass et al., 2007; Wessel et al., 2011, and potentially also Nieuwenhuis et al., 2001, see above) that non-reported errors are more often corrected than reported errors. This is a somewhat unexpected result, provided one interprets error correction as an intentional and conscious act. However, all three AST studies unequivocally report even more prominent effects of error awareness on correction times, i.e., the time from
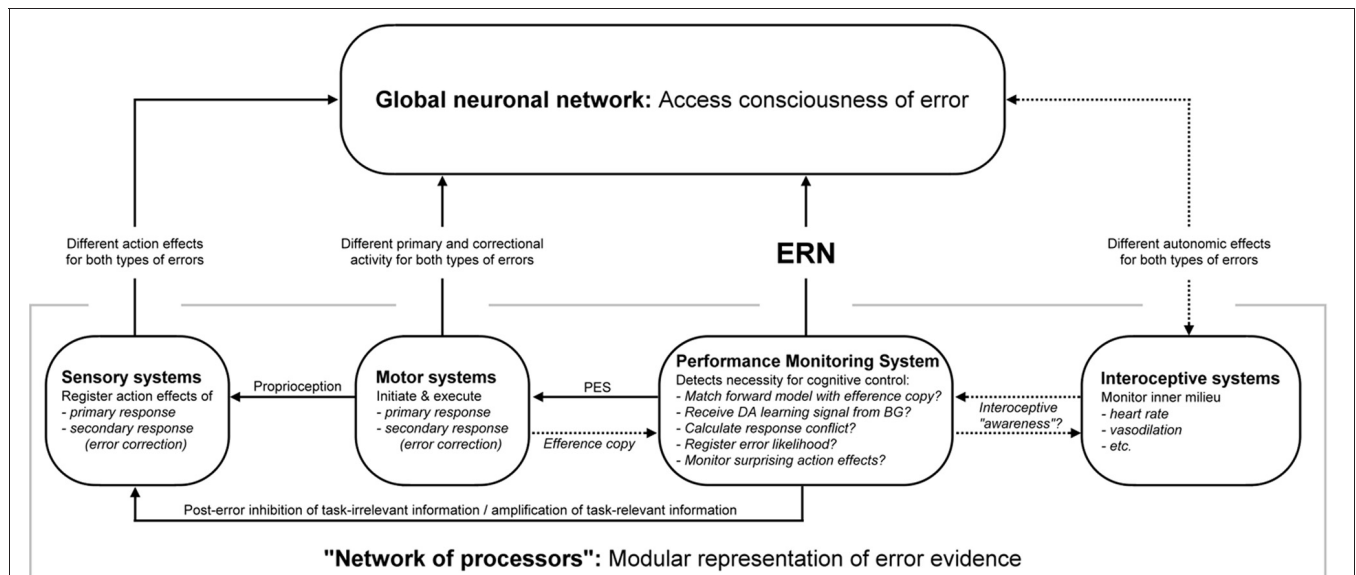
FIGURE 3 | A putative model schematic of emerging error awareness in the human brain, based on the accumulating evidence account of error awareness and the global neuronal workspace model. Information about the accuracy of an action is processed in parallel in different areas that comprise the "network of processors," which feeds forward into the GNW. Note that the flow of information indicated by the arrows is only depicted if potentially meaningful for error awareness. Additional exchange of information is also probable (especially attentional modulation from the GNW to the network of processors). Be aware that the potential functions of the performance monitoring network outlined here represent the main branches of theories that have been put forward, and it doesn't mean that the ERN is a correlate of all these computations, but probably only a subset of them. ERN, error-related negativity; BG, basal ganglia; dACC, dorsal anterior cingulate cortex; RCZ, rostral cingulate zone; PES, post-error slowing; DA, dopamine.

the erroneous to a subsequent corrective saccade, showing much longer correction RTs for aware errors. This means that most non-reported errors were corrected very fast (or vice versa: most fast corrected errors were subsequently not reported), potentially in an automated fashion, making them harder to detect for cognitive systems than the reported errors, which are not only corrected less frequently, but also with longer latencies. In terms of sensory representation, this means that for subsequently reported errors, gaze was directed in the wrong direction for a longer period of time, resulting in more sensory evidence for the cognitive system to detect.

Motor systems
On the motor level, another finding from the AST studies (Nieuwenhuis et al., 2001; Endrass et al., 2007; Wessel et al., 2011) provides a good demonstration of different levels of error-evidence between error types: these studies consistently show larger saccade sizes for reported as compared to non-reported errors. Hence, there is also quantitatively more evidence for inaccuracy of an action on aware errors.

Performance monitoring systems
It is far beyond the scope of this review to speculate as to the exact functional significance of the ERN or its underlying neural generator, the dACC/RCZ, and its associated network of brain regions. However, it does not matter for the purposes of this model what ERN/RCZ activity actually signifies. All four major accounts of ERN/RCZ function (see introduction) have a common theme in that this brain region (RCZ) and its respective neurophysiological signature (ERN) monitor ongoing behavior, potentially with the function of signaling the need for adjustments (Ridderinkhof et al., 2004), or even implementing these adjustments itself.

What could be shown based on the review of the existing literature is that there is a growing amount of evidence that the levels of ERN/RCZ activity differ between reported and non-reported errors, with the former carrying quantitatively more information/activity. So while it is not to be determined what exact function this module serves (detecting mismatch between a forward model and the motor efference copy (Falkenstein et al., 1991; Coles et al., 2001), monitoring response-conflict (Botvinick et al., 2001; Yeung et al., 2004), reflecting a learning signal from the dopaminergic midbrain (Holroyd and Coles, 2002), representing the likelihood of an error on a given trial (Brown and Braver, 2005), or signaling the unsigned reward prediction error, or "surprise" of a given response (Alexander and Brown, 2011; Hayden et al., 2011), it can be said with certainty that this activity differs with respect to subjective error awareness.

Interoceptive systems
One of the most interesting modules in this model is the interoceptive system. It has been shown in at least two studies (O'Connell et al., 2007; Wessel et al., 2011) that the activity of the ANS differs with respect to subjective error awareness. This is particularly interesting with respect to the fact that the insular cortex has been shown to be also sensitive to this factor (Klein et al., 2007, for a review, see: Ullsperger et al., 2010). The insular cortex has been conjectured to reflect the activity of an "interoceptive awareness" system (Critchley et al., 2004; Craig,

2009; Medford and Critchley, 2010). The question of causality (or even temporal order) between the ANS, the insular cortex, and error awareness is not sufficiently clear as of yet. Particularly, this is because of the fact that necessary lesion studies of the insular cortex are hard to conduct. Ischemic stroke damage that is exclusive to the insula, while leaving the prefrontal cognitive controls areas/circuits intact, is very rare given the layout of the cerebral blood supply. Therefore, it can only be speculated whether the differential autonomic activity, which could be picked up by the interoceptive system, contributes to the emergence of error awareness, or whether the awareness of the error leads to an increased activation of the ANS. Nevertheless, it is theoretically possible that this system is another module coding information of relevance for the access of erroneous information to the global neuronal network.

Interaction between different modules

The information coded in these distinct networks is very different in nature, but can be potentially used by the cognitive system in a cumulative fashion, which could then enable the erroneous quality of an action to exceed a threshold necessary for (access-) conscious report. It is notable that these networks, although distinct in nature, also interact with one another in a way that is relevant to error processing. For example, ERN amplitude/RCZ activity has been shown to predict the amount of error-related remedial processes (for a review, see Danielmeier and Ullsperger, 2011). Such processes are evident in both the motor domain (as indicated by post-error slowing (PES), a relative slowing in reaction times following errors as compared to correct trials), as well as in sensory cortices (evident in the attenuation of task-irrelevant information and amplification of task-relevant information following errors). Both these processes have been found to correlate with preceding activity in the ACC/RCZ (PES: King et al., 2010, post-error regulation of sensory areas: Danielmeier et al., 2011). PES has also been found to be predicted by ERN amplitude on the previous error trial (Debener et al., 2005; Wessel and Ullsperger, 2011). Interestingly, these processes could also be mediated by the GNW (or any other part of the cognitive system that mediates error awareness): in studies that examine the relation between ERN/Pe amplitude and error awareness, PES has been consistently found to be exclusive for aware errors (e.g., Nieuwenhuis et al., 2001; Endrass et al., 2007; Wessel et al., 2011), regardless of whether an ERN effect for error awareness is reported. The same is true (to a lesser extent) for Klein et al. (2007) fMRI study. However, it is also possible that the neuronal processes underlying PES happen in the absence of awareness and are triggered by other factors that coincide with greater error awareness. This is later conjecture is backed up by findings from behavioral studies that find PES in the absence of error awareness (Rabbitt, 2002; Logan and Crump, 2010). The PES—error awareness contingency might be exclusive to the AST (which is the paradigm that was used in all studies that report positive findings, see above), where eye-movements (as opposed to button presses) are the primary response domain, and which has been used in all four studies that report greater PES for reported errors. This can potentially give insights into possible variables that give rise to both error awareness and PES at the same time, without the two themselves having a direct, causal connection: in the AST, as seen before, unreported errors are associated with fewer behavioral evidence (smaller saccade sizes), sensory evidence (faster corrections, i.e., less visual evidence of "having looked in the wrong direction"), and proprioceptive evidence for the erroneousness of the action. This lack of evidence compared to reported errors ostensibly ultimately leads to error blindness on these trials. The same might not necessarily be true for button press paradigms, especially with respect to proprioceptive feedback: compared to an eye-movement, an erroneous button press is associated with stronger proprioceptive feedback, but also with all sorts of other sensory evidence (the auditory "click" of the key, the visual feedback of moving the finger), which is the same across both types of errors, unlike in the AST. These same factors (or a subset of them) could in fact be the variables causing PES. More research on the dynamics of the interaction between the different subsystems that carry error-relevant information is needed in order to answer this question.

THE QUESTION OF THRESHOLD: ALL-OR-NOTHING ACCESS AND THE ROLE OF THE Pe

The GNW model postulates access to the GNW as an all-or-nothing process, potentially signified by biological parameters with bimodal distributions, such as the P300 ERP (Dehaene and Changeux, 2004). It has been shown that the P300 does indeed parallel the non-linear properties of subjects' reports of seeing or not seeing a masked stimulus (Del Cul et al., 2007). It has also long been speculated that the error positivity (Pe) signifies processes comparable to the stimulus-locked P300 (Overbeek et al., 2005). Therefore, it is tempting to speculate that the Pe does indeed signify the activity of the GNW (as the P300 seems to do), and, therefore, the actual expression of error awareness. However, in the recent study by Steinhauser and Yeung (2010), the Pe has been found to be more related to the accumulating stimulus input into the error-awareness decision process than the output. It is an interesting question for future research whether the Pe is an input signal into the GNW, which might represent a combination of the input from the network of processors, or whether it is an output signal, reflecting the categorical "all-or-nothing" access to the GNW. What might potentially help is a distinction between the two different parts of the Pe, the late and early Pe (Overbeek et al., 2005; Endrass et al., 2007). The early Pe seems to be largely correlated with the ERN and might potentially signify the activity of the same underlying cortical generator, as is suggested by studies investigating the ERN using independent-component analysis (ICA, Jutten and Herault, 1991), which qualitatively show intact Pe effects when restricting the data to the independent-components underlying the ERN (Debener et al., 2005; Eichele et al., 2010; Wessel and Ullsperger, 2011). The later parts of the Pe seem to reflect a different process that is potentially closer to an actual expression of error awareness (Endrass et al., 2007), and might, therefore, indeed reflect the process that underlies the stimulus-locked P300 and potentially reflects access to the GNW. An early Pe might, therefore, have the properties that Steinhauser and Yeung (2010) describe, i.e., reflecting the cumulative input of error evidence into the GNW, whereas a later part of the Pe could indeed have the bimodal distribution that

would be predicted based on the Pe/p300-equivalency hypothesis and the findings of Del Cul et al. (2007), and signify the actual expression of error awareness. This idea could be tested in future research.

Several predictions from this model, in which the information coded in the network of processors accumulates and is reflected in the amplitude of the Pe, are in line with earlier findings: ERN and Pe amplitude have been found to be significantly correlated on a single-trial level on multiple occasions (e.g., Steinhauser and Yeung, 2010; Hughes and Yeung, 2011). Also, the amplitude of the Pe correlates significantly with skin-conductance changes found following errors (Hajcak et al., 2003), which in turn has been found to be sensitive to subjective error awareness (O'Connell et al., 2007).

FUTURE DIRECTIONS

There are many different areas in which the field of error awareness research could make headway, which are certainly not all related to the specific role of the ERN. I will outline three major strains of research that could significantly contribute to the advancement of the field of error awareness research. Certainly, several other ideas come to mind, such as the assessment of the role of pre-trial states that influence primary task performance (Aston-Jones and Cohen, 2005; Eichele et al., 2008) with respect to their role in error awareness. In the following, I will focus on three general fields of ideas that are either closely related to the research reviewed in this article, or can be directly applied to the research of the role of the ERN in error awareness.

THE QUANTIFICATION OF (ACCESS) CONSCIOUSNESS

As described above, reportability by means of categorical rating procedures is the primarily used index of the degree of "error awareness" on a certain trial.

While this is certainly a valid index of access conscious availability of the accuracy of an action, one could think of more "indirect" quantifications of access consciousness. The issue of reactivity, i.e., interfering with ongoing psychological processes by probing them explicitly, is not as big an issue for the research on error awareness as it is for instance for contingency awareness in implicit learning, where probing explicit memory contents can trigger additional factors that interfere with the processes of interest (cf. Dienes, 2008). However, it is potentially possible that explicitly probing error awareness of every trial alters a generic error monitoring process. Therefore, more indirect measures could be employed. Persaud et al. (2007) recently demonstrated that post-decisional wagering procedures effectively capture awareness of contingencies in an Iowa gambling task. Such measures could be used to get a fine-grain quantification of error awareness as a single-trial measure (e.g., by allowing for a very unconstrained wagering procedure—"Wage anywhere between 1 and 100 cents on your accuracy," or by having subjects bet on their action outcome in case they report their behavior as "unsure" or "don't know"). Correlating these measures with ongoing neuronal activity should allow for specific hypothesis testing and should enable researchers to pull apart the exact mechanics of what really drives the emergence of

error awareness. Also, these measures could allow for the potential quantification of types of consciousness that are not necessarily captured by overt and explicit rating procedures. Research on metacognitive feelings such as feeling of knowing (Koriat et al., 2006; Koriat, 2007) has shown that there are representations of stimuli/internal states that can be both accurate (i.e., greater than chance level), but not available for overt report, potentially getting at what philosophers called "reflexive" or "interoceptive" consciousness (Block, 2001). Another interesting approach that could certainly help elucidating the factors that contribute to error awareness is the quantification of the neuronal processes of stimulus perception from the mechanisms of error monitoring, as has been done in Woodman (2010). In a philosophical framework, it could be argued that this particular study could successfully disentangle phenomenological consciousness of a stimulus from access consciousness of an error. Further experiments along these lines could also help to elucidate the exact processes that are necessary for the emergence of error awareness.

METHODOLOGICAL ADVANCEMENTS AND SINGLE-TRIAL HYPOTHESES

All studies reviewed in this article have measured the ERN using the classic averaging method, according to the logic of event-related potential research. As notable exceptions, Steinhauser and Yeung (2010) and Hughes and Yeung (2011) have used functional logistic classification methods to generate spatial filters that dissociate the ERN from other ongoing brain processes in order to obtain single-trial amplitudes, even though the hypotheses tested were limited to the Pe. Advances in signal processing methods have given rise to many different approaches that can be used to study the single trial properties of ERPs like the ERN. This is particularly important because error awareness studies of ERP data oftentimes deal with the problem that many subjects do not have enough unreported errors to warrant a reliable average. Increasing the signal to noise ratio to the point where a single-trial analysis is possible effectively alleviates this situation.

Independent Component Analysis (ICA, Jutten and Herault, 1991) has been successfully used to study single-trial properties of error-related brain potentials (Debener et al., 2005; Eichele et al., 2010; Wessel and Ullsperger, 2011). Many other techniques are available that yield sufficient single-trial signal-to-noise ratios to enable single-trial research on the ERN. Such methods could be used to test hypotheses that are only hardly testable using averaging procedures: does the amplitude of the ERN on a given trial directly affect the accuracy rating (one would need a continuous or at least non-binary quantification of both ERN and access consciousness to answer this question), as, e.g., Scheffers and Coles (2000) results suggest? Is access to the GNW a continuous phenomenon or is it reflected as an all-or-nothing process in the properties of error-related ERPs? ICA (and other blind source separation or functional source separation techniques) would also enable the dissociation of the ERN/early Pe complex and the late Pe, which could then be used for separate hypothesis testing, e.g., about the association between the central nervous correlates of emerging access consciousness and error awareness. Such questions could be answered by exploiting the single-trial amplitudes

of error-related ERPs, and could thereby significantly promote research in this field.

INDIRECT BENCHMARKS: THE FUNCTIONAL ROLE OF CONSCIOUSNESS IN ERROR PROCESSING

Consciousness, in order for it to be an empirically relevant process, needs to serve a certain function, or as Koriat put it:

> "Self-controlled processes have measurable effects on behavior. Although [. . .] many cognitive processes, including some that are subsumed under the rubric of executive function, occur outside of consciousness, there is also a recognition that the person is not a mere medium through which information flows."
>
> (Koriat, 2007, p. 292)

Koch and Tsuchiya (in: Block, 2007) also discuss functional roles of consciousness, and its effects on overt behavior, and summarize:

> "Consciousness and (top-down controlled attention) are distinct neurobiological processes with distinct functions."
>
> (Koch and Tsuchiya, in Block, 2007, p. 509)

An example for executive function in the context of error awareness research that is independent of (access) consciousness is rapid error correction (see above). Yet it has also already been described that some error-related processes, such as PES, coincide with subjective error awareness, at least in certain paradigms (specifically the AST). If it can be proven that there are indeed behavioral markers in the domain of error processing that are causally dependent on subjective error awareness, this would not only give researchers another indirect index for measuring error awareness, but it would also elucidate the mechanism of the emergence of error awareness itself. PES is a potential candidate for such an index, but it has to be systematically examined under which circumstances PES and access consciousness coincide. Other likely candidates such as the attenuation of task-irrelevant activity and amplification of task-relevant activity found following errors (King et al., 2010; Danielmeier et al., 2011) that potentially are highly dependent on top-down attention need to be studied in a context of error awareness, in order to further outline the potential functional role of error awareness in the adaptive regulation of ongoing behavior.

CONCLUSION

A decade has passed since the first publication of a study on the effects of subjective error awareness on the amplitude of arguably the most prominent index of error-related brain activity, the ERN. A diverse picture emerged in the dozen studies that have been published since that first report, with some studies reporting significantly enlarged ERN amplitudes for reported compared to non-reported errors, and several other studies reporting null effects.

Based on the evidence reviewed and evaluated in this article, it appears safe to conclude that the processes that are reflected in the ERN and the processes involved in the emergence of error awareness are not separate from each other. Whether these processes are linked by a third process that influences both the ERN-underlying process and the emergence of awareness remains to be tested in future studies, and first and foremost needs a definitive identification of the process underlying the ERN. However, it should be evident from central parts of this review that none of the recently proposed factors that have been proposed to explain the differences in ERN amplitude between reported and non-reported errors (e.g. error correction, stimulus misrepresentation) can actually account for these effects.

I propose that the ERN serves as a feed-forward input signal into the systems responsible for error awareness. Alongside the input from many other areas in which error-relevant information is coded, the ultimate emergence of "error awareness" is grounded on the amplitude of this input. This proposition was expressed in terms of a combination of the previously existing AE account of error awareness and a more general model of the mechanisms of emerging access consciousness in the brain. The exact causal and chronological relations should be the focus of future study in this field that combines two of the most exciting areas of research in cognitive neuroscience: cognitive control and the emergence of awareness.

REFERENCES

Alexander, W. H., and Brown, J. W. (2011). Medial prefrontal cortex as an action-outcome predictor. *Nat. Neurosci.* 14, 1338–1344.

Aston-Jones, G., and Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annu. Rev. Neurosci.* 28, 403–450.

Baars, B. J. (1988). *A Cognitive Theory of Consciousness.* (Cambridge, UK: Cambridge University Press).

Bayne, T., and Chalmers, D. (2003). "What is the unity of consciousness," in *The Unity of Consciousness: Binding, Integration, and Dissociation*, ed A. Cleeremans, (Oxford, UK: Oxford University Press), 23–58.

Block, N. (1995). On a confusion about a function of consciousness. *Behav. Brain Sci.* 18, 227–247.

Block, N. (2001). Paradox and cross purposes in recent work on consciousness. *Cognition* 79, 197–219.

Block, N. (2007). Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behav. Brain Sci.* 30, 481–499; discussion 499–548.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., and Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol. Rev.* 108, 624–652.

Brown, J. W., and Braver, T. S. (2005). Learned predictions of error likelihood in the anterior cingulate cortex. *Science* 307, 1118–1121.

Chalmers, D. J. (1995). Facing up to the problem of consciousness. *J. Conscious. Stud.* 2, 200–219.

Coles, M. G., Scheffers, M. K., and Holroyd, C. B. (2001). Why is there an ERN/Ne on correct trials? Response representations, stimulus-related components, and the theory of error-processing. *Biol. Psychol.* 56, 173–189.

Craig, A. D. (2009). How do you feel–now? The anterior insula and human awareness. *Nat. Rev. Neurosci.* 10, 59–70.

Crick, F., and Koch, C. (2003). A framework for consciousness. *Nat. Neurosci.* 6, 119–126.

Critchley, H. D., Wiens, S., Rotshtein, P., Ohman, A., and Dolan, R. J. (2004). Neural systems supporting interoceptive awareness. *Nat. Neurosci.* 7, 189–195.

Danielmeier, C., Eichele, T., Forstmann, B. U., Tittgemeyer, M., and Ullsperger, M. (2011). Posterior medial frontal cortex activity predicts post-error adaptations in task-related visual and motor areas. *J. Neurosci.* 31, 1780–1789.

Danielmeier, C., and Ullsperger, M. (2011). Post-error adjustments. *Front. Psychol.* 2:233. doi: 10.3389/fpsyg.2011.00233

Danielmeier, C., Wessel, J. R., Steinhauser, M., and Ullsperger, M. (2009). Modulation of the error-related negativity by response conflict. *Psychophysiology* 46, 1288–1298.

Debener, S., Ullsperger, M., Siegel, M., Fiehler, K., Von Cramon, D. Y., and Engel, A. K. (2005). Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. *J. Neurosci.* 25, 11730–11737.

Dehaene, S., Artiges, E., Naccache, L., Martelli, C., Viard, A., Schurhoff, F., Recasens, C., Martinot, M. L., Leboyer, M., and Martinot, J. L. (2003). Conscious and subliminal conflicts in normal subjects and patients with schizophrenia: the role of the anterior cingulate. *Proc. Natl. Acad. Sci. U.S.A.* 100, 13722–13727.

Dehaene, S., and Changeux, J. (2004). "Neural mechanisms for access to consciousness," in *The Cognitive Neurosciences III*, ed M. Gazzaniga (Cambridge, MA: MIT Press).

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

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

Del Cul, A., Baillet, S., and Dehaene, S. (2007). Brain dynamics underlying the nonlinear threshold for access to consciousness. *PLoS Biol.* 5:e260. doi: 10.1371/journal.pbio.0050260

Dhar, M., Wiersema, J. R., and Pourtois, G. (2011). Cascade of neural events leading from error commission to subsequent awareness revealed using EEG source imaging. *PLoS One* 6:e19578. doi: 10.1371/journal.pone.0019578

Dienes, Z. (2008). Subjective measures of unconscious knowledge. *Prog. Brain Res.* 168, 49–64.

Eichele, H., Juvodden, H. T., Ullsperger, M., and Eichele, T. (2010). Mal-adaptation of event-related EEG responses preceding performance errors. *Front. Hum. Neurosci.* 4:65. doi: 10.3389/fnhum.2010.00065

Eichele, T., Debener, S., Calhoun, V. D., Specht, K., Engel, A. K., Hugdahl, K., von Cramon, D. Y., and Ullsperger, M. (2008). Prediction of human errors by mal-adaptive changes in event-related brain networks. *Proc. Natl. Acad. Sci. U.S.A.* 105, 6173–6178.

Endrass, T., Franke, C., and Kathmann, N. (2005). Error awareness in a saccade countermanding task. *J. Psychophysiol.* 19, 275–280.

Endrass, T., Reuter, B., and Kathmann, N. (2007). ERP correlates of conscious error recognition: aware and unaware errors in an antisaccade task. *Eur. J. Neurosci.* 26, 1714–1720.

Eriksen, B. A., and Eriksen, C. W. (1974). Effects of noise letters upon identification of a target letter in a nonsearch task. *Percept. Psychophys.* 16, 143–149.

Falkenstein, M., Hohnsbein, J., Hoormann, J., and Blanke, L. (1991). Effects of crossmodal divided attention on late Erp components. 2. Error processing in choice reaction tasks. *Electroencephalogr. Clin. Neurophysiol.* 78, 447–455.

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

Fiehler, K., Ullsperger, M., and Von Cramon, D. Y. (2005). Electrophysiological correlates of error correction. *Psychophysiology* 42, 72–82.

Fodor, J. A. (1985). Precis of the modularity of mind. *Behav. Brain Sci.* 8, 1–5.

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

Gehring, W. J., and Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science* 295, 2279–2282.

Hajcak, G., McDonald, N., and Simons, R. F. (2003). To err is autonomic: error-related brain potentials, ANS activity, and post-error compensatory behavior. *Psychophysiology* 40, 895–903.

Hayden, B. Y., Heilbronner, S. R., Pearson, J. M., and Platt, M. L. (2011). Surprise signals in anterior cingulate cortex: neuronal encoding of unsigned reward prediction errors driving adjustment in behavior. *J. Neurosci.* 31, 4178–4187.

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

Hester, R., Nestor, L., and Garavan, H. (2009). Impaired error awareness and anterior cingulate cortex hypoactivity in chronic cannabis users. *Neuropsychopharmacology* 34, 2450–2458.

Hewig, J., Coles, M. G., Trippe, R. H., Hecht, H., and Miltner, W. H. (2011). Dissociation of Pe and ERN/Ne in the conscious recognition of an error. *Psychophysiology* 48, 1390–1396.

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

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

Hughes, G., and Yeung, N. (2011). Dissociable correlates of response conflict and error awareness in error-related brain activity. *Neuropsychologia* 49, 405–415.

Jutten, C., and Herault, J. (1991). Blind separation of sources. 1. An adaptive algorithm based on neuromimetic architecture. *Signal Process.* 24, 1–10.

King, J. A., Korb, F. M., Von Cramon, D. Y., and Ullsperger, M. (2010). Post-error behavioral adjustments are facilitated by activation and suppression of task-relevant and task-irrelevant information processing. *J. Neurosci.* 30, 12759–12769.

Klein, T. A., Endrass, T., Kathmann, N., Neumann, J., Von Cramon, D. Y., and Ullsperger, M. (2007). Neural correlates of error awareness. *Neuroimage* 34, 1774–1781.

Koriat, A. (2007). "Metacognition and consciousness," in *The Cambridge Handbook of Consciousness*, eds P. D. Zelazo, M. Moscovitch, E. Thompson (Cambridge, UK: Cambridge University press), 289–325.

Koriat, A., Ma'Ayan, H., and Nussinson, R. (2006). The intricate relationships between monitoring and control in metacognition: lessons for the cause-and-effect relation between subjective experience and behavior. *J. Exp. Psychol. Gen.* 135, 36–69.

Logan, G. D., and Crump, M. J. (2010). Cognitive illusions of authorship reveal hierarchical error detection in skilled typists. *Science* 330, 683–686.

Luck, S. J., and Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual-search. *Psychophysiology* 31, 291–308.

Magno, E., Foxe, J. J., Molholm, S., Robertson, I. H., and Garavan, H. (2006). The anterior cingulate and error avoidance. *J. Neurosci.* 26, 4769–4773.

Maier, M., Steinhauser, M., and Hubner, R. (2008). Is the error-related negativity amplitude related to error detectability? Evidence from effects of different error types. *J. Cogn. Neurosci.* 20, 2263–2273.

Mayr, U. (2004). Conflict, consciousness, and control. *Trends Cogn. Sci.* 8, 145–148.

Medford, N., and Critchley, H. D. (2010). Conjoint activity of anterior insular and anterior cingulate cortex: awareness and response. *Brain Struct. Funct.* 214, 535–549.

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

O'Connell, R. G., Bellgrove, M. A., Dockree, P. M., Lau, A., Hester, R., Garavan, H., Fitzgerald, M., Foxe, J. J., and Robertson, I. H. (2009). The neural correlates of deficient error awareness in attention-deficit hyperactivity disorder (ADHD). *Neuropsychologia* 47, 1149–1159.

O'Connell, R. G., Dockree, P. M., Bellgrove, M. A., Kelly, S. P., Hester, R., Garavan, H., Robertson, I. H., and Foxe, J. J. (2007). The role of cingulate cortex in the detection of errors with and without awareness: a high-density electrical mapping study. *Eur. J. Neurosci.* 25, 2571–2579.

Overbeek, T. J. M., Nieuwenhuis, S., and Ridderinkhof, K. R. (2005). Dissociable components of error processing - On the functional significance of the Pe Vis-a-vis the ERN/Ne. *J. Psychophysiol.* 19, 319–329.

Persaud, N., McLeod, P., and Cowey, A. (2007). Post-decision wagering objectively measures awareness. *Nat. Neurosci.* 10, 257–261.

Rabbitt, P. (1966). Errors and error correction in choice-response tasks. *J. Exp. Psychol.* 71, 264.

Rabbitt, P. (1967). Time to detect errors as a function of factors affecting choice-response time. *Acta Psychol. (Amst.)* 27, 131–142.

Rabbitt, P. (2002). Consciousness is slower than you think. *Q. J. Exp. Psychol. A* 55, 1081–1092.

Ridderinkhof, K. R., Ramautar, J. R., and Wijnen, J. G. (2009). To P(E) or not to P(E): a P3-like ERP component reflecting the processing of response errors. *Psychophysiology* 46, 531–538.

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., and Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science* 306, 443–447.

Rodriguez-Fornells, A., Kurzbuch, A. R., and Munte, T. F. (2002). Time course of error detection and correction in humans: neurophysiological evidence. *J. Neurosci.* 22, 9990–9996.

Roger, C., Benar, C. G., Vidal, F., Hasbroucq, T., and Burle, B. (2010). Rostral Cingulate Zone and correct response monitoring: ICA and source localization evidences for the unicity of correct- and error-negativities. *Neuroimage* 51, 391–403.

Rosenthal, D. M. (2002). How many kinds of consciousness? *Conscious. Cogn.* 11, 653–665.

Scheffers, M. K., and Coles, M. G. (2000). Performance monitoring in a confusing world: error-related brain activity, judgments of response accuracy, and types of errors. *J. Exp. Psychol. Hum. Percept. Perform.* 26, 141–151.

Shalgi, S., Barkan, I., and Deouell, L. Y. (2009). On the positive side of error processing: error-awareness positivity revisited. *Eur. J. Neurosci.* 29, 1522–1532.

Steinhauser, M., and Yeung, N. (2010). Decision processes in human performance monitoring. *J. Neurosci.* 30, 15643–15653.

Stephan, K. M., Thaut, M. H., Wunderlich, G., Schicks, W., Tian, B., Tellmann, L., Schmitz, T., Herzog, H., McIntosh, G. C., Seitz, R. J., and Homberg, V. (2002). Conscious and subconscious sensorimotor synchronization–prefrontal cortex and the influence of awareness. *Neuroimage* 15, 345–352.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 18, 643–662.

Ullsperger, M., Harsay, H. A., Wessel, J. R., and Ridderinkhof, K. R. (2010). Conscious perception of errors and its relation to the anterior insula. *Brain Struct. Funct.* 214, 629–643.

Ullsperger, M., and von Cramon, D. Y. (2001). Subprocesses of performance monitoring: a dissociation of error processing and response competition revealed by event-related fMRI and ERPs. *Neuroimage* 14, 1387–1401.

Ullsperger, M., and von Cramon, D. Y. (2006). How does error correction differ from error signaling? An event-related potential study. *Brain Res.* 1105, 102–109.

Van Veen, V., and Carter, C. S. (2002). The timing of action-monitoring processes in the anterior cingulate cortex. *J. Cogn. Neurosci.* 14, 593–602.

Vidal, F., Hasbroucq, T., Grapperon, J., and Bonnet, M. (2000). Is the 'error negativity' specific to errors? *Biol. Psychol.* 51, 109–128.

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

Wessel, J. R., and Ullsperger, M. (2011). Selection of independent components representing event-related brain potentials: a data-driven approach for greater objectivity. *Neuroimage* 54, 2105–2115.

Woodman, G. F. (2010). Masked targets trigger event-related potentials indexing shifts of attention but not error detection. *Psychophysiology* 47, 410–414.

Yeung, N., Botvinick, M. M., and Cohen, J. D. (2004). The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychol. Rev.* 111, 931–959.