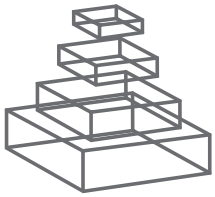
frontiers
RESEARCH TOPICS

ERROR AWARENESS – INSIGHTS FROM COGNITIVE NEUROSCIENCE, PSYCHIATRY AND NEUROLOGY

Topic Editors
Tilmann A. Klein, Markus Ullsperger and Claudia Danielmeier

frontiers in
HUMAN NEUROSCIENCE



frontiers

Cover image provided by Ibbl sarl, Lausanne CH

ISSN 1664-8714
ISBN 978-2-88919-201-4
DOI 10.3389/978-2-88919-201-4

ABOUT FRONTIERS

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

FRONTIERS JOURNAL SERIES

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing.
All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

DEDICATION TO QUALITY

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.
By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

WHAT ARE FRONTIERS RESEARCH TOPICS?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area!
Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

ERROR AWARENESS – INSIGHTS FROM COGNITIVE NEUROSCIENCE, PSYCHIATRY AND NEUROLOGY

Topic Editors:
Tilmann A. Klein[1,2], Markus Ullsperger[3,4] and Claudia Danielmeier[3]
[1] Department of Neurology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
[2] Day Clinic for Cognitive Neurology, University Clinic Leipzig, Leipzig, Germany
[3] Donders Institute for Brain, Cognition, and Behaviour, Radboud University Nijmegen, Nijmegen,Netherlands
[4] Institute of Psychology, Otto-von-Guericke University Magdeburg, Magdeburg, Germany

Everyone who has accidentally poured salt into their coffee instead of sugar and has noticed this only upon drinking is aware of the fundamental difference between aware and unaware errors. Whereas aware errors are consciously perceived upon their commission, unaware errors do not cross the threshold of conscious perception (and might only be noticed retrospectively when their sometimes unpleasant consequences are experienced). Performance monitoring is a key element of human goal-directed behaviour. Detection of performance errors can lead to behavioural modifications in order to reach a desired goal. The question arises whether these error detections need to become conscious in order to allow behavioural modifications. To answer these questions it is necessary to separate aware from unaware errors. While having participants report the awareness of an error is trivial, setting up an experiment that generates a sufficient number of unaware errors is far from trivial. In this Research Topic we will therefore provide a short overview over the history of research on error awareness and the different paradigms used.

Error processing is not limited to changes in brain activity – there are also changes in the autonomic nervous system. Previous studies have shown that pupil diameter, skin conductance response and heart rate are peripheral markers for error processing. These autonomic reactions might either help us to become aware of an error or might represent a consequence of the error becoming aware. Thus, one key question is what comes first: conscious detection of an error or the autonomic reaction?

Since the error processing system is complex, influencing factors can be manifold: from psychiatric illnesses disturbing relevant transmitter systems to structural changes after traumatic brain injury or stroke. One key symptom of major psychiatric and neurologic diseases is unawareness with respect to cognitive or bodily deficits. Is this type of unawareness related to impaired conscious error perception, and does it affect behavioural adaptations following errors?

The aim of the Research Topic is to provide an overview over the current research in error awareness. It aims at widening the angle on its topic by including contributions from cognitive neuroscience, psychiatry, and neurology. Furthermore, linking error awareness to awareness of bodily states seems highly necessary. The brain architecture involved in providing interoceptive awareness will therefore be discussed. One key structure in human interoceptive awareness is the insular cortex. Special emphasize will be put on the anatomy and pathology of this structure.

Table of Contents

frontiers in
HUMAN NEUROSCIENCE

EDITORIAL
published: 05 December 2013
doi: 10.3389/fnhum.2013.00830

Error awareness—insights from cognitive neuroscience, psychiatry and neurology

Tilmann A. Klein[1,2]*, Claudia Danielmeier[3] and Markus Ullsperger[3,4]

[1] *Department of Neurology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany*
[2] *Day Clinic for Cognitive Neurology, University Clinic Leipzig, Leipzig, Germany*
[3] *Centre for Cognition, Donders Institute for Brain, Cognition, and Behaviour, Radboud University Nijmegen, Nijmegen, Netherlands*
[4] *Department of Neuropsychology, Institute of Psychology, Otto-von-Guericke University Magdeburg, Magdeburg, Germany*
**Correspondence: tklein@cbs.mpg.de*

Edited by:
Hauke R. Heekeren, Freie Universität Berlin, Germany

Keywords: performance monitoring, error awareness, cognitive control, error-related negativity (ERN), error positivity (Pe), anterior cingulate cortex (ACC), insular cortex, consciousness

Most people would agree that error processing is essential for goal-directed and adaptive behavior. Within error processing one main distinction can be made between errors that become consciously aware and errors that remain unrecognized. Up until now it is still not completely understood what actually makes the difference between aware and unaware errors. Potential influences on error awareness can be manifold: endogenous factors like lack of attention or lack of expertise or exogenous factors like time pressure or ambiguous task situations. Some psychiatric or neurological diseases are at least at the surface also related to impaired processing of negative action outcomes.

Responses to action errors are not only restricted to brain responses: also vegetative changes can be observed (for example pupil diameter and heart rate). It is still matter of an ongoing debate whether these autonomous changes are cause or consequence of error awareness. Research on error awareness-related brain activity requires experimental paradigms with which error awareness can be manipulated. Given the many different influences on error awareness described above this is not an easy task.

Why should one investigate the distinction between aware and unaware errors then? Is it worthwhile that every error becomes aware? Or isn't it sometimes more economical that mistakes are going unrecognized? There are fields in human life in which high accuracy is necessary and errors need to be addressed: for example physicians should always be aware of irregularities within their professional routine. For this purpose it would of course be highly relevant to know about factors that promote error awareness. Whereupon it is an interesting question in itself whether it needs awareness/consciousness to trigger compensatory behavior.

Also from the clinical perspective it might be relevant to understand error awareness: either to improve awareness (for example in anosognosia) or to limit awareness (for example in obsessive-compulsive or affective disorders).

In order to tackle these practical/clinical questions basic issues need to be addressed: are there separate functional correlates of processing aware and unaware errors? Is there a relationship between electrophysiological correlates of conscious error processing (error-related negativity, ERN) and the electrophysiological response sensitive to error awareness (the error-positivity, PE)? Is there a continuum spanning from aware to unaware errors? And if so, is it just an accumulation of evidence that makes the difference between the two extremes? What cognitive tasks are actually suited best to investigate unaware errors?

The E-Book on error awareness tries to address some of these questions in form of review articles and original research papers. The authors cover different methodological (EEG/ fMRI, temporospatial/ICA analysis approaches), theoretical (ERN vs. PE, evidence accumulation, role of consciousness, saliency), clinical/pharmacological (schizophrenia, serotonin) and functional/structural anatomical (anterior cingulate cortex; insular cortex) topics. Taken together the articles cover large parts of current debate in error awareness research.

The E-Book starts with an article addressing the role of consciousness in cognitive control and decision making. van Gaal et al. (2012) review studies on unconscious information processing thereby showing that unconscious information is well capable of influencing different aspects of cognitive functioning or information processing. Sharpening the theoretical focus into the direction of error processing Harsay et al. (2012) show that there is indeed a functional overlap in brain structures for processes that are related on a theoretical level: (aware) error processing and processing of salient events in the environment. Especially the insular cortex, a region that will be mentioned in some other contributions, too, is engaged in both processes. Steinhauser and Yeung (2012) contribute another interesting theoretical consideration: they propose that error awareness is the result of evidence accumulation which in turn is reflected in an electrophysiological marker within the brain-electrical response to performance errors: the Pe. In his review, Wessel (2012) discusses the question whether error awareness is purely reflected in the error positivity or what the relation between the ERN and error awareness might be. Shalgi and Deouell (2012) follow a similar argument: they show that the amplitude of the ERN is indeed related to error awareness. They relate differences in error awareness to differences in actually reporting erroneous reactions and the confidence with which this report is made. The following two papers present different accounts of evaluating/assessing electrical brain responses to errors: Murphy et al. (2012) used independent component analysis to isolate the Pe within their EEG signal. This analysis enabled them to check for correlations between Pe amplitude and for example latency of the awareness-report response on single trial basis. Endrass et al. (2012) put forward a more spatially oriented analysis: they used principal component analysis in order to characterize spatial factors contributing to the Pe: a centro-parietal positivity correlated significantly with error awareness. Addressing error awareness

with a higher spatial resolution Orr and Hester (2012) report data from functional magnetic resonance imaging (fMRI) showing that error-related activity in the dorsal anterior cingulate cortex is significantly higher in aware as compared to unaware errors. Most interestingly this area is often discussed as being one potential generator for the ERN thereby pointing back to the considerations of Wessel (2012) and Shalgi and Deouell (2012) reflecting on awareness influences on the ERN. The following three articles extend the focus of discussion to the fields of underlying neurochemical mechanisms, anatomical and clinical considerations. Mueller et al. (2012) showed that the neurotransmitter serotonin seems to be important for regulating the coupling between cortical and cardiac responses in performance monitoring. This seems especially relevant given the fact that the insular cortex is heavily involved in detecting bodily changes—a function which might contribute to error awareness. This central role of the Insula in error awareness is discussed in a review by Klein et al. (2013) with a special emphasis on anatomical considerations of Insula subdivisions and related functional aspects. Klein et al. (2013) also provide some information on pathological states that might alter error awareness and might as well be associated with changes in the insular cortex. Clinical aspects are rounded up by Núñez Castellar et al. (2012) showing that schizophrenic patients have more difficulties in reaching adequate performance levels in a performance monitoring task and that related adaptations in behavior following an error are reduced in these patients.

We believe that the collection of excellent scientific contributions gathered in this E-Book provides important new insights into the mechanisms and implications of conscious error perception. In addition, a number of outstanding questions can be derived from the reported findings, for example: Is error awareness a necessary and sufficient pre-requisite of post-error adjustments? Do post-error adjustments differ with respect to their association with error awareness? Can we infer on conscious error perception just based on brain data in the absence of interospective judgments? Which brain lesions and other neurological and psychiatric disorders interfere with error awareness? Hence, this E-Book can and should serve as the basis for new research lines in this field.

REFERENCES

Endrass, T., Klawohn, J., Preuss, J., and Kathmann, N. (2012). Temporospatial dissociation of Pe subcomponents for perceived and unperceived errors. *Front. Hum. Neurosci.* 6:178. doi: 10.3389/fnhum.2012.00178

Harsay, H. A., Spaan, M., Wijnen, J. G., and Ridderinkhof, K. R. (2012). Error awareness and salience processing in the oddball task: shared neural mechanisms. *Front. Hum. Neurosci.* 6:246. doi: 10.3389/fnhum.2012.00246

Klein, T. A., Ullsperger, M., and Danielmeier, C. (2013). Error awareness and the insula: links to neurological and psychiatric diseases. *Front. Hum. Neurosci.* 7:14. doi: 10.3389/fnhum.2013.00014

Mueller, E. M., Evers, E. A., Wacker, J., and van der Veen, F. (2012). Acute tryptophan depletion attenuates brain-heart coupling following external feedback. *Front. Hum. Neurosci.* 6:77. doi: 10.3389/fnhum.2012.00077

Murphy, P. R., Robertson, I. H., Allen, D., Hester, R., and O'Connell, R. G. (2012). An electrophysiological signal that precisely tracks the emergence of error awareness. *Front. Hum. Neurosci.* 6:65. doi: 10.3389/fnhum.2012.00065

Núñez Castellar, E., Houtman, F., Gevers, W., Morrens, M., Vermeylen, S., Sabbe, B., et al. (2012). Increased orienting to unexpected action outcomes in schizophrenia. *Front. Hum. Neurosci.* 6:32. doi: 10.3389/fnhum.2012.00032

Orr, C., and Hester, R. (2012). Error-related anterior cingulate cortex activity and the prediction of conscious error awareness. *Front. Hum. Neurosci.* 6:177. doi: 10.3389/fnhum.2012.00177

Shalgi, S., and Deouell, L. Y. (2012). Is any awareness necessary for an Ne? *Front. Hum. Neurosci.* 6:124. doi: 10.3389/fnhum.2012.00124

Steinhauser, M., and Yeung, N. (2012). Error awareness as evidence accumulation: effects of speed-accuracy trade-off on error signaling. *Front. Hum. Neurosci.* 6:240. doi: 10.3389/fnhum.2012.00240

van Gaal, S., de Lange, F. P., and Cohen, M. X. (2012). The role of consciousness in cognitive control and decision making. *Front. Hum. Neurosci.* 6:121. doi: 10.3389/fnhum.2012.00121

Wessel, J. R. (2012). Error awareness and the error-related negativity: evaluating the first decade of evidence. *Front. Hum. Neurosci.* 6:88. doi: 10.3389/fnhum.2012.00088

Received: 06 November 2013; accepted: 17 November 2013; published online: 05 December 2013.

Citation: Klein TA, Danielmeier C and Ullsperger M (2013) Error awareness—insights from cognitive neuroscience, psychiatry and neurology. Front. Hum. Neurosci. 7:830. doi: 10.3389/fnhum.2013.00830

This article was submitted to the journal Frontiers in Human Neuroscience.

frontiers in
HUMAN NEUROSCIENCE

REVIEW ARTICLE
published: 07 May 2012
doi: 10.3389/fnhum.2012.00121

The role of consciousness in cognitive control and decision making

***Simon van Gaal*[1,2]*, *Floris P. de Lange*[3] *and Michael X Cohen*[4]**

[1] *Institut National de la Santé et de la Recherche Médicale, Cognitive Neuroimaging Unit, Gif-sur-Yvette, France*
[2] *Commissariat à l'Energie Atomique, Neurospin Center, Gif-sur-Yvette, France*
[3] *Donders Institute for Brain, Cognition and Behavior, Radboud University Nijmegen, Nijmegen, Netherlands*
[4] *Department of Psychology, University of Amsterdam, Amsterdam, Netherlands*

Edited by:
Tilmann A. Klein, Max Planck Institute for Human Cognitive and Brain Sciences, Germany

Reviewed by:
Tobias Egner, Duke University, USA
Catherine A. Orr, University of Melbourne, Australia

***Correspondence:**
Simon van Gaal, Institut National de la Santé et de la Recherche Médicale U992/NeuroSpin, Cognitive Neuroimaging Unit, Gif-sur-Yvette, CEA - Saclay, France.
e-mail: simonvangaal@gmail.com

Here we review studies on the complexity and strength of unconscious information processing. We focus on empirical evidence that relates awareness of information to cognitive control processes (e.g., response inhibition, conflict resolution, and task-switching), the life-time of information maintenance (e.g., working memory) and the possibility to integrate multiple pieces of information across space and time. Overall, the results that we review paint a picture of local and specific effects of unconscious information on various (high-level) brain regions, including areas in the prefrontal cortex. Although this neural activation does not elicit any conscious experience, it is functional and capable of influencing many perceptual, cognitive (control) and decision-related processes, sometimes even for relatively long periods of time. However, recent evidence also points out interesting dissociations between conscious and unconscious information processing when it comes to the duration, flexibility and the strategic use of that information for complex operations and decision-making. Based on the available evidence, we conclude that the role of task-relevance of subliminal information and meta-cognitive factors in unconscious cognition need more attention in future work.

Keywords: consciousness, cognitive control, decision-making, awareness, unconscious

INTRODUCTION

Although at first controversial, it is now generally accepted that several perceptual, emotional and cognitive processes can unfold in the absence of awareness. Laboratory examples of this are the processing of subliminal (unconscious) words and numbers (Marcel, 1983; Merikle and Reingold, 1990; Greenwald et al., 1996; Dehaene et al., 1998; Kiefer and Spitzer, 2000), pictures of faces/houses (Sterzer et al., 2008; Kouider et al., 2009), tools (Fang and He, 2005), and emotional material such as angry faces or eye-gaze directions (Whalen et al., 1998, 2004). Going beyond visual perception, subliminal information processing has been demonstrated for auditory (Sadaghiani et al., 2009), somatosensory (Eimer et al., 2002) and olfactory (Li et al., 2007) information. In recent years the number of processes that operate or are influenced unconsciously has increased steadily, and now include reward- and motivation-related processes (Custers and Aarts, 2005; Pessiglione et al., 2007; Capa et al., 2011) as well as decision-making (Bargh and Morsella, 2008; Pessiglione et al., 2008; Custers and Aarts, 2010). One might wonder whether there are any processes that can exclusively be performed on conscious information. In this paper we review studies that explored the boundary conditions of unconscious information processing and specifically highlight those studies that were aimed at testing the role of consciousness in cognitive control, long-term information maintenance and strategic decision-making. Here, we mainly focus on cognitive aspects of unconscious information processing, although neural data is also discussed (see van Gaal and Lamme, in press for a review of the literature in the field from a more neural perspective).

CONSCIOUSNESS AND COGNITIVE CONTROL

Some authors have argued that there might be some (cognitive) processes truly bound to consciousness, although this is strongly debated (for reviews see Umilta, 1988; Dehaene and Naccache, 2001; Jack and Shallice, 2001; Mayr, 2004; Hommel, 2007; Kunde et al., 2012). One of the major candidates for this is cognitive control, a general term for cognitive functions that allow us to rapidly and flexibly adapt our behavior when necessary. Cognitive control functions include error detection and correction mechanisms, conflict resolution, response inhibition, and task-switching. These functions are all strongly associated with the prefrontal cortex, which many consider pivotal for generating awareness (for reviews see Rees, 2007; Dehaene and Changeux, 2011; Lau and Rosenthal, 2011).

Interestingly, some cognitive control processes can be activated by unconscious stimuli. To our knowledge, the first to show that some control processes can be initiated fully automatically and unconsciously were Eimer and Schlaghecken (Eimer and Schlaghecken, 1998; Eimer, 1999). In an impressive set of studies, they showed that unconscious (masked) arrow primes initially facilitated responses, but can also inhibit responses in certain circumstances. In their tasks, subjects generally have to respond to a target-arrow (e.g., $\gg$) that can be preceded by a congruent ($\gg$) or incongruent ($\ll$) masked prime-arrow. When the interval

between the prime and target is short (e.g., 50 ms), subjects respond faster and make fewer errors to congruent than to incongruent trials, as might be expected. However, crucially, when the delay between prime and target was increased (>100 ms), there was no response facilitation but rather automatic inhibition of these responses. This led to the counterintuitive observation that response times (RTs) were faster and error rates lower to incongruent trials compared to congruent trials (note that part of the effect might be explained by lower-level stimulus characteristics, see Lleras and Enns, 2004; Jaskowski and Przekoracka-Krawczyk, 2005; Schlaghecken and Eimer, 2006). More recently, automatic inhibition paradigms have been combined with brain-imaging tools and the results suggest that automatic inhibition relies on activity in the caudate and thalamus (Aron et al., 2003) as well as the supplementary motor areas (Sumner et al., 2007).

Recent studies have demonstrated the possibility to initiate more "voluntary" forms of response inhibition unconsciously, as studied by using the Go/No-Go task and the stop-signal paradigm (Hughes et al., 2009; van Gaal and Lamme, in press). In these tasks, subjects are required to inhibit an already initiated (stop task) or planned response (Go/No-Go task). To illustrate, in one of these experiments, subjects were instructed to respond as fast as possible to the direction of an arrow (the go stimulus), but to withhold this response when the word "STOP" (the "stop stimulus") was presented briefly and quickly after the go-arrow (**Figure 1A**). However, when another word (e.g., "BLUF," the "go-on stimulus") was presented, subjects had to continue responding

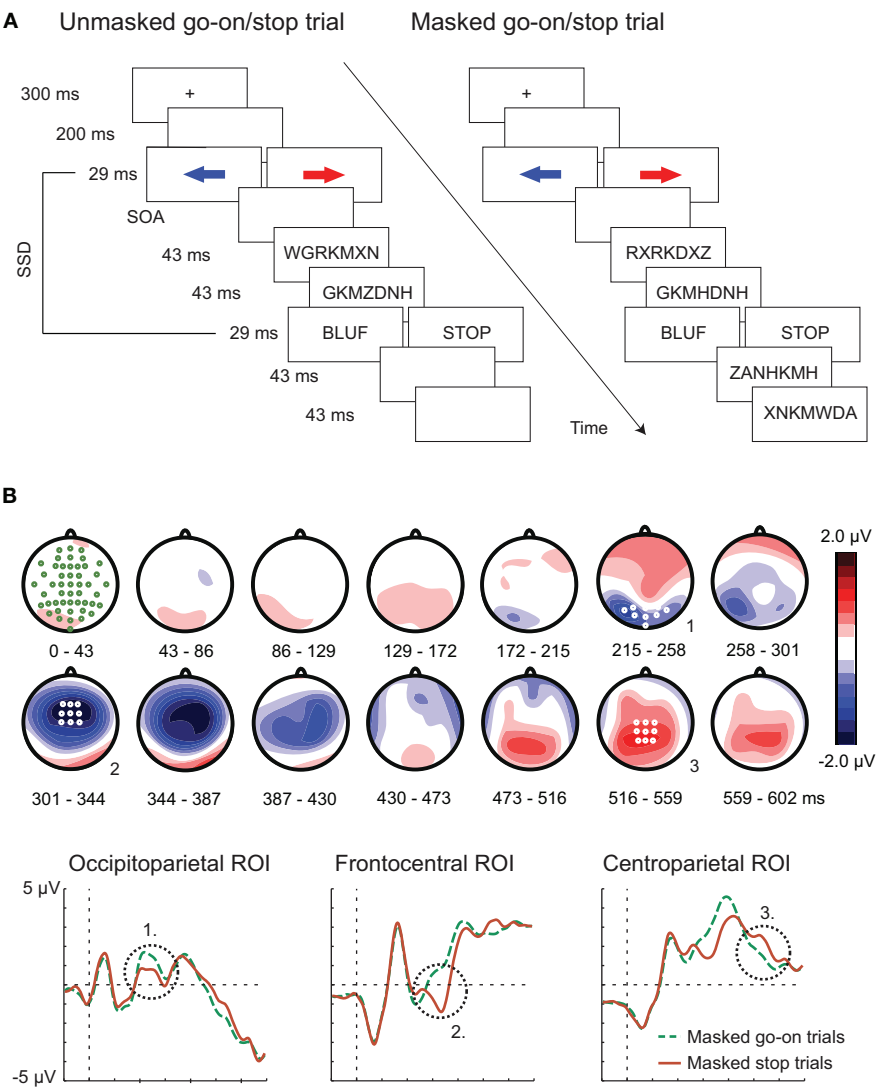


FIGURE 1 | (A) Task-set-up. **(B)** Electrophysiological processing across time of a masked stop-signal (the word "STOP") compared to a control "go-on" condition (e.g., the word "BLUF"). Three neural events can be distinguished: (1) an early event at occipital electrodes, (2) a middle event at fronto-central electrodes (The N2 ERP component), and (3) a late event at centro-parietal electrodes (The P3 ERP component). Adapted with permission from van Gaal et al. (2011).

to the direction of the go-arrow. Crucially, the visibility of the stop/go-on stimulus was manipulated by presenting it in between random letter masks. Therefore, on some trials these stimuli were clearly visible, whereas on other trials they were not. Behaviorally, subjects slowed down their responses to unconscious stop-signals (compared to unconscious go-on signals), as if the STOP signal was briefly processed but not enough to cause a full response inhibition. When electrophysiological responses to unconscious stop- and go-on signals were compared, a cascade of neural events could be observed, starting early at occipital electrodes, swiftly progressing to fronto-central (the N2 ERP component) and centro-parietal electrode sites (the P3 ERP component), later in time (**Figure 1B**). Interestingly, in the conscious condition the magnitude of the N2 ERP component was correlated with the efficiency of inhibitory control across subjects (the stop-signal reaction time) and with the magnitude of slowdown to unconscious stop-signals. Thus, the N2 ERP component likely reflects the initiation of inhibitory control, irrespective of the conscious awareness of the control-initiating stop-signal. The frontal origin of this effect has been confirmed by source reconstruction of the EEG signals (van Gaal et al., 2008) as well as by fMRI (van Gaal et al., 2010b), in similar tasks. In fMRI, RT slowing to unconscious No-Go signals was associated with focal activations in the pre-SMA and inferior frontal cortices, bordering anterior insula (van Gaal et al., 2010b), whereas response inhibition to visible No-Go signals was related to large scale activation in a typically observed fronto-parietal "inhibition network" (Aron, 2007; Simmonds et al., 2008). The strength of activation in the unconscious inhibition network was correlated with the extent of slowdown to unconscious No-Go signals across subjects, suggesting that this activation is functional in a sense that it is related to behavioral effects of cognitive control.

Recent results suggest that several cognitive control functions other than response inhibition can be triggered by unconscious or unnoticed stimuli (for a recent review see van Gaal and Lamme, in press), including task-set preparation (Mattler, 2003; Lau and Passingham, 2007; van Opstal et al., 2010; de Pisapia et al., 2011; Reuss et al., 2011; Zhou and Davis, 2012), conflict detection/resolution (Ursu et al., 2009; D'Ostilio and Garraux, 2012) (but see Dehaene et al., 2003; Bruchmann et al., 2011), motivation (Pessiglione et al., 2007; Aarts et al., 2008; Custers and Aarts, 2010) and error detection (Nieuwenhuis et al., 2001; Hester et al., 2005; Klein et al., 2007; O'Connell et al., 2007; Belopolsky et al., 2008; Cohen et al., 2009; Pavone et al., 2009; Dhar et al., 2011) (but see Woodman, 2010). To illustrate, Lau and Passingham (2007) cued participants consciously to perform either a phonological or semantic judgment on an upcoming word. This conscious instruction cue was always preceded by a conscious or unconscious prime associated with the same or the alternative task (congruent vs. incongruent trials) (see also Mattler, 2003 for a behavioral version of this experiment). When participants were unconsciously primed to perform the phonological task, there was increased activity in a cortical network associated with this task (premotor cortex) and decreased activity in the cortical network associated with the semantic task (inferior frontal cortex and middle temporal gyrus), and vice versa. These results demonstrate that task-related neural networks, incorporating prefrontal cortex, can be modulated unconsciously. Further, the authors showed that unconscious primes triggered stronger activity in the dorsolateral prefrontal cortex compared to conscious primes, irrespective of the specific task being cued. Recently, Zhou and Davis (2012) went one step further and demonstrated that this effect was not caused by low-level perceptual priming and could still be observed when the unconscious cue was not part of the consciously instructed task-set.

Although it has repeatedly been observed that the strength of unconscious information processing increases considerably with practice and learning (Damian, 2001; van den Bussche et al., 2009; van Gaal et al., 2009), it has been shown that strong stimulus-response bindings are not a prerequisite for subliminal processing to occur (for a meta-analysis see van den Bussche et al., 2009). However, primes that are also included as targets ("repeated primes") have a stronger impact and might affect motor responses earlier (have a faster time-course) than primes that are not included as targets ("novel primes") (Finkbeiner and Friedman, 2011). In fact, also for higher-level cognitive control processes, such as response inhibition, stimulus-response mappings can be flexibly changed without abolishing unconscious priming effects. In a task in which a masked stimulus (diamond or square) could be associated with either a Go or No-Go response, but the specific mapping of stimuli onto these actions varied on a trial-by-trial basis (by virtue of a pre-cue), it was recently demonstrated that the same unconscious stimulus could have a substantially different effect on behavior and (prefrontal) brain activity depending on the rapidly changing task-context in which it was presented (Wokke et al., 2011).

In conclusion, several "high-level" (prefrontal) cognitive functions, such as response inhibition and task-switching, have been observed to be influenced and modulated by subliminal stimuli. These activations seem truly functional, because they are associated with behavioral indices of cognitive control. In the next section, we will discuss the influence of top-down factors (e.g., attention, task-set) on the extent of subliminal information processing and whether subliminal information can initiate top-down cognitive task-sets itself.

CONSCIOUS AWARENESS AND TOP-DOWN COGNITIVE CONTROL

Traditionally, it has been assumed that unconscious processes were rather automatic, inflexible, and independent of top-down cognitive control (see Hommel, 2007; Kiefer et al., 2012 for reviews). However, accumulating evidence shows that unconscious information processing is not fully automatic, but can be modulated by several top-down cognitive and attentional factors. Overall, the instructed task-set and subjects' strategy strongly affects the strength, direction and depth of subliminal information processing (Kunde et al., 2003; Greenwald et al., 2003; Ansorge and Neumann, 2005; Kiefer and Martens, 2010; Al-Janabi and Finkbeiner, 2011; O'Connor and Neill, 2011). For example, the top-down instructed task-set, e.g., either to read aloud a visible target word or to categorize it as representing natural or artificial objects, can change the processing route taken by an unconscious (masked) word preceding the target word (Nakamura et al., 2007). Along similar lines, Kiefer and Martens

(2010) recently showed that the N400 ERP component to unrelated prime-target pairs (e.g., masked word "chair" followed by a visible target word "leaf"), compared to related prime-target pairs (e.g., masked word "chair" followed by visible target word "table"), was enhanced when a semantic task-set was induced by a visible cue presented immediately before each trial and was attenuated by a perceptual task-set (see also Martens et al., 2011). Further, attended subliminal stimuli have a stronger impact on behavior than unattended subliminal stimuli, and this is the case for spatial attention (Kentridge et al., 1999, 2004, 2008; Sumner et al., 2006; Bahrami et al., 2008a; Marzouki et al., 2008; Finkbeiner and Palermo, 2009), temporal attention (Naccache et al., 2002; Kiefer and Brendel, 2006; Fabre et al., 2007) and during attentional load (Bahrami et al., 2008b; Martens and Kiefer, 2009).

Task-relevant (attended) stimuli are processed stronger than task-irrelevant (unattended) stimuli, even when unconscious. Ansorge and Neumann (2005) showed that task-relevant prime features (e.g., shape) affected responses to the target only when the shape dimension was response relevant, but not when this feature was task-irrelevant, for example when the color of the target determined the required response (see also Tapia et al., 2010). We recently explored the role of task-relevance of subliminal information using EEG in a task in which subjects had to respond as fast as possible to a black Go annulus, unless it was preceded by a briefly presented gray circle (the no-go stimulus). Due to variations in the SOA between the No-Go circle and Go annulus, on some trials the No-Go circle was perceived consciously, whereas on others it was not. On the current trial, unconscious No-Go circles activated prefrontal control networks (van Gaal et al., 2008), and the extent to which correlated strongly with the amount of RT slowing to these stimuli. Crucially, exactly the same subliminal gray circle did not activate the PFC when it was task-irrelevant, but presented in a highly similar task-context (although it yielded similar early visual responses). This result highlights that the processing route taken by an unconscious stimulus strongly depends on task-relevance (and attention to the stimulus), and that task-irrelevant subliminal stimuli probably decay rapidly while progressing up in the cortical hierarchy. Recently, it has been observed that, under some conditions, cognitive control processes can still be influenced by subliminal stimuli presented outside the direct focus of spatial attention (Rahnev et al., 2012).

The role of attention and other top-down factors for unconscious information processing might depend on type of information to be processed. Recent research suggests that attention might be more crucial for "neutral" stimuli (e.g., numbers: Naccache et al., 2002) than for emotional, arousing or "evolutionary relevant" stimuli. To illustrate, Finkbeiner and Palermo (2009) have found that masked pictures of face stimuli produced priming regardless of whether they were spatially attended (however, this was not the case for subliminal eye-gaze cues: Al-Janabi and Finkbeiner, 2011). In contrast, other non-face stimuli (animals, vegetables) only produced subliminal priming when attended (see also Harry et al., 2012). However, although it seems that the threshold for conscious access is lower for emotional stimuli (Gaillard et al., 2006) and that these produce stronger priming (Brooks et al., 2012), also emotional information processing does not seem to be fully automatic and is also modulated by top-down "attentional sensitization," at least to some extent (Kiefer et al., 2012). In fact, even when emotional pictures (e.g., faces) are presented fully consciously their depth and extent of processing seem to be facilitated by attentional factors (Pessoa et al., 2002, 2003).

Attention itself can also be attracted unconsciously (for review see Mulckhuyse and Theeuwes, 2010), for example by threatening (Lin et al., 2009), emotional (Vuilleumier and Schwartz, 2001; Brooks et al., 2012), erotic (Jiang et al., 2006), or socially relevant stimuli (Sato et al., 2007), but also by lower-level stimulus attributes, such as gamma flicker (Bauer et al., 2009) and stimulus orientation (Rajimerhr, 2004). Recently, it has been shown that individual differences in attentional bias to masked fearful faces are related to gray matter volume in the anterior cingulate cortex (Carlson et al., 2012), suggesting that these attentional effects are truly top-down mediated.

The literature reviewed above illustrates that consciously instructed task-sets and strategies as well as attentional factors strongly influence the processing of subliminal stimuli in various ways. At present, it is still an open and important question whether top-down task-sets can also be triggered by subliminal information. Several studies have reported so-called "top-down context effects." In these experiments, subjects generally perform a masked priming task consisting of congruent and incongruent prime-target pairs. The crucial manipulation in such experiments is the ratio of congruent and incongruent trials within experimental blocks. In blocks in which the prime direction does not predict the direction of the upcoming target (50% congruent and 50% incongruent trials) subjects are generally faster to congruent than to incongruent trials. However, several experiments have consistently revealed that the impact of conflicting stimuli on behavior is larger when incongruent prime-target pairs are infrequent (∼20%) compared to when these are frequent (∼80%), at least when conflicting stimuli are presented consciously (for review see Desender and van den Bussche, 2012). In fact, the effect might even completely reverse in such a way that responses to incongruent prime-target pairs are faster than to congruent pairs (Merikle and Joordens, 1997; Daza et al., 2002), because subjects are able to strategically use the prime information to predict the upcoming target category. Even for conscious trials this might take some time (∼400 ms), suggesting that these strategic effects take some time to build up (Ortells et al., 2003). These conscious strategic effects were recently only observed for spatially attended stimuli, but not for unattended ones (Ortells et al., 2011). At present it is still disputed whether such context effects depend on the conscious awareness of the primes, because several studies have reported an absence of congruency effects when the conflicting stimuli were presented subliminally (Merikle and Joordens, 1997; Daza et al., 2002; van den Bussche et al., 2008; Heinemann et al., 2009). However, other studies have shown that context effects also apply to unconscious prime stimuli (Jaskowski et al., 2003; Bodner and Masson, 2004; Wolbers et al., 2006; Klapp, 2007; Bodner and Mulji, 2010). Interestingly, these context effects initiated by subliminal primes might be related to increased connectivity between the pre-SMA and stimulus-related (LOC) and motor-related (putamen) brain areas (Wolbers et al., 2006),

suggesting that the pre-SMA plays a role in the strategic control over the processing of subliminally presented conflicting stimuli.

Several authors have noted that it is important to examine whether these context effects are truly unconscious, at all processing levels, and which part of the effect might be explained by meta-cognitive (conscious) processes. For example, subjects might become aware of the increased error rate, experienced "difficulty" or "effort" on blocks with high numbers of conflicting trials and thereby might strategically adapt their response strategy or attentional focus (Jaskowski et al., 2003; Kinoshita et al., 2008, 2011, for a more extensive discussion of this issue see Desender and van den Bussche, 2012 and below). Therefore, it is still an open question whether top-down context effects can also be initiated by unconscious stimuli (Dehaene and Naccache, 2001).

Heinemann et al. (2009) studied the role of conflict awareness in a slightly different way, namely by examining the role of context on conflict frequency effects, also referred to as the context-specific proportion congruent effect (see also Crump et al., 2006). They performed a typical masked priming task in which subjects had to categorize target numbers as being larger or smaller than 5. A target was always preceded by a masked prime number that could be congruent or incongruent to the target. Crucially, just before the presentation of the prime-target pair they presented a colored rectangle at the background that determined the congruency context (the colored rectangle disappeared upon presentation of the response feedback). One color was consistently associated with a low interference context (80% congruent trials, 20% incongruent trials), whereas another color was associated with a high interference context (20% congruent trials, 80% incongruent trials). As predicted, for weakly masked primes (visible) the congruency effect (RT incongruent—RT congruent) was significantly smaller in the high interference context than in the low interference context (32 vs. 54 ms). Crucially, they showed that these context-specific congruency effects were absent for strongly masked (poorly visible) trials. The authors concluded that context-specific congruency adaptation requires conscious representation of the conflicting information. Interestingly, previous work suggests that, even when using visible stimuli only, subjects do not have any explicit awareness of the congruency manipulation in similar tasks (Crump and Milliken, 2009). Therefore, it has been suggested that context-specific congruency effects might not depend on explicit knowledge of the congruency proportions, but might require sufficiently strong (i.e., conscious) representations of the prime, target and context (Kunde et al., 2012).

In a recent study, van Opstal et al. (2011a) took a somewhat different approach and demonstrated that context effects might indeed be initiated by subliminal primes. In their task, subjects had to indicate whether two target numbers (e.g., 3–3) were the same or different. These target numbers were always preceded by a masked (subliminal) prime. The crucial comparative prime consisted of a capital letter and a lower-case letter (A-a). In one experiment these primes were mixed with primes consisting of two completely different letters (a-D, the low-similarity context, **Figure 2A**), whereas in another experiment they were mixed with primes consisting of exactly the same letters (a-a, the high-similarity context, **Figure 2B**). In the low-similarity context where a-A primes were relatively similar to a-a primes (compared to a-D primes), a-A primes facilitated a "same" response to the targets (**Figure 2C**). On the other hand, the same prime (a-A) presented in the high-similarity context (containing a-a

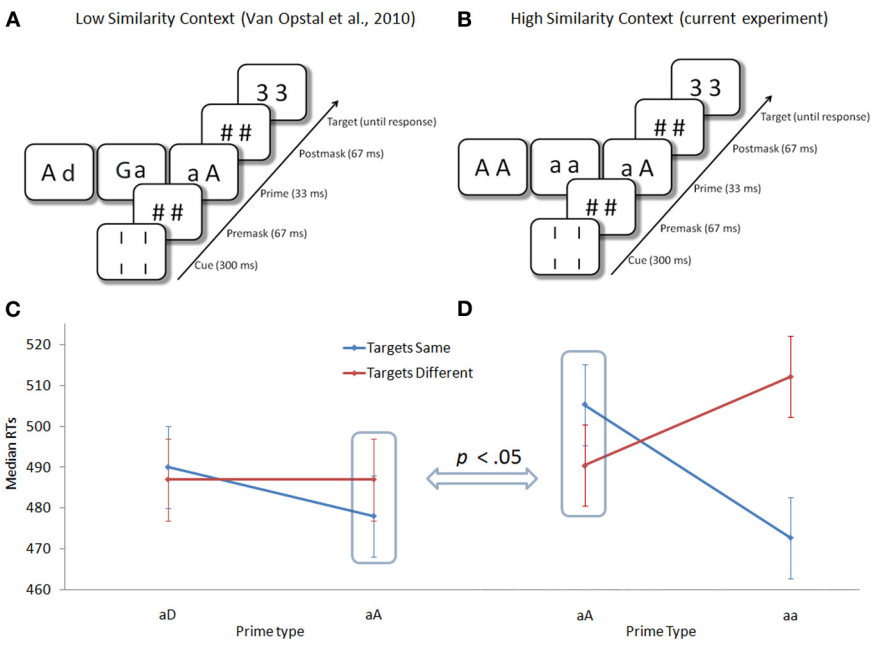


FIGURE 2 | (A,B) Task-set-up. **(C,D)** Response times for the low-similarity and high-similarity context for the different prime types and prime-target congruency. Adapted with permission from van Opstal et al. (2011a).

primes) was relatively different and indeed facilitated a "different" response to the targets (**Figure 2D**). Importantly, RTs were equated across conditions and, therefore, could not (directly) explain the observed effects. This may be an important step in further pushing the boundaries of unconscious information processing and opens the possibility that also a top-down task-set may be enabled unconsciously (Dehaene and Naccache, 2001).

By now it is well established that subliminal information processing (e.g., its depth, extent, and direction) is influenced by several top-down cognitive functions, such as attention, task-set and strategy. However, whether top-down context effects themselves can be initiated or affected by subliminal stimuli is still under scrutiny (see also Kunde et al., 2012). In this respect, the underlying mechanisms, boundary conditions and role of awareness in blockwise congruency effects (Desender and van den Bussche, 2012) and context-specific proportion effects (Heinemann et al., 2009) are interesting avenues for future experimentation. Next, we will discuss another crucial and disputed aspect of subliminal information processing: its alleged short-lived nature.

THE LIFE-TIME OF SUBLIMINAL INFORMATION PROCESSING

The amount of time that unconscious information can influence cognitive processing is strongly debated. Typical masked priming studies have reported a sharp decrease in the processing of masked primes with the passage of time and it has been observed that the influence of unconscious stimuli disappears within approximately half a second (Greenwald et al., 1996; Mattler, 2005). Although this initial evidence strongly pointed into the direction of a severely limited duration of unconscious processes (see also Kunde, 2003), recent studies question this assumption and have shown subtle influences of subliminal information across relatively long periods of time.

A recent fMRI study showed strong and long-lasting effects of unconscious task-relevant stimuli that were used for instrumental conditioning (Pessiglione et al., 2008). Pessiglione and colleagues performed a study in which subjects had to learn stimulus-value associations for stimuli that were presented below the threshold of awareness because of pattern masking. Some masked cues were paired with monetary gains whereas others were coupled with monetary losses. Subjects performed a task in which they could opt for a safe choice that was never rewarded or punished, or they could opt for a risky choice that could be rewarded (+£1), punished (−£1), or remain neutral. Behaviorally, subjects chose the risky response more often following reward predictive (masked) cues compared to punishment predictive (masked) cues. Further, the effect was larger toward the end of the experiment after many trials had been presented. The implicit learning of cue-value associations was related to activation in the ventral striatum. Arguably, one of the most surprising and interesting aspects of these results was that the delay between the masked cue and the eventual decision could be up to 3 s, suggesting a relatively long duration of unconscious information processing (although an RT analysis was not reported, precluding an analysis of which prime processing durations were driving the behavioral effect).

Other long-term subliminal effects were recently reported in various memory paradigms. Although consciousness and (working) memory seem intimately related (Dehaene and Naccache, 2001; Baars and Franklin, 2003), recent studies suggest that at least some components of memory might also operate outside of consciousness. In several experiments, Voss and colleagues demonstrated another form of memory, namely recognition memory, without explicit stimulus awareness (Voss et al., 2008; Voss and Paller, 2009). They used colorful complex geometric shapes (also referred to as "kaleidopscope images") that were shown in a learning phase and were tested in a recognition test afterwards (the test-set is relatively large and the stimuli are highly similar). When new and old items were presented simultaneously in the recognition phase, subjects often make correct forced-choice decisions about these images without any explicit recognition or subjective confidence; in fact, subjects typically felt they were merely guessing (see also Jeneson et al., 2010). Interestingly, subjects' guess responses were more likely to be accurate when the stimuli were initially presented during divided-attention than during full-attention in the study phase, and when subjects were encouraged to guess instead of encouraged responding confidently during the test phase. It seems that instructions that encourage guessing can facilitate responses based on rapid visual information processing, without competing (and distracting) influences from explicit retrieval processes (Voss and Paller, 2010).

Recently Soto et al. (2011) went one step further and specifically tested the relation between working memory and awareness (see also Hassin et al., 2009). They briefly presented either a Gabor cue (16.7 ms) with a specific orientation, or a blank screen, followed by a mask. After a retention interval (2–5 s) a test Gabor stimulus was presented and subjects had to indicate whether the orientation of the masked Gabor cue was tilted clockwise or counter-clockwise with respect to the orientation of the target Gabor. Following this orientation response, subjects had to indicate their subjective awareness of the masked Gabor cue on a 4-point scale (ranging from 1 = "did not see anything," to 4 = "saw the stimulus and its orientation"). For subjectively invisible Gabor cues (all "1" responses) objective orientation comparisons with the target were above chance level (generally just above 55% in several experiments), even when a conscious or unconscious distractor Gabor was presented in the retention interval. Based on this, the authors concluded that "visual memory can encode, maintain and access unconscious items for explicit discrimination goals" (p. 913). In this experimental setup, the authors' interpretation relies on the assumption that the subjective awareness measure was sensitive enough to fully isolate unaware Gabor cues from (partly) aware Gabor cues. At present, this issue needs some further exploration (Overgaard et al., 2006; Block, 2011), mainly because subjects generally used the lower ends of the subjective awareness scale (55% of the trials fell in category 1, 34% in category 2, 5% in category 3 and 6% in category 4) and because objective Gabor detection performance (compared to "nothing" trials) was above chance-level for all "1" responses ($d' = 0.297$). However, if confirmed, the demonstration of unconscious working memory, resistant to distraction, may have large implications for neurobiological and cognitive theories of consciousness.

In all of the discussed studies so far the subliminal stimulus affected (behavioral) responses directly, on the current trial.

However, longer-lasting trial-by-trial modulatory effects of subliminal stimuli are sometimes also reported, and the role of awareness in both cognitive processes might differ considerably (Dehaene and Naccache, 2001; Boy et al., 2010). Typical trial-by-trial modulations are the slowing of responses after errors (post-error slowing) and the reduction of conflict interference after high-conflict compared to low-conflict trials (conflict adaptation). These effects are generally thought to originate from increased prefrontal top-down control triggered by the error/conflict signal (Cohen et al., 2000; Botvinick et al., 2001; Ridderinkhof et al., 2004), although other interpretations have also been proposed (Hommel et al., 2002; Mayr et al., 2003; Notebaert et al., 2009; Schlaghecken and Martini, 2012). It is debated whether unconscious information can elicit strategic, trial-by-trial and long-lasting (top-down) modulations over subsequent stimulus processing (Desender and van den Bussche, 2012). We will discuss the potential role of awareness in conflict adaptation first (Greenwald et al., 1996; Kunde, 2003).

In a seminal study published in 1996, Greenwald and colleagues demonstrated that a subliminal prime-target pair (consisting of valenced words, e.g., bomb-kiss) leaves no "memory trace" that influences responding to the next prime-target pair (Greenwald et al., 1996). Some years later, Kunde (2003) came to the same conclusion. In his experiment, participants performed a speeded two-choice response to a target arrow that was preceded by a smaller arrow (the prime). Because the prime fitted within the contour of the target, the target functioned as a (metacontrast) mask and ensured that participants did not become aware of it when it was presented briefly (14 ms), but they did when it was presented somewhat longer (126 ms). Although in the masked conditions, the prime could not consciously be perceived, RTs were faster and subjects made fewer errors when the prime and target were congruent than when they were incongruent (i.e., an unconscious correspondence effect). In contrast to these same-trial effects, conflict adaptation (when the correspondence effect on trial *n* is smaller when trials were preceded by an incongruent trial compared to a congruent trial on trial *n*-1, Gratton et al., 1992), was only the case when primes were presented consciously, and not when primes were presented subliminally.

In recent years, others have replicated these results using a variety of paradigms; conflict adaptation effects are fully abolished when the conflicting primes are strongly masked (Greenwald et al., 1996; Kunde, 2003; Frings and Wentura, 2008; Boy et al., 2010; Ansorge et al., 2011). However, in some recent studies conflict adaptation has been observed for masked prime stimuli, although the effect is generally small (Bodner and Mulji, 2010; van Gaal et al., 2010a; Francken et al., 2011). Recently, we have suggested that this discrepancy between studies might be due to the timing and the attentional engagement of the subject in between trials. At short trial intervals, the fleeting nature of the subliminal prime stimulus (Dehaene and Naccache, 2001) might cause the effects to dissappear easily, either by overall distraction or by the mere elapsing of time and strong attentional involvement might slow down this process somewhat (note that primes are always task-irrelevant in these tasks). In fact, this might also be the case for conscious stimuli, but at a slower pace (Danielmeier and Ullsperger, 2011; Egner et al., 2011). Recently, Desender and van den Bussche (2012) reviewed a large set of studies regarding the role of awareness in conflict adaptation and highlighted some alternative interpretations of conflict adaptation effects that were driven by subliminal stimuli. They reasoned that, although the stimulus itself might be strongly masked and, therefore, subliminal, the effect it has on behavior and cognition might become conscious and drive conflict adaption. Subjects might for example be able to monitor their RTs (Marti et al., 2011) and because responses to incongruent trials are generally slower than to congruent trials, subjects might become aware of the conflict or difficulty by this means. Recently, it has been observed that our "sense of control" is larger following congruent than incongruent trials (or when action selection is "smooth and easy") (Wenke et al., 2010). In the case of unconscious conflict, we might sense an increased difficulty that calls for an increase of our sense of control over behavior, leading to trial-by-trial behavioral adaptations, such as conflict adaptation and post-error slowing.

Neuroimaging has revealed that unnoticed errors do trigger some aspects of error monitoring, such as ACC activity and Error-related Negativity (ERN) ERP modulations, but not others, such as insula activation and Error Positivity (Pe) ERP modulations, although evidence is mixed (Nieuwenhuis et al., 2001; Hester et al., 2005; Overbeek et al., 2005; Klein et al., 2007; O'Connell et al., 2007; Pavone et al., 2009; Ullsperger et al., 2010; Woodman, 2010; Dhar et al., 2011). To examine the behavioral consequences of aware and unaware errors, research has focused mainly on post-error slowing: the behaviorally observed slowing that occurs after the commission of an error (compared to a correct response), potentially as a strategy to prevent future errors. It has been observed that unnoticed (or unconscious) errors sometimes elicit small post-error slowing effects (Cohen et al., 2009) whereas sometimes they do not (Nieuwenhuis et al., 2001; Endrass et al., 2007; Klein et al., 2007; van Gaal et al., 2009). Recently, we have explored the trial-by-trial effects of errors made to unconsciously presented stimuli (responses to subliminal No-Go stimuli) (Cohen et al., 2009). To do so, we separated responses on trials with a conscious No-Go circle (termed "conscious errors"), responses on trials with an unconscious No-Go circle ("unconscious errors"), and responses on go trials (**Figure 3A**). Subjects slowed down their responses considerably after conscious errors (∼20 ms) and very subtly (but significantly) after unconscious errors (∼3 ms). Spectral granger causality analyses revealed that conscious errors elicited top-down modulations from frontal electrodes to occipital electrodes leading up to the next trial (in the inter-trial-interval) (see also King et al., 2010 for error-related top-down modulations of motor and sensory regions using fMRI). Crucially, these top-down modulations were also observed after unconscious errors (**Figure 3B**), thus suggesting that unconscious events can elicit an "automatic" feedback loop in the absence of stimulus awareness. Importantly, in this task, RTs on the previous trial on which the unconscious error was made were equal, ruling out the possibility that these longer-lasting trial-by-trial effects were due to the conscious monitoring of RTs on the previous trials (Marti et al., 2011). However, whether these effects can be explained by other metacognitive processes such as experienced "effort" or "difficulty" is

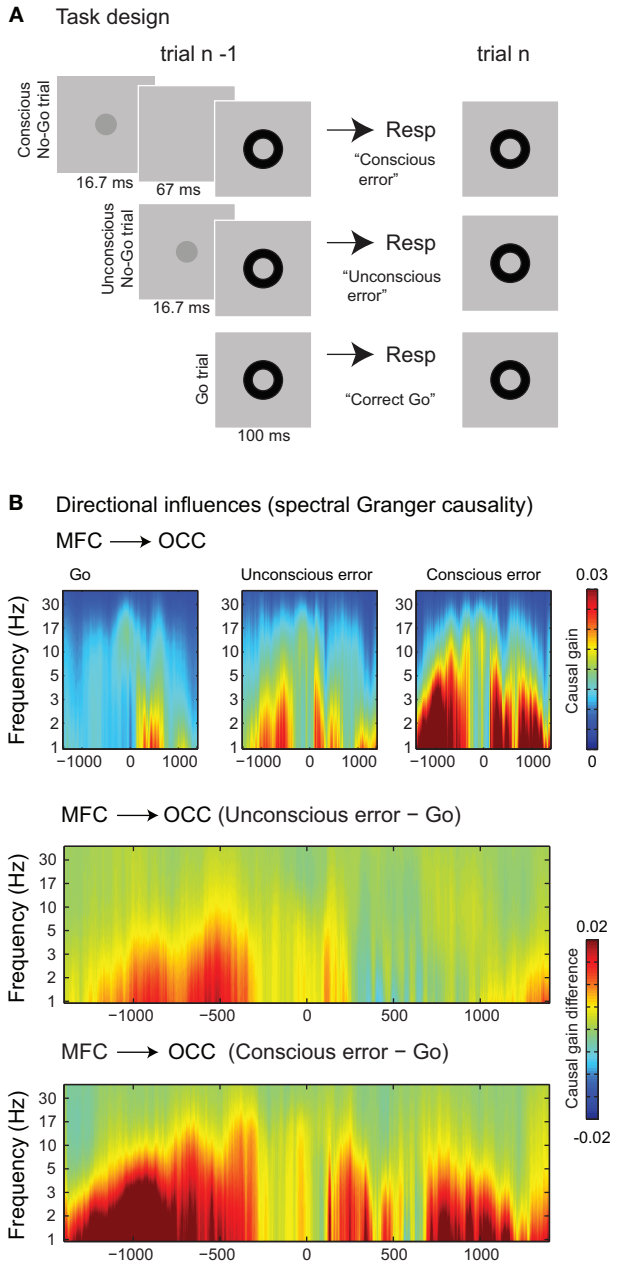


FIGURE 3 | (A) Task-set-up. **(B)** Response errors committed to trials containing a conscious No-Go signal and unconscious No-Go signal elicit long-lasting top-down influences from medial frontal (MFC) to occipital (OCC) regions (electrodes) in a broad frequency band (∼1–12 Hz), leading up to the next trial. Effects for each of the three conditions separately (upper panel) and the difference plots (lower panel) are shown. Time 0 is the presentation of the Go-signal (black annulus). Plots are separated by the significance of the event on the previous trial (correct Go response, conscious error, or unconscious error). Adapted with permission from Cohen et al. (2009).

an avenue for future research (Wenke et al., 2010; Desender and van den Bussche, 2012). It is interesting to note that although unaware errors seem to elicit relatively strong neural activity on the current trial (mainly in the medial frontal cortex) which are accompanied and followed by short-lived increases in top-down interactions that might drive automatic behavioral adaptations at relatively short inter-trial intervals (Cohen et al., 2009), long-term behavioral adaptations on the next trial are generally weak. Therefore, although speculative, error awareness might be beneficial for broader longer-lasting control adaptations that might be associated with activation in the anterior insula (Ullsperger et al., 2010) and reflected in the (late part) of the Pe ERP component (Nieuwenhuis et al., 2001; Endrass et al., 2007).

Overall, the discussed studies suggest that, under some specific conditions, subliminal information might durably affect behavior and brain activity, sometimes even beyond several seconds (Gaillard et al., 2007; Pessiglione et al., 2008; Capa et al., 2010; Soto et al., 2011). Occasionally indirect consequences of subliminal (affective) information can be measured for several minutes (Gaillard et al., 2007) and up to even as long as 24 h after its presentation (Sweeny et al., 2009). Although some of these effects may be caused by a form of learning and, therefore, synaptic changes rather than long-lasting (and active) subliminal activation (Gaillard et al., 2007), other effects might be truly mediated by the active maintenance of information across several seconds of time. In the next section, we will focus on the role of stimulus awareness in the integration of information across time and space.

CONSCIOUSNESS AND DECISION-MAKING

The literature we reviewed so far shows that unconscious information can affect high-level processes, and might even act on aspects of cognitive control and (working) memory. Lastly, we will discuss studies that investigated whether and how unconscious information can be accumulated across time or space for perception and decision-making. Active information integration is considered one of the hallmarks of consciousness by many contemporary models of consciousness (Tononi and Edelman, 1998; Engel and Singer, 2001; Crick and Koch, 2003; Baars, 2005; Seth et al., 2008).

Several recent studies have observed that, under some conditions, subliminal information can be accumulated and integrated spatially across the visual field (van Opstal et al., 2011b). In one of van Opstal's experiments, subjects were presented with 4 spatially separated numbers (primes) that were preceded and followed by masks that prevented conscious perception of the primes. A target, also consisting of four digits, followed the prime rapidly and subjects had to indicate whether the mean of the 4 target digits was more or less than 5. Interestingly, the mean of the subliminal primes affected RTs and accuracy to target responses, suggesting that so-called "ensemble statistics" might be extracted unconsciously. Other recent evidence also suggests that multiple unconscious stimuli can be integrated across space, for example when visual scenes are presented in the absence of awareness because of continuous flash suppression (Mudrik et al., 2011). Also, expert chess players (but not novices) are able to extract whether a subliminal (masked) simplified chess configuration entails a checking configuration or not. However, this was only the case for highly familiar chess configurations, and was not present in a task that required the integration of local features, namely field color (black or white) and chess piece (rook or knight).

This suggests that experts have created chunks of common chess configurations in long-term memory (which novices have not) and, therefore, that they might not actively have to integrate individual stimulus features (Kiesel et al., 2009). Generally, the extent of practice might be crucial and partly explain why evidence in the field is somewhat mixed. Others have shown that the integration of local features into global shapes does require stimulus awareness, for example when stimuli are rendered invisible due to counter-phase flickering of stimulus contrast (Schwarzkopf and Rees, 2010).

Unconscious information also seems to be integrated or accumulated across time, at least to some degree. Previous studies have shown that subliminal information can be accumulated linearly over a few hundreds of milliseconds (Jaskowski et al., 2003; Vorberg et al., 2003; Wentura and Frings, 2005; Del Cul et al., 2007; Frings et al., 2008). For example, Jaskowski et al. (2003) nicely showed that increasing the number of primes presented before a target increases the behavioral priming effect. In their task, subjects were required to respond to the spatial location of a square with horizontal gaps presented together with a square without such gaps. Targets could be preceded by either 1, 2, 3 or 4 primes (presented for 35 ms each) which were smaller copies of the target. Because the squares in every next stimulus were slightly larger than the previous ones, they masked the preceding stimulus. They showed that each of the 4 primes had an influence on the response to target, and that with increasing number of primes the priming effect was larger. Similarly, Vorberg et al. (2003) have shown that when the time between prime and target is increased (from 14 to 86 ms in steps of 14 ms) the behavioral priming effect increases monotonically. Subjects had to respond to the direction of a metacontrast target arrow that was preceded by a smaller version of it. Importantly, because the stimuli were presented outside the focus of attention (below and above fixation), there were no SOA-related changes in prime awareness (see also Schmidt et al., 2010). Together, these results suggest that subliminal information can be accumulated over short periods of time (<150 ms) and increasingly impact behavior.

However, while the accumulation of information may be possible irrespective of the level of awareness over short periods of time, recent studies have shown that awareness might play an important role when the time across which information has to be accumulated is increased. de Lange et al. (2011) performed a task in which subjects had to accumulate multiple pieces of evidence across 1.5 s. On each trial, subjects were presented a stream of five arrows, each of which could point to the left or right with equal probability. They had to quickly decide on the direction of the majority of arrows, guessing if necessary (**Figure 4A**). The strength of the evidence could range from one (low evidence, e.g., two left and three right arrows) to five (high evidence, e.g., five right arrows). The visibility of the arrows was manipulated by masking them with an effective metacontrast mask leading to arrows near the threshold of awareness (low visibility condition) or with an equiluminant but less effective "pseudo" mask (leading to high visibility). On each trial, all arrows were either of low- or high visibility. Qualitative differences in perception were confirmed by objective as well as subjective discrimination measurements (leading to low vs. high visibility arrows, instead of conscious vs. unconscious arrows). Importantly, stimulus and mask duration were identical for both conditions, which allowed the comparison of behavioral performance of evidence accumulation (and the underlying neural responses) without confounding stimulus visibility with basic task parameters (e.g., signal strength) (Lau, 2008; Francken et al., 2011). Behaviorally, subjects were able to accumulate evidence

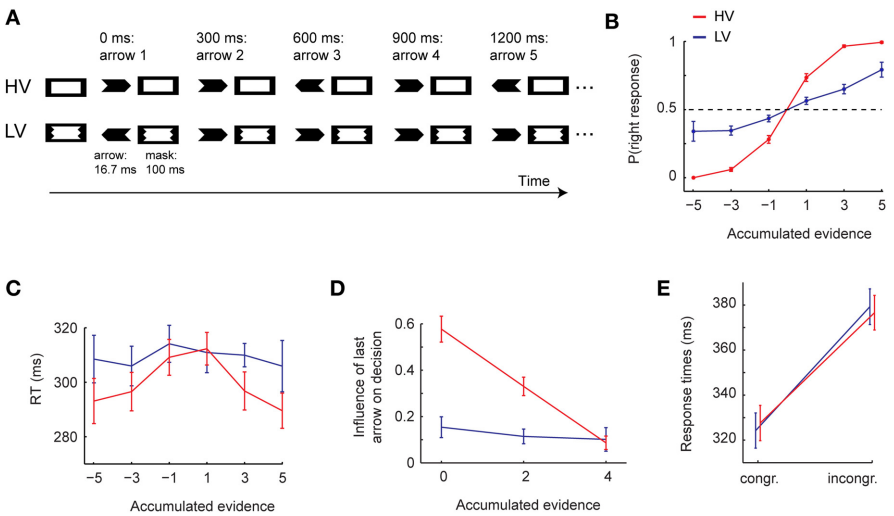


FIGURE 4 | (A) Task-set-up. **(B)** Decision-making performance as a function of accumulated evidence. Negative and positive numbers denote evidence for a left- and right-hand response, respectively, (number of right-pointing arrows minus number of left-pointing arrows). **(C)** Reaction times as a function of accumulated evidence. **(D)** Influence of the last arrow on the final decision as a function of the amount of previously accumulated evidence. **(E)** Priming strength of single high visibility or low visibility arrow as measured in a distinct masked priming task, in terms of response times. HV, high visibility condition; LV, low visibility condition. Adapted with permission from de Lange et al. (2011).

over time for both visibility conditions (**Figure 4B**). However, there were marked qualitative differences in *how* information was accumulated for the different levels of awareness. First, decision-making speed was modulated by the amount of accumulated evidence, but only for high-visible stimuli (**Figure 4C**). Second, once enough evidence had been gathered, participants strategically reduced the impact of new incoming stimuli (**Figure 4D**). Crucially, by using the same stimulus parameters but now in a masked priming task, it was observed that the amount of bottom-up information provided by the arrows was the same for both conditions, as reflected in an equal size of the behavioral priming effect for both visibility conditions (**Figure 4E**). Thus, although unconscious evidence may be accumulated in a linear fashion, i.e., adding and subtracting new information without any regard to the history of prior accumulated evidence, non-linearities in evidence accumulation (for example, reducing the weight of new information under conditions of high certainty, Kiani et al., 2008; de Lange et al., 2010) may be present only for fully consciously perceived information. This qualitative difference (linear vs. non-linear integration) was also observed in concurrently measured neural recordings: occipito-parietal regions that were involved in the accumulation of the sensory evidence showed a "linear" stereotypic response when presented with near-threshold information, but modulated their activity strategically during the task for clearly visible information. These results suggest that the level of awareness of information changes decision-making: while accumulation of evidence is already possible for low visibility information, high visibility allows evidence to be accumulated up to a much higher-level, leading to important changes in strategic top-down decision-making.

Similarly, Sackur and Dehaene (2009) demonstrated a qualitative processing difference as a function of stimulus awareness when studying sequential two-step arithmetic tasks with masked and visible digits. In their task, participants were instructed to add/subtract 2 from one masked or unmasked number (numbers were part of a restricted set: 2, 4, 6, 8). Thereafter, subjects were required to indicate whether the outcome of this first operation was larger or smaller than 5. Interestingly, subjects were only able to perform a "chained task" of addition followed by comparison when the target digits were presented consciously, but not when they were presented subliminally. However, they could perform each individual computation above chance when the digits were presented subliminally (see also Garcia-Orza et al., 2009; Ric and Muller, in press). The authors have suggested that this qualitative difference can be explained by assuming that subliminal digits do not yield enough evidence to reach a threshold for the first computational step of the chained task and that this did not allow subjects to go past the first processing stage and deploy further strategies.

Bijleveld et al. (2010) reported an interesting related observation. They presented subjects with a poorly or strongly masked high- or low-reward cue (50 cents or 1 cent, respectively) that was quickly followed by a simple mathematical operation (e.g., $3 + 5 + 9 = 16$). Subjects had to indicate whether the expression was true or not. Two manipulations were crucial: (1) only correct responses were rewarded, and (2) on each trial the anticipated reward declined with time. They showed that subliminal high-reward cues (compared to low-reward cues) *sped up* the time people took to perform the mathematical operation, probably reflecting the increased investment of effort, but did not change the overall accuracy of participants' responses. In contrast, conscious high-reward cues (compared to low-reward cues) *slowed down* the time to perform the computation, but crucially, with the benefit of increasing the overall accuracy of the responses. In a second experiment, they showed that conscious and unconscious reward cues both elicit increases in effort only (speeding of responses), when the possibility to change the speed-accuracy balance is eliminated. Therefore, valuable rewards seem to enhance the effort put in a task at hand irrespective of the awareness of the reward (see also Pessiglione et al., 2007). However, the authors concluded that conscious rewards might impact the balance between speed and accuracy (preferring accuracy over speed), whereas unconscious reward cues do not, thereby providing a qualitative difference between the pursuit of rewards of which one is conscious vs. unconscious.

Although the sequential and spatial integration of multiple elements was generally considered a hallmark of consciousness, these recent results challenge that assumption and showed that some parts of this integration process might also operate outside of awareness. Further, it seems that, at least under some conditions, awareness is beneficial for enabling strategic changes in decision-making (see below for a more extensive discussion on why this might be the case).

CONCLUSIONS, OPEN QUESTIONS AND FUTURE DIRECTIONS

In this article, we have reviewed recent studies that have focused on the complexity and strength of unconscious information processing in relation to cognitive control (e.g., response inhibition, conflict resolution, and task-switching), the life-time of information maintenance (e.g., working memory, recognition memory) and the possibility to integrate multiple pieces of information across space and time. Unconscious information has been shown to affect various perceptual and high-level cognitive functions and the associated brain areas, including prefrontal cortex. In some cases, unconscious information has been observed to affect behavior and brain activity for relatively long periods of time. Overall, these recent results highlight the power of unconscious information processing, going beyond specific expectations formulated in traditional theoretical models of consciousness and the cognitive functions thought to require consciousness (for reviews see Umilta, 1988; Dehaene and Naccache, 2001; Jack and Shallice, 2001; Mayr, 2004; Hommel, 2007).

Based on this, one can conclude that the potential function of consciousness might not be related to the initiation of cognitive control functions by specific stimuli that signal the need for increased control (e.g., stop-signals, task-switching cues). These cognitive control operations are probably triggered by a fast feedforward, and unconscious, early sweep of information processing that reaches even regions in the prefrontal cortex (van Gaal and Lamme, in press). This unconscious fast feedforward sweep can directly affect (the speed of) ongoing cognitive processes. However, recent evidence also points out interesting dissociations between conscious and unconscious information

processing when it comes to the duration, the flexibility, and the strategic use of information for complex operations and decision-making (Sackur and Dehaene, 2009; de Lange et al., 2011). Although recent evidence has clearly pushed the boundaries regarding the duration of unconscious effects, the general observation is that unconscious events are much less able to elicit (long-term) future behavioral adaptations than conscious events (e.g., post-error slowing, conflict adaptation). Why might this be the case? Theoretical models of consciousness suggest that conscious awareness is related to long-lasting recurrent interactions between (distant) brain regions (Lamme, 2006; Dehaene and Changeux, 2011). This might enable the exchange of information between several spatially separated cognitive modules, which seems to break the automaticity of information processing (Sackur and Dehaene, 2009). Awareness might be beneficial for enabling flexible and durable information processing strategies that are not directly triggered by a specific stimulus, for example when information has to be integrated across longer periods of time to bias information acquisition (de Lange et al., 2011) or signal the need for performance adjustments (Desender and van den Bussche, 2012). Recently, Kunde et al. (2012) suggested that awareness might be dispensable when cognitive control is signaled *explicitly* (by specific control-eliciting stimuli) but not when it has to be inferred *implicitly* (by the context, or history of events). This may prove to be a very useful and plausible distinction and needs further experimentation and exploration. Especially, the boundary conditions of implicitly signaled cognitive control are a promising avenue for future research and might clarify why unconsciously signaled (implicit) control operations are observed in some occasions but not in others.

At present, we believe that two aspects of subliminal information processing deserve more attention in future work, namely the role of task-relevance of unconscious information and the potential source of some of the subliminally triggered effects. Typically, masked priming studies have used subliminal stimuli that should be (actively) ignored for optimal task performance (e.g., Dehaene et al., 1998; Kunde, 2003; Vorberg et al., 2003). However, as discussed above, it seems that task-relevant unconscious information has much stronger and longer-lasting effects on behavior and brain activity than task-irrelevant information (e.g., Ansorge and Neumann, 2005; Pessiglione et al., 2008; van Gaal et al., 2008; Soto et al., 2011), which might explain part of the discrepancy in the overall findings in the field. More generally, the *significance* of the unconscious events seems a crucial factor, but at this point, future studies are needed to specifically test this prediction.

Second, recent evidence suggests that some trial-by-trial effects can be triggered by unconsciously presented stimuli, although some of the effects are limited and are generally relatively small. In some cases, these complex and relatively long-lasting effects (e.g., congruency effects, trial-by-trial modulations) might be explained by meta-cognitive effects, which need to be carefully controlled for to foster interpretation about the source of the observed "unconscious" phenomena (Jaskowski et al., 2003; Kinoshita et al., 2008; Wenke et al., 2010; see Desender and van den Bussche, 2012 for an extensive review on this issue). One of the main goals of future research could be to test whether and to what extent these long-term effects are caused by subjects becoming meta-cognitively aware of the conflict/error signals, although they are unaware of the initial source of it.

FUNDING

This work was supported by the Netherlands Organization for Scientific Research (NWO-Rubicon awarded to Simon van Gaal; NWO-VENI awarded to Floris P. de Lange; NWO-VIDI awarded to Michael X Cohen).

REFERENCES

Aarts, H., Custers, R., and Marien, H. (2008). Preparing and motivating behavior outside of awareness. *Science* 319, 1639–1639.

Al-Janabi, S., and Finkbeiner, M. (2011). Effective processing of masked eye gaze requires volitional control. *Exp. Brain Res.* 216, 433–443.

Ansorge, U., Fuchs, I., Khalid, S., and Kunde, W. (2011). No conflict control in the absence of awareness. *Psychol. Res.* 5, 351–365.

Ansorge, U., and Neumann, O. (2005). Intentions determine the effect of invisible metacontrast-masked primes: evidence for top-down contingencies in a peripheral cuing task. *J. Exp. Psychol. Hum. Percept. Perform.* 31, 762–777.

Aron, A. R. (2007). The neural basis of inhibition in cognitive control. *Neuroscientist* 13, 214–228.

Aron, A. R., Schlaghecken, F., Fletcher, P. C., Bullmore, E. T., Eimer, M., Barker, R., Sahakian, B. J., and Robbins, T. W. (2003). Inhibition of subliminally primed responses is mediated by the caudate and thalamus: evidence from functional MRI and Huntington's disease. *Brain* 126, 713–723.

Baars, B. J. (2005). Global workspace theory of consciousness: toward a cognitive neuroscience of human experience. *Prog. Brain Res.* 150, 45–53.

Baars, B. J., and Franklin, S. (2003). How conscious experience and working memory interact. *Trends Cogn. Sci.* 7, 166–172.

Bahrami, B., Carmel, D., Walsh, V., Rees, G., and Lavie, N. (2008a). Spatial attention can modulate unconscious orientation processing. *Perception* 37, 1520–1528.

Bahrami, B., Carmel, D., Walsh, V., Rees, G., and Lavie, N. (2008b). Unconscious orientation processing depends on perceptual load. *J. Vis.* 8, 12.1–12.10.

Bargh, J. A., and Morsella, E. (2008). The unconscious mind. *Perspect. Psychol. Sci.* 3, 73–79.

Bauer, F., Cheadle, S. W., Parton, A., Muller, H. J., and Usher, M. (2009). Gamma flicker triggers attentional selection without awareness. *Proc. Natl. Acad. Sci. U.S.A.* 106, 1666–1671.

Belopolsky, A. V., Kramer, A. F., and Theeuwes, J. (2008). The role of awareness in processing of oculomotor capture: evidence from event-related potentials. *J. Cogn. Neurosci.* 20, 2285–2297.

Bijleveld, E., Custers, R., and Aarts, H. (2010). Unconscious reward cues increase invested effort, but do not change speed-accuracy trade-offs. *Cognition* 115, 330–335.

Block, N. (2011). Perceptual consciousness overflows cognitive access. *Trends Cogn. Sci.* 15, 567–575.

Bodner, G. E., and Masson, M. E. (2004). Beyond binary judgments: prime validity modulates masked repetition priming in the naming task. *Mem. Cogn.* 32, 1–11.

Bodner, G. E., and Mulji, R. (2010). Prime proportion affects masked priming of fixed and free-choice responses. *Exp. Psychol.* 57, 360–366.

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.

Boy, F., Husain, M., and Sumner, P. (2010). Unconscious inhibition separates two forms of cognitive control. *Proc. Natl. Acad. Sci. U.S.A.* 107, 11134–11139.

Brooks, S. J., Savov, V., Allzen, E., Benedict, C., Fredriksson, R., and Schioth, H. B. (2012). Exposure to subliminal arousing stimuli induces robust activation in the amygdala, hippocampus, anterior cingulate, insular cortex and primary visual cortex: a systematic meta-analysis

of fMRI studies. *Neuroimage* 59, 2962–2973.

Bruchmann, M., Herper, K., Konrad, C., Pantev, C., and Huster, R. J. (2011). Individualized EEG source reconstruction of Stroop interference with masked color words. *Neuroimage* 49, 1800–1809.

Capa, R. L., Bustin, G. M., Cleeremans, A., and Hansenne, M. (2011). Conscious and unconscious reward cues can affect a critical component of executive control. *Exp. Psychol.* 58, 370–375.

Capa, R. L., Cleeremans, A., Bustin, G. M., and Hansenne, M. (2010). Long-lasting effect of subliminal processes on cardiovascular responses and performance. *Int. J. Psychophysiol.* 81, 22–30.

Carlson, J. M., Beacher, F., Reinke, K. S., Habib, R., Harmon-Jones, E., Mujica-Parodi, L. R., and Hajcak, G. (2012). Nonconscious attention bias to threat is correlated with anterior cingulate cortex gray matter volume: a voxel-based morphometry result and replication. *Neuroimage* 59, 1713–1718.

Cohen, J. D., Botvinick, M., and Carter, C. S. (2000). Anterior cingulate and prefrontal cortex: who's in control? *Nat. Neurosci.* 3, 421–423.

Cohen, M. X., van Gaal, S., Ridderinkhof, K. R., and Lamme, V. A. F. (2009). Unconscious errors enhance prefrontal-occipital oscillatory synchrony. *Front. Hum. Neurosci.* 3:54. doi: 10.3389/neuro.09.054.2009

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

Crump, M. J., Gong, Z., and Milliken, B. (2006). The context-specific proportion congruent Stroop effect: location as a contextual cue. *Psychon. Bull. Rev.* 13, 316–321.

Crump, M. J., and Milliken, B. (2009). The flexibility of context-specific control: evidence for context-driven generalization of item-specific control settings. *Q. J. Exp. Psychol. (Hove)* 62, 1523–1532.

Custers, R., and Aarts, H. (2005). Positive affect as implicit motivator: on the nonconscious operation of behavioral goals. *J. Pers. Soc. Psychol.* 89, 129–142.

Custers, R., and Aarts, H. (2010). The unconscious will: how the pursuit of goals operates outside of conscious awareness. *Science* 329, 47–50.

D'Ostilio, K., and Garraux, G. (2012). Dissociation between unconscious motor response facilitation and conflict in medial frontal areas. *Eur. J. Neurosci.* 35, 332–340.

Damian, M. F. (2001). Congruity effects evoked by subliminally presented primes: automaticity rather than semantic processing. *J. Exp. Psychol. Hum. Percept. Perform.* 27, 154–165.

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

Daza, M. T., Ortells, J. J., and Fox, E. (2002). Perception without awareness: further evidence from a Stroop priming task. *Percept. Psychophys.* 64, 1316–1324.

de Pisapia, N., Turatto, M., Lin, P., Jovicich, J., and Caramazza, A. (2011). Unconscious priming instructions modulate activity in default and executive networks of the human brain. *Cereb. Cortex* 3, 639–649.

Dehaene, S., Artiges, E., Naccache, L., Martelli, C., Viard, A., Schurhoff, F., Recasens, C., Martinot, M. L. P., 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., Changeux, J.-P., Naccache, L., Le Clec'H, G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., van de Moortele, P.-F., and Le Bihan, D. (1998). Imaging unconscious semantic priming. *Nature* 395, 597–600.

Dehaene, S., and Changeux, J. P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron* 70, 200–227.

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

de Lange, F. P., Jensen, O., and Dehaene, S. (2010). Accumulation of evidence during sequential decision making: the importance of top-down factors. *J. Neurosci.* 30, 731–738.

de Lange, F. P., van Gaal, S., Lamme, V. A., and Dehaene, S. (2011). How awareness changes the relative weights of evidence during human decision-making. *PLoS Biol.* 9:e1001203. doi: 10.1371/journal.pbio.1001203

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

Desender, K., and van den Bussche, E. (2012). Is consciousness necessary for conflict adaptation? A state of the art. *Front. Hum. Neurosci.* 6:3. doi: 10.3389/fnhum.2012.00003

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

Egner, T., Ely, S., and Grinband, J. (2011). Going, going, gone: characterizing the time-course of congruency sequence effects. *Front. Psychol.* 1:154. doi: 10.3389/fpsyg.2010.00154

Eimer, M. (1999). Facilitatory and inhibitory effects of masked prime stimuli on motor activation and behavioural performance. *Acta Psychol. (Amst.)* 101, 293–313.

Eimer, M., Maravita, A., van Velzen, J., Husain, M., and Driver, J. (2002). The electrophysiology of tactile extinction: ERP correlates of unconscious somatosensory processing. *Neuropsychologia* 40, 2438–2447.

Eimer, M., and Schlaghecken, F. (1998). Effects of masked stimuli on motor activation: behavioral and electrophysiological evidence. *J. Exp. Psychol. Hum. Percept. Perform.* 24, 1737–1747.

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.

Engel, A. K., and Singer, W. (2001). Temporal binding and the neural correlates of sensory awareness. *Trends Cogn. Sci.* 5, 16–25.

Fabre, L., Lemaire, P., and Grainger, J. (2007). Attentional modulation of masked repetition and categorical priming in young and older adults. *Cognition* 105, 513–532.

Fang, F., and He, S. (2005). Cortical responses to invisible objects in the human dorsal and ventral pathways. *Nat. Neurosci.* 8, 1380–1385.

Finkbeiner, M., and Friedman, J. (2011). The flexibility of nonconsciously deployed cognitive processes: evidence from masked congruence priming. *PLoS One* 6:e17095. doi: 10.1371/journal.pone.0017095

Finkbeiner, M., and Palermo, R. (2009). The role of spatial attention in nonconscious processing: a comparison of face and nonface stimuli. *Psychol. Sci.* 20, 42–51.

Francken, J. C., van Gaal, S., and de Lange, F. P. (2011). Immediate and long-term priming effects are independent of prime awareness. *Conscious. Cogn.* 20, 1793–1800.

Frings, C., Bermeitinger, C., and Wentura, D. (2008). Center-surround or spreading inhibition: which mechanism caused the negative effect from repeated masked semantic primes? *Exp. Psychol.* 55, 234–242.

Frings, C., and Wentura, D. (2008). Trial-by-trial effects in the affective priming paradigm. *Acta Psychol. (Amst.)* 128, 318–323.

Gaillard, R., Cohen, L., Adam, C., Clemenceau, S., Hasboun, D., Baulac, M., Willer, J. C., Dehaene, S., and Naccache, L. (2007). Subliminal words durably affect neuronal activity. *Neuroreport* 18, 1527–1531.

Gaillard, R., Del Cul, A., Naccache, L., Vinckier, F., Cohen, L., and Dehaene, S. (2006). Nonconscious semantic processing of emotional words modulates conscious acces. *Proc. Natl. Acad. Sci. U.S.A.* 103, 7524–7529.

Garcia-Orza, J., Damas-Lopez, J., Matas, A., and Rodriguez, J. M. (2009). "2 × 3" primes naming "6": evidence from masked priming. *Atten. Percept. Psychophys.* 71, 471–480.

Gratton, G., Coles, M. G., and Donchin, E. (1992). Optimizing the use of information: strategic control of activation of responses. *J. Exp. Psychol. Gen.* 121, 480–506.

Greenwald, A. G., Abrams, R. L., Naccache, L., and Dehaene, S. (2003). Long-term semantic memory versus contextual memory in unconscious number processing. *J. Exp. Psychol. Learn. Mem. Cogn.* 29, 235–247.

Greenwald, A. G., Draine, S. C., and Abrams, R. L. (1996). Three cognitive markers of unconscious semantic activation. *Science* 273, 1699–1702.

Harry, B., Davis, C., and Kim, J. (2012). Subliminal acces to face representations does not rely on attention. *Conscious. Cogn.* 21, 573–583.

Hassin, R. R., Bargh, J. A., Engell, A. D., and McCulloch, K. C. (2009). Implicit working memory. *Conscious. Cogn.* 18, 665–678.

Heinemann, A., Kunde, W., and Kiesel, A. (2009). Context-specific prime-congruency effects: on the role of conscious stimulus representations for cognitive control. *Conscious. Cogn.* 18, 966–976.

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.

Hommel, B. (2007). Consciousness and control: not identical twins. *J. Conscious. Stud.* 14, 155–167.

Hommel, B., Ridderinkhof, K. R., and Theeuwes, J. (2002). Cognitive control of attention and action: issues and trends. *Psychol. Res.* 66, 215–219.

Hughes, G., Velmans, M., and de Fockert, J. (2009). Unconscious priming of a no-go response. *Psychophysiology* 46, 1258–1269.

Jack, A. I., and Shallice, T. (2001). Introspective physicalism as an approach to the science of consciousness. *Cognition* 79, 161–196.

Jaskowski, P., Skalska, B., and Verleger, R. (2003). How the self controls its "automatic pilot" when processing subliminal information. *J. Cogn. Neurosci.* 15, 911–920.

Jaskowski, P., and Przekoracka-Krawczyk, A. (2005). On the role of mask structure in subliminal priming. *Acta Neurobiol. Exp. (Wars)* 65, 409–417.

Jeneson, A., Kirwan, C. B., and Squire, L. R. (2010). Recognition without awareness: an elusive phenomenon. *Learn. Mem.* 17, 454–459.

Jiang, Y., Costello, P., Fang, F., Huang, M., and He, S. (2006). A gender- and sexual orientation-dependent spatial attentional effect of invisible images. *Proc. Natl. Acad. Sci. U.S.A.* 103, 17048–17052.

Kentridge, R. W., Heywood, C. A., and Weiskrantz, L. (1999). Attention without awareness in blindsight. *Proc. Biol. Sci.* 266, 1805–1811.

Kentridge, R. W., Heywood, C. A., and Weiskrantz, L. (2004). Spatial attention speeds discrimination without awareness in blindsight. *Neuropsychologia* 42, 831–835.

Kentridge, R. W., Nijboer, T. C., and Heywood, C. A. (2008). Attended but unseen: visual attention is not sufficient for visual awareness. *Neuropsychologia* 46, 864–869.

Kiani, R., Hanks, T. D., and Shadlen, M. N. (2008). Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. *J. Neurosci.* 28, 3017–3029.

Kiefer, M., Adams, S. C., and Zovko, M. (2012). Attentional sensitization of unconscious visual processing: top-down influences on masked priming. *Adv. Cogn. Psychol.* 8, 50–61.

Kiefer, M., and Brendel, D. (2006). Attentional modulation of unconscious "automatic" processes: evidence from event-related potentials in a masked priming paradigm. *J. Cogn. Neurosci.* 18, 184–198.

Kiefer, M., and Martens, U. (2010). Attentional sensitization of unconscious cognition: task sets modulate subsequent masked semantic priming. *J. Exp. Psychol. Gen.* 139, 464–489.

Kiefer, M., and Spitzer, M. (2000). Time course of conscious and unconscious semantic brain activation. *Neuroreport* 11, 2401–2407.

Kiesel, A., Kunde, W., Pohl, C., Berner, M. P., and Hoffmann, J. (2009). Playing chess unconsciously. *J. Exp. Psychol. Learn. Mem. Cogn.* 35, 292–298.

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.

Kinoshita, S., Forster, K. I., and Mozer, M. C. (2008). Unconscious cognition isn't that smart: modulation of masked repetition priming effect in the word naming task. *Cognition* 107, 623–649.

Kinoshita, S., Mozer, M. C., and Forster, K. I. (2011). Dynamic adaptation to history of trial difficulty explains the effect of congruency proportion on masked priming. *J. Exp. Psychol. Gen.* 140, 622–636.

Klapp, S. T. (2007). Nonconscious control mimics a purposeful strategy: strength of Stroop-like interference is automatically modulated by proportion of compatible trials. *J. Exp. Psychol. Hum. Percept. Perform.* 33, 1366–1376.

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.

Kouider, S., Eger, E., Dolan, R., and Henson, R. N. (2009). Activity in face-responsive brain regions is modulated by invisible, attended faces: evidence from masked priming. *Cereb. Cortex* 19, 13–23.

Kunde, W. (2003). Sequential modulations of stimulus-response correspondence effects depend on awareness of response conflict. *Psychon. Bull. Rev.* 10, 198–205.

Kunde, W., Kiesel, A., and Hoffmann, J. (2003). Conscious control over the content of unconscious cognition. *Cognition* 88, 223–242.

Kunde, W., Reuss, H., and Kiesel, A. (2012). Consciousness and cognitive control. *Adv. Cogn. Psychol.* 8, 9–18.

Lamme, V. A. F. (2006). Towards a true neural stance on consciousness. *Trends Cogn. Sci.* 10, 494–501.

Lau, H., and Rosenthal, D. (2011). Empirical support for higher-order theories of conscious awareness. *Trends Cogn. Sci.* 15, 365–373.

Lau, H. C. (2008). A higher order Bayesian decision theory of consciousness. *Prog. Brain Res.* 168, 35–48.

Lau, H. C., and Passingham, R. E. (2007). Unconscious activation of the cognitive control system in the human prefrontal cortex. *J. Neurosci.* 27, 5805–5811.

Li, W., Moallem, I., Paller, K. A., and Gottfried, J. A. (2007). Subliminal smells can guide social preferences. *Psychol. Sci.* 18, 1044–1049.

Lin, J. Y., Murray, S. O., and Boynton, G. M. (2009). Capture of attention to threatening stmuli without perceptual awareness. *Curr. Biol.* 19, 1118–1122.

Lleras, A., and Enns, J. T. (2004). Negative compatibility or object updating? A cautionary tale of mask-dependent priming. *J. Exp. Psychol. Gen.* 133, 475–493.

Marcel, A. J. (1983). Conscious and unconscious perception: experiments on visual masking and word recognition. *Cogn. Psychol.* 15, 197–237.

Martens, U., Ansorge, U., and Kiefer, M. (2011). Controlling the unconscious: attentional task sets modulate subliminal semantic and visuomotor processes differentially. *Psychol. Sci.* 22, 282–291.

Martens, U., and Kiefer, M. (2009). Specifying attentional top-down influences on subsequent unconscious semantic processing. *Adv. Cogn. Psychol.* 5, 56–68.

Marti, S., Sackur, J., Sigman, M., and Dehaene, S. (2011). Mapping introspection's blind spot: reconstruction of dual-task phenomenology using quantified introspection. *Cognition* 115, 303–313.

Marzouki, Y., Midgley, K. J., Holcomb, P. J., and Grainger, J. (2008). An ERP investigation of the modulation of subliminal priming by exogenous cues. *Brain Res.* 1231, 86–92.

Mattler, U. (2003). Priming of mental operations by masked stimuli. *Percept. Psychophys.* 65, 167–187.

Mattler, U. (2005). Inhibition and decay of motor and nonmotor priming. *Percept. Psychophys.* 67, 285–300.

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

Mayr, U., Awh, E., and Laurey, P. (2003). Conflict adaptation effects in the absence of executive control. *Nat. Neurosci.* 6, 450–452.

Merikle, P. M., and Joordens, S. (1997). Parallels between perception without attention and perception without awareness. *Conscious. Cogn.* 6, 219–236.

Merikle, P. M., and Reingold, E. M. (1990). Recognition and lexical decision without detection: unconscious perception? *J. Exp. Psychol. Hum. Percept. Perform.* 16, 574–583.

Mudrik, L., Breska, A., Lamy, D., and Deouell, L.Y. (2011). Integration without awareness: expanding the limits of unconscious processing. *Psychol. Sci.* 22, 764–770.

Mulckhuyse, M., and Theeuwes, J. (2010). Unconscious attentional orienting to exogenous cues: a review of the literature. *Acta Psychol.* 134, 299–309.

Naccache, L., Blandin, E., and Dehaene, S. (2002). Unconscious masked priming depends on temporal attention. *Psychol. Sci.* 13, 416–424.

Nakamura, K., Dehaene, S., Jobert, A., Le Bihan, D., and Kouider, S. (2007). Task-specific change of unconscious neural priming in the cerebral language network. *Proc. Natl. Acad. Sci. U.S.A.* 104, 19643–19648.

Nieuwenhuis, S., Ridderinkhof, K. R., Blom, J., Band, G. P. H., 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.

Notebaert, W., Houtman, F., van Opstal, F., Gevers, W., Fias, W., and Verguts, T. (2009). Post-error slowing: an orienting account. *Cognition* 111, 275–279.

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.

O'Connor, P. A., and Neill, W. T. (2011). Does subliminal priming of free response choices depend on task set or automatic response activation? *Conscious. Cogn.* 20, 280–287.

Ortells, J. J., Fox, E., Noguera, C., and Abad, M. J. (2003). Repetition priming effects from attended vs. ignored single words in a semantic categorization task. *Acta Psychol. (Amst.)* 114, 185–210.

Ortells, J. J., Frings, C., and Plaza-Ayllon, V. (2011). Influence of spatial attention on conscious and unconscious word priming. *Conscious. Cogn.* 21, 117–138.

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.

Overgaard, M., Rote, J., Mouridsen, K., and Ramsoy, T. Z. (2006). Is conscious perception gradual or dichotomous? A comparison of report methodologies during a visual task. *Conscious. Cogn.* 15, 700–708.

Pavone, E. F., Marzi, C. A., and Girelli, M. (2009). Does subliminal visual perception have an error-monitoring system? *Eur. J. Neurosci.* 30, 1424–1431.

Pessiglione, M., Petrovic, P., Daunizeau, J., Palminteri, S., Dolan, J. D., and Frith, C. D. (2008). Subliminal instrumental conditioning demonstrated in the human brain. *Neuron* 59, 561–567.

Pessiglione, M., Schmidt, L., Draganski, B., Kalisch, R., Lau, H., Dolan, R. J., and Frith, C. D. (2007). How the brain translates money into force: a neuroimaging study of subliminal motivation. *Science* 316, 904–906.

Pessoa, L., Kastner, S., and Ungerleider, L. G. (2003) Neuroimaging studeis of attention: from modulations of sensory processing to top-down control. *J. Neurosci.* 23, 3990–3998.

Pessoa, L., McKenna, M., Gutierrez, E., and Ungerleider, L. G. (2002). Neural processing of emotional faces requires attention. *Proc. Natl. Acad. Sci. U.S.A.* 99, 11458–11463.

Rahnev, D. A., Huang, E., and Lau, H. (2012). Subliminal stimuli in the near absence of attention influence top-down cognitive control. *Atten. Percept. Psychophys*. 74, 521–532.

Rajimerhr, R. (2004). Unconscious orientation processing. *Neuron* 41, 663–673.

Rees, G. (2007). Neural correlates of the contents of visual awareness in humans. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 362, 877–886.

Reuss, H., Kiesel, A., Kunde, W., and Hommel, B. (2011). Unconscious activation of task sets. *Conscious. Cogn.* 20, 556–567.

Ric, F., and Muller, D. (in press). Unconscious addition: when we unconsciously initiate and follow arithmetic rules. *J. Exp. Psychol. Gen.*

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.

Sackur, J., and Dehaene, S. (2009). The cognitive architecture for chaining of two mental operations. *Cognition* 111, 187–211.

Sadaghiani, S., Hesselmann, G., and Kleinschmidt, A. (2009). Distributed and antagonistic contributions of ongoing activity fluctuations to auditory stimulus detection. *J. Neurosci.* 29, 13410–13417.

Sato, W., Okada, T., and Toichi, M. (2007). Attentional shif by gaze is triggered without awareness. *Exp. Brain Res.* 183, 87–94.

Schlaghecken, F., and Eimer, M. (2006). Active masks and active inhibition: a comment on Lleras and Enns (2004) and on Verleger, Jaskowski, Aydemir, van der Lubbe, and Groen (2004). *J. Exp. Psychol. Gen.* 135, 484–494.

Schlaghecken, F., and Martini, P. (2012). Context, not conflict, drives cognitive control. *J. Exp. Psychol. Hum. Percept. Perform.* 38, 272–278.

Schmidt, F., Haberkamp, A., and Schmidt, T. (2010). Dos and don'ts in response priming research. *Adv. Cogn. Psychol.* 7, 120–131.

Schwarzkopf, D. S., and Rees, G. (2010). Interpreting local visual features as a global shape requires awareness. *Proc. R. Soc. Biol.* 278, 2207–2215.

Seth, A. K., Dienes, Z., Cleeremans, A., Overgaard, M., and Pessoa, L. (2008). Measuring consciousness: relating behavioural and neurophysiological approaches. *Trends Cogn. Sci.* 12, 314–321.

Simmonds, D. J., Pekar, J. J., and Mostofsky, S. H. (2008). Meta-analysis of Go/No-go tasks demonstrating that fMRI activation associated with response inhibition is task-dependent. *Neuropsychologia* 46, 224–232.

Soto, D., Mantyla, T., and Silvanto, J. (2011). Working memory without consciousness. *Curr. Biol.* 21, R912–R913.

Sterzer, P., Haynes, J. D., and Rees, G. (2008). Fine-scale activity patterns in high-level visual areas encode the category of invisible objects. *J. Vis.* 8, 10.1–10.12.

Sumner, P., Nachev, P., Morris, P., Peters, A. M., Jackson, S. R., Kennard, C., and Husain, M. (2007). Human medial frontal cortex mediates unconscious inhibition of voluntary action. *Neuron* 54, 697–711.

Sumner, P., Tsai, P., Yu, K., and Nachev, P. (2006). Attentional modulation of sensorimotor processes in the absence of perceptual awareness. *Proc. Natl. Acad. Sci. U.S.A.* 103, 10520–10525.

Sweeny, T. D., Grabowecky, M., Suzuki, S., and Paller, K. A. (2009). Long-lasting effects of subliminal affective priming from facial expressions. *Conscious. Cogn.* 18, 929–938.

Tapia, E., Breitmeyer, B. G., and Shooner, C. R. (2010). Role of task-directed attention in nonconscious and conscious response priming by form and color. *J. Exp. Psychol. Hum. Percept. Perform.* 36, 74–87.

Tononi, G., and Edelman, G. M. (1998). Consciousness and complexity. *Science* 282, 1846–1851.

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.

Umilta, C. (1988). "The control operations of consciousness," in *Consciousness in Contemporary Science,* eds A. J. Marcel, and E. Bisiach (Oxford: Oxford University Press), 334–356.

Ursu, S., Clark, K. A., Aizenstein, H. J., Stenger, V. A., and Carter, C. S. (2009). Conflict-related activity in the caudal anterior cingulate cortex in the absence of awareness. *Biol. Psychol.* 80, 279–286.

van den Bussche, E., Segers, G., and Reynvoet, B. (2008). Conscious and unconscious proportion effects in masked priming. *Conscious. Cogn.* 17, 1345–1358.

van den Bussche, E., van den Noortgate, W., and Reynvoet, B. (2009). Mechanisms of masked priming: a meta-analysis. *Psychol. Bull.* 135, 452–477.

van Gaal, S., Lamme, V. A., and Ridderinkhof, K. R. (2010a). Unconsciously triggered conflict adaptation. *PloS One* 5:e11508. doi: 10.1371/journal.pone.0011508

van Gaal, S., Ridderinkhof, K. R., Scholte, H. S., and Lamme, V. A. F. (2010b). Unconscious activation of the prefrontal no-go network. *J. Neurosci.* 30, 4143–4150.

van Gaal, S., Ridderinkhof, K. R., Fahrenfort, J. J., Scholte, H. S., and Lamme, V. A. F. (2008). Frontal cortex mediates unconsciously triggered inhibitory control. *J. Neurosci.* 28, 8053–8062.

van Gaal, S., Ridderinkhof, K. R., van den Wildenberg, W. P. M., and Lamme, V. A. F. (2009). Dissociating consciousness from inhibitory control: evidence for unconsciously triggered inhibitory control in the stop-signal paradigm. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 1129–1139.

van Gaal, S., Scholte, H. S., Lamme, V. A. F., Fahrenfort, J. J., and Ridderinkhof, K. R. (2011). Pre-SMA gray-matter density predicts individual differences in action selection in the face of conscious and unconscious response conflict. *J. Cogn. Neurosci.* 23, 382–390.

van Gaal, S., and Lamme, V. A. F. (in press). Unconscious high-level information processing: implications for neurobiological theories of consciousness. *Neuroscientist*.

van Opstal, F., Calderon, C. B., Gevers, W., and Verguts, T. (2011a). Setting the stage subliminally: unconscious context effects. *Conscious. Cogn.* 20, 1860–1864.

van Opstal, F., de Lange, F. P., and Dehaene, S. (2011b). Rapid parallel semantic processing of numbers without awareness. *Cognition* 120, 136–147.

van Opstal, F., Gevers, W., Osman, M., and Verguts, T. (2010). Unconscious task application. *Conscious. Cogn.* 19, 999–1006.

Vorberg, D., Mattler, U., Heinecke, A., Schmidt, T., and Schwarzbach, J. (2003). Different time courses for visual perception and action priming. *Proc. Natl. Acad. Sci. U.S.A.* 100, 6275–6280.

Voss, J. L., Baym, C. L., and Paller, K. A. (2008). Accurate forced-choice recognition without awareness of memory retrieval. *Learn. Mem.* 15, 454–459.

Voss, J. L., and Paller, K. A. (2009). An electrophysiological signature of unconscious recognition memory. *Nat. Neurosci.* 12, 349–355.

Voss, J. L., and Paller, K. A. (2010). What makes recognition without awareness appear to be elusive? Strategic factors that influence the accuracy of guesses. *Learn. Mem.* 17, 460–468.

Vuilleumier, P., and Schwartz, S. (2001). Beware and be aware: capture of spatial attention by fear-related stimuli in neglect. *Neuroreport* 12, 1119–1122.

Wenke, D., Fleming, S. M., and Haggard, P. (2010). Subliminal priming of actions influences sense of control over effects of action. *Cognition* 115, 26–38.

Wentura, D., and Frings, C. (2005). Repeated masked category primes interfere with related exemplars: new evidence for negative semantic priming. *J. Exp. Psychol. Learn. Mem. Cogn.* 31, 108–120.

Whalen, P. J., Kagan, J., Cook, R. G., Davis, F. C., Kim, H., Polis, S., McLaren, D. G., Somerville, L. H., McLean, A. A., Maxwell, J. S., and Johnstone, T. (2004). Human amygdala responsivity to masked fearful eye whites. *Science* 306, 2061–2061.

Whalen, P. J., Rauch, S. L., Etcoff, N. L., McInerney, S. C., Lee, M. B., and Jenike, M. A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *J. Neurosci.* 18, 411–418.

Wokke, M. E., van Gaal, S., Scholte, H. S., Ridderinkhof, K. R., and Lamme, V. A. (2011). The flexible nature of unconscious cognition. *PLoS One* 6:e25729. doi: 10.1371/journal.pone.0025729

Wolbers, T., Schoell, E. D., Verleger, R., Kraft, S., McNamara, A., Jaskowski, P., and Buchel, C. (2006). Changes in connectivity profiles as a mechanism for strategic control over interfering subliminal information. *Cereb. Cortex* 16, 857–864.

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

Zhou, F. A., and Davis, G. (2012). Unconscious priming of task sets: the role of spatial attention. *Atten. Percept. Psychophys*. 74, 105–114.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 29 February 2012; accepted: 17 April 2012; published online: 07 May 2012.

*Citation: van Gaal S, de Lange FP and Cohen MX (2012) The role of consciousness in cognitive control and decision making. Front. Hum. Neurosci. **6**:121. doi: 10.3389/fnhum.2012.00121*

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 27 August 2012
doi: 10.3389/fnhum.2012.00246

Error awareness and salience processing in the oddball task: shared neural mechanisms

Helga A. Harsay*[1]*, Marcus Spaan*[1]*, Jasper G. Wijnen*[1] *and K. Richard Ridderinkhof*[1,2]*

[1] *Department of Psychology, Amsterdam Center for the Study of Adaptive Control in Brain and Behavior, University of Amsterdam, Amsterdam, Netherlands*
[2] *Cognitive Science Center Amsterdam, University of Amsterdam, Amsterdam, Netherlands*

Edited by:
Tilmann A. Klein, Max Planck Institute for Human Cognitive and Brain Sciences, Germany

Reviewed by:
Floris P. De Lange, Radboud University Nijmegen, Netherlands
Erik M. Mueller, Marburg University, Germany

****Correspondence:***
K. Richard Ridderinkhof, Department of Psychology, University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam, Netherlands.
e-mail: k.r.ridderinkhof@uva.nl

A body of work suggests similarities in the way we become aware of an error and process motivationally salient events. Yet, evidence for a shared neural mechanism has not been provided. A within subject investigation of the brain regions involved in error awareness and salience processing has not been reported. While the neural response to motivationally salient events is classically studied during target detection after longer target-to-target intervals in an oddball task and engages a widespread insula-thalamo-cortical brain network, error awareness has recently been linked to, most prominently, anterior insula cortex. Here we explore whether the anterior insula activation for error awareness is related to salience processing, by testing for activation overlap in subjects undergoing two different task settings. Using a within subjects design, we show activation overlap in six major brain areas during aware errors in an antisaccade task and during target detection after longer target-to-target intervals in an oddball task: anterior insula, anterior cingulate, supplementary motor area, thalamus, brainstem, and parietal lobe. Within subject analyses shows that the insula is engaged in both error awareness and the processing of salience, and that the anterior insula is more involved in both processes than the posterior insula. The results of a fine-grained spatial pattern overlap analysis between active clusters in the same subjects indicates that even if the anterior insula is activated for both error awareness and salience processing, the two types of processes might tend to activate non-identical neural ensembles on a finer-grained spatial level. Together, these outcomes suggest a similar functional phenomenon in the two different task settings. Error awareness and salience processing share a functional anatomy, with a tendency toward subregional dorsal and ventral specialization within the anterior insula.

Keywords: anterior insula, error awareness, oddball processing, salience, magnetic resonance imaging, eyetracking

INTRODUCTION

When we interact with our environment, neural activity enabling goal-directed behavior is formed and continuously updated in order to adjust new action based on the experience of previous actions. As human behavior is susceptible to occasional errors, the ability to become aware of such errors keeps us from repeating the inadequate actions. This protects us from potentially harmful situations. How the brain instigates the ability to become aware of errors is yet unknown. Initial neuroimaging evidence suggests that, while error processing *per se* engages both anterior insula cortex (AIC) and anterior cingulate cortex (ACC), error awareness engages only the AIC (Klein et al., 2007). The functional significance of this AIC activation is however unclear.

One potential way to further our understanding of error awareness has been suggested by event-related potential (ERP)-work. ERP-studies on error awareness suggest neural similarities in the way we become aware of an error and attend to salient events (Ridderinkhof et al., 2009). The ability to attend to salient events is a basic ability that helps us to attend to meaningful events that have motivational importance. This ability is typically studied in a classical oddball paradigm, requiring the detection of distinct infrequent target stimuli or oddballs which are embedded in a series of frequently presented non-target or standard stimuli (Duncan-Johnson et al., 1984). The neural circuits that mediate oddball processing are well delineated by ERP as well as neuroimaging work (Kiehl and Liddle, 2003; Kiehl et al., 2005; Stevens et al., 2005). Whether the similarity as present in ERP work is also apparent in the neuroimaging manifestations of error awareness and oddball processing is yet unknown.

Here we set out to provide a test of the hypothesis that the AIC engaged during error awareness is also recruited (in the same subjects during the same session) in an oddball task during manipulations of oddball stimuli known to affect the processing of motivational salience.

THE SALIENCE SYSTEM

Generally, the insula is viewed as a dynamic interactive structure. It is well-placed to evaluate the motivational or emotional salience of certain events and is acting as an interface between external information and internal motivational states (Mesulam and Mufson, 1982a,b; Mufson and Mesulam, 1982; Craig, 2002, 2009; Seeley et al., 2007). Differences have been found in structural connectivity and in evoked responses to specific tasks across subregions of the

insula (Dupont et al., 2003; Mutschler et al., 2009; Deen et al., 2011).

In order to appreciate the activation of the anterior subregion of the insula during error awareness, we adopt a systems perspective that considers complex and multi-faceted functions to arise from the dynamic interactions of larger scale brain systems connected to this anterior subregion (Bressler and Menon, 2010). This principled theoretical perspective may aptly guide our exploration of how activation in the AIC can promote as well as constrain the emergence of salience signaling in both error awareness and the parametrical oddball task.

The use of various neuroimaging techniques has helped characterize a number of large-scale brain systems. Such systems may be configured dynamically and transiently, in response to current task demands, whereas other systems may be more fundamental and constant, so as to deal consistently and generically with common or recurrent demands. One of these networks comprises the dorsal ACC and the AIC/frontal operculum, a consistently observed functional network, described as a salience or control network (Dosenbach et al., 2007, 2008). This AIC–ACC network was initially thought to be task-specific, involved in the initiation and maintenance of task set, in task control such as monitoring, error feedback, and in subsequent performance adjustments. When a similar AIC–ACC network was subsequently identified in task-free states, it became termed the salience network (Menon and Uddin, 2010), thought to be involved in orienting to homeostatically relevant (salient) intrapersonal and extrapersonal events. The AIC and ACC often act in concert, as supported by findings of reciprocal projections in monkeys. Resting-state fMRI studies also indicate functional connectivity between anterior insula and the ACC (Taylor et al., 2009).

Not surprisingly, then, the AIC and ACC are often found to be co-activated in functional neuroimaging studies, in particular in response to the degree of subjective salience across domains (Sridharan et al., 2008; Craig, 2009). Co-activation of these core components of the salience network has been associated with orienting to, and facilitating the processing of personally and motivationally salient information, in the broad spectrum of emotional, social, cognitive, sensorimotor, homeostatic, and sympathetic efferent and interoceptive autonomic domains. Within the salience network, the AIC appears more specialized in receiving multimodal sensory input, whereas the ACC is connected more to action selection and action execution systems in cortical and subcortical brain regions, allowing the salience network to influence not only attention (to facilitate the further processing of salient signals) but also adaptive action in response to such signals.

THE CORE FUNCTION OF THE SALIENCE SYSTEM

Identifying motivationally salient stimuli has been proposed as the core function of the salience system; once a stimulus activates the salience system, it will have preferential access to the brain's attentional and working memory resources (Menon and Uddin, 2010). That is, once sensory areas detect a salient stimulus, this signal is transmitted to the salience system which in turn generates a control signal to engage brain areas mediating attentional, working memory, and action selection processes (while disengaging the default mode network). Critically, these switching mechanisms help focus attention on stimuli that signal deviant events or undesirable outcomes, as a result of which they take on added significance or saliency (Ullsperger et al., 2010).

Orienting to salient events or states that are associated with motivational significance could take various guises. One may orient attention to extraneous stimuli that call for action updating in order to secure valued outcomes and avoid undesired outcomes (stimuli that are novel, infrequent, deviant, unexpected, threatening, etcetera; or that serve as instructed targets or distracters); one may become receptive to induced emotions or affective states that call for approach or avoidance; or one may seek to monitor one's internal and external milieu for signals that register as a risk for undesirable outcomes (e.g., slips of action, performance errors; response capture, action conflict; negative feedback, punishment, lack of expected reward). In general, the salience system appears to be central to monitoring for specifically those motivationally important changes that require autonomic regulation (Critchley, 2009).

The AIC and the ACC have direct anatomical connections to the autonomic nervous system, mostly via brainstem nuclei that provide feedback on bodily states and changes in autonomic arousal (Craig, 2002). In particular, these cortical areas have robust connectivity to the locus coeruleus/norepinephrine (LC/NE) system involved in boosting and maintaining phasic and tonic arousal (Aston-Jones and Cohen, 2005). The LC is the main NE-generating nucleus in the brainstem, and the LC/NE system is central to regulating the sympathetic discharge and the inhibition of parasympathetic tone in arousal responses. Indeed, salient events are consistently associated with increased pupil-dilation response and skin conductance and with decelerated heart-rate, the more so for more unexpected events such as errors (Critchley, 2005). Taken together, this new understanding of the AIC within the context of the salience system provides a starting point to study communalities in inter-individual differences in error awareness and in the ability to selectively attend to motivational relevant events, as discussed in the next sections.

ERRORS AS SALIENCE SIGNALS

Empirical (Notebaert et al., 2009) and theoretical work (Ullsperger et al., 2010) has emphasized notable parallels between the processing of errors and of other rare/deviant/novel stimuli (or otherwise potentially significant or motivationally relevant events). Erroneous outcomes and other performance problems can be seen as salient events because of their infrequent occurrence and their usefulness as learning signals. They trigger a reflex-like orienting response in the salience network, which is accompanied by a cascade of central and autonomic nervous system reactions associated with increased autonomic arousal as needed to recruit the mental and physical resources required for adaptive action. This reflex-like orienting signal in the salience networks may act as an internal monitoring signal, timely informing the organism of behavioral changes that need to be made.

Meta-analyses have shown that the AIC and ACC are consistently reported to be co-activated during errors and other instances when performance monitoring becomes necessary (Ridderinkhof et al., 2004; Klein et al., 2007). Consistent with these observations, indices of autonomic arousal co-vary with conflicts, errors, and

feedback. For instance, error commission results in robust heart-rate deceleration and enhanced pupil-dilation responsivity, and these changes (that represent the recruitment of arousal so as to prepare the organism for adaptive action) tend to correlate with activity in the AIC and ACC.

ERROR AWARENESS VERSUS ERROR BLINDNESS

Error signals sometimes go unnoticed – they might need an appropriate potential in order for them to alert and engage the salience system and tip the balance between other related large-scale brain systems. For example, in order to be amplified into an orienting reaction in the salience network, error signals might need to surpass a certain energy threshold, or be accompanied by sufficient levels of physiological arousal. Performance errors are almost routinely registered in ACC, even if the individual does not consciously recognize the error as such (Nieuwenhuis et al., 2001; Endrass et al., 2007) but subsequent post-error slowing and changes in autonomic activity are observed only when subjects were aware of their error (Overbeek et al., 2005; Wessel et al., 2011). Error awareness has been found to engage specifically the right AIC but seems to place demands on bilateral anterior insula when applying a less conservative threshold (Klein et al., 2007). Specifically neurons situated in the anterior part of the insula are hypothesized to play a role in error awareness (Ullsperger et al., 2010). Activation of these anterior neurons is also observed during interoceptive awareness and the regulation of the body's homeostasis (Critchley et al., 2005), whereas neurons in the posterior part of the insula are thought to be involved in somatosensory or proprioceptive perception (Craig, 2002).

ERROR AWARENESS VIS-À-VIS ORIENTING TO ODDBALLS

Event-related potential (ERP) studies have highlighted two electrocortical components that can be observed when people make errors: the error(-related) negativity (N_E or ERN) and the error positivity (P_E; Falkenstein et al., 1999). The N_E is believed to reflect activity in the dorsal ACC when the detection of a performance error signals the loss of anticipated reward and the need for adjustments to achieve action goals; the P_E appears to reflect the conscious recognition of the fact that an error was committed (for review, see Overbeek et al., 2005). A perspective on the functional significance of the P_E in terms of error salience or motivational significance suggests that the P_E reflects processes similar to those expressed in another ERP component, the classical P3b (Polich, 2007). The events that give rise to a P3b can vary widely (from salient, novel, or rare stimuli to the absence of expected stimuli) but appear to have in common that they are motivationally significant, that is, they should motivate the individual to initiate or change a course of action in order to keep performance at an optimal level (Ridderinkhof et al., 2009). According to recent views, the P3b comprises the electrocortical expression of the response of the LC/NE system to the preliminary outcome of internal decision making processes and the consequent effects of the noradrenergic potentiation of information processing (Aston-Jones and Cohen, 2005; Nieuwenhuis et al., 2005).

A robust finding is that P3 amplitude is inversely related to target probability in oddball tasks (requiring the detection of distinct infrequent target stimuli or oddballs which are embedded in a series of frequently presented non-target or standard stimuli; e.g. (Duncan-Johnson et al., 1984). Moreover, P3s to oddballs are more enhanced when the target stimulus is embedded in a train of non-target stimuli rather than in a train of other targets (Squires et al., 1976). Rather than being attributable to target probability *per se*, these P3 effects are crucially mediated by target-to-target interval (TTI) duration (Croft et al., 2003). The effect of TTI on P3b amplitude was observed to co-vary with the amplitude of the P_E (Ridderinkhof et al., 2009), supporting the notion that the P_E and P3b reflect similar neurocognitive processes possibly involved in the conscious processing of motivationally significant events. In an earlier combined neuroimaging ERP study, Horovitz et al. (2002) found similar parametric effects of TTI on P3 amplitude.

Several groups have examined brain regions critical for identifying and responding to oddball-targets (Horovitz et al., 2002; Liebenthal et al., 2003; Kiehl et al., 2005). Areas sensitive to the parametric effects of TTI were found in ACC and AIC (as well as parietal cortex and the thalamus), confirming the suggestion that regions implicated in generating the P3 (Soltani and Knight, 2000; Stevens et al., 2005) coincide with the observed activations in AIC in error awareness (Klein et al., 2007).

CURRENT AIMS

The studies reviewed above strongly suggest a role for the AIC in orienting to salient events, such as errors (when recognized as such) and relevant infrequent events (when occurring unexpectedly). The current study aims to test the involvement of the AIC in both processes directly. The notion that conscious detection of an error triggers an orienting response toward a motivationally significant event, similar to the orienting response to a rare target stimulus, would gain considerable support if it could be shown that the hemodynamic response during error awareness overlaps with the parametric effect of TTI during an oddball task. The orienting response toward the detection of a deviant target was examined using an oddball task, using a TTI manipulation known to parametrically affect specifically the processing of motivational salience (Nieuwenhuis et al., 2005). Thus, here the TTI manipulation was introduced into the oddball task to tap the process salience processing.

We aim to explore whether the AIC activation for error awareness is related to salience processing by testing for activation overlap in subjects undergoing two different task settings: in the same scanning session, the same subjects completed an antisaccade task with self-evaluation of each antisaccadic response; a task frequently used to study error awareness as it typically elicits a considerable number of performance errors, of which approximately 50% remain unaware (Nieuwenhuis et al., 2001). The advantage of acquisition of both the antisaccade and the oddball task in one scanning session is that brain activation on these two tasks can be compared not only at the group-level, but can also be tested within each participant's brain activation. This yields a more precise comparison of the exact spatial distribution of the brain activation between the two cognitive processes. We predict that the hemodynamic response during aware (but not unaware) errors in the AIC overlaps with the oddball TTI effect. Specifically, we hypothesize that AIC of an individual, who engages to a higher degree in

consciously detected errors also engages to a higher degree in the processing of deviant targets after a longer interval.

MATERIALS AND METHODS

PARTICIPANTS

Fourteen healthy right-handed volunteers (12 females, mean age 21.2 ± 1.79)[1] with normal or corrected-to-normal vision participated in the experiment after giving written informed consent according to the Helsinki Declaration. They were paid 50 Euros for participation. None of the participants had a history of neurological or psychiatric disorders or eye-problems nor was taking medications influencing the central nervous- or cardiovascular systems. Participants were administered two tasks (antisaccade and oddball in counterbalanced order) within one scanning session.

TASKS

Oddball task

The orienting response was examined in an oddball task, using a TTI manipulation shown to parametrically affect salience processing specifically as reflected in the P3 (Ridderinkhof et al., 2009). The oddball task comprised a series of non-target and target stimuli that were presented for 100 ms on a computer screen in white uppercase letters (Os and Xs respectively, $2.5\,\text{cm} \times 2.5\,\text{cm} = 1.16^\circ \times 1.16^\circ$) against a gray background. Between stimuli a white fixation cross appeared ($0.30\,\text{cm} \times 0.30\,\text{cm}, 0.14^\circ \times 0.14^\circ$) for 1400 ms. Three experimental blocks, each lasting 8.15 min, were presented to the subject, each of which contained 300 non-targets and 30 target stimuli. The sequence of target and non-target trials was varied in such a way that 15 TTI (the number of non-targets between two targets) were created. These TTIs ranged from 3 to 17 non-targets between targets. The sequence of these 15 TTI conditions within blocks was determined randomly by the computer. Participants were instructed to react as quickly and accurately as possible to targets only using a button of an fMRI-compatible response box with their index finger. No reaction was required to the presentation of non-targets. For the fMRI analysis of the effect of inter target interval on BOLD signal, the 15 TTIs were divided *post hoc* into three TTI conditions TTI-1, TTI-2, and TTI-3. TTI-1 comprised 3–7 non-targets between targets, TTI-2 comprised 8–12 non-targets, and TTI-3 comprised 13–17 non-targets. The temporal order of stimuli is depicted in **Figure 1**.

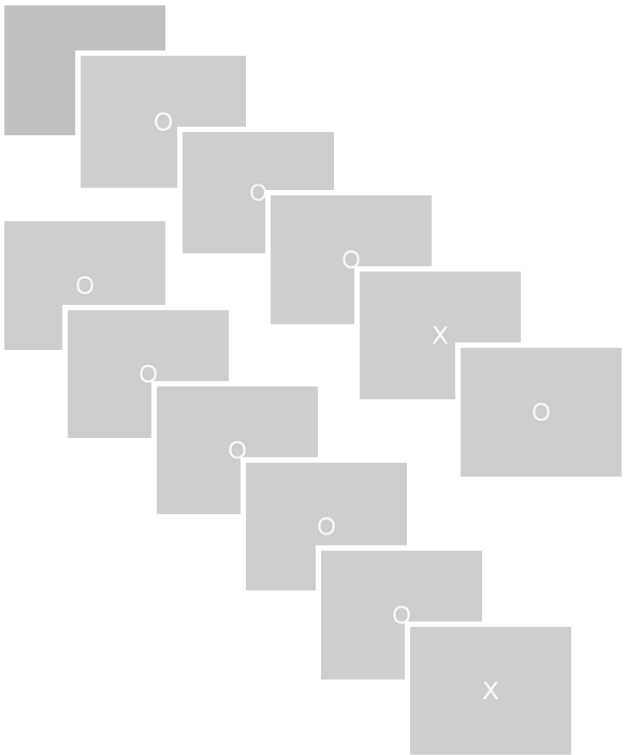


FIGURE 1 | Oddball task. A series of non-target (0) and target (X) images was presented against a gray background. The target- and non-target ranged from 3 to 17 non-targets between targets. Participants were to react as quickly and accurately as possible to targets only by pressing a button with their index finger.

Antisaccade task

We examined unaware and aware errors in an antisaccade task with self-evaluation of each antisaccadic response, a task that typically elicits a considerable number of performance errors, of which approximately 50% remain unaware (Nieuwenhuis et al., 2001). Participants were instructed to fixate on a central target and generate an immediate eye movement away from an abrupt peripheral target to its mirror location on the opposite side of the screen without making an eye movement to the peripheral target itself.

A trial was classified as an error, when the participant looked at the peripheral target, even when this error was immediately corrected. To increase the error rate, a brief precue was presented at the position where the gaze should be directed to (Fischer and Weber, 1996; Klein et al., 2007). To reduce predictability, the precue was presented at the position of the following peripheral stimulus in 33% of the trials.

After the eye movement, participants were to indicate with a button-press whether their antisaccadic response was correct (immediate eye movement to the other side of the screen) or incorrect (initial eye movement toward the target). The erroneous responses participants had rated as incorrect were classified as aware errors and erroneous responses rated as correct were classified as unaware errors. If the erroneous eye movement was redirected to the correct (opposite) side of the screen, the response was labeled "corrected error."

Participants completed 3 blocks of 100 antisaccade-trials, each lasting 11 min. For assessment of the pupil response, light flux was calibrated to equal luminance across trials with the program Colorfacts 7 and the color calibration system "EyeOneMonitor[2]"

[1] Subsample of Harsay et al. Shifts between default mode and task-focused brain networks during error awareness (*article submitted for publication*). To ensure a sufficient number of errors only participants with a minimum of 15 errors in the aware/unaware condition, with a false alarm rate lower than the aware/unaware error rates, and a maximum of 5% of self-rated uncertainty (based on a 1–100% post-experimental self-rating scale of uncertainty in performance evaluation) participated in the scanning session.

[2] www.datacolor.eu

and tested for equal pupil luminance response across precue conditions. There was no significant difference in pupil-dilation between trials with (0.4 ± 1.1) and without precue [0.4 ± 1.2; $t(22) = 0.01$, $p < 0.995$]. Light in the scanning environment was constrained to video presentation of stimuli against a black background. The trial started with a central fixation dot surrounded by two square outlines (each subtending 3.8° visual angle; distance from fixation 12.4°; display-duration 1000 ms). After a 150–300-ms jittered fixation gap, the peripheral target (a white circle subtending 2.9°) was unpredictably presented for 117 ms in the left or the right square. To induce erroneous responses a precue was presented in 50% of the trials, briefly (50 ms) thickening the outlines of the square at the opposite side of the target and validly indicating the target location. After a response window (of 880 ms) a cross appeared (for 500 ms) in the correct square indicating the correct gaze direction. Participants were to evaluate their performance (within 1500 ms) by pressing one of two buttons of an fMRI-compatible response box. On trial number 20, 40, 60, and 80, an instruction screen (duration: 2 s) appeared, reminding participants to keep saccading at fast pace. A black screen with jittered duration (16, 500, 1000, 1500 ms) was displayed between trials and 10% of the trials were "null events" (fixation-only trials of 5952 ms). The temporal order of stimulus presentation is displayed in **Figure 2**.

BEHAVIORAL DATA ACQUISITION AND ANALYSIS

Oddball task

Stimuli were presented on a 66 cm × 88 cm screen, placed at a 4-m viewing distance at the front end of the scanner and seen through a mirror above the participants' heads. Stimuli were presented and button-press responses (from an MRI compatible response box) were recorded with a presentation PC (Neurobehavioral Systems[3], Albany, NY, USA), that was connected to the MRI-scanner allowing for the time locking of stimuli, responses, and fMRI image acquisition. Two generalized linear model repeated-measures analyses of variance (ANOVAs) were used to investigate accuracy and reaction time in the oddball task. The independent variable for this analysis was TTI. TTIs were collapsed together into three TTI bins (3–7, 8–12, and 13–17 consecutive non-targets).

Antisaccade task

Oculomotor, pupil, and button-press responses were recorded with two interconnected PCs: an eye-tracker PC (ViewPoint Eye-Tracker, Arrington Research)[4] and a presentation PC (Neurobehavioral Systems, see text footnote 3, Albany, NY, USA). Both PCs were connected to the MRI-scanner allowing for the time locking of stimuli, responses, and fMRI image acquisition. The participant's left eye was continuously monitored with an MRI compatible infrared oculographic limbus tracker (Resonance Technology, Inc.)[5] attached to the head coil and placed 3 cm beneath the participant's left eye. The eye-tracker registered eye movements, aspect ratio, and diameter of the pupil with a sampling rate of 60 Hz along with scanner pulses and stimulus onsets. Before the scan, a nine-point calibration was performed and calibrated eye position was slip corrected during the task to eliminate slow drifts. Calibration

[3]www.neurobs.com
[4]www.arringtonresearch.com
[5]www.mrivideo.com

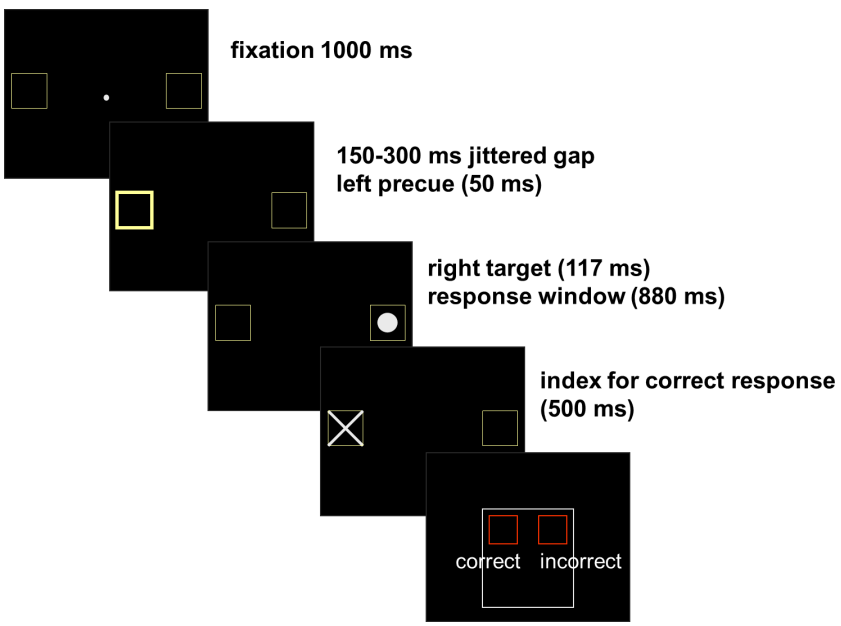


FIGURE 2 | Antisaccade task: participants were instructed to fixate on a central target and generate an immediate eye movement away from an abrupt peripheral target to its mirror location on the opposite side of the screen without making an eye movement to the peripheral target itself. After the response a cross appeared in the correct square indicating the correct gaze direction. Participants were to evaluate their performance by pressing one of two buttons of an fMRI-compatible response box. An initial eye movement toward the peripheral target was classified as an error. The erroneous responses participants had rated as incorrect were classified as aware errors and erroneous responses rated as correct were classified as unaware errors.

and stimuli were presented on a 66 cm × 88 cm screen, placed at a 4-m viewing distance at the front end of the scanner and seen through a mirror above the participants' heads. Saccade onsets, amplitudes, and directions were detected with in-house Java-based software[6] using minimum amplitude (>1.5°) and velocity (>30°/s) criteria and were subsequently double-checked by two raters. In line with common definitions (Fischer et al., 1993) we excluded trials in which subjects initiated saccades faster than 80 ms after target appearance [3.3 ± 4.1% (SD) of all trials], trials in which subjects were looking away from fixation during target presentation (2.7 ± 3.9%), blinked during target appearance (0.6 ± 1.2%) and trials for which the eye movement data were not interpretable due to poor quality of the eye-tracker signal (5.0 ± 4.3%).

fMRI ACQUISITION AND ANALYSIS

Acquisition

Functional images during the oddball- and the antisaccade task were acquired in the same subjects in the same scan session on a Philips (Philips, the Netherlands) 3 T MRI system equipped with echo planar imaging (EPI) capabilities using a standard head coil for radio frequency transmission and signal reception. Functional scans of the entire brain were acquired with a single-shot, gradient-recalled EPI sequence parallel to the AC–PC plane (TE/TR = 28/2000 ms; 30 axial slices; slice thickness 3 mm; interslice gap 0.3 mm; voxel size 3 mm × 3 mm × 3 mm; FOV = 222 mm × 2 mm; 96 × 96 in-plane resolution/matrix size, 90° flip angle). The first two volumes were discarded to allow for T1 equilibration effects. The duration of the oddball task was three times 8.15 min (245 scans per scanblok), the antisaccade task was three times 11 min (335 scans per scanblok). High-resolution anatomical images were subsequently acquired using a 3-D T1-weighted scan in steady state sequence (TE/TR = 4.6/9.69 ms; 182 sagittal slices; slice thickness 1.2, interslice gap 0.3 mm; voxel size 1 mm × 1 mm × 1 mm cubic; FOV = 25 cm × 2 cm; 256 × 2 in-plane resolution, 8° flip angle, sagittal orientation).

Preprocessing and GLM

Preprocessing of the functional data and calculation of the contrast images for statistical analysis was done with FEAT (FMRI Expert Analysis Tool) Version 5.63, a part of FSL (FMRIB's Software Library)[7]. Functional images were realigned to compensate for small head movements, slice-time corrected, spatially smoothed with a 5-mm full-width half-maximum Gaussian kernel, filtered in the temporal domain using a high-pass filter with a cutoff frequency of 1/50 Hz to correct for baseline drifts in the signal and prewhitened (Woolrich et al., 2009). For each experimental run of each participant, the overall activity was modeled as evoked by the targets (which were associated with one of three TTI conditions: TTI-1, TTI-2, and TTI-3; see task description), and by the correct responses and error commissions in the antisaccade task (two levels: aware errors versus unaware errors). The three levels (TTI-1, TTI-2, and TTI-3) in the oddball task were statistically compared first by fitting a linear model describing a linear signal increase from TTI-1 to TTI-2 to TTI-3), and second by subtracting TTI-1 from TTI-2, TTI-1 from TT-3, and TTI-2 from TTI-3. Each regressor in the oddball task and in the antisaccade task was convolved by a prototypical synthetic hemodynamic response function and its first derivative. To remove any artifactual signal changes due to head motion, six parameters describing the head movements (three translations, three rotations) were included as confounds in the model. In the second-stage analysis participants were treated as a fixed factor to concatenate the three experimental runs. Contrasts pertaining to the main effects constituted the data for the third-stage (mixed effect) analysis, where the significance of observations was determined across the group of 14 subjects using FLAME 1 and 2 (FMRIB's Local Analysis of Mixed Effects; Smith et al., 2004). For each whole-brain comparison of the target interval conditions in the oddball task we computed the initial statistical test with FSL-FEAT (FMRIB's Software Library; see text footnote 7), and thresholded the resulting z statistic image to show which voxels or clusters of voxels are activated at a particular significance level. We selected cluster thresholding, and used a z statistic threshold to define contiguous clusters. Each cluster's estimated significance level, corrected for whole-brain multiple comparisons using Gaussian random field theory (GRFT), and was compared with the cluster probability threshold. Significant clusters were then used to mask the original z statistic image for later production of color blobs. A cluster of voxels was considered significantly active if it passed the threshold of $z = 2.3$ and $p = 0.0.05$. This method of thresholding is an alternative to voxel-based correction, and is normally more sensitive to activation.

Comparative analyses

Participants had completed both the antisaccade and the oddball task within one scanning session. The advantage of this acquisition is that brain activation on these two tasks can be compared not only at the group-level, but can also be tested within each participant's brain activation, i.e., in his native functional space. For a given participant this native functional space is an image with brain activation acquired on that particular subject. The image is not yet transformed into a standard reference image, as for example the MNI brain from the Montreal Neurological Institute that defined a standard brain by using a large series of MRI scans on normal controls, representative of the population. This yields a more precise comparison of brain activation between the two cognitive processes. Four types of comparison were applied: spatial overlap analysis at the group-level, contrast masking analysis, Region of interest (ROI)-based correlation analysis and ROI-based ANOVA-analyses of average regression weights across tasks, within subjects.

Step 1: group-level spatial overlap analysis

In step 1 we plotted mean group activation during aware (compared to unaware) errors in the antisaccade task on top of the mean group activation that was elicited by oddballs and sensitive to parametric TTI effects in the oddball task. This yielded a map illustrating the spatial localization of brain areas showing increased amplitude of the hemodynamic response to aware errors and to target stimuli with a parametrically increasing TTI (**Figure 3**-Overlap).

[6] www.java.com

[7] www.fmrib.ox.ac.uk/fsl

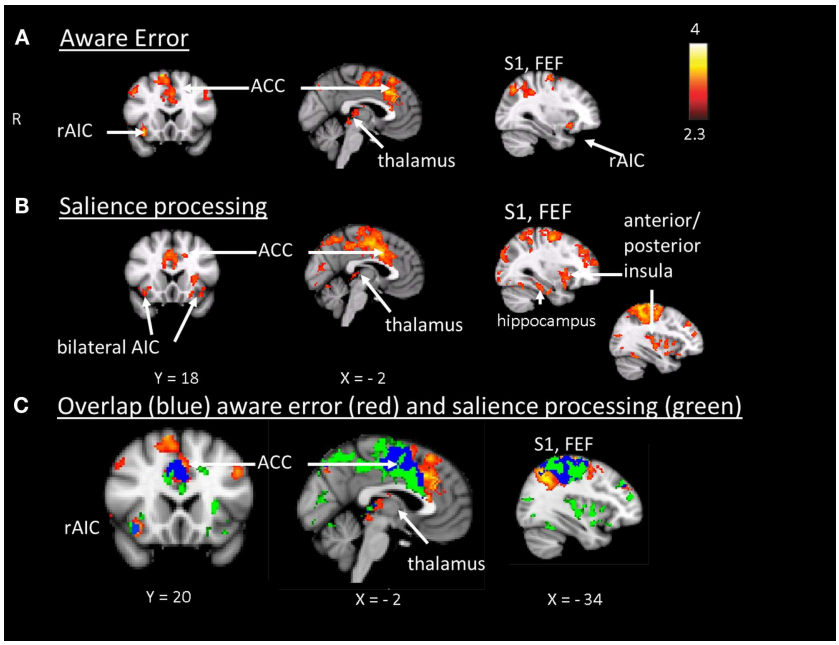


FIGURE 3 | (A) Aware error: statistical parametrical map of difference in BOLD activation between aware and unaware errors. Red and yellow voxels represent clusters of significant BOLD signal increase. **(B)** Salience processing: statistical parametrical map of difference in BOLD activation for the parametrical oddball. Red and yellow voxels represent clusters of significant BOLD signal increase. Renderings (on MNI stereotactic space) are thresholded at $z = 2.3$ and $p = 0.05$. **(C)** Overlap: plotted overlap between BOLD activation in the same subjects and the same scan session during aware errors and during the salience processing Note: R, right; ACC, anterior cingulate cortex; AIC, anterior insula cortex; FEF, frontal eyefields; S1, somatosensory cortex.

Step 2: contrast masking analysis on single subject level

In step 2 we tested within each participant at the whole-brain level for overlapping clusters of activation between the aware versus unaware contrast and for oddball-target detection (which was associated with parametric TTI effects). To this purpose we applied FMRIB's Local Analysis of contrast masking (Smith et al., 2004).

With the FSL-function of contrast masking, one can set up the masking of contrasts by other contrasts. After thresholding of all contrasts has taken place one can further threshold a given z statistic image by masking it with non-zeroed voxels from other contrasts. Non-zeroed voxels are voxels with have passed the cluster threshold of $z = 2.3$ and $p = 0.05$ in the contrast. This means that of the voxel clusters, which passed thresholding in the first contrast of interest, only those, which also survived thresholding in the other contrasts, are kept. Aim of this analysis is to detect overlapping clusters of voxels that survive within one participant both the threshold for the awareness contrast and the threshold for oddball-target detection (which was associated with parametric TTI effects).

First the initial statistical test was carried out for the error awareness task. The resulting z statistic images were thresholded to show which contiguous clusters of voxels were activated in each participant at the statistic threshold of $z = 2.3$ and $p = 0.05$ in the contrast aware versus unaware error (Smith et al., 2007). The result was a thresholded z statistic image for aware as compared to unaware errors, that constituted all contiguous clusters of voxels that had survived the cluster threshold of $z = 2.3$. In the next step the contrast for the oddball-target detection which was associated with activation after the longest target tot target interval) was computed at statistic threshold of $z = 2.3$ and $p = 0.05$, within the "mask" of the error awareness contrast. This means that of the oddball-target clusters which passed z-thresholding, only those which also survived z-thresholding in the aware versus unaware contrast are kept.

Thus, we constrained our search to activation in the aware versus unaware contrast which was also sensitive to oddball-targets which were associated with the longest TTI.

The result is a conservative analysis: brain structures with few or distributed active voxels will not survive thresholding. The resulting spatial overlap maps of each subject were subsequently fed into a group-level analysis. For this mixed effect analysis, FLAME 1 and 2 (FMRIB's Local Analysis of Mixed Effects; Smith et al., 2004) was used, in which the significance of activation common to error awareness and oddball-target detection associated with parametric TTI effects was computed across the group of all 14 subjects. We report a cluster-corrected threshold of $p < 0.05$ corrected for whole-brain multiple comparisons (using GRFT). The result is a precise spatial map depicting "error awareness areas" that are also sensitive to oddball-target detection associated with parametric TTI effects (see **Figure 4**).

Step 3: ROI-based correlation analysis

The AIC has been found associated to error awareness more consistently than the ACC. Hence, the AIC constituted an *a priori* ROI. Specifically, we were interested in determining whether those individuals who engaged the AIC to a greater extent during

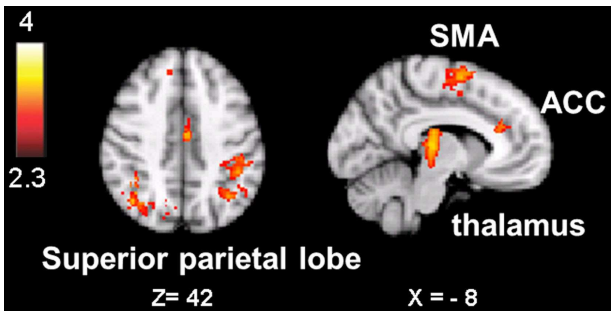


FIGURE 4 | Spatial overlap map of clusters of activation on group-level that survived, within each participant's native functional space, both the threshold for the awareness contrast and the threshold for salience processing. Analyses were we constrained by creating for each individual an "error awareness brain mask" within which activation was sensitive to salience processing. Renderings (on MNI stereotactic space) are thresholded at $z = 2.3$ and $p = 0.05$. Note: ACC, anterior cingulate cortex; SMA, supplementary motor area.

consciously detected errors also engaged this area more strongly when processing deviant targets after longer TTI's. Therefore, in step 3 we extracted the hemodynamic response of each participant's AIC to error awareness (aware error-unaware error) and to parametric TTI effects during target detection (linearly increasing parametric hemodynamic response across TTI, TTI-2, and TT-3), and correlated these two extracted difference scores across participants. As error awareness has previously been shown to engage only the anterior part of the insula and may furthermore engage the left and right AIC differentially (Klein et al., 2007). ROIs were defined for the anterior and posterior insula and for the left and right hemisphere separately. To test for the specificity of the AIC in orienting to salience, we contrasted the AIC to posterior insula cortex (PIC) activation.

Definition of ROIs was based on the MNI structural atlas of the FSL-atlas toolbox and available literature on neurosurgical landmarks (Mazziotta et al., 1995; Ture et al., 1999; Eickhoff et al., 2007). For anterior and posterior masks, coordinates were taken from Brooks et al. (2005). The vertical border between anterior and posterior portions of the insula was chosen such that the AIC seed subtended the three principal short insular gyri (anterior, middle, posterior) and the accessory and transverse insular gyri, all anterior to the insular sulcus (border for right and left insular cortex at $y = 1.3$, see **Figure 5A**). Percent signal change in bilateral AIC seeds was extracted for each subject for the aware versus unaware error contrast from the antisaccade task and for the TTI-3 minus TTI-1 contrast from the oddball task. Since we were interested specifically in whether the insula was engaged, within participants, in error awareness as well as in oddball processing, we computed bivariate correlations between percent signal change in AIC (and PIC) during aware errors and during interval-related target detection across participants. The predicted correlations were tested one-sided. Pearson correlation coefficients and p-values are presented. The resulting correlation maps show the relation between signal changes derived from bilateral AIC (and PIC) seeds as induced by aware (compared to unaware) errors and by target detection at long (compared to shorter) TTI's.

Step 4: ROI-based ANOVA-analyses of average regression weights (across tasks, within subjects, for each voxel). Within subjects and across tasks (the error awareness task and the oddball task) we performed analysis of regression weights using FSL's Featquery signal change processing tool (Analysis group, FMRIB, Oxford, UK)[8]. Featquery was conducted to interrogate signal change of *a priori* ROIs, previously defined by the literature reviewed, the anterior and posterior parts of insula cortex. After transforming the anterior and posterior insula masks into the native low resolution space, Featquery extracted regression weights (parameter estimates) and converted them to percent signal change values. This is achieved by scaling the P_E values by (100*) the peak-peak height of the regressor and then by dividing by the mean image from fil filtered_func_data. This analysis yielded mean statistical values of signal change across the time series with the anterior and posterior insula. In the next step we fed these values into group-level ANOVA analysis (SPSS)[9] to compare activation in anterior insula and posterior insula within subjects across tasks.

RESULTS

BEHAVIOR

Antisaccade task

Mean error rate was 27.5 (SD: 15.5%). Erroneous responses were initiated faster than correct responses [190 versus 282 ms; $t(13) = 7.1$; $p < 0.001$]. Participants were aware of roughly half of their errors; the other half went unnoticed [13.6 versus 13.9%; $t(13) = -0.063$; $p = 0.95$]. In 74.8% of errors, participants immediately corrected their erroneous response with an eye movement to the correct location. Unaware and aware errors were similar in mean latency [186 versus 194 ms; $t(13) = 0.40$; $p = 0.69$]. Yet, unaware errors were corrected significantly more often than aware errors [93.1 versus 63.5%; $t(13) = 2.8$; $p < 0.013$]. False alarm rates below 5.0% indicated that participants rarely reported an error when they had made a correct antisaccade.

Oddball task

For the oddball task, mean reaction time for correct target detection responses was 317 ms. RT did not vary as a function of TTI, $F(2,26) = 1.49$, $p = 0.25$. Overall accuracy of target detection was 98.7% and did not vary systematically as a function of TTI, $F(2,26) = 0.034$, $p = 0.97$.

fMRI ACTIVATION PATTERNS

The antisaccade task: aware versus unaware errors

Compared to unaware errors, aware errors yielded significantly increased activation in right AIC, dorsal ACC, bilateral pre- and postcentral gyrus (somatosensory cortex), bilateral frontal eyefields, superior parietal lobules, and bilateral thalamus (see **Figure 3A**; also **Figure A1**; **Table A1** in Appendix).

The oddball task and salience processing: interval effects on target detection

The parametric effect of interval length (TTI) on the detection of an oddball-target was observed in a number of areas, including

[8]http://www.fmrib.ox.ac.uk/fls/
[9]http://www.spss.com

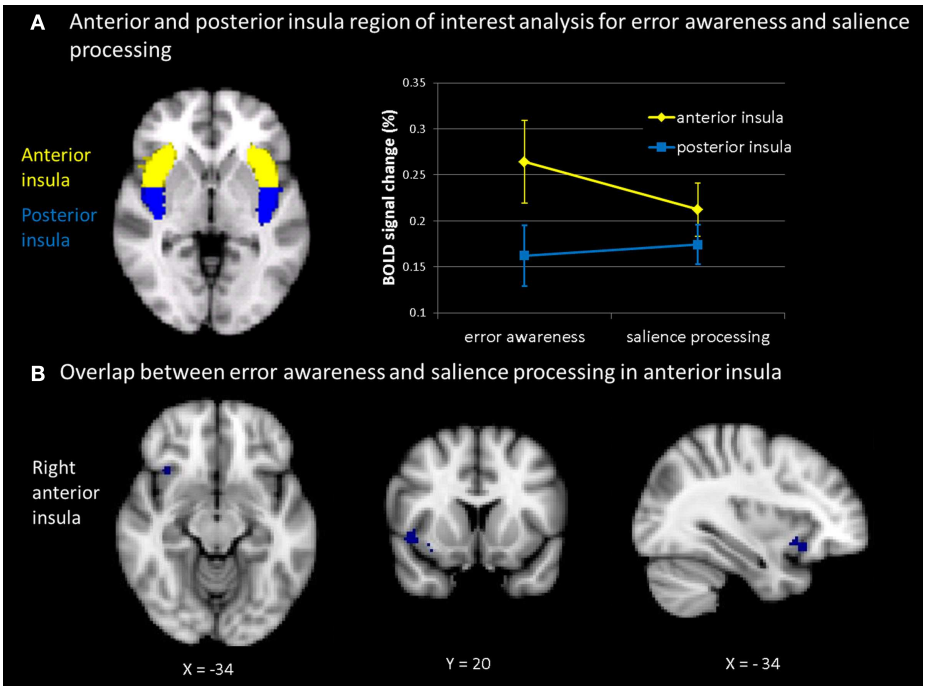


FIGURE 5 | (A) Mean percent BOLD signal change within subjects across tasks (for the contrast aware errors as compared to unaware errors; and for the contrast of salience processing, i.e., linear signal increase across inter-target interval) in anterior insula and posterior insula (thresholded at $z = 2.3$ and $p = 0.05$). Participants showed during both error awareness and salience processing a significantly higher percent signal change in the anterior than in the posterior insula. The main effect of task indicated no differences in percent signal change in the insula between error awareness and the processing of motivationally significant events **(B)**: plotted overlap in the anterior insula within subjects across tasks (for the contrast aware errors as compared to unaware errors; and for the contrast of salience processing, i.e., linear signal increase across inter-target interval).

AIC and PIC, dorsal ACC, supplementary motor area, pre- and postcentral gyri (somatosensory cortex), inferior and superior parietal lobules, and the thalamus, mostly bilateral (see **Figure 3B**; also **Figure A2**; **Table A2** in Appendix). TTI-3 minus TTI-1 contrast analysis (subtracting the shortest interval length from the longest interval length) yielded highly similar activation patterns (see **Figure A3** in Appendix). These regions have been observed previously to be active not only during target detection but also as a parametric effect of target interval (Kiehl and Liddle, 2003; Kiehl et al., 2005; Stevens et al., 2005).

Group-level spatial overlap analysis

The hemodynamic response during aware errors (compared to unaware) errors showed commonalities and differences with salience processing (see **Figure 3**-Overlap). Overlapping activation was observed in the right AIC, dorsal ACC, somatosensory cortex and precentral gyrus (frontal eyefields), thalamus, and brainstem. Compared to error awareness, salience processing additionally yielded increased activation in the left AIC, bilateral PIC, hippocampus, and inferior and superior parietal lobules.

Contrast masking analysis

Active voxels which passed, in each individual participant, cluster-corrected thresholding both for the aware-unaware contrast and for salience processing were found in the supplementary motor area, dorsal ACC, inferior and superior parietal lobule (supramarginal gyrus, postcentral gyrus), and thalamus, mostly bilateral (see **Figure 4**; **Table A3** in Appendix), as well as in the precuneus and lateral occipital gyrus (not shown). Notably, this analysis did not reveal overlapping voxels of activation in the AIC.

ROI-based correlation analysis

Mean percent signal change in the right and left anterior and posterior insula was extracted for all individual participants during aware (compared to unaware) errors. During error awareness, engagement of the left and right anterior part of the insula were strongly correlated ($r = 0.82$; $p = 0.0001$). Moreover, as depicted in **Figure 5**, participants who showed stronger engagement of the right and left anterior part of insula cortex during error awareness also showed stronger engagement of the right and left anterior part of insula cortex during target detection after longer compared to shorter TTI's ($r = 0.50$; $p = 0.03$, one-sided). This association was observed in both the right ($r = 0.44$; $p = 0.05$, one-sided) and the left AIC ($r = 0.51$; $p = 0.03$, one-sided). Thus, participants who activated the AIC to a greater extent to aware compared to unaware errors also activated the anterior part of the insula to a greater extent to motivational salience (target stimuli after a longer compared to shorter sequence of non-target stimuli). Bilateral anterior and posterior insula activation during error awareness failed to correlate with oddball processing (target compared to standard stimuli) *per se* (right AIC: $r = 0.29$, $p = 0.16$, left AIC: $r = 0.33$, $p = 0.13$, right PIC: $r = 0.10$, $p = 0.37$, left PIC: $r = 0.27$, $p = 0.18$). These correlation coefficients for error awareness and

motivational salience in the anterior insula were larger than the correlation coefficients in the anterior insula for error awareness and oddball processing *per se*. However, using the Fisher *r*-to-*z* transformation and the Meng test of two correlations with one variable in common from the same sample (Meng et al., 1992), the difference between these correlation coefficients of motivational salience and oddball processing did not reach significance in both tests (Fisher: $z = 0.41$, $p = 0.34$; Meng: $z = 0.038$, $p = 0.485$ for right AIC; $z = 0.52$, $p = 0.3$, Meng: $z = 0.043$, $p = 0.483$ for left AIC). Furthermore, the observed association was only observed in the anterior part of the insula; activation in the posterior part of the right and left insula showed no significant association between error awareness and TTI effects (bilateral PIC: $r = 0.18$; right PIC: $r = -0.04$; left PIC: $r = 0.43$). The difference between the correlation coefficients of error awareness and motivational salience in anterior insula ($r = 0.5$) and in posterior insula ($r = 0.18$) however failed to reach significance in both the Fisher and the Meng test (Fisher: $z = 0.92$, $p = 0.18$ (one tailed); Meng: $z = 0.08$, $p = 0.468$). In conclusion, we observe a tendency toward higher correlations between the two processes error awareness and motivational salience in the anterior insula than in the posterior insula and a tendency toward higher correlation between motivational salience and error awareness, than between oddball processing and error awareness, but the difference between the correlation coefficients does not reach significance level.

ROI-based ANOVA-analyses of average regression weights across tasks, within subjects

As can be seen in **Figure 5A**, a main effect of insula indicated differences in percent signal change between anterior and posterior insula [$F(1,13) = 38.717$, $p < 0.0001$]. The percent signal change values were analyzed using a mixed 2×2 ANOVA design with two within subjects variables (insula with two levels anterior and posterior; task with two levels error awareness and motivational significance). Participants showed during both error awareness and the processing of motivational significance a significantly higher percent signal change in the anterior insula than in the posterior insula. The main effect of task indicated no differences in percent signal change in the insula between error awareness and the processing of motivationally significant events [$F(1,13) = 0.777$, $p < 0.394$]. The test for interaction indicated that error awareness is associated with a marginally higher percent signal change in the anterior insula and a lower percent signal change in the posterior insula than the processing of motivational significance in the oddball task, trending toward significance [$F(1,13) = 23.989$, $p < 0.067$].

DISCUSSION

We report that error awareness shares anterior insula and cortico-thalamic circuits with target detection as modulated by TTI in a visual oddball task (referred to as "salience processing" in the remainder of the text).

Error awareness and salience processing showed activation overlap in six major brain areas: anterior insula, anterior cingulate, supplementary motor area, thalamus, brainstem, and parietal lobe. The findings of individual differences analysis of the *a priori* ROI AIC revealed that participants who activated the AIC to a higher degree to error awareness, also activated the AIC to a higher degree to salience processing. Within the AIC, interesting topographic differences were visible: error awareness activated predominantly the ventral AIC, whereas salience processing seemed to activate the AIC to a larger extent with maxima in the dorsal AIC, and with activation extending to PIC. The fine-grained contrast masking analysis within each participant's brain activation confirmed this observation: within AIC non-identical neural ensembles seem to be robustly activated within the same subjects during error awareness and salience processing. Robust direct spatial overlap was visible in the dorsal ACC, the supplementary motor area, the thalamus, and the parietal lobes. The results of the ROI-based ANOVA-analyses of average regression weights show that within subjects the insula shows significant percent signal change in both error awareness and salience processing (no significant main effect of task), and that the anterior part of the insula is significantly more involved in both processes than the posterior part (significant main effect of insular sub regions). Furthermore there is a tendency toward more AIC involvement and less PIC involvement in error awareness than in salience processing in the oddball task.

Together, these outcomes suggest a similar functional phenomenon in the two different task settings. In particular, they show a shared functional insula-cortico-thalamic anatomy for error awareness and salience processing, with some subregional anterior posterior specialization within the insula, and ventral dorsal specialization within the anterior insula.

The advantage of the current approach lies in the acquisition of both the "error awareness antisaccade task" and the oddball task in one scanning session in the same subjects. Overlap in brain activation on these two tasks can be compared not only at the group-level, but can also be tested within the brain activation of each participant. This yields a more precise comparison of the exact spatial distribution of the brain activation between the two cognitive processes.

One potential disadvantage of this approach lies in high stringent thresholds applied to extract only voxel clusters that are robustly involved in both tasks in each participant's brain activation. This threshold was chosen to account for noise in the individual data, but may lead to false negative results in small brain structures with activation in small voxel clusters.

Therefore, in order to gain a comprehensive picture of the overlap, three comparisons have been computed, and will be discussed below: (1) A whole-brain comparison with the plotted overlap of group-level activation patterns (for aware errors as compared to unaware errors plotted on parametrical effects of long as opposed to short ITIs in the oddball task); (2) a comparison showing the whole-brain group-level result of spatial overlap calculations (contrast masking of parametric target detection with activation clusters of error awareness) within each participant's brain activation; (3) ROI analyses focused on AIC, the *a priori* structure of interest in error awareness. In the following sections we will discuss first the findings on the whole-brain level, and second the findings that focus on the AIC. The AIC findings are placed in the context of current views on its role within larger scale functional and structural brain networks.

WIDELY DISTRIBUTED OVERLAP

Error awareness and salience processing showed overlap in six major brain areas. In general, this widespread overlap suggests that both error awareness and salience processing seem to engage multiple, spatially distributed processing systems. The most parsimonious cognitive interpretation of the widespread overlap is that it reflects a greater capture or orienting of attention (for aware errors as compared to unaware errors, as well as for long as opposed to short TTIs). The widespread overlap is reminiscent of similar patterns reported by Kiehl and co-workers in relation to reflexive or automatic orienting processes, that have been shown to reliably activate an extensive neural network (Kiehl and Liddle, 2003; Kiehl et al., 2005; Stevens et al., 2005).

Halgren et al. (1980, 1995a,b) have argued that a widely distributed response to salient events may be "adaptive" in an evolutionary sense. Activating many potentially useful areas, despite the low probability that these regions are all immediately functionally necessary, may lead to superior incidental learning and performance. The results of the "contrast masking comparison" of error awareness and salience processing, with stringent thresholds, included regions believed to mediate attentional control, particularly for salient stimuli, including thalamus, ACC, supplementary motor area, and superior parietal lobule (Corbetta and Shulman, 2002). As the ACC is known to increase in activation during conditions involving conflict monitoring (Ridderinkhof et al., 2004), the current activation overlap may reflect a cumulative process of increasing target expectancy and error monitoring (Squires et al., 1976; Kiehl et al., 2000). Additionally, a number of studies suggest a role of the dorsal ACC and the thalamus in the generation of peripheral arousal (Critchley, 2005). Single cell recordings show that the thalamus seems to be the first to react to an oddball stimulus to elicit an arousal response via its connections with the dorsal ACC (Klostermann et al., 2006). The robust overlap in dorsal ACC and the thalamus in the current data-sets may reflect a central correlate of peripheral system reactions associated with the increased autonomic arousal as needed to recruit the mental and physical resources required for adaptive action to the detection of errors and rare targets. Taken these two lines of thoughts together, the overlapping signal in the dorsal anterior cingulate and thalamus may act as an reflex-like orienting and monitoring signal, timely informing, and preparing the central and the peripheral neural system for behavioral changes that need to be made.

In line with this interpretation, a number of studies suggested that anterior cingulate activation might influence norepinephrine modulation of P3 in oddball tasks (Nieuwenhuis et al., 2005). Nieuwenhuis and colleagues proposed that phasic norepinephrine activity as mediated by the brainstem nucleus Locus Coeruleus may serve to enhance future top-down mediated selective attention for salient stimuli. Nieuwenhuis et al. (2005) predict that TTI would enhance neural response in brain areas active during task-relevant target processing, whereas no modulation would be seen to motivationally insignificant stimuli, which is the pattern of results found in this study in both the detection of motivationally significant rare targets as well as the conscious detection of errors. This is in line with the striking parallels in the association of the P_E with error awareness and the association of the P3b with salience processing (Ridderinkhof et al., 2009). Notably, both the P3b and the P_E have been proposed to be related to phasic activity of the Locus Coeruleus/norepinephrine system as has the orienting response (Overbeek et al., 2005).

The imminent question whether the observed robust hemodynamic overlap within the participants' native functional space in the thalamus, the parietal lobes, the dorsal ACC, and the supplementary motor area in the current study reflects activation in brain regions implicated in both generating the P3b (Soltani and Knight, 2000; Stevens et al., 2005) and the P_E (Klein et al., 2007) remains unclear. With the present fMRI study, it is not possible to definitely state whether the hemodynamics are related to one or another ERP "component." Even with the currently observed overlap, it remains possible that the neural source of the P3 elicited by the parametric oddball task may differ from the neural source of the P_E elicited by error awareness. The degree to which this overlap in hemodynamic change reflects common neural sources of the P3 and the P_E is not certain and might be effectively addressed by future studies combining fMRI and ERP technologies. Another suggestion for future studies could be to experimentally manipulate both the awareness of the error and the motivational significance of the event within one task. If error awareness would trigger insula activation also in the absence of motivational salience, it would make sense to describe these as different processes with a shared functional anatomy. As such manipulations are not be workable in the classical error awareness antisaccade task, used by this and previous studies on error awareness (Nieuwenhuis et al., 2001; Endrass et al., 2007; Klein et al., 2007), one may transfer the question to perceptual awareness tasks (such as used by van Gaal et al., 2010) or to tasks investigating the reward prediction error (Schoenbaum et al., 1998; Schultz, 1998).

In the error awareness task we observed a higher proportion of behavioral adjustments of the antisaccadic response (i.e., reversing an initial prosaccade into a timely antisaccade) after unaware errors than after aware errors. In speculation, this could have an impact on the BOLD-contrast aware versus unaware errors. Behavioral corrections after unaware errors may theoretically be associated with several neural processes. First, they may be associated with BOLD signal related to oculomotor behavior in the neural oculomotor circuits, specifically in the frontal eyefields and intraparietal sulcus (Connolly et al., 2005). As the corrective oculomotor response after unaware errors is, however, only slightly different in terms of saccadic control and is occurring at high pace, fMRI signal may have failed to pick up these slight oculomotor differences on the small amount of trials. Second, following evidence from unconscious inhibitory control (van Gaal et al., 2010) a higher proportion of oculomotor adjustments after unaware errors as compared to aware errors may imply a higher level of unconscious inhibitory control. A higher level of inhibitory control has previously been associated with a higher level of activation in the inferior frontal cortex and the pre-supplementary motor area. In the current study the activation pattern for signal change on unaware errors as compared to aware errors did not show increased activation in the inferior frontal cortex, the pre-supplementary motor area, the frontal eyefields and the intraparietal sulcus. This suggests that no consistent and significant BOLD signal related to a higher proportion of behavioral adjustments after unaware errors was picked up, due perhaps to the

too incidental and inconsistent occurrence and the slight oculomotor differences between corrected and uncorrected trials. This suggests that the contrast aware versus unaware errors should not be confounded by signal during unaware errors which is related to a higher proportion of behavioral adjustments after unaware errors.

ANTERIOR INSULA

The current results may provide information about the functional significance of AIC activation during error awareness. The plotted overlap of group-level activation patterns showed that the AIC responded to both error awareness and salience processing, while the PIC responded only to salience processing. Individual difference ROI analysis suggested that higher activation in AIC (but not in PIC) during salience processing predicted higher AIC activation to error awareness. In the individual difference analysis we however only observed a tendency toward higher correlations between error awareness and salience processing specifically in the AIC. The difference between the correlation coefficients of this significant correlation in the AIC on the one hand and the non-significant correlations on the other hand (between error awareness and salience processing in the PIC; and between error awareness and oddball processing *per se*) did not reach significance level. The results from the individual difference analysis in the AIC therefore lack specificity. The results of the ROI-based ANOVA-analyses of average regression weights indicated that within subjects the insula shows significant percent signal change in both error awareness and the processing of motivational significance (no significant main effect of task), but that the anterior insula is significantly more involved in both processes than the posterior insula (significant main effect of insular sub regions). Furthermore there is a tendency toward more AIC involvement and less PIC involvement in error awareness than in the processing of motivational significance in the oddball task.

In the light of previous findings on structural and functional connectivity of AIC, the currently observed similarity in AIC activation during error awareness and salience processing suggests that neural activity during both cognitive processes has direct access to similar larger scale neural systems. In contrast to the posterior part of the insula with few structural frontal projections, the AIC has been shown to be associated with strong frontal connectivity in studies of human probabilistic tractography. This anterior insula-frontal structural connectivity has been associated with the emotional salience and the cognitive control network linked to the implementation of goal-directed behavior (Cloutman et al., 2012). A recent investigation of insula-based resting-state fMRI has revealed similar results: whereas the posterior insula was functionally connected with primary and secondary somatomotor cortices; the dorsal anterior to middle insula was connected with dorsal ACC, along with other regions of the control network; and a ventral anterior region was primarily connected with pregenual ACC (Deen et al., 2011). Thus, error awareness and salience processing activate anterior subdivisions of the insula which seem ideally situated to communicate and integrate information within the salience network. This hypothesis might be effectively addressed by future studies combining structural and functional connectivity of AIC.

VENTRAL VERSUS DORSAL ANTERIOR INSULA

The overlap of error awareness and the monitoring of motivationally significant events in the ventral AIC as visible in the plotted overlap of the two group-level data-sets, did not survive the conservative thresholds of the spatial overlap analysis in the "contrast masking analysis." The contrast masking analysis however may have its methodological drawbacks, as the between-subject jitter in activation may prevent detecting activation overlap at the group-level, in particular in structures with a more distributed activation pattern of smaller voxel clusters. In the current study, it seems indeed that activation overlaps in larger and more continuous voxel clusters (such as thalamus and ACC) survived the "contrast masking" approach, whereas the smaller and more distributed voxel clusters in the insula may not have survived. While for the insula the contrast masking approach may have a drawback, the fact that contrast masking performed adequate on most of the structures involved at the whole-brain level, led us to include the contrast masking, approach into the paper. For the insula cortex, we have supplemented the contrast masking analyses with additional ROI analysis. The ROI analysis showed that the AIC is significantly involved in both task settings. The contrast masking analysis suggested further that error awareness activated predominantly the ventral AIC, whereas salience processing (the TTI effect) seemed to activate the AIC with maxima in the dorsal AIC. A functional dorsal-ventral distinction within anterior insula has not yet received much emphasis in the experimental literature on cognitive control, but has recently been addressed in a meta-analysis (Ullsperger et al., 2010). In a refined meta-analysis of 55 fMRI studies Ullsperger et al. focused on the patterns of co-activation of AIC and ACC across conditions that call for adjustments. They found that conditions of pre-response conflict (arising when a stimulus elicits competing response tendencies) and decision uncertainty (referring to situations when information about the correct response is underdetermined) primarily activated the dorsal part of AIC. Both conditions indicate an increased risk of imminent error, but the error might still be countermanded if the conflict is resolved or the uncertainty is reduced in time. By contrast, action slips and negative feedback cannot be repaired, but do call for remedial actions compensating the failure and/or subsequent adjustments improving future performance; these conditions predominantly activated the ventral part of AIC. Thus, the dorsal and ventral subregions of the AIC appear to play partially different roles in conditions that call for adjustments. The dorsal AIC appears to be involved in signaling increased risk (and hence the anticipation of imminent errors); the ventral AIC appears to register prediction error. Thus the dorsal AIC appears important in prospective control (recruiting the necessary effort to pre-empt potential risks and failures), whereas the ventral subdivision appears more important for reactive processing (monitoring for the need to undertake remedial action and homeostatic regulation; Lamm and Singer, 2010; Ullsperger et al., 2010). The current results seem to be in line with this proactive/reactive account of dorsal/ventral anterior insula.

Here, dorsal AIC activation during salience processing may reflect increased prospective control, due to the increased effort necessary to recruit sufficient resources to stay alert until the next target stimulus. In experimental research, the effect of a fore period

on the reaction to a target stimulus has often been used as an independent variable of primary interest (Los et al., 2001). The focus of interest is the process of attaining and maintaining a state of potential action toward a future target event. Reaction time in reaction to a target stimulus, following a preparational period is commonly accepted as behavioral index for the efficiency of preparation (Jennings et al., 1998). A fast reaction should index that the participant is optimally ready to respond whereas a slow reaction would index that the participant is unprepared. Thus, on targets in a prepared state we should see low reaction times. Here, in the oddball task, reaction time after the three different TTIs did not differ significantly, suggesting that subjects attained a preparation state across longer intervals that sufficed to maintain reaction time. Functionally, the preparation state can take many different guises, ranging from the simple presetting of a motor response to complex cognitive preparation. Whatever form preparation takes, though, it is always oriented toward some goal, and takes time to reach a level that is optimal for that goal. The preparation process across trial-to-trial interval has been described to rely on the principle of "trace conditioning" (Los et al., 2001; Los and Schut, 2008). Trace conditioning refers to an inverted u shaped function describing a high preparatory state that is quickly attained but hard to maintain over time, wherein the participant aims at synchronizing the preparation peak and the imperative moment in order to produce a fast response. The most characteristic for the trace conditioning model of preparation is that the response preparation declines if its corresponding critical target occurs prior to the expected moment of the response, but remains unchanged if its critical target occurs after the expected moment (Los et al., 2001). In the current study the similar reaction times after all three interval conditions suggest that optimal preparation has been maintained across longer intervals until the expected (oddball) target occurred. In speculation, an initially increased and then maintained level of preparation aimed to rapidly respond to an anticipated stimulus may partly be reflected in the higher activation to targets after longer as compared to shorter TTIs with equal reaction speed. Hence, dorsal AIC activation during salience processing may reflect increased prospective control, due to the increased effort necessary to recruit sufficient resources to stay alert until the next target after a longer interval. This remains however speculative and can adequately be addressed by experimental paradigms that allow for measuring BOLD signal during the interval. The ventral AIC activation during error awareness in contrast may reflect reactive control due to the need to take remedial action.

Additionally, the increased ventral AIC activation during specifically error awareness (as compared to salience processing), may reflect physiological arousal related to an aversive affective response. Error awareness has been related to increases in peripheral physiological response (O'Connell et al., 2009; Wessel et al., 2011). Consistently, a recent meta-analysis found that peak coordinates from studies linking brain activation to peripheral physiological responses related to emotional experiences, such as heart-rate or galvanic skin response, tended to lie in ventral AIC (Mutschler et al., 2009).

Following this thought, the functional activation of ventral AIC during error awareness may also relate to the experienced valence of a salient event as an error is likely to be experienced as more unpleasant than a parametric oddball-target. Both dorsal and ventral AIC activation has been observed in response to unpleasant or disgusting odorants and aversive tastes (Zald et al., 1998; Wicker et al., 2003) and disgusting images (Calder et al., 2007). The ventral AIC, in particular, has been consistently found to be modulated by the hedonic valence of olfactory and gustatory stimuli (Royet et al., 2003). Ventral AIC activations to disgusting stimuli may reflect affective response to disgusting stimuli, while the dorsal AIC is involved in linking this affective response to attentional or executive mechanisms, similar to such divisions in pain processing (Baliki et al., 2009). The current results seem to support this functional affective/cognitive distinction of dorsal/ventral insula.

Another proposal is that the AIC contributes to the conscious error processing by generating a form of orienting response toward the error (Ullsperger et al., 2010). The current results partly encourage this proposal. The direct activation overlap in the dorsal ACC during error awareness and oddball processing might point to the generation of autonomic arousal processes in both tasks. As described above, the dorsal ACC has been consistently related to the generation of peripheral arousal. The AIC in turn has been related to the mapping of the arousal response (Critchley et al., 2005). The currently observed activation of the AIC during both error awareness and oddball processing may reflect the AIC mapping of the dorsal ACC arousal response. By mapping the arousal response the AIC may ascribe emotional significance to deviant targets and perceived errors and initiate the integration of the salient information into decision making processes to guide behavioral responses. In this context, errors may be homeostatically more salient and experienced emotionally as more aversive than a rare/deviant oddball-target. Thus, the activation in specifically the ventral AIC to aware errors might relate to increased peripheral arousal linked to an aversive affective response to the error. This aversive arousing component may be functional in the sense that it may increase the likelihood that the neural and peripheral system takes immediate remedial action.

A potentially informative next step for future research seems to be functional connectivity analysis of coordinated activity between ACC and ventral versus dorsal AIC during error awareness and oddball processing. As of yet, network research has not yet been able to consistently dissociate ventral versus dorsal AIC function based on its network profile in humans (Cloutman et al., 2012) In general agreement with insula patterns of structural connectivity in the macaque (Mesulam and Mufson, 1982a,b; Mufson and Mesulam, 1982) studies of human functional connectivity revealed ventral AIC to be correlated mostly with dorsal ACC, while dorsal and posterior insula correlated with more posterior parts of ACC (Deen et al., 2011). In humans however, in contrast to the consistency with which AIC–ACC functional connectivity has been identified using human resting-state measures, white matter connections between the two areas in the human brain have been failed to be demonstrated or only inconsistently observed via tractographic methods, if at all (van den Heuvel et al., 2009). Future studies combining measures of peripheral arousal with neural network analysis may show if the dorsal and ventral AIC form distinct pathways by which different aspects of salient neural signal, such as peripheral arousal or valence, can differentially mediate cognitive control and behavior.

REFERENCES

Aston-Jones, G., and Cohen, J. D. (2005). Adaptive gain and the role of the locus coeruleus-norepinephrine system in optimal performance. *J. Comp. Neurol.* 493, 99–110.

Baliki, M. N., Geha, P. Y., and Apkarian, A. V. (2009). Parsing pain perception between nociceptive representation and magnitude estimation. *J. Neurophysiol.* 101, 875–887.

Bressler, S. L., and Menon, V. (2010). Large-scale brain networks in cognition: emerging methods and principles. *Trends Cogn. Sci. (Regul. Ed.)* 14, 277–290.

Brooks, J. C., Zambreanu, L., Godinez, A., Craig, A. D., and Tracey, I. (2005). Somatotopic organisation of the human insula to painful heat studied with high resolution functional imaging. *Neuroimage* 27, 201–209.

Calder, A. J., Beaver, J. D., Davis, M. H., van Ditzhuijzen, J., Keane, J., and Lawrence, A. D. (2007). Disgust sensitivity predicts the insula and pallidal response to pictures of disgusting foods. *Eur. J. Neurosci.* 25, 3422–3428.

Cloutman, L. L., Binney, R. J., Drakesmith, M., Parker, G. J., and Lambon Ralph, M. A. (2012). The variation of function across the human insula mirrors its patterns of structural connectivity: evidence from in vivo probabilistic tractography. *Neuroimage* 59, 3514–3521.

Connolly, J. D., Goodale, M. A., Goltz, H. C., and Munoz, D. P. (2005). fMRI activation in the human frontal eye field is correlated with saccadic reaction time. *J. Neurophysiol.* 94, 605–611.

Corbetta, M., and Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215.

Craig, A. D. (2002). How do you feel? Interoception: the sense of the physiological condition of the body. *Nat. Rev. Neurosci.* 3, 655–666.

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

Critchley, H. D. (2005). Neural mechanisms of autonomic, affective, and cognitive integration. *J. Comp. Neurol.* 493, 154–166.

Critchley, H. D. (2009). Psychophysiology of neural, cognitive and affective integration: fMRI and autonomic indicants. *Int. J. Psychophysiol.* 73, 88–94.

Critchley, H. D., Tang, J., Glaser, D., Butterworth, B., and Dolan, R. J. (2005). Anterior cingulate activity during error and autonomic response. *Neuroimage* 27, 885–895.

Croft, R. J., Gonsalvez, C. J., Gabriel, C., and Barry, R. J. (2003). Target-to-target interval versus probability effects on P300 in one- and two-tone tasks. *Psychophysiology* 40, 322–328.

Deen, B., Pitskel, N. B., and Pelphrey, K. A. (2011). Three systems of insular functional connectivity identified with cluster analysis. *Cereb. Cortex* 21, 1498–1506.

Dosenbach, N. U., Fair, D. A., Cohen, A. L., Schlaggar, B. L., and Petersen, S. E. (2008). A dual-networks architecture of top-down control. *Trends Cogn. Sci. (Regul. Ed.)* 12, 99–105.

Dosenbach, N. U., Fair, D. A., Miezin, F. M., Cohen, A. L., Wenger, K. K., Dosenbach, R. A., Fox, M. D., Snyder, A. Z., Vincent, J. L., Raichle, M. E., Schlaggar, B. L., and Petersen, S. E. (2007). Distinct brain networks for adaptive and stable task control in humans. *Proc. Natl. Acad. Sci. U.S.A.* 104, 11073–11078.

Duncan-Johnson, C. C., Roth, W. T., and Kopell, B. S. (1984). Effects of stimulus sequence on P300 and reaction time in schizophrenics. A preliminary report. *Ann. N. Y. Acad. Sci.* 425, 570–577.

Dupont, S., Bouilleret, V., Hasboun, D., Semah, F., and Baulac, M. (2003). Functional anatomy of the insula: new insights from imaging. *Surg. Radiol. Anat.* 25, 113–119.

Eickhoff, S. B., Paus, T., Caspers, S., Grosbras, M. H., Evans, A. C., Zilles, K., and Amunts, K. (2007). Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. *Neuroimage* 36, 511–521.

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

Falkenstein, M., Hoormann, J., and Hohnsbein, J. (1999). ERP components in Go/Nogo tasks and their relation to inhibition. *Acta Psychol. (Amst.)* 101, 267–291.

Fischer, B., and Weber, H. (1996). Effects of procues on error rate and reaction times of antisaccades in human subjects. *Exp. Brain Res.* 109, 507–512.

Fischer, B., Weber, H., Biscaldi, M., Aiple, F., Otto, P., and Stuhr, V. (1993). Separate populations of visually guided saccades in humans: reaction times and amplitudes. *Exp. Brain Res.* 92, 528–541.

Halgren, E., Baudena, P., Clarke, J. M., Heit, G., Liegeois, C., Chauvel, P., and Musolino, A. (1995a). Intracerebral potentials to rare target and distractor auditory and visual stimuli. I. Superior temporal plane and parietal lobe. *Electroencephalogr. Clin. Neurophysiol.* 94, 191–220.

Halgren, E., Baudena, P., Clarke, J. M., Heit, G., Marinkovic, K., Devaux, B., Vignal, J. P., and Biraben, A. (1995b). Intracerebral potentials to rare target and distractor auditory and visual stimuli. II. Medial, lateral and posterior temporal lobe. *Electroencephalogr. Clin. Neurophysiol.* 94, 229–250.

Halgren, E., Squires, N. K., Wilson, C. L., Rohrbaugh, J. W., Babb, T. L., and Crandall, P. H. (1980). Endogenous potentials generated in the human hippocampal formation and amygdala by infrequent events. *Science* 210, 803–805.

Horovitz, S. G., Skudlarski, P., and Gore, J. C. (2002). Correlations and dissociations between BOLD signal and P300 amplitude in an auditory oddball task: a parametric approach to combining fMRI and ERP. *Magn. Reson. Imaging* 20, 319–325.

Jennings, J. R., van der Molen, M. W., and Steinhauer, S. R. (1998). Preparing the heart, eye, and brain: foreperiod length effects in a nonaging paradigm. *Psychophysiology* 35, 90–98.

Kiehl, K. A., and Liddle, P. F. (2003). Reproducibility of the hemodynamic response to auditory oddball stimuli: a six-week test-retest study. *Hum. Brain Mapp.* 18, 42–52.

Kiehl, K. A., Liddle, P. F., and Hopfinger, J. B. (2000). Error processing and the rostral anterior cingulate: an event-related fMRI study. *Psychophysiology* 37, 216–223.

Kiehl, K. A., Stevens, M. C., Laurens, K. R., Pearlson, G., Calhoun, V. D., and Liddle, P. F. (2005). An adaptive reflexive processing model of neurocognitive function: supporting evidence from a large scale ($n = 100$) fMRI study of an auditory oddball task. *Neuroimage* 25, 899–915.

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.

Klostermann, F., Wahl, M., Marzinzik, F., Schneider, G. H., Kupsch, A., and Curio, G. (2006). Mental chronometry of target detection: human thalamus leads cortex. *Brain* 129, 923–931.

Lamm, C., and Singer, T. (2010). The role of anterior insular cortex in social emotions. *Brain Struct. Funct.* 214, 579–591.

Liebenthal, E., Ellingson, M. L., Spanaki, M. V., Prieto, T. E., Ropella, K. M., and Binder, J. R. (2003). Simultaneous ERP and fMRI of the auditory cortex in a passive oddball paradigm. *Neuroimage* 19, 1395–1404.

Los, S. A., Knol, D., and Boers, R. M. (2001). The foreperiod effect revisited: conditioning as a basis for nonspecific preparation. *Acta Psychol. (Amst.)* 6, 121–145.

Los, S. A., and Schut, M. L. (2008). The effective time course of preparation. *Cogn. Psychol.* 57, 20–55.

Mazziotta, J. C., Toga, A. W., Evans, A., Fox, P., and Lancaster, J. (1995). A probabilistic atlas of the human brain: theory and rationale for its development. The International Consortium for Brain Mapping (ICBM). *Neuroimage* 2, 89–101.

Meng, X., Rosenthal, R., and Rubin, D. B. (1992). Comparing correlated correlation coefficients. *Psychol. Bull.* 111, 172–175.

Menon, V., and Uddin, L. Q. (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Struct. Funct.* 214, 655–667.

Mesulam, M. M., and Mufson, E. J. (1982a). Insula of the old world monkey. I. Architectonics in the insulo-orbito-temporal component of the paralimbic brain. *J. Comp. Neurol.* 212, 1–22.

Mesulam, M. M., and Mufson, E. J. (1982b). Insula of the old world monkey. III: efferent cortical output and comments on function. *J. Comp. Neurol.* 212, 38–52.

Mufson, E. J., and Mesulam, M. M. (1982). Insula of the old world monkey. II: afferent cortical input and comments on the claustrum. *J. Comp. Neurol.* 212, 23–37.

Mutschler, I., Wieckhorst, B., Kowalevski, S., Derix, J., Wentlandt, J., Schulze-Bonhage, A., and Ball, T. (2009). Functional organization of the human anterior insular cortex. *Neurosci. Lett.* 457, 66–70.

Nieuwenhuis, S., Aston-Jones, G., and Cohen, J. D. (2005). Decision making, the P3, and the locus coeruleus-norepinephrine system. *Psychol. Bull.* 131, 510–532.

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.

Notebaert, W., Houtman, F., Van Opstal, F., Gevers, W., Fias, W., and Verguts, T. (2009). Post-error slowing: an orienting account. *Cognition* 111, 275–279.

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.

Overbeek, T. J. M., Nieuwenhuis, S., and Ridderinkhof, K. R. (2005). Dissociable components of error processing. *J. Psychophysiol.* 19, 319–329.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118, 2128–2148.

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.

Royet, J. P., Plailly, J., Delon-Martin, C., Kareken, D. A., and Segebarth, C. (2003). fMRI of emotional responses to odors: influence of hedonic valence and judgment, handedness, and gender. *Neuroimage* 20, 713–728.

Schoenbaum, G., Chiba, A. A., and Gallagher, M. (1998). Orbitofrontal cortex and basolateral amygdala encode outcomes during learning. *Nat. Neurosci.* 1, 155–159.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *J. Neurophysiol.* 80, 1–27.

Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., Reiss, A. L., and Greicius, M. D. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *J. Neurosci.* 27, 2349–2356.

Smith, S. M., Jenkinson, M., Beckmann, C. F., Miller, K., and Woolrich, M. W. (2007). Meaningful design and contrast estimability in fMRI. *Neuroimage* 34, 127–136.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I., Flitney, D. E., Niazy, R. K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J. M., and Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23(Suppl. 1), S208–S219.

Soltani, M., and Knight, R. T. (2000). Neural origins of the P300. *Crit. Rev. Neurobiol.* 14, 199–224.

Squires, K. C., Wickens, C., Squires, N. K., and Donchin, E. (1976). The effect of stimulus sequence on the waveform of the cortical event-related potential. *Science* 193, 1142–1146.

Sridharan, D., Levitin, D. J., and Menon, V. (2008). A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *Proc. Natl. Acad. Sci. U.S.A.* 105, 12569–12574.

Stevens, M. C., Calhoun, V. D., and Kiehl, K. A. (2005). fMRI in an oddball task: effects of target-to-target interval. *Psychophysiology* 42, 636–642.

Taylor, K. S., Seminowicz, D. A., and Davis, K. D. (2009). Two systems of resting state connectivity between the insula and cingulate cortex. *Hum. Brain Mapp.* 30, 2731–2745.

Ture, U., Yasargil, D. C., Al-Mefty, O., and Yasargil, M. G. (1999). Topographic anatomy of the insular region. *J. Neurosurg.* 90, 720–733.

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

van den Heuvel, M. P., Mandl, R. C., Kahn, R. S., and Hulshoff Pol, H. E. (2009). Functionally linked resting-state networks reflect the underlying structural connectivity architecture of the human brain. *Hum. Brain Mapp.* 30, 3127–3141.

van Gaal, S., Ridderinkhof, K. R., Scholte, H. S., and Lamme, V. A. F. (2010). Unconscious activation of the prefrontal no-go network. *J. Neurosci.* 30, 4143–4150.

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.

Wicker, B., Keysers, C., Plailly, J., Royet, J. P., Gallese, V., and Rizzolatti, G. (2003). Both of us disgusted in My insula: the common neural basis of seeing and feeling disgust. *Neuron* 40, 655–664.

Woolrich, M. W., Jbabdi, S., Patenaude, B., Chappell, M., Makni, S., Behrens, T., Beckmann, C., Jenkinson, M., and Smith, S. M. (2009). Bayesian analysis of neuroimaging data in FSL. *Neuroimage* 45, S173–S186.

Zald, D. H., Lee, J. T., Fluegel, K. W., and Pardo, J. V. (1998). Aversive gustatory stimulation activates limbic circuits in humans. *Brain* 121(Pt 6), 1143–1154.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 29 February 2012; accepted: 07 August 2012; published online: 27 August 2012.

Citation: Harsay HA, Spaan M, Wijnen JG and Ridderinkhof KR (2012) Error awareness and salience processing in the oddball task: shared neural mechanisms. Front. Hum. Neurosci. ***6****:246. doi: 10.3389/fnhum.2012.00246*

APPENDIX

Table A1 | Complete list of brain regions showing significant BOLD activation during aware errors as compared to unaware errors.

Brain region	*X*	*Y*	*Z*	Max *z*
R anterior insula cortex	34	18	−12	3.65
R mid insula cortex	50	8	−4	3.63
R postcentral gyrus (somatosensory cortex BA2R)	54	−26	44	3.04
L postcentral gyrus (somatosensory cortex BA2L, BA1L, BA3bL)	−46	−28	50	4.86
R thalamus	10	−24	10	3.15
L thalamus	−8	−22	8	4.39
R brain stem	8	−30	−8	3.21
R rostral anterior cingulate cortex	2	26	16	3.18
L rostral anterior cingulate cortex	−2	26	16	3.18
R dorsal anterior cingulate cortex	4	20	36	3.03
L dorsal anterior cingulate cortex	−4	32	32	3.64
R supplementary motor cortex (BA6R)	6	8	56	3.63
L supplementary motor cortex (BA6L)	−6	6	60	3.04
R precuneus cortex	4	−68	42	3.08
R inferior frontal gyrus	52	12	20	3.06
L inferior frontal gyrus	−48	12	22	2.48
R frontal eyefields BA8R, BA6R	20	−4	70	3.63
L frontal eyefields BA8L, BAL	−28	−26	70	4.22
R anterior intraparietal sulcus	−50	−44	50	4.23
L anterior intraparietal sulcus	40	−48	50	3.63
R parietal occipital junction (superior parietal lobe/lateral occipital lobe)	36	−58	40	3.72
L parietal occipital junction (superior parietal lobe/lateral occipital lobe)	−32	−60	40	4.30

Aware errors > unaware errors, BOLD activation cluster-corrected at $z = 2.3$, $p = 0.001$. Coordinates are given in MNI space.

Table A2 | Complete list of brain regions showing significant BOLD activation during target detection as a function of parametrically increasing interval length. during odd 3 as compared to odd 1.

Brain region	*X*	*Y*	*Z*	Max *z*
R anterior insula cortex	34	14	−4	2.94
L anterior insula cortex	−40	16	−6	2.94
L mid insula cortex	−40	−12	10	2.82
L postcentral gyrus	−42	−34	50	3.09
R thalamus	14	−22	8	3.39
L thalamus	−10	−22	8	2.91
R brain stem	10	−22	−12	2.80
Supplementary motor area	2	−12	64	2.85
Dorsal anterior cingulate cortex	2	8	44	3.28
Supplementary motor area	2	−12	64	2.85
Precuneus cortex	2	−64	56	2.71
L dorsolateral prefrontal cortex	−38	50	20	2.67
R inferior frontal gyrus	52	12	20	3.06
L inferior frontal gyrus	−48	12	22	2.48
R premotor cortex (frontal eyefields BA8R, BA6R)	20	−4	70	3.63
L premotor cortex (frontal eyefields BA8L, BAL)	−28	−26	70	4.22
R anterior intraparietal sulcus	−50	−44	50	4.23
L anterior intraparietal sulcus	40	−48	50	3.63
R parietal occipital junction (superior parietal lobe/lateral occipital lobe)	36	−58	40	3.72
L parietal occipital junction (superior parietal lobe/lateral occipital lobe)	−32	−60	40	4.30

Activation during oddball detection as a function of parametrically increasing interval length BOLD activation cluster-corrected at z = 2.3, p = 0.05. Coordinates are given in MNI space.

Table A3 | Spatial overlap map of clusters of activation that survived, within each participant's native space, both the threshold for the awareness contrast and the threshold for parametric TTI effects during target detection.

Brain region	*X*	*Y*	*Z*	Max *z*
R thalamus	8	−26	0	2.89
L thalamus	−6	−26	4	3.25
R supplementary motor area	6	8	64	3.95
L supplementary motor area	−8	6	64	3.18
R dorsal ACC	2	18	26	2.39
L dorsal ACC	−8	34	24	3.21
Precunus	14	−64	46	2.82
L somatosensory cortex	−46	−32	46	3.56
R lateral occipital cortex	40	−64	44	3.59
L lateral occipital cortex	−38	−64	52	3.56

Local maxima of activation of all significant clusters (at z = 2.3, p = 0.05, cluster-corrected) varying with aware errors and with the interval effect on target detection. All coordinates are given in MNI space.

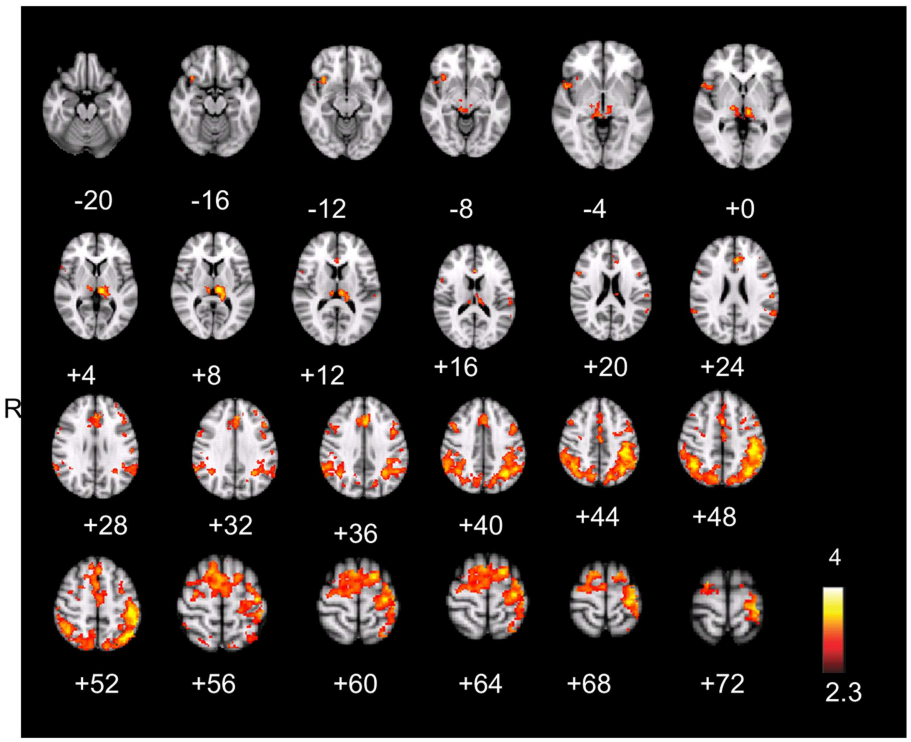


FIGURE A1 | Neural activation on aware errors. Statistical parametrical map of difference in BOLD activation between aware and unaware errors. Red and yellow voxels represent clusters of significant BOLD signal increase across all subjects. For a full list of activated regions ($z > 2.3$, whole-brain cluster-corrected, $p < 0.05$), see **Table A1**.

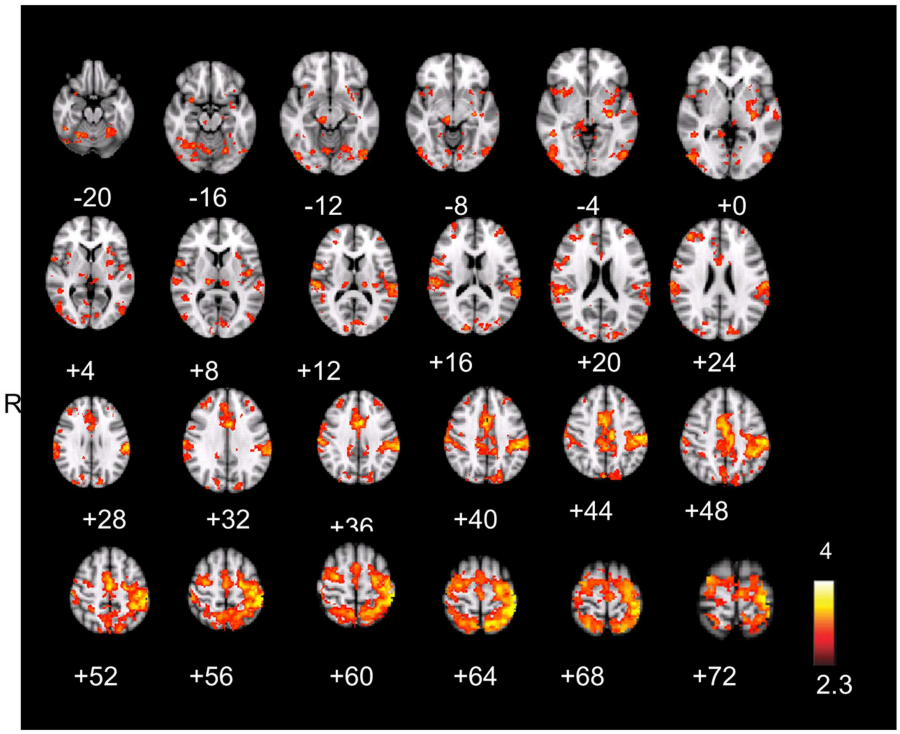


FIGURE A2 | Illustration of brain areas showing increasing amplitude of the hemodynamic response to target stimuli with longer target interval. Target interval effects were found in numerous brain structures, including bilateral thalamus, bilateral anterior insula, dorsal anterior cingulate, supplementary motor area, dorsolateral prefrontal cortex, bilateral middle temporal gyri, bilateral pre- and postcentral gyri (somatosensory cortex), bilateral inferior and superior parietal lobules, parietal occipital junction, superior/middle and inferior frontal gyrus, precuneus, and bilateral cerebellum. The legend shows *z*-score value associated with the color map. The statistical parametric map has a threshold of $z > 2.6$; $p < 0.05$ (cluster-corrected). For a full list of activated regions, see **Table A2**.

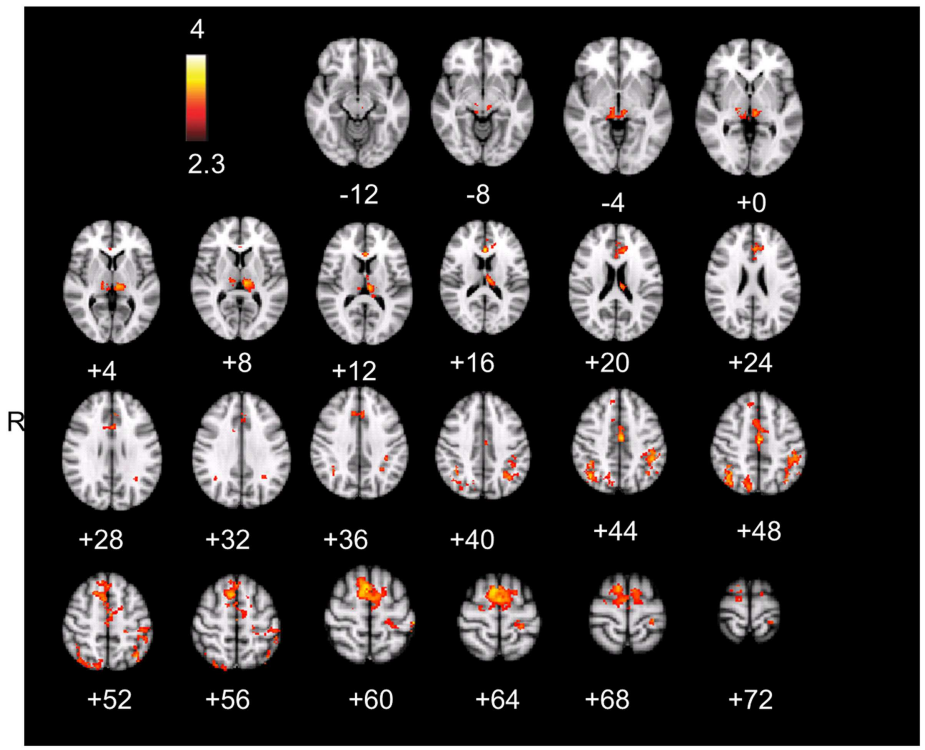


FIGURE A3 | Statistical parametrical map of hemodynamic response varying in each individual with aware errors and with the interval effect on target detection. Red and yellow voxels represent clusters of significant BOLD signal which passed the thresholding in the target interval contrast (ITI3–ITI1) and also survived thresholding in the awareness contrast (aware versus unaware errors). Four major brain areas were involved in both contrasts: bilateral thalamus, supplementary motor area, rostral cingulate, and in bilateral parietal lobule. Furthermore, overlapping activations were found in the precuneus and lateral occipital gyrus.

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 13 August 2012
doi: 10.3389/fnhum.2012.00240

Error awareness as evidence accumulation: effects of speed-accuracy trade-off on error signaling

***Marco Steinhauser*[1,2]* *and Nick Yeung*[3]**

[1] *Department of Psychology, Catholic University of Eichstätt-Ingolstadt, Eichstätt, Germany*
[2] *Department of Psychology, University of Konstanz, Konstanz, Germany*
[3] *Department of Experimental Psychology, University of Oxford, Oxford, UK*

Edited by:
Tilmann A. Klein, Max Planck Institute for Human Cognitive and Brain Sciences, Germany

Reviewed by:
Shani Shalgi, The Hebrew University of Jerusalem, Israel
Peter Murphy, Trinity College Dublin, Ireland

***Correspondence:**
Marco Steinhauser, Department of Psychology, Catholic University of Eichstätt-Ingolstadt, 85072 Eichstätt, Germany.
e-mail: marco.steinhauser@ku.de

Errors in choice tasks have been shown to elicit a cascade of characteristic components in the human event-related potential (ERPs)—the error-related negativity (Ne/ERN) and the error positivity (Pe). Despite the large number of studies concerned with these components, it is still unclear how they relate to error awareness as measured by overt error signaling responses. In the present study, we considered error awareness as a decision process in which evidence for an error is accumulated until a decision criterion is reached, and hypothesized that the Pe is a correlate of the accumulated decision evidence. To test the prediction that the amplitude of the Pe varies as a function of the strength and latency of the accumulated evidence for an error, we manipulated the speed-accuracy trade-off (SAT) in a brightness discrimination task while participants signaled the occurrence of errors. Based on a previous modeling study, we predicted that lower speed pressure should be associated with weaker evidence for an error and, thus, with smaller Pe amplitudes. As predicted, average Pe amplitude was decreased and error signaling was impaired in a low speed pressure condition compared to a high speed pressure condition. In further analyses, we derived single-trial Pe amplitudes using a logistic regression approach. Single-trial amplitudes robustly predicted the occurrence of signaling responses on a trial-by-trial basis. These results confirm the predictions of the evidence accumulation account, supporting the notion that the Pe reflects accumulated evidence for an error and that this evidence drives the emergence of error awareness.

Keywords: error awareness, performance monitoring, event-related potentials, single-trial analysis, error-related negativity, error positivity

Continuous monitoring of action outcomes is crucial for achieving optimal performance. Evidence for a performance monitoring system involved in error detection has been provided by studies examining event-related potentials (ERPs). In these studies, errors in simple choice tasks have been shown to elicit a negative deflection at fronto-central electrodes called the error negativity (Ne, Falkenstein et al., 1990) or error-related negativity (ERN, Gehring et al., 1993), that is followed by a positive deflection at posterior electrodes called the error positivity (Pe, Falkenstein et al., 1990). Whereas early theories suggested that the Ne/ERN directly reflects error detection (Falkenstein et al., 1990; Gehring et al., 1993), it has recently been proposed that the Ne/ERN is related to other aspects of error processing like response conflict (Yeung et al., 2004) or reinforcement learning (Holroyd and Coles, 2002). In contrast, the Pe has been suggested to be a correlate of conscious error processing or error awareness (e.g., Falkenstein et al., 2000). The goal of the present study was to contribute to a deeper understanding of how error awareness is achieved by investigating the relationship between the Pe and behavioral measures of error awareness.

In recent years, the neural correlates of error awareness have been investigated in a number of studies (for an overview, see Ullsperger et al., 2010). A frequently-used method for measuring error awareness is the so-called error signaling paradigm, initially introduced by Rabbitt and colleagues (Rabbitt, 1968, 2002). In this paradigm, participants perform a speeded choice task (the primary task). After each response, they have to press a signaling key whenever they think that they have made an error. Using this paradigm, several studies have investigated the relation between signaling responses and the amplitude of error-related ERP components. Whereas the majority of studies have reported an increased Pe for signaled errors relative to unsignaled errors (Nieuwenhuis et al., 2001; Endrass et al., 2005, 2007; Overbeek et al., 2005; O'Connell et al., 2007; Shalgi et al., 2009; Steinhauser and Yeung, 2010; Dhar et al., 2011; Hewig et al., 2011; Hughes and Yeung, 2011; Wessel et al., 2011; Murphy et al., 2012), only a few studies found such a result for the Ne/ERN (e.g., Maier et al., 2008; Steinhauser and Yeung, 2010; Wessel et al., 2011).

ERROR AWARENESS AS EVIDENCE ACCUMULATION

Although these findings suggest a relationship between the Pe and error awareness, they are less informative regarding the specific role of the Pe in the emergence of awareness. To address this question, we recently proposed that error awareness can be conceptualized as a decision process, in which the available evidence that an error has occurred is accumulated until a decision criterion is reached (Steinhauser and Yeung, 2010). Within this framework, we asked whether error-related brain activity reflects

the accumulated evidence that an error has occurred, or the output of this decision. By varying the decision criterion of error signaling, we were able to test specific predictions associated with each hypothesis. We found that although a higher decision criterion led to fewer signaled errors, it was not associated with a reduced Pe amplitude. This finding implies that average Pe amplitude does not reflect the number of signaled errors, and thus, the output of the decision process. We further found that a higher decision criterion was associated with a larger Pe amplitude if signaled errors were considered. This result reflects the fact that with a high criterion, more evidence for an error is required to exceed this criterion, which is consistent with the assumption that the Pe reflects the accumulated evidence that an error has occurred. Further support for this conclusion was provided by single-trial analyses. Using a logistic regression approach (Parra et al., 2002, 2005), we derived a single-trial measure of the Pe amplitude. As predicted by our evidence accumulation account, this "error signal" could be used to robustly predict whether or not an error would be followed by an error signaling response. Taken together, these results suggest that the Pe does not reflect whether an error was consciously detected or not but rather reflects the accumulation of evidence for an error that precedes the emergence of awareness. Whether a given amount of evidence (i.e., a given Pe amplitude) on a trial leads to error awareness depends on the decision criterion.

COMPUTATIONAL ACCOUNTS OF ERROR DETECTION

Whereas our previous study provides a framework for explaining the relation between the Pe and error awareness, it did not specify the process that delivers the internal evidence for an error, nor did it make assumptions about the nature of this evidence. Potential answers to these questions have been provided by theories of error detection in decision-making (for an overview, see Yeung and Summerfield, 2012). In recent years, two accounts have been proposed which themselves are based on evidence accumulation models: the response monitoring account (Steinhauser et al., 2008; see also Rabbitt and Vyas, 1981) and the conflict monitoring account (Yeung et al., 2004). These accounts share the assumption that response selection in choice tasks occurs when evidence for a response exceeds a response criterion. A crucial feature of evidence accumulation models is their strong self-correction tendency. After an error has occurred due to noise in the accumulation process, continued evaluation of the stimulus usually ensures that accumulated evidence for the correct response eventually exceeds that for the incorrect response. The two accounts mainly differ with respect to which aspect of self correction provides the diagnostic feature that underlies error detection:

(a) The response monitoring account (Steinhauser et al., 2008) assumes that performance monitoring registers that a second response (i.e., an internal correction response) exceeds the primary task's response criterion. However, when Steinhauser et al. (2008) fitted a model of this account to empirical data, it turned out that only about 60% of trials with an internal correction response also led to a signaling response. This suggests that an internal correction response does not directly trigger error awareness (which implies that the response criterion of the primary task does not correspond to the decision criterion associated with error awareness). It rather provides the internal evidence for an error, which forms the basis of the error decision, and which could lead to error awareness or not[1].

(b) The conflict monitoring account (Yeung et al., 2004) assumes that performance monitoring registers response conflict which occurs when strong evidence is accumulated for multiple responses—a condition that necessarily accompanies self correction. This response conflict is accumulated until it reaches another criterion, which then leads to error awareness. Accordingly, this account assumes that response conflict rather than an internal correction response provides the internal evidence for an error.

In a simulation study, Steinhauser et al. (2008) investigated whether these two accounts can predict the latencies and frequencies of error signaling responses in an experiment in which the speed-accuracy trade-off (SAT) of the primary task was manipulated. To derive predictions, response monitoring and conflict monitoring were implemented in a connectionist model. Following standard theories of SAT (for an overview, see Bogacz et al., 2010), the effects of speed pressure were simulated by varying the primary task's response criterion. For such a case, one might expect that slower responding is beneficial for performance monitoring, for instance, because it leads to a better representation of the correct response (e.g., Falkenstein et al., 2000). In contrast to this intuition, the simulations revealed that both accounts predict the opposite: with an increased response criterion and, thus, slow responding, both accounts predicted that fewer errors were signaled and that the latency of error signaling was increased. The analysis of simulation data revealed that, for both accounts, this pattern was due to the fact that evidence for an error was weaker: response monitoring predicted that an increased primary task's response criterion reduces the probability and prolongs the time until an internal correction response exceeded this criterion. Similarly, conflict monitoring predicted that an increased primary task's response criterion reduces and delays response conflict after an error. The latter result obtains because a larger response criterion implies that, at the time of the error response, there is a larger difference between the accumulated evidence for the incorrect response alternative and that for the correct response alternative. This impairs the emergence of response conflict after the error, because with this larger initial difference, the self-correction tendency of the primary task's response selection process requires more time until enough evidence is accumulated for the correct response to cause a response

[1] Steinhauser et al. (2008) discussed the alternative idea that participants simply forgot to give a signaling response on some trials (comparable to a goal neglect; De Jong et al., 1999). However, further evidence that the decision criterion associated with error awareness does not correspond to the response criterion associated with primary task comes from Steinhauser and Yeung (2010). In this study, a manipulation of the decision criterion did not lead to a significant shift of speed-accuracy trade-off in the primary task. This suggests that, if the response monitoring account is valid, one has to assume an additional decision stage that leads to error awareness.

conflict with the already accumulated evidence for the incorrect response.

The experimental data by Steinhauser et al. (2008) confirmed these predictions by showing that low speed pressure, and thus a high response criterion, led to fewer signaled errors and delayed signaling responses (for a similar result, see Shalgi et al., 2007). Because the quantitative fit of the response monitoring model was much better than that of the conflict monitoring model, it was concluded that, at least in this experiment, error signaling was driven by response monitoring. Most importantly for the present study, however, this finding demonstrates that response monitoring and conflict monitoring not only provide specific assumptions about the nature of the internal evidence for an error, they also make specific predictions how this evidence is influenced by experimental variables like SAT.

THE PRESENT STUDY

In the present study, we used the model predictions of Steinhauser et al. (2008) to test a crucial prediction of our evidence accumulation account of error awareness. Whereas our previous study (Steinhauser and Yeung, 2010) manipulated the decision process itself, we now manipulated the evidence feeding into this decision, and asked whether the amplitude of the Pe varies as a function of the strength and latency of the evidence. To achieve this, we manipulated the SAT of a primary task and investigated its influence on error signaling and the Pe. Following the simulation results of Steinhauser et al. (2008), we predicted that low speed pressure should be associated with weaker evidence for an error. As a consequence, if the evidence accumulation account is valid and the Pe reflects the evidence for an error, then low speed pressure should also imply a reduced Pe amplitude[2].

Interestingly, previous studies investigating the effects of SAT on error-related brain activity have typically found the opposite result: Ne/ERN and Pe amplitudes in these studies were increased when accuracy was prioritized over speed (e.g., Gehring et al., 1993; Arbel and Donchin, 2009). However, these studies used paradigms in which SAT shifts were associated with changes in selective attention (for a discussion, see Yeung et al., 2004), and in error significance. In the present study, we manipulated the SAT in a brightness discrimination task in which no selective attention was necessary because no distractor stimuli were used (Steinhauser and Yeung, 2010). Moreover, SAT was manipulated by means of a speed pressure instruction without emphasizing accuracy, and thus, without affecting the subjective significance of an error.

Speed pressure was varied across two conditions, a low speed pressure (lowSP) condition and a high speed pressure (highSP) condition. According to the model predictions of Steinhauser et al. (2008), lower speed pressure should result in weaker evidence for the occurrence of an error. As discussed above, this change is not a direct consequence of the reduced response speed, but rather reflects the increased response criterion in the primary task. If evidence is weaker in the lowSP condition, signaling responses should be less frequent, and the latency of these signaling responses should be increased. Moreover, decreased accumulated evidence in the lowSP condition should be reflected in a smaller Pe. Similar to Steinhauser and Yeung (2010), we tested these predictions for the Pe for all error trials as well as for signaled error trials only. If we found similar effects for all errors and for signaled errors, this would show that changes in Pe amplitudes are actually due to changes in the strength of accumulated evidence rather than changes in the number of signaled errors.

METHOD

PARTICIPANTS

Eighteen right-handed participants (12 female) between 19 and 24 years of age (mean 21.1) with normal or corrected-to-normal vision participated in the study. Participants were recruited at the University of Konstanz for course credit or a payment of 6 Euro per hour, and were paid an additional performance-dependent bonus.

TASK AND PROCEDURE

We used the paradigm introduced by Steinhauser and Yeung (2010), in which participants first performed a brightness discrimination task and then were prompted to make a signaling response when they thought they had made an error. All stimuli were presented on a screen with a resolution of 1080 by 1024 pixels and at a viewing distance of 60 cm. The stimuli in the primary task consisted of two boxes presented on a black background above and below a white fixation cross. Each box consisted of a 64-by-64 array of randomly arranged white and black pixels, with new arrays generated on each trial. Discrimination difficulty depended on the relative proportions of white and black pixels in the two boxes. In contrast to Steinhauser and Yeung (2010), the difficulty level was set to a constant value throughout the experiment, with 55% white pixels in the brighter box compared with 45% in the darker box.

Figure 1 depicts a sample trial. First, a white fixation cross was centrally presented for 500 ms. Then, the stimulus of the primary task appeared for 160 ms, followed by a blank screen. The primary task response was provided by pressing one of two keys on a standard keyboard: the "T" key with the left index finger when the upper box was brighter and the "G" key with the right index finger when the lower box was brighter. 500 ms after the response, the word "error?" was centrally presented for 1000 ms. During that time, participants were instructed to press the space bar with their right thumb if they thought that they had committed an error in the primary task. Then another blank screen appeared for 500 ms, followed by a feedback screen presented for 1000 ms.

The feedback screen indicated the accuracy of both the primary task response and the error signaling response. If the primary task response was correct, and was not followed by an error signaling response, the feedback indicated "yes, correct" in green (correct rejection). If the primary task response was correct, but was followed by an erroneous error signal, the feedback indicated "no, correct" in red (false alarm). If an incorrect primary task response was followed by an error signaling response,

[2]This reasoning relies on the assumption that evidence accumulation does not stop when the decision criterion is reached and, thus, that the amount of accumulated evidence can differ across two conditions with the same decision criterion, even when considering only those trials on which the criterion is exceeded.

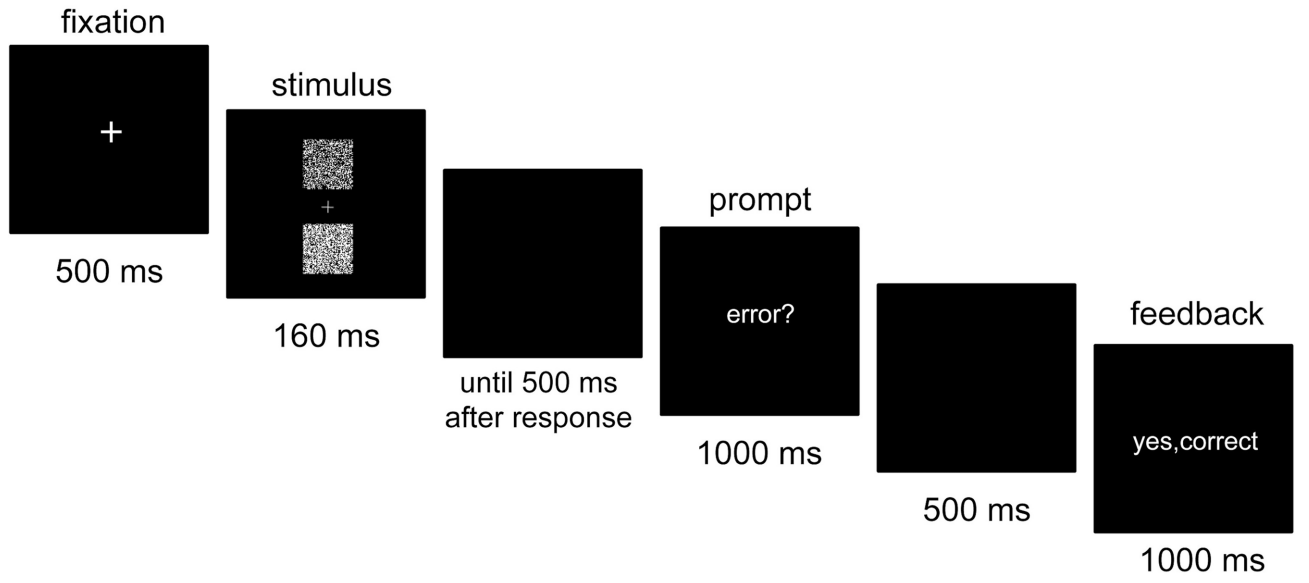


FIGURE 1 | Sequence of stimulus events in a trial. Participants were first required to indicate which of two boxes in the stimulus was brighter. Following the error prompt, they pressed a signaling key if they judged that their primary task response was an error.

the feedback indicated "yes, error" in green (hit). Finally, if an incorrect primary task response remained unsignaled, the feedback indicated "no, error" in red (miss). In experimental blocks, the feedback screen additionally indicated the amount of win or loss (e.g., "+2" or "−2") in this trial.

The experiment consisted of three parts: a practice part, a low speed pressure part, and a high speed pressure part. The practice part consisted of five blocks: first, three blocks of 30 trials were conducted in which only the primary task was practiced and no feedback was provided after each trial. Participants were instructed to respond as quickly and as accurately as possible. After each block (in this and the following parts), feedback about mean RT and error rate was provided. If error rate in these blocks fell below 20%, participants were instructed to increase response speed. Then, two further practice blocks were conducted in which the error signaling procedure was introduced. In these blocks, trial feedback as described above (but without indicating wins and losses) was presented. If the error rate in the final practice block fell below 20%, another practice block was conducted and participants were instructed to increase response speed. This was repeated until the required error rate was achieved. Note that although we applied an accuracy criterion in this part, only response times but never error rates were mentioned during instructions in this and the following parts. We did this to ensure that instructions did not influence the subjective significance of errors.

After the practice part, half of the participants continued with the low speed pressure part and then with the high speed pressure part. This order was reversed in the other half of the participants. Low and high speed pressure was induced only by means of instruction. Participants were instructed not to exceed an individually determined criterion RT. If the mean RT during a block was larger than this criterion RT, participants were instructed to increase response speed. In the low speed pressure blocks, criterion RT was the mean RT from the last practice block plus 50 ms. In the high speed pressure block, criterion RT was the mean RT from the last practice block minus 50 ms. Each part started with two practice blocks of 30 trials, in which participants could adapt to the instructed speed pressure. These practice blocks were followed by four experimental blocks of 60 trials each, resulting in 240 experimental trials in each speed pressure condition. In experimental trials, participants earned money for correct error signaling. They won 2 points each time they signaled on error trials (hits) or withheld from signaling on correct trials (correct rejections). They lost 2 points each time they signaled on correct trials (false alarms) or failed to signal after errors (misses). At the end of the experiment, points were converted into a monetary reward (1 point = 1 Eurocent). In the present study, this reward scheme served no specific purpose beyond encouraging accurate error signaling, but this feature makes the design comparable to our previous study (Steinhauser and Yeung, 2010).

DATA ACQUISITION

The electroencephalogram (EEG) was recorded using a BIOSEMI Active-Two system (BioSemi, Amsterdam, The Netherlands) with 64 Ag-AgCl electrodes from channels Fp1, AF7, AF3, F1, F3, F5, F7, FT7, FC5, FC3, FC1, C1, C3, C5, T7, TP7, CP5, CP3, CP1, P1, P3, P5, P7, P9, PO7, PO3, O1, Iz, Oz, POz, Pz, CPz, Fpz, Fp2, AF8, AF4, AFz, Fz, F2, F4, F6, F8, FT8, FC6, FC4, FC2, FCz, Cz, C2, C4, C6, T8, TP8, CP6, CP4, CP2, P2, P4, P6, P8, P10, PO8, PO4, O2 as well as the left and right mastoid. The Common Mode Sense (CMS) and Driven Right Leg (DRL) electrodes were used as reference and ground electrodes. Vertical and horizontal electrooculogram (EOG) was recorded from electrodes above and below the right eye and on the outer canthi of both eyes. All

electrodes were off-line re-referenced to averaged mastoids. EEG and EOG data were continuously recorded at a sampling rate of 1024 Hz, and were re-sampled to 512 Hz offline.

DATA ANALYSIS

For analysis of primary task performance and ERP data, trials from each condition were categorized as correct responses and errors. For analysis of error signaling performance, trials from each condition were categorized as correct rejections, false alarms, misses, and hits. The absolute frequencies were used to calculate the hit rate, *H* (= proportion of hits among all errors), and the false alarm rate, *FA* (= proportion of false alarms among all correct trials), for both conditions. We then estimated two parameters from Signal Detection Theory (Green and Swets, 1966; Macmillan and Creelman, 1991): the detection criterion *c*, and the sensitivity *d'*. Signaling latencies were calculated as the difference between the primary task response and the signaling response. In this way, occasional signaling responses that occurred prior to the signaling prompt were assigned a positive latency[3].

EEG data were analyzed using EEGLAB v6.01 (Delorme and Makeig, 2004) and custom routines written in MatLab 7.0.4 (The Mathworks, Natick, MA). The data were band-pass filtered excluding activity below 1 Hz and above 30 Hz (waveforms in figures were additionally filtered with a 15 Hz low-pass filter). Epochs were extracted ranging from 500 ms before and 1000 ms after the response. Artifacts were removed using standard routines implemented in EEGLAB v6.01: first, large artifacts were identified by computing the joint probability of each epoch and excluding epochs that deviated more than five standard deviations from the distribution mean. Second, ocular artifacts were corrected by an eye movement correction procedure (Automatic Artifact Removal Version 1.3, http://kasku.org/aar/) based on a linear regression approach (Gratton et al., 1983). Baseline activity was removed by subtracting the average voltage in an interval between 400 ms and 100 ms before the response. This baseline was chosen because it precedes the onset of the Ne/ERN.

After artifact removal, the resulting waveforms included an average of 153 correct trials (range: 113–204) and 75 error trials (range: 26–113) in the highSP condition and 176 correct trials (range: 119–225) and 48 error trials (range: 9–111)[4] in the lowSP condition. If only signaled errors were considered, there were an average of 139 correct trials (range: 89–202) and 62 error trials (range: 21–93) in the highSP condition and 163 correct trials (range: 106–224) and 37 error trials (range: 8–94) in the lowSP condition. Error-related brain activity was quantified by computing the mean amplitude of the waveform for errors in a time interval that captures the main portion of the component of interest. For the Ne/ERN, an interval from –50 to 50 ms relative to the response was used. For the Pe, an interval from 150 to 400 ms after the response was used. All components were quantified for each channel. However, statistical analysis was applied only to data from channel FCz for the Ne/ERN (for which the Ne/ERN is typically largest) and from channel POz for the Pe. The latter was chosen because the error signal found by Steinhauser and Yeung (2010) was maximal at this channel, a finding that was replicated in the present study.

Because we found that ERP differences between conditions partially reflected RT differences (due to differential influence of stimulus-locked components on response-locked data), analyses were also applied to a subset of RT matched trials. To achieve RT matching, we first identified the condition with the fewest trials (i.e., errors/lowSP) and then matched all other conditions (errors/highSP, corrects/lowSP, corrects/highSP) to this condition using the following algorithm: First, a trial from the error/lowSP condition was randomly drawn (without replacement). Second, from each other condition, the trial providing the closest match to the RT of the drawn trial was selected (without replacement) and assigned to the RT-matched sample. These steps were repeated until all trials from the error/lowSP condition were drawn. Note that only artifact-free trials were included. Therefore, mean RT of the error/lowSP condition deviates slightly from the value obtained in the initial analysis of behavioral data.

In addition, we aimed to replicate the findings by Steinhauser and Yeung (2010) that Pe amplitude predicts error signaling on a trial-by-trial level. To achieve this, a single-trial analysis was conducted using the linear integration method introduced by Parra et al. (2002) to measure error-related EEG activity with improved signal-to-noise ratio. The rationale of this method is to extract a specific spatial component of the ERP by constructing a classifier that maximally discriminates between two conditions differing in this component. Specifically, with *x(t)* being the vector of electrode activity at time *t*, we used logistic regression to compute a spatial weighting coefficient *v* such that the component

$$y(t) = v^T x(t)$$

is maximally discriminating between two different conditions. In the present case, we used this method to discriminate between error and correct-response trials in order to estimate error-related EEG activity on individual trials (independent of speed pressure condition). As input, we used *T* samples from each of the *N* baseline-corrected ERP epochs, resulting in a training set of size *NT*. After finding the optimal *v*, we estimated the error signal, $\bar{y}_k$, on each trial *k* by averaging across the *T* samples from each trial. This value ranges between 0 and 1, with higher values indicating a higher probability that the trial was an error.

To visualize the spatial distribution of weights of the discriminating component, we computed the coupling coefficient vector

$$a = \frac{Xy}{y^T y},$$

with time *t* being a dimension of the matrix *X* and the vector *y*. Coupling coefficients represent the activity at each electrode site that correlates with the discriminating component, and thus can be thought of as the "sensor projection" of that component (Parra et al., 2002, 2005).

[3]Because the onset of the signaling prompt was fully predictable, participants sometimes initiated a signaling response that preceded the prompt. 9.6% of signaling responses were too early and preceded the prompt by an average time of 72 ms. We did not exclude these trials because these responses occurred still out of the time range of the Pe and thus did not contaminate our data.

[4]The results did not change when participants with low trial numbers were excluded.

The analysis was applied to the same time range (250–350 ms after the response) as in Steinhauser and Yeung (2010). First, classifier sensitivity was quantified in terms of *Az*-score, which corresponds to the area under the Receiver Operating Characteristic curve, with 0.5 indicating chance-level classification and 1 indicating perfect discrimination. *Az*-scores were computed for each window using split-half cross-validation, i.e., the classifier was trained on half of the trials and was then used to predict the category (correct or error) on the remaining trials. This procedure was repeated for each half of 10 random splits, and the average of these 20 values was taken as the overall sensitivity for a specific window and participant. To test whether sensitivity significantly exceeded chance level, a permutation test was applied (e.g., Philiastides et al., 2010; Steinhauser and Yeung, 2010). For each participant, a test distribution under the Null hypothesis was generated by recomputing *Az*-scores with random assignment of the correct/error categories. This procedure was repeated 100 times for each of the 20 subsets of trials from which each *Az*-score was computed. The resulting 2000 values represented the test distribution, and were used to determine critical *Az*-values associated with significance levels of 0.05 and 0.01. Overall critical *Az*-values were computed by averaging across participants.

Following Steinhauser and Yeung (2010), we used the error signal y_k as a neural correlate of the accumulated evidence that an error has occurred, and investigated whether this error signal can be used to predict error signaling on a trial-by-trial basis. To this end, we first calculated the mean error signal separately for each trial by averaging across values from the 20 split-half samples. Prediction of the occurrence of a signaling response was achieved using a logit regression analysis with a binary dependent variable (signaled error vs. unsignaled error) and a continuous independent variable (mean error signal). The resulting beta values were analyzed using *t*-tests and repeated measurement ANOVAs.

RESULTS

BEHAVIORAL DATA

Behavioral data are presented in **Table 1**. We first analyzed primary task performance to examine whether our manipulation of SAT was successful. As expected, the lowSP condition was associated with decreased error rates, $F_{(1,\ 17)} = 21.0$, $p < 0.001$, increased correct RTs, $F_{(1,\ 17)} = 14.0$, $p < 0.01$, and increased error RTs, $F_{(1,\ 17)} = 8.75$, $p < 0.01$, indicating that the speed pressure manipulation led to a shift in SAT.

As predicted, this SAT shift in the primary task also influenced error signaling. Signaling RT was increased in the lowSP condition, $F_{(1,\ 17)} = 5.07$, $p < 0.05$. Moreover, the frequency of signaling responses was numerically reduced. Although this effect was not significant for the false alarm rates and was only marginally significant for the hit rates, $F_{(1,\ 17)} = 3.71$, $p = 0.07$, the estimated detection criterion—a measure that combines the two rates—was significantly increased in the lowSP condition[5], $F_{(1,\ 17)} = 4.53$, $p < 0.05$. In contrast, detection sensitivity *d'* did not differ reliably across the two conditions ($F < 1$).

Table 1 | Behavioral performance.

	HighSP		LowSP			
	Mean	**SE**	**Mean**	**SE**	$F_{(1, 17)}$	p
PRIMARY TASK PERFORMANCE						
Error rate (%)	33.1	0.03	21.5	0.03	21.0	<0.001
RT correct (ms)	338	23	412	22	14.0	<0.01
RT error (ms)	320	25	410	43	8.75	<0.01
ERROR SIGNALING PERFORMANCE						
Latency (ms)	703	33	756	29	5.07	<0.05
Hit rate (%)	81.7	3.6	76.6	0.05	3.71	0.07
False alarm rate (%)	1.8	0.3	1.5	0.03	0.79	0.39
Criterion c	0.56	0.09	0.73	0.11	4.53	<0.05
Sensitivity d'	3.23	0.14	3.15	0.19	0.78	0.39

Primary task error rates and response times (RTs), error signaling rates and latency, and estimated signal detection parameters for the two speed pressure conditions.

Note: SE, standard error of the mean; lowSP, low speed pressure; highSP, high speed pressure.

EVENT-RELATED POTENTIALS

The behavioral data followed a similar pattern to the one obtained in Steinhauser et al. (2008): low speed pressure for the primary task led to longer signaling RTs and a lower frequency of signaling responses. In a next step, we examined whether these behavioral effects were reflected in specific changes in the Pe and the Ne/ERN. Based on the simulations of Steinhauser et al. (2008), we predicted that weaker evidence for an error would be evident with low speed pressure. Provided that the accumulated evidence for an error is reflected by the Pe amplitude, we should therefore observe a reduced Pe amplitude in the lowSP condition. Moreover, this effect should be obtained across all error trials as well as for signaled error trials specifically. If such an effect were obtained only if all error trials were included, it could simply reflect the decreased rate of signaled errors in the lowSP condition.

Figure 2 presents waveforms at two characteristic channels, FCz and POz, for all trials (**Figures 2A,C**) and for trials that were correctly signaled (i.e., signaled errors and unsignaled correct trials, **Figures 2B,D**). The waveforms reveal strong differences between speed pressure conditions. At least for correct trials, however, these differences seem to reflect RT differences between these conditions: waveforms for correct trials in the highSP condition are delayed relative to those in the lowSP condition. This might reflect that, due to shorter RTs in the highSP condition, stimulus-locked components occur later relative to the response in this condition (Coles et al., 2001; Maier et al., 2010)—a conclusion receiving support from the observation that this effect disappeared when RT matched data were analysed (see below). At first glance, such an effect does not seem to be responsible for differences between the waveforms for error trials, at least in the time

[5]This detection criterion does not necessarily reflect the decision criterion we assumed for the error signaling process but only represents a parameter of the signal detection analysis. An increased signal detection criterion could also reflect that the underlying signal is decreased for both correct and error signals.

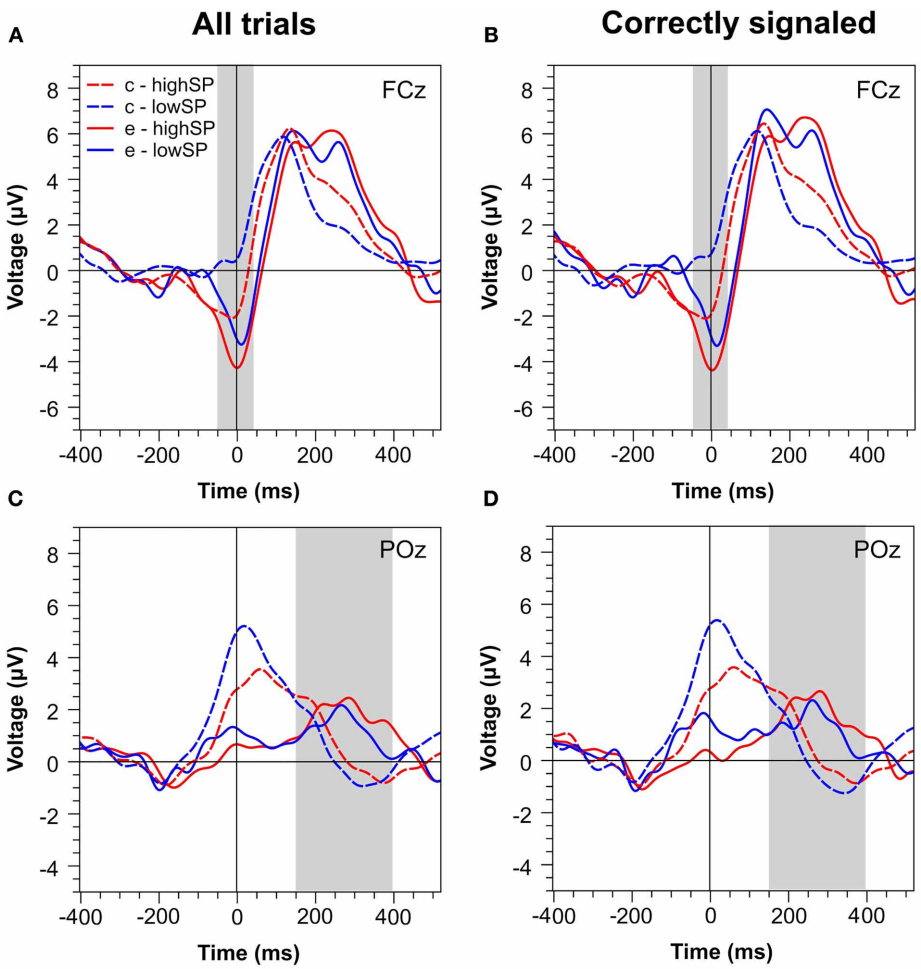


FIGURE 2 | Response-locked ERPs separately for the highSP and lowSP conditions. (A,C) Mean ERP waveforms at electrodes FCz and POz for all errors and correct responses. **(B,D)** Mean ERP waveforms at electrodes FCz and POz for signaled errors and unsignaled correct responses. Zero indicates the time of the response. HighSP = high speed pressure. LowSP = low speed pressure.

range of the Pe. However, to prevent bias of our analysis by the differential contribution of RT effects to correct and error trials, we directly compared error trials between our conditions.

Figure 3 plots the spatial distribution of the difference wave between the lowSP and the highSP condition for error trials in the time range of the Ne/ERN and the Pe. For the Pe, the data reveal differences with a broad central spatial distribution, irrespective of whether all trials or only correctly signaled trials were considered. As predicted, Pe amplitude was decreased for the lowSP condition relative to the highSP condition, and this difference was significant for all trials (1.25 μV vs. 1.81 μV), $F_{(1,\ 17)} = 7.57$, $p < 0.05$, as well as for correctly signaled trials (1.27 μV vs. 1.83 μV), $F_{(1,\ 17)} = 5.45$, $p < 0.05$, at channel POz. For the Ne/ERN, we obtained a difference in the same direction at channel FCz which was marginally significant for all trials (–2.19 μV vs. –3.29 μV), $F_{(1,\ 17)} = 3.19$, $p < 0.10$, as well as for correctly signaled trials (–2.15 μV vs. –3.41 μV), $F_{(1,\ 17)} = 3.40$, $p < 0.10$.

As already mentioned, differences between waveforms in our speed pressure conditions partially reflect RT differences. Although this seems to hold mainly for correct trials, we cannot exclude the possibility that RT differences also influenced the waveforms on error trials. To rule out that our results reflect a confound with between-condition differences in RT, we reanalyzed the data after matching RTs between errors and correct trials in the two speed pressure conditions. Note that although RT matching produces trial sets with similar RTs, these trial sets still differ with respect to whether participants were under low speed pressure or high speed pressure (i.e., whether participants adopted a high or low response criterion), thus leaving our experimental manipulation intact. **Table 2** illustrates latencies calculated for the RT matched trials. Due to the strong RT differences, matching was not perfect. Whereas RT differences between correct trials of the two speed pressure conditions were not significant anymore, $F_{(1,\ 17)} = 1.52$, $p = 0.23$, a marginally significant speed pressure effect of 47 ms remained for error trials, $F_{(1,\ 17)} = 4.02$, $p < 0.10$. Interestingly, equalizing RTs of the primary task also abolished the effects of speed pressure effects on signaling latency, $F < 1$. This finding might indicate that signaling latencies

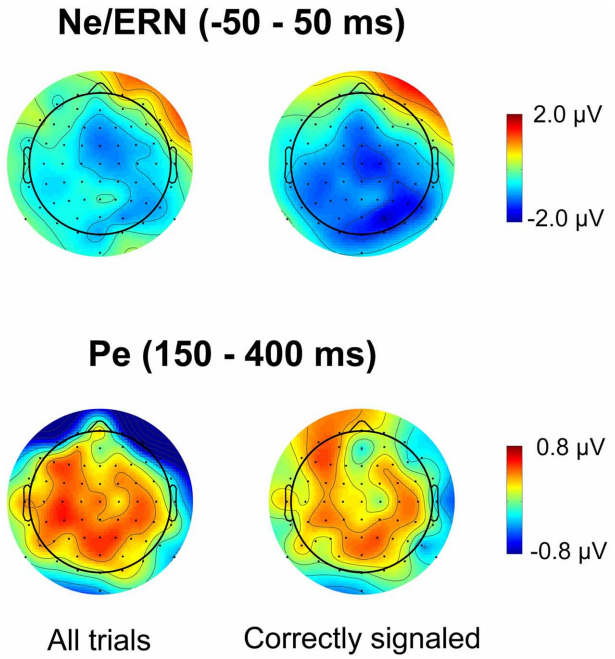


FIGURE 3 | Spatial distribution of ERPs for the difference between errors in the lowSP condition and errors in the highSP condition. *Upper row:* Time period of the Ne/ERN (−50 – 50 ms). *Lower row:* Time period of the Pe (150–400 ms). *Left column:* Data from all errors. *Right column:* Data from signaled errors. HighSP = high speed pressure. LowSP = low speed pressure.

Table 2 | Behavioral performance in matched conditions.

	HighSP		LowSP			
	Mean	SE	Mean	SE	$F_{(1, 17)}$	p
PRIMARY TASK PERFORMANCE						
RT correct. (ms)	375	33	393	41	1.52	0.23
RT error (ms)	346	29	393	42	4.02	0.06
ERROR SIGNALING PERFORMANCE						
Latency (ms)	676	41	702	51	0.18	0.68

Primary task response times (RTs) and error signaling latency after matching RTs for the two speed pressure conditions.

Note: SE, standard error of the mean; lowSP, low speed pressure; highSP, high speed pressure.

and primary task RTs are additionally correlated due to other variables than response criterion. If RT matching eliminated differences in response criterion, this should have eliminated any Pe differences, which was not case, as we will see in the next analysis.

Figures 4 and **5** present waveforms and spatial distributions for the RT-matched data. Although RT matching was imperfect, effects such as the shifted ERP latencies for correct trials disappeared, suggesting that these effects were due to RT differences in the primary task. Crucially, however, amplitude differences in the Pe range of error trials between the speed pressure conditions were even slightly increased after RT matching. Again, the Pe was decreased for the lowSP condition relative to the highSP condition, and this difference was significant for all trials (1.26 μV vs. 2.21 μV), $F_{(1, 17)} = 7.97$, $p < 0.05$, as well as for correctly signaled trials (1.27 μV vs. 2.20 μV), $F_{(1, 17)} = 5.44$, $p < 0.05$, at channel POz. For the Ne/ERN, we now obtained a nonsignificant difference at channel FCz for all trials (−2.19 μV vs. −3.03 μV), $F_{(1, 17)} = 2.03$, $p = 0.17$, as well as for correctly signaled trials (−2.15 μV vs. −3.01 μV), $F_{(1, 17)} = 1.61$, $p = 0.22$.

Taken together, the analyses of response-locked ERPs suggest that less frequent and slower error signaling in the lowSP condition was associated with a reduced Pe amplitude. This reduced amplitude was obtained if all error trials were considered as well as if only correctly signaled error trials were considered, and thus seems not to reflect the decreased rate of signaled trials in the lowSP condition. In contrast, the Ne/ERN did not differ significantly between speed pressure conditions, a result which once again speaks against a direct relation between the Ne/ERN and error signaling. In the following analyses, we apply single-trial analysis to further investigate the relation between the Pe and error signaling.

SINGLE-TRIAL ANALYSIS

To obtain single-trial estimates of the Pe, we trained a classifier to differentiate between errors and correct trials, and used its prediction value as a single-trial measure of the error signal (Parra et al., 2002). The classifier significantly discriminated between correct and error trials ($Az = 0.612$; critical value for $p = 0.05 : 0.576$; critical value for $p = 0.01 : 0.606$). **Figure 6A** illustrates the spatial distribution of the discriminating component. It reveals a posterior distribution of weights with a peak around electrode POz, which replicates the results of Steinhauser and Yeung (2010). In a next step, we extracted the mean error signal to obtain an estimate of the single-trial Pe amplitude. In further analyses, two participants had to be excluded because they had either no signaled ($n = 1$) or no unsignaled artifact-free error trial ($n = 1$) in one of the conditions. The mean error signal for the remaining participants was significantly larger for signaled errors than for unsignaled errors (**Figure 6B**), $F_{(1, 15)} = 7.62$, $p < 0.05$. The logit regression analysis revealed that the error signal significantly predicted the occurrence of error signaling (beta = 4.99), $F_{(1, 15)} = 5.94$, $p < 0.05$. Both results demonstrate that the strength of the error signal predicts whether a signaling response is elicited—a crucial prediction of the evidence accumulation account—and thus replicates the findings of Steinhauser and Yeung (2010).

DISCUSSION

In a recent study (Steinhauser and Yeung, 2010), we proposed that error awareness—as measured by error signaling—can be described as a decision process in which evidence is accumulated until a criterion is reached. We showed that the Pe, a posterior positive ERP wave following errors, reflects the accumulated evidence that an error has occurred (rather than the outcome of such a decision). The goal of the present study was to test a specific

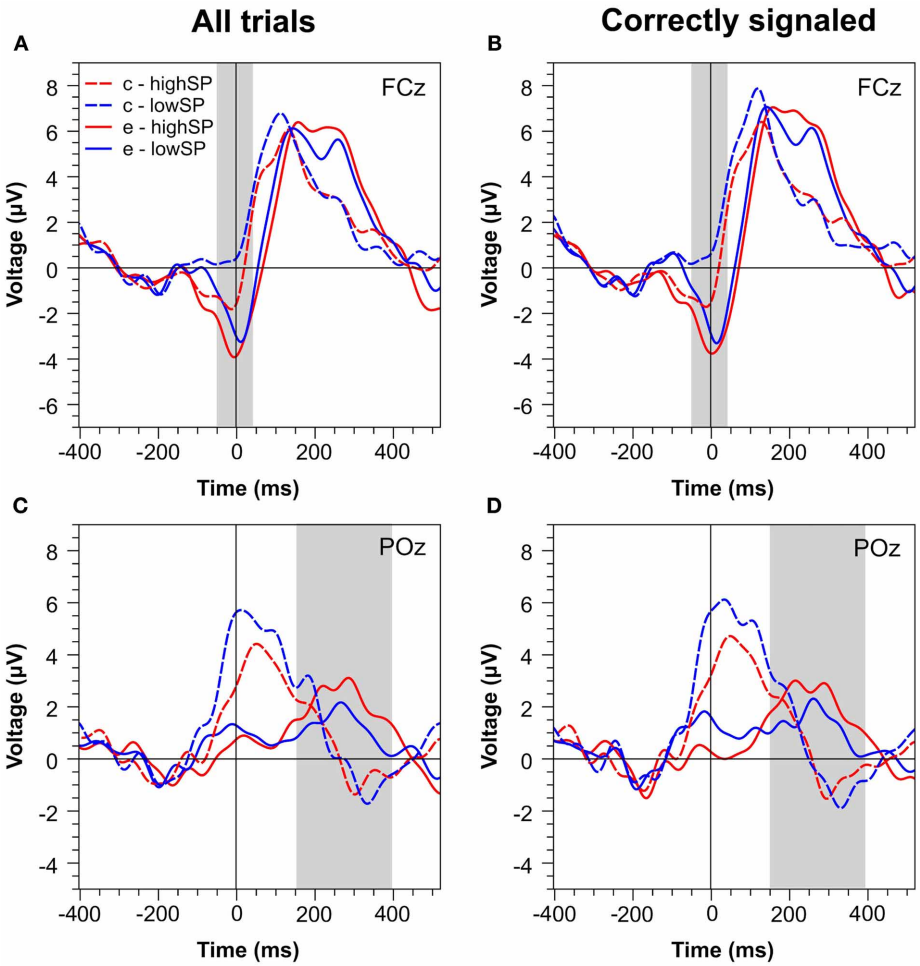


FIGURE 4 | Response-locked ERPs for RT matched data separately for the highSP and lowSP conditions. (A,C) Mean ERP waveforms at electrodes FCz and POz for all errors and correct responses. **(B,D)** Mean ERP waveforms at electrodes FCz and POz for signaled errors and unsignaled correct responses. Zero indicates the time of the response. HighSP = high speed pressure. LowSP = low speed pressure.

prediction of this evidence accumulation account: that the amplitude of the Pe should vary as a function of the accumulated evidence for an error. To this end, we investigated the effects of manipulating the SAT on error signaling and the Pe. We predicted that low speed pressure in the primary task should be associated with delayed and weaker evidence for an error and, thus, with smaller Pe amplitudes. This prediction was derived from a recent modeling study (Steinhauser et al., 2008) showing that two theoretical accounts of error detection—conflict monitoring and response monitoring—predict that error detection should be impaired when speed pressure is low as compared to when speed pressure is high[6].

To manipulate SAT without producing confounding effects of selective attention and error significance, we used a brightness discrimination task with error signaling (Steinhauser and Yeung, 2010) and instructed participants to respond within a short or a long RT limit, thus exerting high speed pressure (highSP condition) or low speed pressure (lowSP condition). Replicating findings of Steinhauser et al. (2008), the lowSP condition was associated not only with increased RTs and decreased error rates in the primary task, but also with longer signaling latencies and decreased signaling frequencies (see also Shalgi et al., 2007). Crucially, these behavioral effects were accompanied by corresponding effects in the average amplitude of the Pe. The lowSP condition exhibited a decreased Pe amplitude relative to the highSP condition. This effect was obtained irrespective of whether all error trials were analyzed or only signaled error trials, which demonstrates that this effect does not reflect the decreased rate of signaled errors in the lowSP condition. Furthermore, this effect was not reduced after matching RTs between conditions, which demonstrates that it is not due to RT differences between the speed pressure conditions. Although RT matching only reduced differences between error RTs from 90 to 47 ms rather than eliminating it, this should have reduced the Pe effect

[6]Although these predictions were derived from simulations of a flanker task, they can be generalized to any speeded choice task. The effects of response criterion on response conflict/internal correction responses should be similar for any choice task with a self-correction tendency, irrespective of whether distractor stimuli are present or not.

if the effect was entirely due to RT differences. In contrast, the same effect of speed pressure on Pe was obtained when RTs were matched. Taken together, these results provide support for a crucial assumption of our evidence accumulation account of error awareness. Steinhauser et al. (2008) predicted that, with low speed pressure, performance monitoring provides less evidence for an error. The present study demonstrates that this reduced evidence is reflected in reduced Pe amplitudes, suggesting a relation between the Pe and the accumulated evidence for an error.

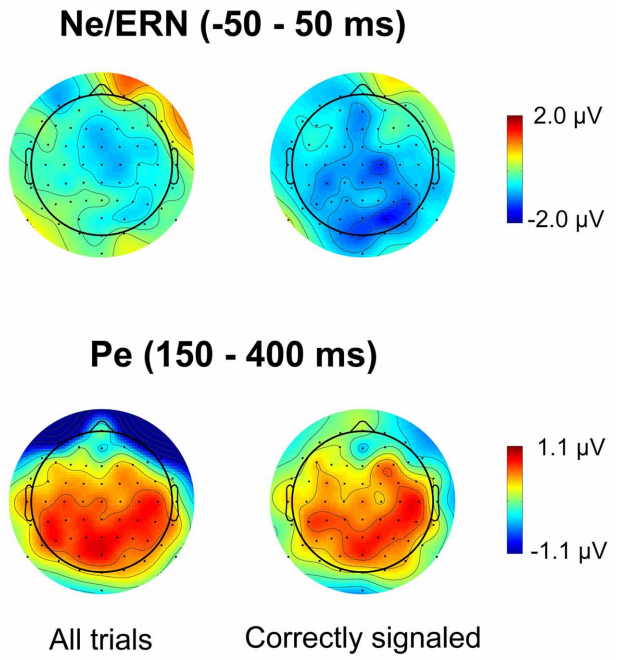


FIGURE 5 | Spatial distribution of ERPs in RT matched data for the difference between errors in the lowSP condition and errors in the highSP condition. *Upper row:* Time period of the Ne/ERN (−50 – 50 ms). *Lower row:* Time period of the Pe (150–400 ms). *Left column:* Data from all errors. *Right column:* Data from signaled errors. HighSP = high speed pressure. LowSP = low speed pressure.

In further analyses, we tested another prediction of the evidence accumulation account by investigating whether the Pe amplitude can be used to predict error signaling on a trial-by-trial basis. As a single-trial measure of error-related brain activity, we used the "error signal," that is, the prediction value of a logistic regression classifier (Parra et al., 2002, 2005) that discriminated between correct and error trials. **Figure 6A** suggests that the classifier is associated with the typical posterior distribution of the Pe. Replicating the results by Steinhauser and Yeung (2010), the error signal extracted in the time range of the Pe was predictive of the error signaling response. The mean error signal was larger for signaled errors than for unsignaled errors, and the error signal on single trials significantly predicted whether an error would be signaled or not.

The results of the single-trial analysis replicate the findings of Steinhauser and Yeung (2010) by showing that the Pe amplitude is a valid predictor of the occurrence of signaling responses. Recently, another study extended these results by showing that the latency of the single trial Pe can also be used to predict the latency of the error signaling response (Murphy et al., 2012). Such a finding is fully in line with the idea that the Pe is related to an evidence accumulation process. In the present study, we did not focus on signaling latencies because, as in Steinhauser and Yeung (2010), we used a prompting procedure that delays error signaling in order to avoid the time range of the Pe becoming contaminated by motor activity. Signaling latencies are less informative under these conditions because this procedure eliminates variance of signaling latencies (although not fully, as indicated by the significant effect of speed pressure on mean latencies). Given that Murphy et al. (2012) used independent component analysis and that their component has a more anterior distribution than that obtained in our studies, future research will have to show whether both components really reflect the same activity.

The combined results from the present study and our previous work (Steinhauser et al., 2008; Steinhauser and Yeung, 2010), suggest that the emergence of error awareness proceeds in at least two stages. First, internal evidence for an error is provided by an implicit performance monitoring mechanism registering errors

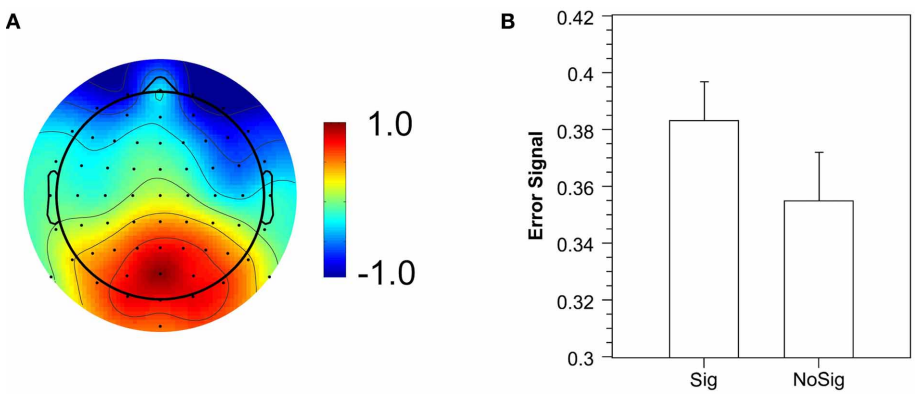


FIGURE 6 | Error signal extracted by single-trial analysis. (A) Spatial distribution of the error signal as illustrated by normalized coupling coefficients. **(B)** Mean error signal for signaled (Sig) and unsignaled (NoSig) errors.

immediately after error commission. This mechanism could be based on conflict monitoring (Yeung et al., 2004) or response monitoring (Rabbitt and Vyas, 1981; Steinhauser et al., 2008), or both. Second, the output of this process either directly generates the evidence reflected in Pe amplitude, or it causes affective responses providing this evidence. The latter is suggested by studies showing that error awareness is correlated with activity related to autonomic responses (e.g., Klein et al., 2007; Wessel et al., 2011; for an overview, see Ullsperger et al., 2010). This evidence then feeds a decision which forms the basis of error awareness and which is observed in the error signaling response.

Although SAT has been a well-known empirical phenomenon for many years (Wickelgren, 1977), it is still not fully understood. Recent evidence suggests that the brain adapts to increased speed pressure by increasing baseline activity in associative areas and the pre-supplementary motor area (pre-SMA), which is computationally equivalent to a decrease in the response criterion (Forstmann et al., 2008; Bogacz et al., 2010). The present study replicates the finding that manipulating SAT of the primary task also affects the frequency and latency of error signaling (Shalgi et al., 2007; Steinhauser et al., 2008), and additionally shows that low speed pressure decreases Pe amplitude. We assumed that this effect is mediated by the effects of SAT on conflict monitoring and/or response monitoring (Steinhauser et al., 2008). Whereas conflict monitoring assumes that an error is detected by registering conflict between an incorrect response and subsequent corrective activity (Yeung et al., 2004), response monitoring assumes that an error is detected by registering that the internal correction response has exceeded the response criterion (Steinhauser et al., 2008). Despite these differences, these two accounts share the prediction that an increased response criterion in the primary task (associated with the lowSP condition) should impair the emergence of internal evidence for an error: An increased response criterion should delay the occurrence of response conflict in the conflict monitoring model, and should delay the internal correction response exceeding this criterion in the response monitoring model. Accordingly, both models can account for the finding that Pe amplitude is reduced in the lowSP condition.

In other studies investigating the effect of SAT on error processing, various alternative accounts have been proposed. Shalgi et al. (2007) explained the effects of SAT on error signaling by assuming that arousal, and thus sustained attention, is reduced under low speed pressure, and that this is the reason why error signaling is also impaired. However, these authors manipulated SAT in a go/no-go task by either exerting speed pressure (speed condition) or by instructing participants to synchronize their response to a late stimulus offset (accuracy condition), and they argued that reduced sustained attention is a direct consequence of the monotonous rhythm induced by responding to stimulus offset (Shalgi et al., 2007, p. 122). In the present paradigm, we used a more traditional SAT manipulation, such that there is no reason why sustained attention should be reduced in the lowSP condition. Reduced sustained attention should have negative effects on both speed and accuracy rather than influencing the SAT. Instead, it is possible that a change of response criterion has contributed to the results of Shalgi et al. (2007).

Several studies have investigated the effects of SAT on error-related brain activity and found the opposite results to the present study; that is, they found that low speed pressure leads to both an increased Ne/ERN and an increased Pe (e.g., Gehring et al., 1993; Arbel and Donchin, 2009). Effects like these have typically been explained by assuming that errors are generally less significant under high speed pressure (Gehring et al., 1993), or that speed pressure impairs the determination of the correct response (Falkenstein et al., 2000). However, these explanations cannot account for the absence of such a finding in the present study. The differences across studies could reflect differences in the primary task and in the method of manipulating SAT. Previous studies have mostly used a flanker task in which selective attention is necessary to respond to the target while ignoring distractors, and SAT was manipulated by emphasizing either speed or accuracy. Yeung et al. (2004) proposed that the SAT effect on the Ne/ERN in the flanker task is due to increased selective attention when accuracy is prioritized over speed (which increases response conflict after errors). This assumption can explain why we did not find such an effect in the present paradigm, in which selective attention is less relevant and instructions focus exclusively on speed rather than accuracy. This interpretation further implies that the present results reflect the pure effect of response criterion on error processing without being contaminated by effects of attention.

The absence of an SAT effect on the Ne/ERN not only contradicts previous explanations of such a finding, it also seems to violate another prediction by Steinhauser et al. (2008). Although Ne/ERN amplitudes were not directly simulated in this study, they found that the conflict monitoring model predicted a reduced level of post-error response conflict when speed pressure was reduced. Given that the conflict monitoring framework (Yeung et al., 2004) postulates a relation between post-error response conflict and the Ne/ERN amplitude, we should have obtained smaller Ne/ERN amplitudes in the lowSP condition as compared to the highSP condition. Indeed our data showed a numerically smaller Ne/ERN amplitude in the lowSP condition which reached marginal significance only when response times were not matched. This could indicate that our manipulation was simply not strong enough to reveal an SAT effect on response conflict, and thus, on the Ne/ERN amplitude. The fact that the same manipulation revealed a significant effect on the Pe could reflect that the accumulated evidence for an error reflected by the Pe is not only provided by conflict monitoring but also (and maybe even stronger) by response monitoring (Steinhauser et al., 2008). If one assumes that the Ne/ERN is more related to conflict monitoring than to response monitoring, this could also explain why single-trial amplitudes of the Pe and the Ne/ERN are only weakly correlated across trials (Steinhauser and Yeung, 2010; Hughes and Yeung, 2011).

ACKNOWLEDGMENTS

We thank Jonas Matuschek for his help with data collection. This work was supported by a grant to Marco Steinhauser from the Deutsche Forschungsgemeinschaft (DFG: STE 1708/1).

REFERENCES

Arbel, Y., and Donchin, E. (2009). Parsing the componential structure of post-error ERPs: a principal component analysis of ERPs following errors. *Psychophysiology* 46, 1179–1189.

Bogacz, R., Wagenmakers, E.-J., Forstmann, B. U., and Nieuwenhuis, S. (2010). The neural basis of the speed-accuracy tradeoff. *Trends Neurosci*. 33, 10–16.

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.

De Jong, R., Berendsen, E., and Cools, R. (1999). Goal neglect and inhibitory limitations: dissociable causes of interference effects in conflict situations. *Acta Psychol.* 101, 379–394.

Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *J. Neurosci. Methods* 134, 9–21.

Dhar, M., Wiersma, 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

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.

Falkenstein, M., Hohnsbein, J., Hoormann, J., and Blanke, L. (1990). "Effects of errors in choice reaction tasks on the ERP under focused and divided attention," in *Psychophysiological Brain Research,* eds C. H. M. Brunia, A. W. K. Gaillard, and A. Kok (Tilburg, Netherlands: Tilburg University Press), 192–195.

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.

Forstmann, B. U., Dutilh, G., Brown, S., Neumann, J., von Cramon, D. Y., Ridderinkhof, K. R., and Wagenmakers, E. J. (2008). Striatum and pre-SMA facilitate decision making under time pressure. *Proc. Natl. Acad. Sci. U.S.A.* 105, 17538–17542.

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

Gratton, G., Coles, M. G., and Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalogr. Clin. Neurophysiol.* 55, 468–484.

Green, D. M., and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. New York, NY: Wiley.

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.

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

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

Macmillan, N. A., and Creelman, C. D. (1991). *Detection Theory: A User's Guide*. Cambridge, UK: Cambridge University Press.

Maier, M., Steinhauser, M., and Hübner, 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.

Maier, M., Steinhauser, M., and Hübner, R. (2010). Effects of response set size on error-related brain activity. *Exp. Brain Res.* 202, 571–581.

Murphy, P. R., Robertson, I. H., Allen, D., Hester, R., and O'Connell, R. G. (2012). An electrophysiological signal that precisely tracks the emergence of error awareness. *Front. Hum. Neurosci.* 6:65. doi: 10.3389/fnhum.2012.00065.

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., 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.

Parra, L., Alvino, C., Tang, A., Pearlmutter, B., Yeung, N., Osman, A., and Sajda, P. (2002). Linear spatial integration for single-trial detection in encephalography. *Neuroimage* 17, 223–230.

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

Philiastides, M. G., Biele, G., Vavatzanidis, N., Kazzer, P., and Heekeren, H. R. (2010). Temporal dynamics of prediction error processing during reward-based decision making. *Neuroimage* 53, 221–232.

Rabbitt, P. (1968). Three kinds of error-signalling responses in a serial choice task. *Q. J. Exp. Psychol.* 20, 179–188.

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

Rabbitt, P., and Vyas, S. (1981). Processing a display even after you make a response to it: how perceptual errors can be corrected. *Q. J. Exp. Psychol. Hum. Exp. Psychol.* 33, 223–239.

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

Shalgi, S., O'Connell, R. G., Deouell, L. Y., and Robertson, I. H. (2007). Absent minded but accurate: delaying responses increases accuracy but decreases error awareness. *Exp. Brain Res.* 182, 119–124.

Steinhauser, M., Maier, M., and Hübner, R. (2008). Modeling behavioral measures of error detection in choice tasks: response monitoring versus conflict monitoring. *J. Exp. Psychol. Hum. Percept. Perform.* 34, 158–176.

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

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.

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.

Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychol.* 41, 67–85.

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, 939–959.

Yeung, N., and Summerfield, C. (2012). Metacognition in human decision-making: confidence and error-monitoring. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 1310–1321.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 02 March 2012; accepted: 26 July 2012; published online: 13 August 2012.

Citation: Steinhauser M and Yeung N (2012) Error awareness as evidence accumulation: effects of speed-accuracy trade-off on error signaling. Front. Hum. Neurosci. ***6****:240. doi: 10.3389/fnhum. 2012.00240*

frontiers in
HUMAN NEUROSCIENCE

REVIEW ARTICLE
published: 17 April 2012
doi: 10.3389/fnhum.2012.00088

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*

Edited by:
Claudia Danielmeier, Radboud University Nijmegen, Netherlands

Reviewed by:
Robert Hester, University of Melbourne, Australia
Egbert Hartstra, Donders Institute for Brain, Cognition and Behaviour, Netherlands

***Correspondence:**
Jan R. Wessel, Psychology Department, University of California, 3133 McGill Hall, 9500 Gilman Drive, La Jolla, San Diego, CA 92093, USA.
e-mail: jwessel@ucsd.edu

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**).

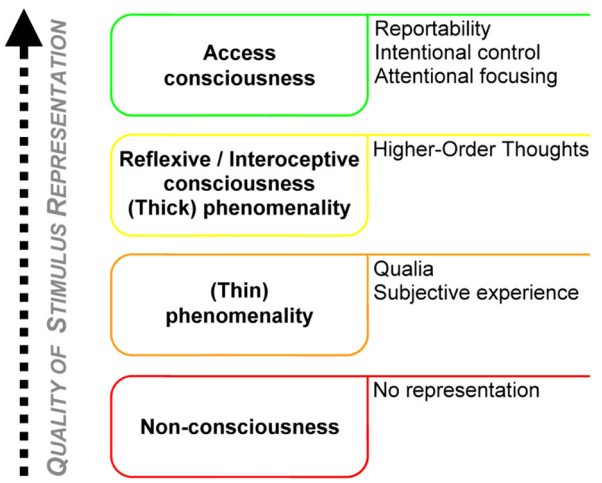


FIGURE 1 | Typology of transitive consciousness, based on different theoretical accounts from the philosophy of mind (see text for further details). Right column outlines defining properties of the different types of consciousness.

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

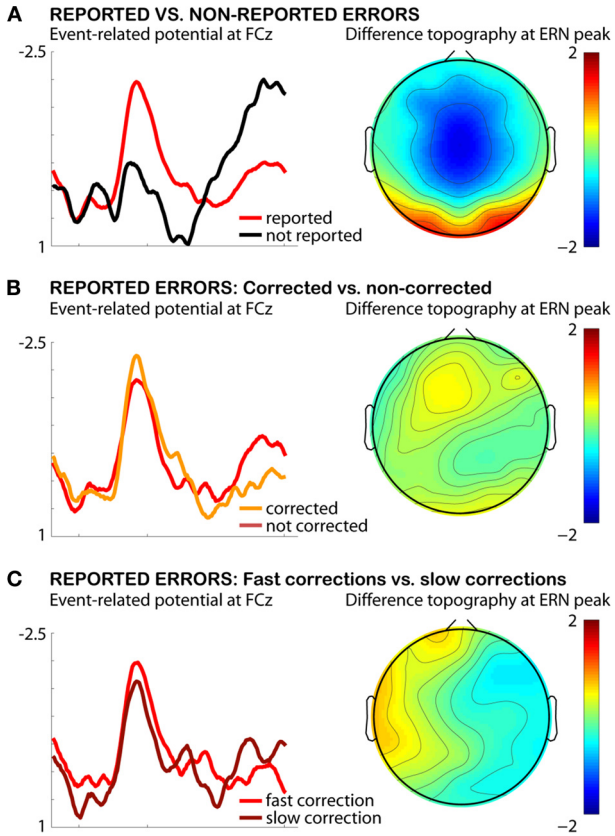


FIGURE 2 | Testing the error-correction hypothesis of ERN amplitude in the AST. Depicted are the combined data from both experiments in Wessel et al. (2011), limited to the 24 subjects that exhibited enough errors to warrant statistical comparison. **(A)** Difference between reported and non-reported 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): $t_{(23)} = 0.6739$, $p > 0.5$.

[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.

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

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

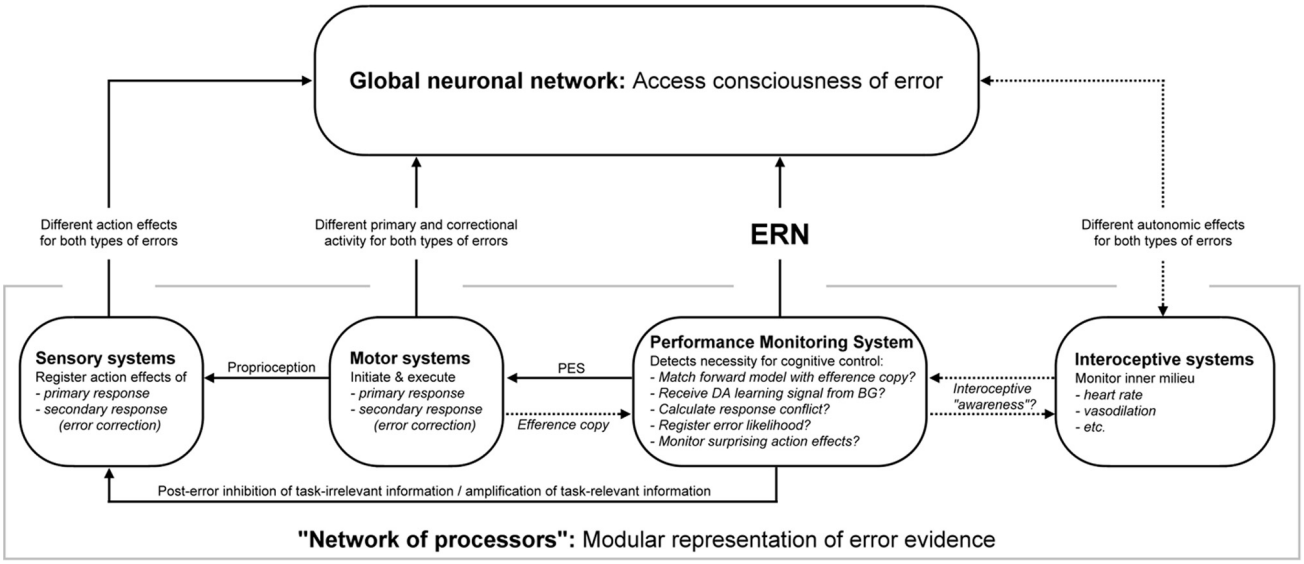


FIGURE 3 | A putative model schematic of emerging error awareness in the human brain, based on the accumulating evidence account of error awareness and the global neuronal workspace model. Information about the accuracy of an action is processed in parallel in different areas that comprise the "network 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.

ACKNOWLEDGMENTS

I would like to acknowledge the work of Dr. Claudia Danielmeier and Dr. Markus Ullsperger on the original study (Wessel et al., 2011), of which I used a subset of data to test the error correction account of the ERN, and for previous discussions about many of the issues presented in this article. This work was partially funded by a grant by the Gertrud Reemtsma Foundation for Brain Research.

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). Maladaptation 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 maladaptive 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.

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 24 October 2011; accepted: 27 March 2012; published online: 17 April 2012.

Citation: Wessel JR (2012) Error awareness and the error-related negativity: evaluating the first decade of evidence. Front. Hum. Neurosci. ***6****:88. doi: 10.3389/fnhum.2012.00088*

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 04 May 2012
doi: 10.3389/fnhum.2012.00124

Is any awareness necessary for an Ne?

***Shani Shalgi*[1]* and *Leon Y. Deouell*[2]**

[1] *Department of Cognitive Science, The Hebrew University of Jerusalem, Jerusalem, Israel*
[2] *Department of Psychology, Interdisciplinary Center for Neural Computation, and Edmond and Lily Safra Center for Brain Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel*

Edited by:
Markus Ullsperger, Max Planck Institute for Neurological Research, Germany

Reviewed by:
Michael Falkenstein, Leibniz Research Centre für Working Environment and Human Factors, Germany
Tanja Endrass, Humboldt University Berlin, Germany

***Correspondence:**
Shani Shalgi, The Department of Cognitive Science, The Hebrew University of Jerusalem, Jerusalem 91905, Israel.
e-mail: shanish@mscc.huji.ac.il

The Error-Related Negativity (Ne or ERN) is a reliable electrophysiological index of error processing, which has been found to be independent of whether a subject is aware of an error or not. A large Ne was equally seen after errors that were consciously detected (Aware errors) and those that were not (Unaware errors), compared to a small negativity for correct responses (CRN). This suggests a dissociation between an automatic, preconscious error processing mechanism and subjective evaluation. A common concern regarding this finding is that subjects could have been somewhat aware of their errors, but did not report them due to lack of confidence. Here we tested this possibility directly using a betting paradigm which allowed us to separate occasions in which the subjects were confident of their response and trials in which they were unsure. In a choice reaction time task, subjects directly judged the accuracy of each response (correct or error) and then bet on this judgment using a high, medium, or low amount of money. The bets were used to determine the level of confidence the subjects had of their response. The average across all subjects regardless of confidence (betting) measure replicated the reported finding of an equal Ne for Aware and Unaware errors which was larger than the CRN. However, when Ne measurement was confined to high confidence (high bet) trials in confident subjects, a prominent Ne was seen only for Aware errors, while confident Unaware errors (i.e., error trials on which subjects made high bets that they were correct) elicited a response that did not differ from the CRN elicited by truly correct answers. In contrast, for low confidence trials in unconfident subjects, an intermediate and equal Ne/CRN was elicited by Correct responses, Aware and Unaware errors. These results provide direct evidence that the Ne is related to error awareness, and suggest the amplitude of the Ne/CRN depends on individual differences in error reporting and confidence.

Keywords: error processing, error awareness, Ne, Pe, wagering, confidence

INTRODUCTION

In everyday life, we are sometimes acutely aware of having made an error (the notorious "oops" sensation), but at other times we are oblivious of our errors (e.g., when we make typographical errors). In many cases, effective performance of a task requires that errors will be promptly detected and swiftly corrected (e.g., driving or machine operation). Understanding the conditions that lead to error awareness is important for understanding the executive functions that guide goal-directed activity and learning. One of the most intriguing questions in the area of error awareness is why some errors reach awareness while others stay unnoticed. Yet before this question can be answered, it is crucial to determine whether unnoticed errors are processed as errors at all. Using electrophysiology, it has repeatedly been shown that errors in various tasks and modalities are processed as errors by the brain even if we are unaware of making them (Nieuwenhuis et al., 2001; Endrass et al., 2005, 2007; O'Connell et al., 2007, 2009; Pavone et al., 2009; Shalgi et al., 2009; Dhar et al., 2011; Hughes and Yeung, 2011). In each of these studies, awareness of errors was reported on a single trial basis, and a large Error-Related Negativity (Ne; Falkenstein et al., 1991; or ERN; Gehring et al., 1993), a fronto-central ERP component locked to the incorrect response, was equally seen after errors that were reported (Aware errors) and those that weren't (Unaware errors). In most cases, a much smaller response-locked negativity was elicited by correct responses (Correct-Related Negativity, CRN; Ford, 1999; Vidal et al., 2000; Coles et al., 2001). In contrast to the insensitivity of the Ne to awareness, a later centro-parietal Error Positivity wave (Pe; Falkenstein et al., 1991) was elicited only after Aware errors. Two additional event-related fMRI studies by Hester et al. (2005) and Klein et al. (2007) further supported these findings by showing that the Anterior-Cingulate region associated with the generation of the Ne (Dehaene et al., 1994; Brazdil et al., 2002; Debener et al., 2005) does not show a difference between Aware and Unaware errors (see also Stemmer et al., 2004), while Aware errors only were associated with larger bilateral activation of prefrontal and parietal regions (Hester et al., 2005) or left anterior Insula activity (Klein et al., 2007). These findings led to the widespread notion that error monitoring, as reflected by the Ne, goes on regardless of conscious awareness of making errors (e.g.,

Simons, 2009). However, findings from a handful of studies that showed that the Ne was smaller for ostensibly unaware errors raise doubts regarding whether the Ne is truly insensitive to awareness.

Maier et al. (2008, Experiment 2), using a masked Flanker task, and Steinhauser and Yeung (2010), using a brightness discrimination task, both asked subjects to judge their own accuracy on a single-trial basis, and found a larger Ne for Aware than for Unaware errors. Woodman (2010) used a visual-search task in which in some trials the targets were rendered invisible by substitution masking. Although the participants were not required to indicate awareness of their errors, errors in masked-target trials were assumed to be Unaware errors because the participants were at chance in detecting the presence of the target. An Ne was generated only by errors committed during the non-masked trials, even though masked targets also attracted attention, as evidenced by the N2pc component. Praamstra et al. (2003) induced errors in a cued tapping task by small or large shifts (15 or 50 ms) in the cue timing. While they also did not require their subjects to report awareness of an error, they assumed that the small shifts are below perception threshold whereas the large shifts are consciously detected. Since only large shifts were followed by an Ne, the results were once again taken to suggest that awareness of the errors is a prerequisite for Ne elicitation.

A major critique for the experiments that showed that the Ne was independent of conscious awareness pertains to the determination of the level of error awareness. Were subjects really unaware of the error or was their criterion for reporting an error too high? This criterion may be influenced by individual tendencies, as well as by task instructions and demand characteristics artifacts. If subjects are in fact under-reporting their awareness, the Unaware Error bin could be contaminated by (unreported) Aware errors which have a larger Ne, leading to an inflated Ne for Unaware errors.

This possibility was in fact directly addressed by Scheffers and Coles (2000) who used certainty ratings instead of error-reporting on a trial-by-trial basis. Their subjects performed a Flanker task and immediately afterwards had to indicate their confidence on a five-point scale from "sure correct," through "don't know," to "sure incorrect." The size of the Ne depended on the subjective confidence of making an error. These results have recently been replicated by Hewig et al. (2011) and suggest that the Ne is an index of subjective, rather than objective error monitoring. This conjecture is supported by studies which used indirect measure of error certainty, such as post-test questionnaires and reaction times (RTs), under the assumption that slower responses indicate uncertainty (Pailing and Segalowitz, 2004a; Wessel et al., 2011). All the above studies that measured the Pe component (Steinhauser and Yeung, 2010; Woodman, 2010; Hewig et al., 2011; Wessel et al., 2011), found that it was larger for Aware errors than Unaware errors. Taken together, these studies support the premise that the Ne, like the Pe, is an index of subjective error processing. They contradict a large body of evidence (including our own, Shalgi et al., 2009 and Shalgi and Deouell, 2010) that showed no effect of awareness on the Ne.

Here, we directly addressed this contradiction by using a paradigm in which we previously found a similar Ne for Aware and Unaware errors (Shalgi and Deouell, 2010), together with a wagering paradigm (see Ullsperger et al., 2010) which was shown to be a more objective measure of awareness then direct introspection (Persaud et al., 2007). Instead of directly asking participants to assess their own certainty, the subjects are required to make a bet of a small, medium, or large amount of money on their accuracy judgment (rather than on the actual response). If the participants' decision about their accuracy is correct, they win this money; otherwise, they lose it. The optimal strategy is to bet high whenever they feel they are not just guessing. In other words, participants are willing "to put their money where their mouth is" (Koch and Preuschoff, 2007). In the current study, we employed this betting paradigm to separate trials in which the participants were sure of their response from trials in which they were unsure without explicitly asking for confidence ratings. We expected that if the Ne is indeed independent of awareness, it should be elicited even when the subjects are highly confident that they are right. Conversely, if Ne is only elicited by Aware errors, it should be eliminated if subjects are confident that they had not made an error.

To summarize, previous studies suggested that the Ne is independent from error awareness, while other findings showed that the Ne is affected by subjective confidence. We aimed to replicate both findings in the same experiment, with better confidence assessment, in order to explain the discrepancy between the former studies.

MATERIALS AND METHODS

SUBJECTS

The participants were 22 students (14 males, all right handed as a prerequisite due to the lateralized aspect of the task), aged 23–32 (mean = 25.73, SD = 2.7), with reportedly normal or corrected-to-normal vision and no history of neurological disorders. Two additional subjects were tested but were excluded from analysis; one due to a very small number of Unaware errors (see "Analysis of Behavior") and the other due to a misunderstanding of task instructions. All students were paid a minimum flat fee or given course credits for their participation in the study. In addition, they could make up to an extra 25 NIS (∼7$, the equivalent of lunch at the campus cafeteria) by wagering correctly on their responses in the experiment (see "Procedure"), but they could never lose money (i.e., get less than the minimum fee). Written consent was obtained after the experimental procedures were explained. The study was approved by the Ethics Committee of the Faculty of Social Science at the Hebrew University.

STIMULI AND APPARATUS

Participants sat in a dimly lit, sound attenuated and echo-reduced chamber (Eckel C-26, UK). Stimuli included seven white shapes (diamond, triangle, pentagon, parallelogram, trapezoid, circle, pie) of four different sizes (subtending visual angles of 1.2°–1.8°) on a gray background, presented on a Viewsonic G75f CRT monitor, 100 cm from the subjects' eyes. A row of three shapes appeared on every trial (see **Figure 1**). The lateral shapes and the central shape were never the same size. Shapes were spaced 5.08° apart. A cushioned tray was placed on the subjects' lap. On the right of the tray was a two-button response box with the buttons "YES" and "NO," their side counterbalanced across participants.

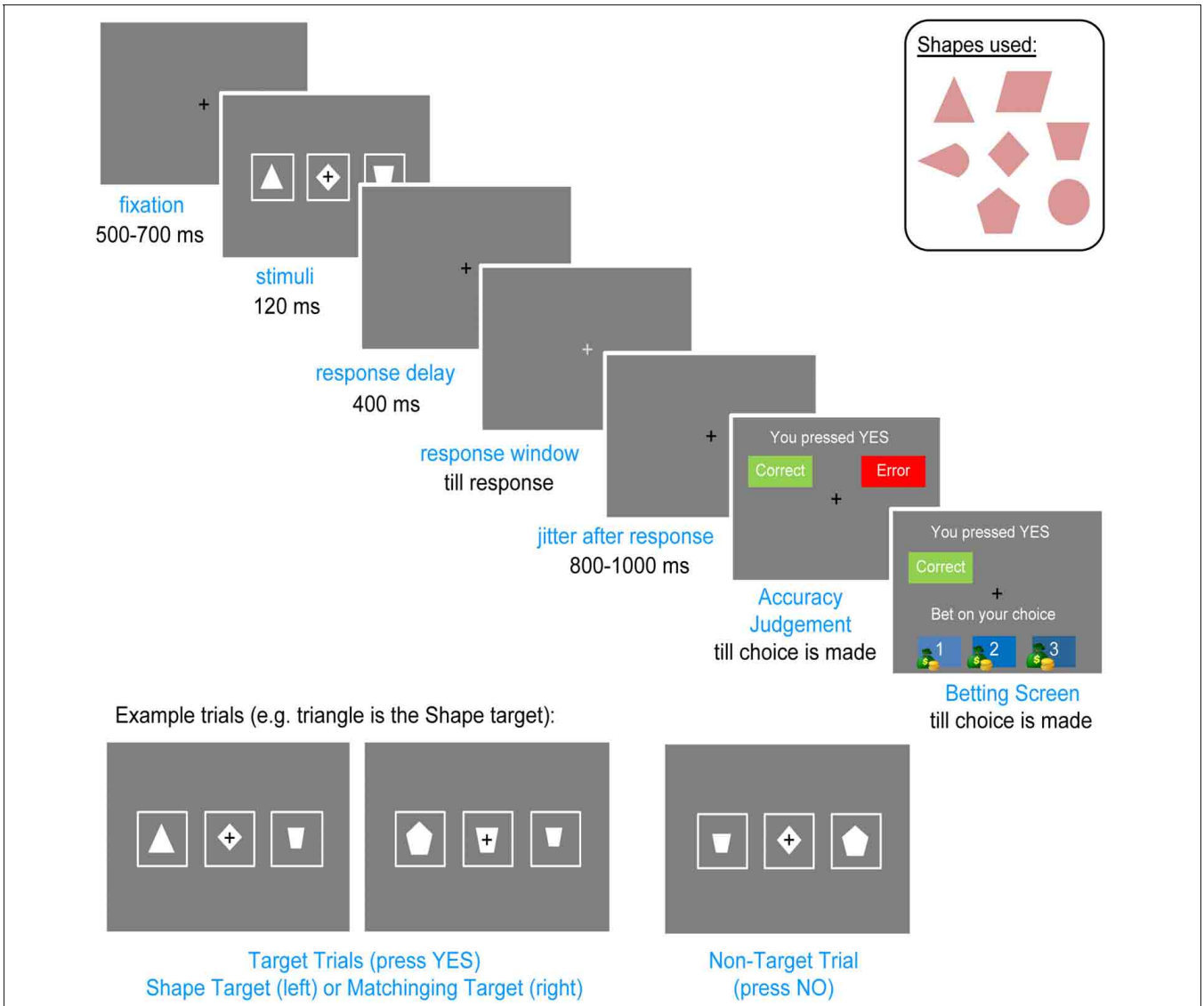


FIGURE 1 | The trial procedure and some example trials.

On the left was a computer mouse used to make the secondary responses (see "Procedure").

PROCEDURE

We developed the Lateralized Error Awareness Task (LEAT) to enable the separation of left and right errors and to obtain an awareness index of each error (Shalgi and Deouell, 2010). The LEAT is a manual choice-reaction time task in which participants are presented with three horizontally aligned shapes in each trial. In the primary task, the participants are instructed to press the "YES" button if one of the shapes is a designated target shape (Shape target), or if one of the lateral shapes (left or right) is the same shape as the central shape (Matching target), disregarding shape sizes, and to press the "NO" button otherwise.

A small black fixation cross appeared constantly at the center of the screen. Each trial started with the three shapes presented for 120 ms. Four hundred milliseconds after their disappearance, the color of the central fixation changed from black to light gray. The participants were instructed to respond as fast as possible only once the fixation had changed its color, creating a response delay in order to increase the number of Unaware errors (see Shalgi et al., 2007). There was no time limit on the response, so participants were forced to make a choice in each trial, but if their RT was longer than 1320 ms, a screen was displayed which reminded the participants to respond when the color of the fixation changes. Shape presentation time was too short to allow lateral eye movements; participants were required to fixate on the central cross and were advised that this will allow for best performance.

Five to seven hundred milliseconds after the primary response, an "Accuracy Judgment" screen appeared in which the choice the participant had just made ("YES" or "NO") was displayed,

together with two on-screen buttons labeled "Correct" and "Error," and the participant had to click one of the buttons using the mouse controlled with their left hand. The left hand was used for mouse responses as the right (dominant) hand was always kept on the response box for the primary speeded response. Immediately after the participants made their judgment, a "Betting" screen appeared which informed the participant of the judgment they had just made ("Correct" or "Error") and asked them to bet on this decision by clicking the mouse on one of three credits: 1 (small bet), 2 (medium bet), or 3 (high bet). The participants were not told how much money each credit was worth, only that a larger number of credits meant more money (as an added incentive, an image of a small, medium, or large money bag was displayed beside each credit amount, see **Figure 1**). The participants were told that if they had made a correct accuracy judgment they would win the number of credits they had bet on, but in case of an incorrect accuracy judgment they would lose that same amount. Both the Accuracy Judgment and the Betting screens were displayed until the participants made their choice. The black fixation cross stayed on constantly during these screens. After placing a bet, the fixation screen was displayed for 800–1000 ms, followed by the next trial.

Each block consisted of 88 trials, of which 44 were target trials (50%), equally divided between Shape and Matching targets. Each target shape or matching shape appeared equally on the left or on the right (the target shape never appeared as the central shape). The order of the stimuli was randomized across participants. To complicate the task and elicit more errors, every 22 trials (i.e., four times per block), a different shape (one of the seven possible shapes at random) was designated as the Shape target, using a screen that announced the change and displayed the new target shape at the center of the screen for two seconds. Each block lasted approximately six minutes. The results of each block were not displayed to the participants so they could not use them to change or adjust their strategy.

Two practice blocks of 15 trials each preceded the experiment, both with neither the Accuracy judgment screen nor the Betting screen. In the first practice, the participants were introduced only to the Matching targets in order to acquaint them with the different shapes and sizes and to gradually ease them into the task. In the second practice, the Shape target condition was introduced. When the participants were confident in their understanding of the task, they performed 16 consecutive experimental blocks. Due to the length of the experiment, after eight blocks the participants were taken out of the recording chamber for a 5–10 minute rest, without removing the electrode cap.

Following the testing session, participants filled a questionnaire in which they rated task difficulty and estimated their own accuracy. The amount of credits won at the end of the experiment was converted to cash using a pre-formed conversion table.

ANALYSIS OF BEHAVIOR

Errors judged as "Error" in the Accuracy Judgment screen were considered Aware errors, and Errors judged as "Correct" were considered Unaware errors. Correct responses judged as "Error" were classified as False Errors. The percentage of False Errors for each participant was very low, ranging between 0.1–6% (mean 2.1%, SD 1.6%), and therefore, this type of error was not further analyzed. Trials with primary RTs faster than 50 ms or slower than 2500 ms were excluded from analysis. Accuracy was calculated separately for each bet category (1, 2, or 3). Error Awareness rate was calculated as the number of Aware errors divided by the total number of errors. RTs were calculated across blocks separately for the different response types and for the different bet categories (there were not enough error trials to divide response type by bet). The different measures were compared separately using repeated measures ANOVAs. Greenhouse–Geisser correction was applied when necessary and the uncorrected degrees of freedom are presented along with the Greenhouse–Geisser epsilon where it was <1 (Picton et al., 2000). Contrasts were performed using paired sample *t*-tests.

EEG RECORDING

EEG was recorded continuously with Ag/AgCl electrodes from 64 scalp electrodes, left and right mastoid sites, and the tip of the nose, using a BioSemi Active 2 system (Biosemi, Netherlands). Blinks and eye movements were monitored using four additional EOG electrodes located at the outer canthus of the right and left eyes and above and below the center of the right eye. The EEG was continuously sampled at 512 Hz and stored for off-line analysis. Analsyis was conducted using BrainVision Analyzer 2 (Brain Products, Germany), and Matlab R2011a (The MathWorks Inc., Natick, MA). For two participants, data from one malfunctioning electrode (P3 or CP3) was replaced by using spherical spline interpolation (Perrin et al., 1989; Perrin, 1990, implemented by Analyzer 2). The EEG data was digitally referenced to the nose and filtered with a bandpass of 0.5–20 Hz (zero-phase 24 dB/octave Butterworth filter). Blink artifacts were removed using the Independent Component Analysis (ICA) method (Jung et al., 2000; as implemented in Analyzer 2). Note that although we corrected for blink artifacts, we first verified that the participants did not blink during stimulus presentation. This was the case for all subjects and trials, with negligible exceptions. Segments contaminated by other artifacts were discarded (rejection criteria: $>100\,\mu V$ absolute difference between samples within segments of 100 ms; absolute amplitude beyond the $\pm 100\,\mu V$ range). Artifact-free EEG data was parsed into 900 ms response-locked segments starting 400 ms before the response. An average of 93.6% segments survived artifact rejection (range per participant, per response type 75.7–100%). The inclusion criterion was at least 20 Unaware Error segments which survived artifact rejection.

Response-locked segments were averaged separately for Correct responses, Aware errors, and Unaware errors, and within each category for bets 1 and 3. The 2 (medium) bet was excluded from analysis as it was only used to get a cleaner measure of the participants' confidence in the 1 and 3 bet bins. By including this bet option, we could be more confident that when participants chose to bet on 1 they were unsure of their decision and when they chose to bet on 3 they were confident, as all the intermediate cases would fall into the 2 bin. Potentials were measured relative to a 400–200 ms pre-response baseline period (O'Connell et al., 2007; Shalgi et al., 2009). This baseline period did not include the immediate pre-response period as the neural activity leading to

the response naturally starts earlier than the final motor outcome of the button press.

ERP ANALYSIS

The Error Negativity: The Ne amplitude and latency were detected individually for each participant and each subcategory and measured in the response-locked averages as the mean amplitude ±4 ms around the most negative peak at the interval of 0–150 ms post-response at electrode FCz.

The Error Positivity: As the Pe is a more sustained parietal wave, it was pre-defined based on our previous studies as the average amplitude at electrode Pz between 300 and 500 ms post response in the response locked averages. Since inspection of the grand average waveforms (see **Figure 3**) suggested that the Pe began at around 200 ms, we also performed all analyses of the Pe on the 200–500 ms timeframe, and obtained the same qualitative results. Therefore, only the results of the first analysis will be reported.

To simulate traditional procedures, the Ne and Pe amplitudes were entered separately into One-Way repeated measures ANOVAs for Response Type (Correct, Aware Error, Unaware Error), across all subjects and bets. Greenhouse–Geisser correction was applied where necessary. Contrasts were performed using paired sample *t*-tests.

The cardinal question of this study was whether the Ne is affected by awareness when subjects are highly confident of being aware or unaware of making an error. To that end, we isolated the group of 12 participants who had enough (defined a-priori as more than 20 EEG segments available for analysis) Aware and Unaware error trials with bet 3, and repeated the above One-Way ANOVA analysis with bet 3 trials only. A similar analysis for bet 1 responses was conducted for those eight subjects who had enough Aware and Unaware errors with bet 1. Three subjects were included in both groups, four participants did not have enough segments to be included in any group and one participant was excluded from this analysis due to noisy data (see "Analysis of the Response-Locked Waveforms After Subtracting the Stimulus-locked Waveform" in Appendix for the number of trials in each group).

Hardly any subjects bet enough on both 1 and 3 for both Aware and Unaware errors to allow a full within subject analysis. Therefore, in order to compare the awareness effect between confident and less confident subjects, we omitted the three subjects who were included in both groups above from the high confidence group, which now included only those nine subjects who had enough bet 3 errors of both types *but less than 20 bet 1 errors*. The Ne and Pe of the High Confidence (HiC) and Low Confidence (LoC) groups were then compared across all bets using a mixed ANOVA with Group as the between-subject factor and Response Type as the within-subject factor. Contrasts were performed using independent-samples *t*-tests.

RESULTS

BEHAVIORAL RESULTS

Accuracy and error awareness

The behavioral results are summarized in **Figure 2**. Although the subjects rated the task as relatively difficult (mean 3.6/5, SD 0.85), accuracy in the task was quite high (85.9%, SD 4.6%). Overall awareness of errors was 50.1% (SD 14.7%).

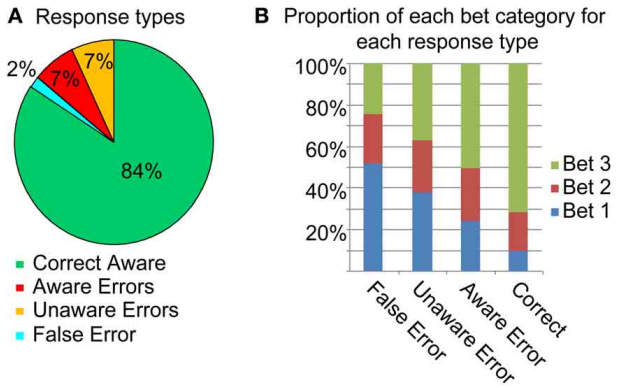


FIGURE 2 | Behavioral results of all 22 participants.

Overall, bet 3 was the popular bet (67.9%), followed by bet 2 (19.6%), and finally bet 1 (12.3%). Rational calculations predict that due to the high accuracy in the task, a strategy of always betting on 3 would yield the highest result, and participants might have intuitively made this assumption. **Figure 2B** describes the betting patterns for the different response types. When the participants were correct, they were usually certain about their answer (bet 3). Uncertainty went up with the different types of errors: when participants made an Aware Error they tended to bet less on 3 than when they were correct, and even less so when they made an Unaware error. They were least certain when they (rarely) mistakenly judged a correct response as an error. Primary response accuracies for the three bet categories were 52.8% (SD 12.4) for bet 1, 75.1% (SD 12.2) for bet 2 and 91.1% (SD 4) for bet 3 [$F_{(2,\ 38)} = 152.9$, $p < 0.001$, $\varepsilon = 0.95$]; two subjects were excluded from this analysis as they did not have any bet 2 or any bet 1 choices. Accuracy for bet 1 was lower than for bet 2 [$t(19) = 10.55$, $p < 0.001$], which was in turn lower than the accuracy for bet 3 [$t(19) = 7.89$, $p < 0.001$].

Reaction times

Reaction times in this task were relatively long (across bets: Correct responses 814.7 ms, SD 42.8; Aware errors 885.8 ms, SD 64; Unaware errors 889 ms, SD 64.9) due to the task instructions to delay responses until the fixation changed its color. Nevertheless, the RTs depended on accuracy [$F_{(2,\ 42)} = 54.67$, $p < 0.001$, $\varepsilon = 0.98$] as the RTs for Correct responses were shorter than those for errors [$t(21) = 11.34$, $p < 0.001$]; there was no difference between RTs for Aware and Unaware errors $p = 0.71$. There was also a significant difference between RTs of responses followed by the different bet categories [$F_{(2,\ 38)} = 66.61$, $p < 0.001$, $\varepsilon = 0.99$]: RTs followed by bet 1 (952.2 ms, SD = 84.9) were significantly longer than RTs followed by bet 2 [885 ms, SD = 98.1; $t(19) = 5.47$, $p < 0.001$] and these in turn were significantly longer than RTs followed by bet 3 [796.9 ms SD = 41.4; $t(20) = 6.33$, $p < 0.001$]. This supports the assumption that participants bet on 1 when they were less confident of their response (slower RTs due to hesitation) and on 3 when they were confident.

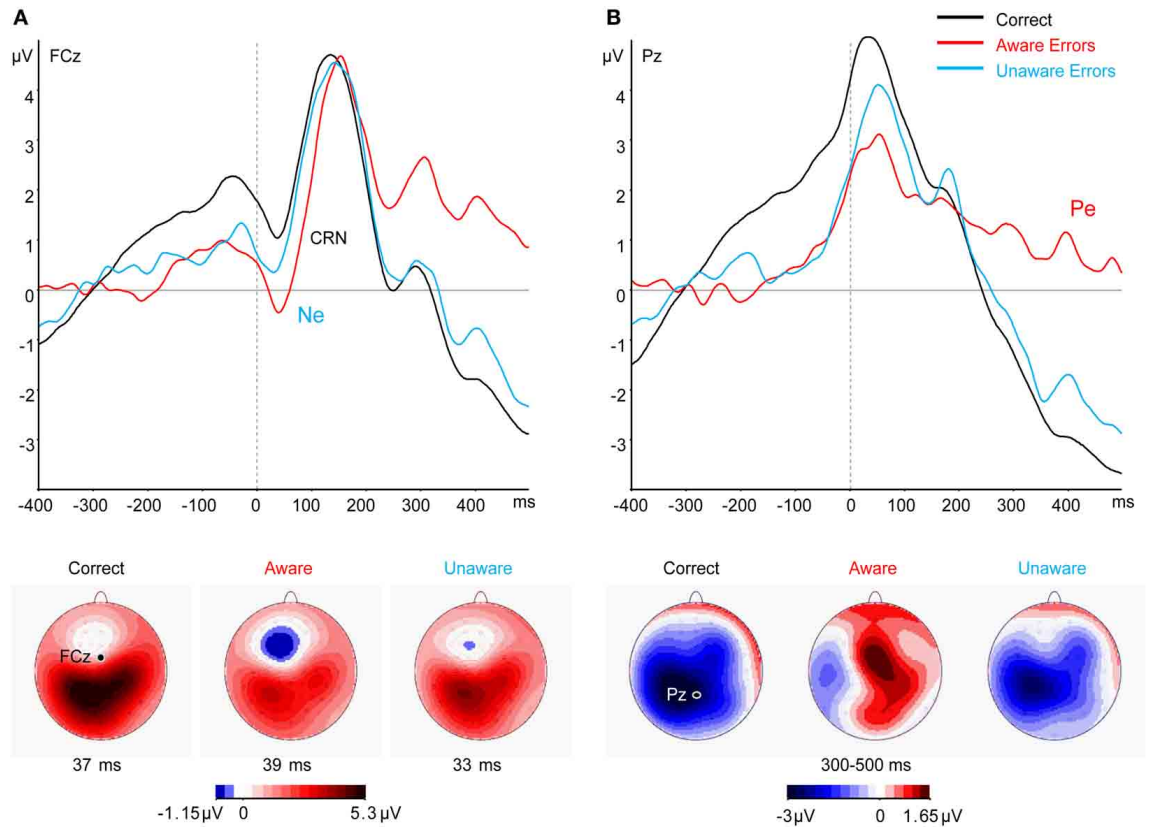


FIGURE 3 | Grand average ($N = 22$) response-locked ERPs to correct and error responses across all bets at electrode (A) FCz and (B) Pz. The bottom panel shows the scalp topography at the time of **(A)** the CRN/Ne peak and **(B)** the Pe.

Debriefing

Upon debriefing in the post-test questionnaire, all participants noted they made their bets according to their level of certainty. Taken together with the behavioral results, we can be quite sure that each bet category corresponds to a certainty rating.

ERP RESULTS

Awareness effects regardless of confidence

Replicating the results of traditional studies not assessing confidence, a clear Ne was elicited across subjects and bets by both Aware and Unaware errors, beginning before the response and peaking after the response (**Figure 3A**). A CRN can be seen at about the same time for Correct responses. A significant effect in a One-Way repeated measures ANOVA for Response Type was found for the Ne amplitude [$F_{(2,\ 42)} = 10.2$, $p < 0.001$, $\varepsilon = 0.959$]. This stemmed from a difference between errors and Correct responses [$t(21) = 4.57$, $p < 0.001$] while there was no difference between the Ne amplitude for Aware and Unaware errors [$t(21) = 1.79$, $p = 0.09$]. Note that while the peak-to-trough CRN seems to be around the same size or even larger than the Ne, supplementary analysis (see "Analysis of the Response-locked Waveforms After Subtracting the Stimulus-locked Waveform" in Appendix) which corrected for the overlapping effect of the stimulus-locked P3, revealed that the CRN is indeed smaller than the Ne. **Figure 3B** illustrates the results at electrode Pz, where the Pe can be seen as the slow positivity following the response for Aware errors only. A significant effect in a One-Way repeated measures ANOVA for Response Type was found for the Pe amplitude [$F_{(2,\ 42)} = 35.32$, $p < 0.001$, $\varepsilon = 0.82$], which stemmed from the difference between Aware errors and Correct responses [$t(21) = 6.67$, $p < 0.001$] and the difference between Aware errors and Unaware errors [$t(21) = 6.47$, $p < 0.001$].

Awareness effects with confident accuracy judgments

The most critical question in our study is whether awareness affects the Ne when subjects are highly confident of having made an error or highly confident that they have not made one, as indicated by their willingness to put a high bet on this decision. Note that when participants make an error and bet the highest bet on having made one, they are confident that an error has been made, that is, they demonstrate high error awareness. Similarly, when subjects place the highest bet on having made a correct response, when in fact they have committed an error, they demonstrate clear unawareness of the error. Thus, we isolated 12 subjects who had enough bet 3 trials for both Aware and Unaware errors (see "Methods"), and measured the Ne and Pe for their bet 3 trials only. Contrary to the previous results, the Ne

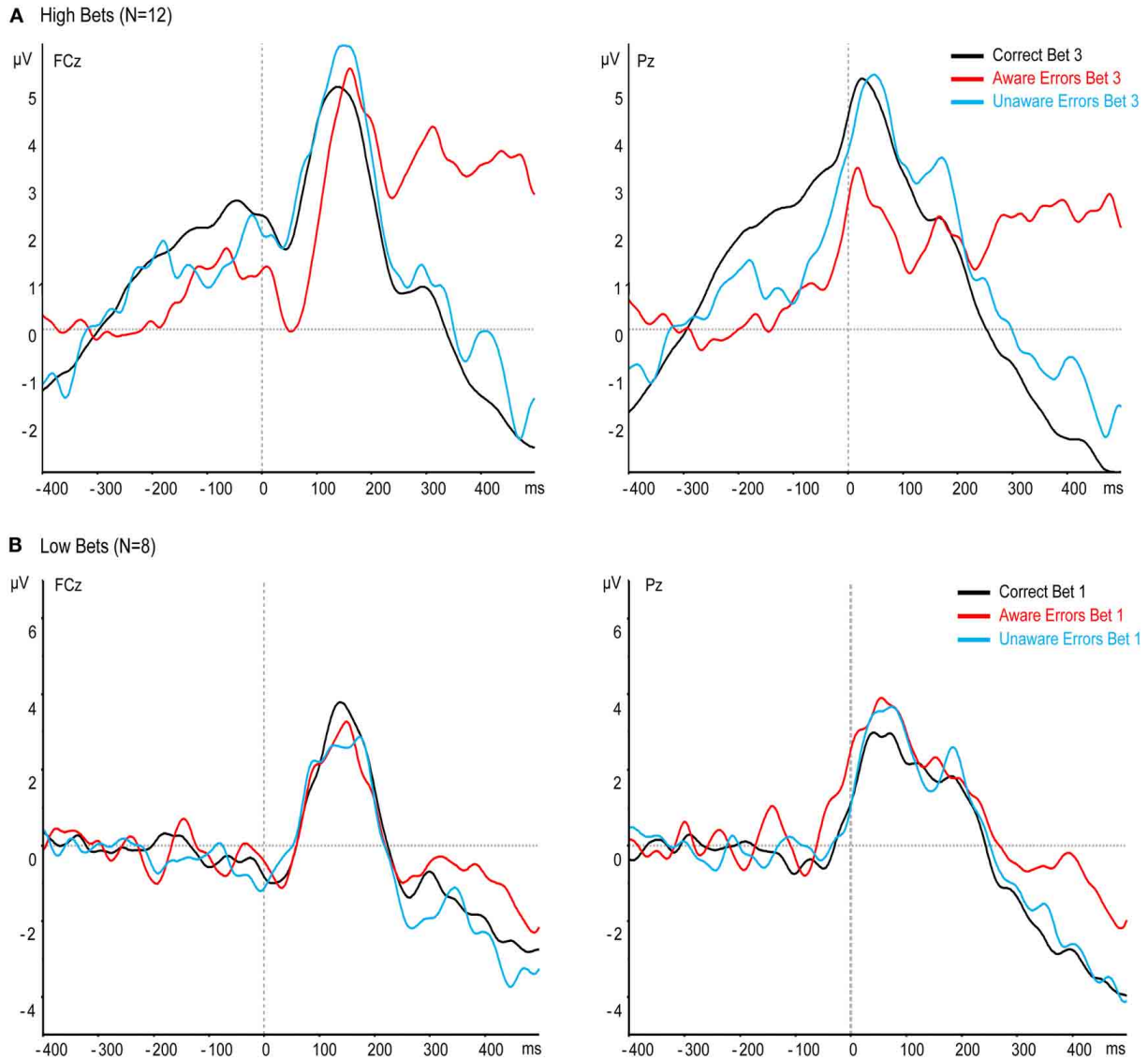


FIGURE 4 | Effect of confidence on response-locked ERPs at electrodes FCz and Pz, when only high (A) or low (B) bets are averaged.

was larger for Aware errors than for Unaware errors, and Unaware errors did not differ from Correct responses (**Figure 4A**), main effect: $[F_{(2,\ 22)} = 8.65,\ p < 0.005,\ \varepsilon = 0.77]$, Aware errors vs. Unaware errors: $[t(11) = 3.26, p < 0.01]$, Corrects vs. Unaware errors: $[t(11) = 0.42,\ p = 0.68]$. Thus, when only clear cases of subjective error awareness or unawareness are taken into consideration, the Ne is found to be strongly affected by error awareness. The Pe too was large only for Aware errors in the bet 3 trials in this group $[F_{(2,\ 22)} = 20.35, p < 0.001, \varepsilon = 0.67]$, Correct Responses vs. Aware errors: $[t(11) = 4.67,\ p < 0.005]$, Aware vs. Unaware: $[t(11) = 5.36, p < 0.001]$. Thus, both the Ne and the Pe are larger for Aware errors when subjects are confident of their awareness.

Awareness effects with low-confidence accuracy judgments

To examine the awareness effect when subjects are not sure about their performance, we isolated eight subjects who had enough bet 1 trials for both Aware and Unaware errors (see "Methods"), and measured the Ne and Pe for their bet 1 trials only. When subjects were only willing to place the lowest possible bet on their decision, we surmise that they were not sure of whether they had made an error or not. In other words, they may have had only a trace of awareness of an error when they had made one, and conversely may have had some trace of awareness of an error even if one had not been committed. Indeed, the Ne and Pe were similar for all responses—Correct, Aware errors, and Unaware errors (**Figure 4B**), [Ne: $F_{(2,\ 14)} < 1, p = 0.47, \varepsilon = 0.67$; Pe: $F_{(2,\ 14)} = 2.77,\ p = 0.13,\ \varepsilon = 0.6$]. This finding supports the notion that the Ne is affected by the level of subjective error awareness.

Since there were many Correct responses with both bets in this group of subjects, we could directly compare between bet 1 and bet 3. "Analysis of the Correct Response-locked Waveforms

of the Low Confidence Group After Subtracting the Stimulus-locked Waveform"in Appendix shows the CRN and Pe of Correct responses for the two bets after correcting for the overlapping effect of the stimulus locked P3. The difference between the CRN for bet 1 and bet 3 was significant [$t(7) = 2.38$, $p < 0.05$]. However, the Pe for Correct responses did not significantly differ between bet 1 and 3 [$t(7) = 1$, $p = 0.35$].

Direct comparison between low and high confidence groups

As noted in the methods section, the number of errors in each response-type and bet precluded a complete within-subject analysis of high and low bets. However, we could compare high and low confidence subjects, collapsing across their bets (see "Methods"). For this analysis, high confidence subjects were those nine subjects who had made enough bet 3 aware and unaware errors but few (<20) bet 1 errors (HiC group), while low confidence subjects were the eight subjects who had enough bet 1 aware and unaware errors (LoC group). As the groups did not overlap, they could be compared directly (**Table 1**, **Figures 5** and **6**). The groups differed significantly in primary task accuracy [$t(15) = 2.88$, $p < 0.05$] but not in error awareness rates [$t(15) = 1.16$, $p = 0.26$] or subjective rating of task difficulty [$t(15) = 1.2$, $p = 0.25$]. As may be expected by the group division criteria, the betting patterns of the groups (**Figure 5**) were significantly different; a significant main effect of the within-subject factor Bet was moderated by an interaction with the between-subject factor Group in a mixed ANOVA [main effect: $F_{(2,\ 30)} = 28.7$, $p < 0.001$, $\varepsilon = 0.71$; interaction: $F_{(2,\ 30)} = 6.98$, $p < 0.01$]. The interaction stemmed from different usage of the 1 and 3 bets between groups.

Figure 6 shows a dramatic effect of confidence in the Ne period. While the Ne for Unaware errors aligns with the Ne for Aware errors in the LoC group, there is practically no Ne for Unaware errors in the HiC group. That is, the response for Unaware errors is similar to the CRN elicited by Correct responses. A mixed ANOVA with the within factor Response Type and between factor Group for the Ne amplitude confirmed these results, showing a significant effect of Response Type [$F_{(2,\ 30)} = 8.8$, $p < 0.005$, $\varepsilon = 0.95$] which was moderated by a significant interaction [$F_{(2,\ 30)} = 3.55$, $p < 0.05$]. To elucidate this interaction, we compared the Ne amplitude between groups separately for each response type. There was no significant difference between the Ne for Correct responses [$t(15) = 0.77$, $p = 0.45$] or for Aware errors [$t(15) = 0.12$, $p = 0.9$], but there was a significantly larger Ne for Unaware errors in the LoC group compared to the HiC group [$t(15) = 2.26$, $p < 0.05$].

Table 1 | Comparison of the behavioral results of the High and Low Confidence Groups.

	High Confidence ($N = 9$)	Low Confidence ($N = 8$)	*P*-value
Accuracy	87.9% (4%)	82.5% (3.65%)	0.011*
Awareness	51.9% (11%)	45% (12.9%)	0.264
Difficulty rating	3.33 (1.12)	3.87 (0.64)	0.248
RT	833.4 (19.95)	823.5 (52.83)	0.074
%Bet 1	5.92% (4.5%)	22.39% (18.2%)	0.019*
%Bet 2	12.61 (9.6%)	27.96% (20.95%)	0.066
%Bet 3	81.21 (13.9%)	49.41% (29.2%)	0.01*

*$p < 0.05$.

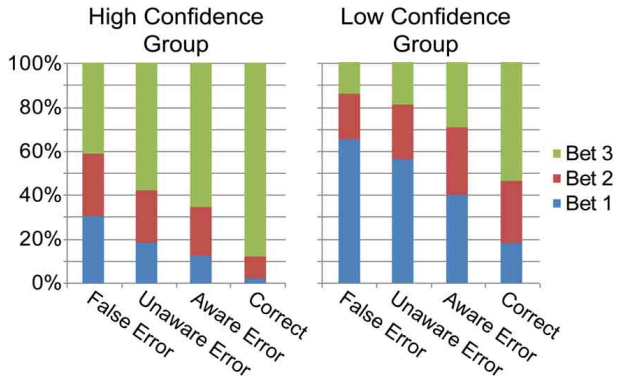


FIGURE 5 | Proportion of each bet category for each response type for the High and Low Confidence groups.

The same mixed ANOVA for the amplitude of the Pe showed a significant effect of Response Type [$F_{(2,\ 30)} = 23.91$, $p < 0.001$, $\varepsilon = 0.81$] but no interaction [$F_{(2,\ 30)} < 1$, $p = 0.98$]. Follow-up contrasts on the main effect of Response Type revealed a difference between the Pe for Aware and Unaware errors [$t(15) = 5.21$, $p < 0.001$] and no difference between the Pe for Correct and Unaware errors [$t(15) = 1.75$, $p = 0.1$].

DISCUSSION

Optimizing goal-directed behavior in imperfect conditions likely relies on both implicit and explicit error detection processes. In order to understand the unique contribution of each type of process, an essential initial step is to find physiological indices of aware and unaware error processing. Previous research linked the Ne component, occurring within the first 100 ms after the response, to processing of errors independent of consciousness, while awareness of errors influenced the amplitude of the Pe component at latencies longer than 250 ms post-response (Nieuwenhuis et al., 2001; Endrass et al., 2005, 2007; O'Connell et al., 2007, 2009; Pavone et al., 2009; Shalgi et al., 2009; Dhar et al., 2011; Hughes and Yeung, 2011). This was based on the finding that an Ne was similarly elicited whether subjects were aware of making an error or not, while a large Pe was restricted to errors of which the subjects were aware. Here, we replicated this result across subjects in a new visual choice-reaction time task. A significant Ne was elicited both by errors the subjects reported (Aware errors) and by errors the subjects did not report (Unaware errors). However, awareness is not necessarily a binary situation, and subjects may differ in their criterion for reporting awareness. That is, some subjects may not report awareness of an error even if they have some notion of making one, but are not absolutely sure. In this situation, some putative unaware errors would be contaminated with "partial awareness," and the elicitation of Ne for these errors would not be evidence for processing of errors without awareness.

In an attempt to get a stricter selection of unaware errors and clarify this question, we used a wagering paradigm, asking

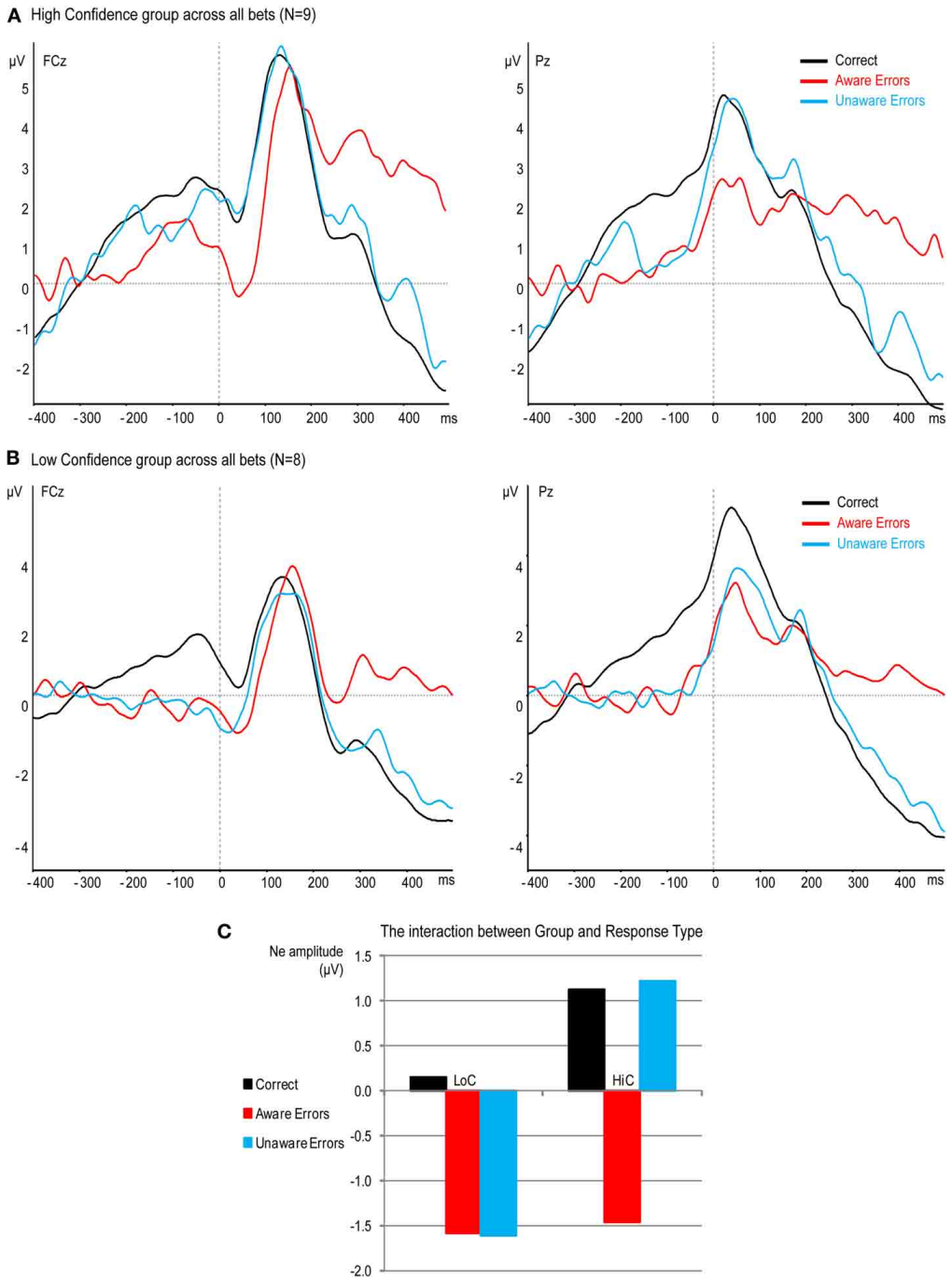


FIGURE 6 | Effect of confidence on response-locked ERPs at electrodes FCz and Pz. (A) High confidence group **(B)** Low confidence group. **(C)** The interaction between Group and Response Type.

subjects to indicate not only whether they had made an error or not, but also to bet money on their answer to this question. The pivotal working assumption was that if subjects report that they have not made an error (when one was made) yet they have some awareness that an error might have been committed, they would be reluctant to bet a large sum of money on their report. Indeed, subjects' debriefing, as well as RTs, suggest that they placed their bets based on their confidence. Based on this working hypothesis, the critical question was whether an Ne would be elicited when subjects were confident that they had not made an error, so much that they were willing to risk losing money. That is, in a condition mitigating the risk of partial awareness. The answer to this question was clear: when participants were highly confident that they had made an error (as indicated by a high bet), a large

Ne was elicited, but when they were highly confident that they did not make an Error, the Ne was comparable to the CRN. That is, there was no indication of error processing without awareness in this case. On the other hand, subjects who were more hesitant about their accuracy-judgment responses had an equal Ne for Aware and Unaware errors.

Thus, under stricter selection of Unaware errors, our results support the premise that the Ne *is* related to subjective awareness of an error. Unlike previous studies that showed this effect (e.g., Scheffers and Coles, 2000; Hewig et al., 2011), our findings cannot be explained by the difficulty of performing the task, i.e., of making mistakes rather than slips (Reason, 1990). Mistakes are errors in which the participant cannot solve the task, for example, when representation of the stimulus is degraded, or when the task is too cognitively demanding, and therefore, a correct response representation cannot be accurately formed. Slips are errors in which the correct response representation can be (easily) evoked, yet an error is made due to premature responding or a momentary lack of attention to the task (Shalgi et al., 2007). Defined this way, if errors are made due to mistakes, no error processing can be expected a-priori, as the correct response is unknown. Thus, the lack of Ne for Unaware errors in Scheffers and Coles (2000; who used reduced stimulus quality) and in Hewig et al. (2011; who used a digit entering task, which relies on a participant's digit span) studies could be attributed to the possibility that errors were of the mistake kind. This possibility is reduced in our current task, as the stimuli were clearly presented and accuracy in the task was high (>85%). The current results, therefore, strengthens the notion that the Ne is related to subjective awareness of an error rather than to error processing *per se*.

The current results go beyond this conclusion to explain the source of difference between studies that show no awareness effect on the Ne (Nieuwenhuis et al., 2001; Endrass et al., 2005, 2007; O'Connell et al., 2007, 2009; Pavone et al., 2009; Shalgi et al., 2009; Dhar et al., 2011; Hughes and Yeung, 2011) and the studies that show such an effect (Scheffers and Coles, 2000; Praamstra et al., 2003; Pailing and Segalowitz, 2004a; Maier et al., 2008; Steinhauser and Yeung, 2010; Woodman, 2010; Hewig et al., 2011; Wessel et al., 2011). As described above, when we separated our subjects to high and low confidence groups, we obtained both results, depending on individual subjects' error reporting tendencies: across all bets (which parallels the standard procedure for Ne calculation), there was a strong Ne error awareness effect for high confidence subjects, but not for low confidence subjects. Error-reporting experiments likely include a mixture of subjects who are more confident about their performance in the task (i.e., have less instances of partial awareness) and subjects who are less confident. When a large enough group of subjects is unsure about their own accuracy, both the Aware Error and the Unaware Error bins are contaminated by trials that belong to the other bin. A reported error, labeled "Aware," might be accompanied by a relatively low level of awareness, lowering the average Ne of the Aware errors bin, while an unreported error, labeled "Unaware," may in fact be accompanied by some awareness and therefore add to the average amplitude of the Unaware errors bin. Consequently, the grand average results, which do not take into account these individual error-reporting differences may not show a significant Ne awareness effect. Indeed, as noted also by Wessel et al. (2011), several previous studies which did not report a statistically significant difference between the Ne to Aware and Unaware errors, did show a numerically smaller Ne in the latter case, as we also show here in the grand average across all subjects and bets. Moreover, we would assume that in difficult tasks, in which the confidence about response accuracy is lower for all subjects compared to their confidence in an easier task, we would be less likely to obtain an Ne awareness effect, as more trials are expected to be in the "gray zone" of awareness.

Error awareness has previously been ascribed mainly to the Pe. Indeed, the Pe, and not the Ne, has already been used to examine the incidence of error awareness in patient studies (Jonkman et al., 2007; Larson and Perlstein, 2009; O'Connell et al., 2009; Wiersema et al., 2009). Our study replicates the sensitivity of the Pe to awareness of errors: Aware errors elicit a higher sustained positivity around 300 ms post-response. In what we see as the critical condition, in which subjects were highly confident of their decision, the Pe for Unaware errors could not be differentiated from the response elicited in Correct trials, as was the case for the Ne. That is, both Ne and Pe were elicited only by Aware errors. The results from the less confident subjects are less clear, however. Like the case of the Ne, the Pe awareness effect was not significant for low bets. Yet unlike in the case of the Ne, the interaction between the high and low confidence groups and error awareness was not significant for the Pe. Also, although one could surmise that for Correct responses with low confidence some Pe would emerge (reflecting some level of false error awareness), no significant difference was found between correct answers followed by bet 1 and 3 in this group (**Figure A2D**). Since in all previous studies of error awareness, in which a mix of confident and less confident subjects must have been included, a Pe awareness effect was nevertheless found, we can conjecture that the Pe awareness effect reflects a more binary decision regarding the error, whereas the Ne is more sensitive to the level of confidence. We previously proposed that the Pe may not even be a truly response-locked component, but a manifestation of a delayed evaluation (reflected by a P3b) of the stimulus preceding the response (Shalgi et al., 2009). The continuation of that inquiry will be reported elsewhere.

The main shortcoming of our study, like the studies of Scheffers and Coles (2000) and Hewig et al. (2011), is the difficulty of directly comparing between bet 1 and bet 3 responses in the same subjects. Although all our subjects made use of the full range of the betting scale, only three subjects out of 22 (13.6%) made enough 1 and 3 bets for both Aware and Unaware errors. A much larger pool of subjects would have been required to find a substantial group with enough error trials for within-subject comparisons. This obstacle, may suggest in fact that some individuals tend to either be mostly confident about their responses or mostly unsure of them. One might argue that the confidence report, as reflected by the bets, is not really related to differences in the level of error awareness between the subjects, but rather reflects subjects' tendency to be more decisive or daring when making their bets, or their perceived demand characteristics. However, we argue that if this were the case, we would not have found the dissociation we report here at the time window of the Ne (i.e., when the response was made), between high and low

confident subjects, and between high and low bets. It would be interesting to study what other traits these individual differences correlate with, and whether we could use these traits to create more homogenous groups for the study of error processing with or without awareness. Recent studies have pointed to the effect of individual differences in demographic characteristics (such as age and sex), personality traits (such as absentmindedness, impulsivity, affective style, anxiety), affective and motivational variables and behavioral performance (e.g., response speed) on error processing (e.g., Hester et al., 2004; Pailing and Segalowitz, 2004b; Boksem et al., 2006; Chang et al., 2010). Attention to individual differences may be critical for advancing our understanding of the role of the ERP components in error processing.

Can error monitoring occur without conscious recognition? Some studies show evidence for error processing without awareness. van Gaal et al. (2010) showed that behavioral conflict adaptation can occur when the response conflict occurs without awareness. Cohen et al. (2009) recently showed using EEG that error monitoring can occur when subjects are not only unaware of their errors, but also of the actual stimuli; when subjects performed a Go/No-Go task, errors following a masked No-Go signal elicited directional synchrony between the ACC and the occipital cortex, akin to the synchrony found when the No-Go signal was consciously perceived. Similarly, Ursu et al. (2009) showed, using fMRI, that the ACC, considered to be the locus of error processing and the origin of the Ne, was active when subjects were presented with a response conflict of which they were seemingly unaware (because a target's location violates an implicitly learned sequence), or when they made errors of which they were unaware. In contrast, other studies suggest that awareness is necessary for conflict adaptation (Kunde, 2003) or for the activation of the ACC (e.g., Dehaene et al., 2003). Whereas unconscious error processing may take place, the question is at what level and what are the prerequisites for error awareness to emerge. The evidence from our study, which controlled error awareness more closely, does not support error processing without awareness at the stage immediately after an error was made, as indexed by the Ne. We suggest that future studies of error awareness must address the methods used for awareness testing.

ACKNOWLEDGMENTS

We wish to thank Shany Grossman, Assaf Breska, and Tali Shrem for their help in the production of this manuscript. Supported by the National Institute of Psychobiology in Israel founded by the Charles E. Smith family.

REFERENCES

Boksem, M. A., Tops, M., Wester, A. E., Meijman, T. F., and Lorist, M. M. (2006). Error-related ERP components and individual differences in punishment and reward sensitivity. *Brain Res.* 1101, 92–101.

Brazdil, M., Roman, R., Falkenstein, M., Daniel, P., Jurak, P., and Rektor, I. (2002). Error processing – evidence from intracerebral ERP recordings. *Exp. Brain Res.* 146, 460–466.

Chang, W. P., Davies, P. L., and Gavin, W. J. (2010). Individual differences in error monitoring in healthy adults: psychological symptoms and antisocial personality characteristics. *Eur. J. Neurosci.* 32, 1388–1396.

Cohen, M. X., van Gaal, S., Ridderinkhof, K. R., and Lamme, V. A. (2009). Unconscious errors enhance prefrontal-occipital oscillatory synchrony. *Front. Hum. Neurosci.* 3:54. doi: 10.3389/neuro.09.054.2009

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.

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

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

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.

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

Ford, J. M. (1999). Schizophrenia: the broken P300 and beyond. *Psychophysiology* 36, 667–682.

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

Hester, R., Fassbender, C., and Garavan, H. (2004). Individual differences in error processing: a review and reanalysis of three event-related fMRI studies using the GO/NOGO task. *Cereb. Cortex* 14, 986–994.

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.

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.

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

Jonkman, L. M., van Melis, J. J., Kemner, C., and Markus, C. R. (2007). Methylphenidate improves deficient error evaluation in children with ADHD: an event-related brain potential study. *Biol. Psychol.* 76, 217–229.

Jung, T. P., Makeig, S., Humphries, C., Lee, T. W., Mckeown, M. J., Iragui, V., and Sejnowski, T. J. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology* 37, 163–178.

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.

Koch, C., and Preuschoff, K. (2007). Betting the house on consciousness. *Nat. Neurosci.* 10, 140–141.

Kunde, W. (2003). Sequential modulations of stimulus-response correspondence effects depend on awareness of response conflict. *Psychon. Bull. Rev.* 10, 198–205.

Larson, M. J., and Perlstein, W. M. (2009). Awareness of deficits and error processing after traumatic brain injury. *Neuroreport* 20, 1486–1490.

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.

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.

Pailing, P. E., and Segalowitz, S. J. (2004a). The effects of uncertainty in error monitoring on associated ERPs. *Brain Cogn.* 56, 215–233.

Pailing, P. E., and Segalowitz, S. J. (2004b). The error-related negativity as a state and trait measure: motivation, personality, and ERPs in response to errors. *Psychophysiology* 41, 84–95.

Pavone, E. F., Marzi, C. A., and Girelli, M. (2009). Does subliminal visual perception have an error-monitoring system? *Eur. J. Neurosci.* 30, 1424–1431.

Perrin, F. (1990). Correction. *Electroencephalogr. Clin. Neurophysiol.* 76, 565.

Perrin, F., Pernier, J., Bertrand, O., and Echallier, J. F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalogr. Clin. Neurophysiol.* 72, 184–187.

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

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R. Jr., Miller, G. A., Ritter, W., Ruchkin, D. S., Rugg, M. D., and Taylor, M. J. (2000). Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. *Psychophysiology* 37, 127–152.

Praamstra, P., Turgeon, M., Hesse, C. W., Wing, A. M., and Perryer, L. (2003). Neurophysiological correlates of error correction in sensorimotor-synchronization. *Neuroimage* 20, 1283–1297.

Reason, J. T. (1990). *Human Error*. New York, NY: Cambridge University Press.

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.

Shalgi, S., and Deouell, L. Y. (2010). "Is there a hemispatial bias in detecting errors and in error awareness?", in *The Israel Society for Neuroscience 19th Annual Meeting Eilat: Journal of Molecular Neuroscience*, vol. 45. (Eilat, Israel: Humana Press), 1–137.

Shalgi, S., O'Connell, R. G., Deouell, L. Y., and Robertson, I. H. (2007). Absent minded but accurate: delaying responses increases accuracy but decreases error awareness. *Exp. Brain Res.* 182, 119–124.

Simons, R. F. (2009). The way of our errors: theme and variations. *Psychophysiology* 47, 1–14.

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

Stemmer, B., Segalowitz, S. J., Witzke, W., and Schonle, P. W. (2004). Error detection in patients with lesions to the medial prefrontal cortex: an ERP study. *Neuropsychologia* 42, 118–130.

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.

Ursu, S., Clark, K. A., Aizenstein, H. J., Stenger, V. A., and Carter, C. S. (2009). Conflict-related activity in the caudal anterior cingulate cortex in the absence of awareness. *Biol. Psychol.* 80, 279–286.

van Gaal, S., Lamme, V. A., and Ridderinkhof, K. R. (2010). Unconsciously triggered conflict adaptation. *PLoS One* 5:e11508. doi: 10.1371/journal.pone.0011508

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.

Wiersema, J. R., van der Meere, J. J., and Roeyers, H. (2009). ERP correlates of error monitoring in adult ADHD. *J. Neural. Transm.* 116, 371–379.

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

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 26 January 2012; accepted: 18 April 2012; published online: 04 May 2012.

Citation: Shalgi S and Deouell LY (2012) Is any awareness necessary for an Ne?. Front. Hum. Neurosci. ***6****:124. doi: 10.3389/fnhum.2012.00124*

APPENDIX

NUMBER OF EEG SEGMENTS IN THE HIGH AND LOW CONFIDENCE GROUPS

The High Confidence (HiC) group was comprised of the nine subjects who had more than 20 bet 3 segments for both Aware and Unaware Errors and less than 20 bet 1 segments for both error types. The Low Confidence (LoC) Group was comprised of the eight subjects who had more than 20 bet 1 segments for both Aware and Unaware errors. **Table A1** summarizes the number of trials in each group.

ANALYSIS OF THE RESPONSE-LOCKED WAVEFORMS AFTER SUBTRACTING THE STIMULUS-LOCKED WAVEFORM

In the response-locked waveform, the peak-to-trough CRN appears to be riding on a slow positive wave starting well before the response, which creates an impression that the onset-to-peak CRN is around the same size or even larger than the Ne (**Figure A1A**). We hypothesized that the slow positivity is a residual of the jittered stimulus-locked target P3 (**Figure A1C**). To investigate this hypothesis, we calculated for each subject the mean ERP of the stimulus-locked response, separately for each response type. Then, for each single response-locked segment we subtracted the relevant condition's stimulus-locked ERP, after aligning the two waveforms according to that trial's RT. Trials with RTs longer than 1000 ms could not be included in this procedure as the stimulus-locked grand average was not long enough to be aligned with the response-locked ERP, resulting in the elimination of between 2 and 31 error trials per subject. The grand average result is presented in **Figure A1B**, which shows the elimination of the slow positivity, and confirms that the CRN is in fact smaller than the Ne. Statistical analysis of the resulting Ne amplitudes replicates the finding from the uncorrected waveforms reported in the Results section of the paper, namely that the Ne for errors is significantly different from the CRN but not different between Aware and Unaware Errors—there was a significant effect in a one-way repeated measures ANOVA for Response Type [$F_{(2,\ 42)} = 8.17$, $p < 0.005$, $\varepsilon = 0.76$] which stemmed from a difference between errors and correct responses [$t(21) = 5.27, p < 0.001$], while there was no difference between the Ne amplitude for Aware and Unaware Errors [$t(21) = 1.47$, $p = 0.16$].

Table A1 | Mean number and range of error trials of subjects included in the High and Low Confidence groups.

	High Confidence	Low Confidence
Aware Errors (all bets)	71.1 (29–118)	119.6 (55–176)
Bet 1	7.9 (0–18)	47.9 (28–73)
Bet 3	45.3 (27–70)	33.5 (2–65)
Unaware Errors (all bets)	85.6 (39–141)	110.9 (61–159)
Bet 1	15 (0–27)	59.1 (30–148)
Bet 3	48.9 (21–95)	23.3 (1–87)

ANALYSIS OF THE CORRECT RESPONSE-LOCKED WAVEFORMS OF THE LOW CONFIDENCE GROUP AFTER SUBTRACTING THE STIMULUS-LOCKED WAVEFORM

In the Low Confidence group, there were enough Correct responses with bet 1 and bet 3 for comparison within subjects between bets. However, the response-locked grand average of the Correct response on which subjects bet 3 seems to be riding on a positive pre-response wave (see **Figures A2A** and **A2C**). As in Appendix section B, we subtracted the mean of the stimulus-locked response (separately for bet 1 and bet 3) of each subject from each response-locked trial according to its RT. Trials with RTs longer than 1000 ms could not be included in this analysis, resulting in the elimination of between 6 and 128 correct trials per subject (this elimination did not significantly reduce the number of single trials for each subject and bet—range after elimination: 31–960). The grand average result for electrodes FCz and Pz is shown in **Figures A2B** and **A2D**. Statistical analysis of the resulting CRN and Pe amplitudes shows that the CRN for bet 3 was significantly smaller than the CRN for bet 1 in these subjects [$t(7) = 2.38$, $p < 0.05$] but there was no difference in Pe [$t(7) = 1.01, p = 0.35$].

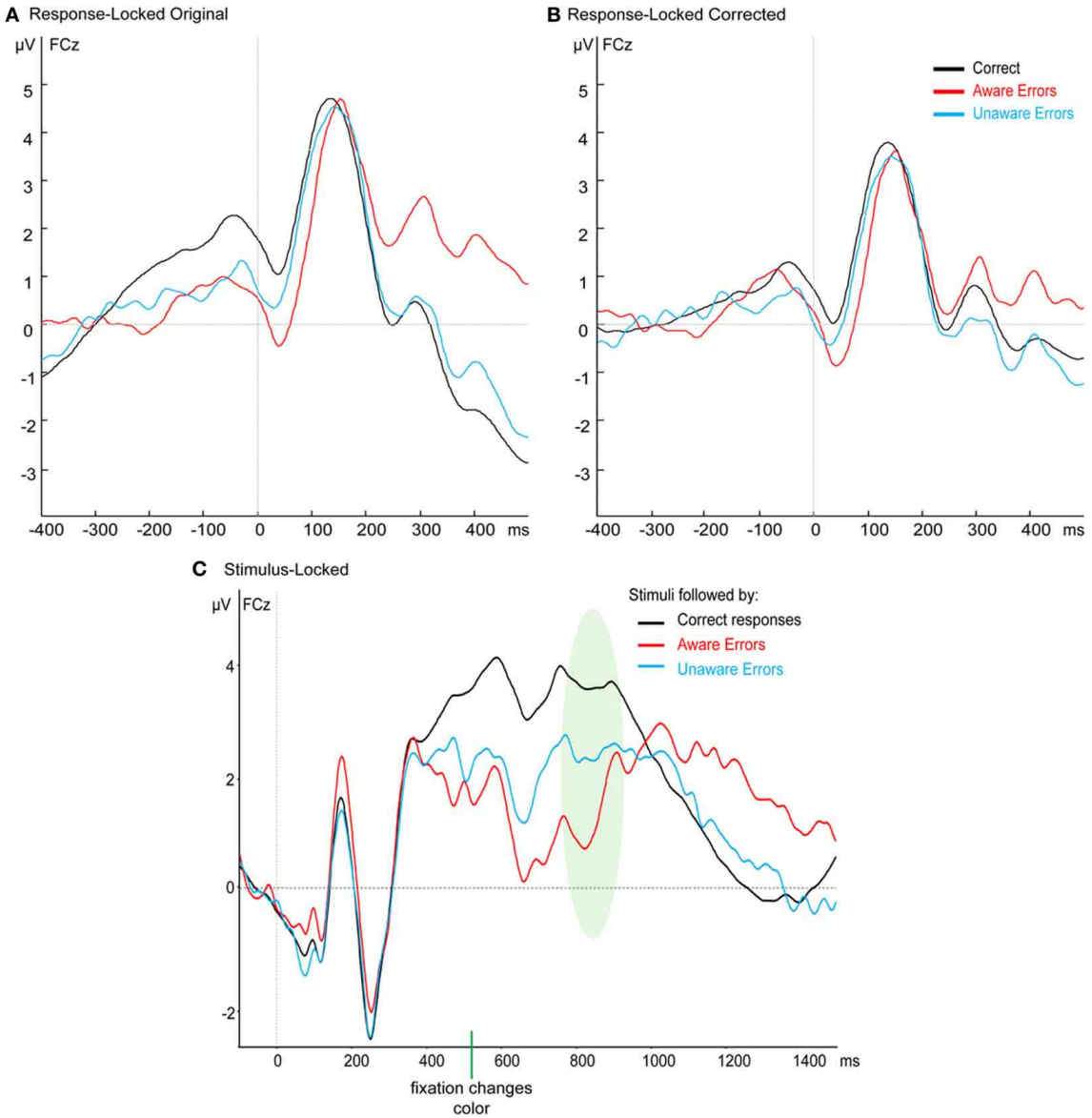


FIGURE A1 | Grand average waveforms at FCz of **(A)** the original response-locked results, **(B)** the response-locked results after the subtraction of the stimulus-locked ERPs, and **(C)** the stimulus-locked waveforms. The shaded area in the stimulus-locked waveforms represents the average reaction time ± one standard deviation.

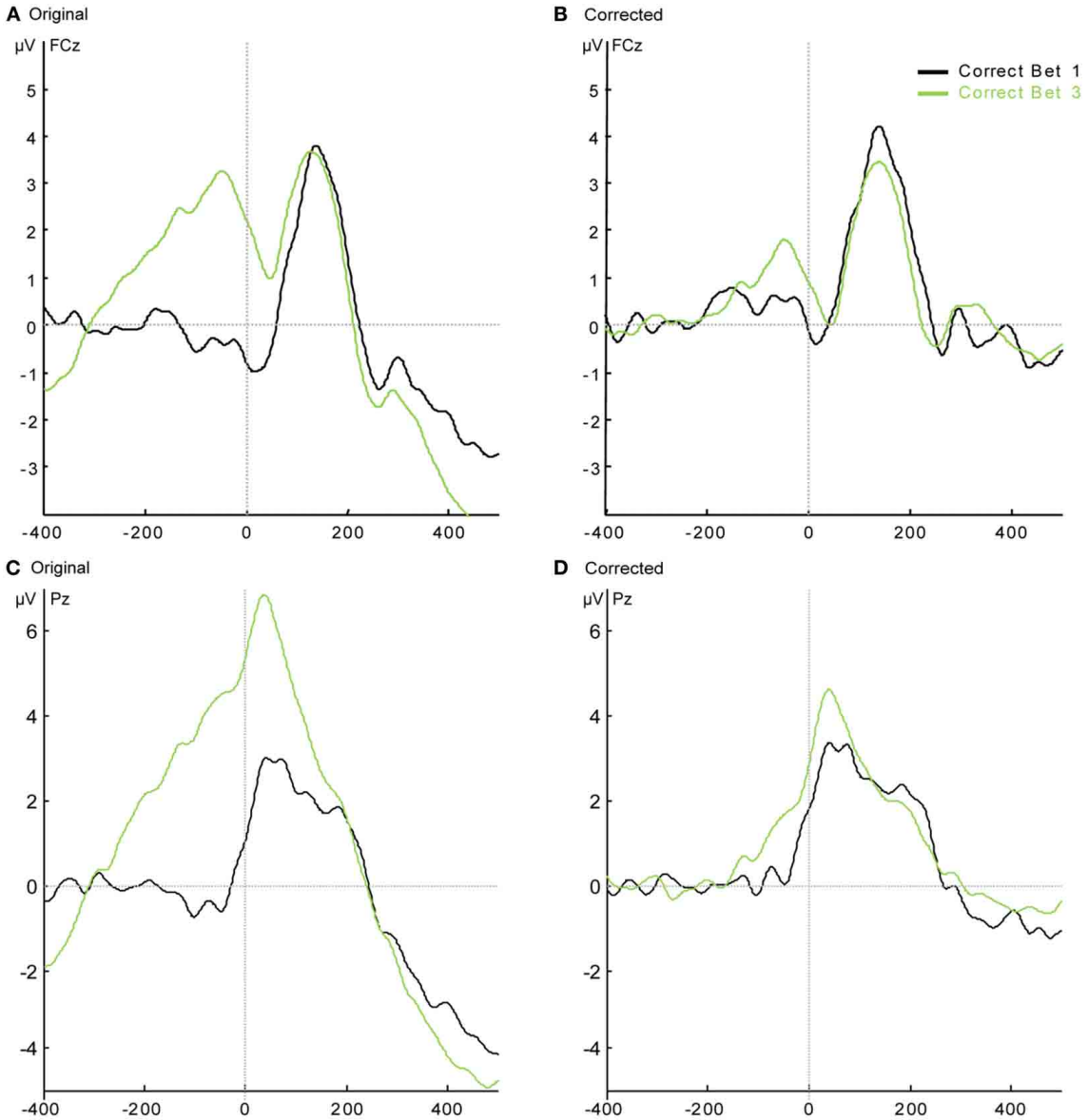


FIGURE A2 | Response-locked grand average waveforms of correct responses of the Low Confidence group. (A) The original results at FCz; **(B)** the results after the subtraction of the stimulus-locked ERPs at FCz. **(C)** The original results at Pz; **(D)** the results after the subtraction of the stimulus-locked ERPs at Pz.

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 28 March 2012
doi: 10.3389/fnhum.2012.00065

An electrophysiological signal that precisely tracks the emergence of error awareness

***Peter R. Murphy*[1]*, *Ian H. Robertson*[1], *Darren Allen*[1], *Robert Hester*[2] *and* *Redmond G. O'Connell*[1]**

[1] *Trinity College Institute of Neuroscience and School of Psychology, Trinity College Dublin, Dublin, Ireland*
[2] *Department of Psychological Sciences, University of Melbourne, Melbourne, VIC, Australia*

Edited by:
Tilmann A. Klein, Max Planck Institute for Human Cognitive and Brain Sciences, Germany

Reviewed by:
Marco Steinhauser, University of Konstanz, Germany
Sander Nieuwenhuis, Leiden University, Netherlands

***Correspondence:**
Peter R. Murphy, Trinity College Institute of Neuroscience, Lloyd Building, Trinity College Dublin, Dublin 2, Ireland.
e-mail: murphyp7@tcd.ie

Recent electrophysiological research has sought to elucidate the neural mechanisms necessary for the conscious awareness of action errors. Much of this work has focused on the error positivity (Pe), a neural signal that is specifically elicited by errors that have been consciously perceived. While awareness appears to be an essential prerequisite for eliciting the Pe, the precise functional role of this component has not been identified. Twenty-nine participants performed a novel variant of the Go/No-go Error Awareness Task (EAT) in which awareness of commission errors was indicated via a separate speeded manual response. Independent component analysis (ICA) was used to isolate the Pe from other stimulus- and response-evoked signals. Single-trial analysis revealed that Pe peak latency was highly correlated with the latency at which awareness was indicated. Furthermore, the Pe was more closely related to the timing of awareness than it was to the initial erroneous response. This finding was confirmed in a separate study which derived IC weights from a control condition in which no indication of awareness was required, thus ruling out motor confounds. A receiver-operating-characteristic (ROC) curve analysis showed that the Pe could reliably predict whether an error would be consciously perceived up to 400 ms before the average awareness response. Finally, Pe latency and amplitude were found to be significantly correlated with overall error awareness levels between subjects. Our data show for the first time that the temporal dynamics of the Pe trace the emergence of error awareness. These findings have important implications for interpreting the results of clinical EEG studies of error processing.

Keywords: error positivity, EEG, error awareness, error processing, performance monitoring

INTRODUCTION

The ability to detect errors is an essential prerequisite for adaptive behavior, signaling that performance levels are inadequate to achieve current goals. Our understanding of the neural networks involved in such a process has greatly increased in recent decades (Botvinick et al., 2004; Ridderinkhof et al., 2004; Yeung et al., 2004), but researchers have not typically made the important distinction between error detection and conscious error awareness. Diminished error awareness greatly limits the extent to which corrective behavior can be initiated and maintained in the long-term and has been linked to loss of insight and symptom severity in several clinical populations (Mintz et al., 2004; O'Keeffe et al., 2004; Larson and Perlstein, 2009; O'Connell et al., 2009a; Perez et al., in press). Early scalp electroencephalography (EEG) studies in humans revealed the existence of two distinct event-related signals associated with the processing of action errors: the error-related negativity (ERN; Falkenstein et al., 1990; Gehring et al., 1993), a fronto-central deflection peaking 20–100 ms after the erroneous response, and the error positivity (Pe; Falkenstein et al., 1991, 1995), a late (300–500 ms) positive wave that is maximal at centro-parietal electrodes. Of the relatively small number of studies that asked participants to explicitly signal any errors they made, a majority have reported that the amplitude of the ERN is unaffected by error awareness whereas the Pe is only present on error trials that are consciously perceived as such (Nieuwenhuis et al., 2001; Endrass et al., 2005; Overbeek et al., 2005; O'Connell et al., 2007; Shalgi et al., 2009; Dhar et al., 2011).

Extensive research has suggested that the ERN signal may reflect an early detection mechanism sensitive to response conflict (van Veen and Carter, 2002; Yeung et al., 2004), or changes in probability or expectation of reward (Holroyd and Coles, 2002; Holroyd et al., 2004). Despite its robust link to conscious error perception, however, the specific neural mechanism represented by the Pe remains a matter of considerable debate (Overbeek et al., 2005; Ridderinkhof et al., 2009). The Pe has variously been suggested to reflect conscious recognition that an error has occurred (Nieuwenhuis et al., 2001; Endrass et al., 2005), a P3b-like potential in response to the motivational significance of an error (Leuthold and Sommer, 1999; Ridderinkhof et al., 2009), delayed stimulus processing (Shalgi et al., 2009) and, most recently, the accumulation of evidence that an error has occurred (Steinhauser and Yeung, 2010; see also Ullsperger et al., 2010; Wessel et al., 2011). Although much of the literature supports the proposal that the presence of a Pe component is an important prerequisite to awareness of an error, this characterization is more descriptive than mechanistic (Ridderinkhof et al., 2009). Disambiguating whether the Pe may reflect processes that contribute to, or result

from, the emergence of error awareness is a critical question that is only beginning to be addressed (see Steinhauser and Yeung, 2010). If the Pe does reflect the emergence of error awareness then one clear and testable prediction which follows is that the temporal dynamics of this component should closely relate to the time at which an error is perceived.

Behavioral studies have measured the timing of error detection using two different types of speeded response (Rabbitt, 1990, 2002; Steinhauser et al., 2008). In the first kind, errors must be signaled via a response which is not mapped to any stimulus (Error Signaling Response, ESR). In the second, errors are signaled by pressing the correct response after error commission (Error Correction Response, ECR). There is evidence to suggest that the ESR is a more appropriate measure when investigating error awareness: error correction is typically associated with significantly faster response times (RTs) than error signaling (some ECRs occur as early as 40 ms after initial error commission; Rabbitt, 1966a,b), and it is unclear whether all ECRs reflect true error detection or merely a delayed activation of the correct response without explicit detection of the initial error (Rabbitt, 2002). By contrast, a correct ESR requires a switching of response sets which is *contingent* on detection of the initial response as erroneous. Correct ESRs are, therefore, unambiguously characterized by correct error detection and, presumably, by awareness that an error has been committed. For this reason, we employed a speeded awareness press analogous to the ESR as our marker of the timing of error awareness.

Although behavioral studies of the error detection process have frequently employed such speeded awareness responses, almost all event-related potential (ERP) studies to date have enforced a delay between error commission and the subsequent indication of error awareness. One reason for the imposition of such a delay is that the additional motor activity related to error signaling would introduce topographical and morphological distortions within the latency range of the Pe (e.g., Colebatch, 2007). This aspect of experimental design has precluded any investigation of the relationship between the Pe and the timing of error awareness. One notable exception compared error-evoked ERPs from a condition requiring speeded ESRs after error commission to those from a condition requiring ECRs (Ullsperger and von Cramon, 2006), but the relationship between error signaling latency and the Pe component was not examined.

Another limitation of the majority of Pe studies to date has been a tendency to analyze the *average* ERP signal only (though see Steinhauser and Yeung, 2010; Hughes and Yeung, 2011 for recent exceptions), which discards much of the variability in the original EEG and is not always representative of trends at the single-trial level of analysis (Arieli et al., 1996; Debener et al., 2006; Eichele et al., 2010; Bland et al., 2011). Analysis of the Pe has, therefore, typically been limited to a constrained latency window as derived from the average ERP, which fails to take much of the inherent variability in component amplitude and latency into account. Hence, while it is clear that the Pe is exclusively elicited by consciously perceived errors, previous research has been largely unable to go beyond this "binary" characterization to examine the extent to which variation in Pe amplitude and latency relates to fluctuations in the timing of awareness. This presents a significant problem in interpreting the findings of clinical studies in which group differences in Pe morphology have been reported (e.g., Brazil et al., 2009; O'Connell et al., 2009a; Olvet et al., 2010; Larson et al., 2011; Luijten et al., 2011; Perez et al., in press; Peterburs et al., 2012).

The present study attempted to address these issues by quantifying the relationship between Pe latency and the precise timing of error awareness. Trial-by-trial variations in the timing of awareness were measured via the introduction of a speeded awareness press (analogous to an ESR) to a previously validated error awareness paradigm, the Error Awareness Task (EAT) (Hester et al., 2005, 2012; O'Connell et al., 2007; Shalgi et al., 2007, 2009). To eliminate possible motor confounds from our analysis we utilized independent component analysis (ICA) to decompose the EEG into orthogonal independent components (ICs) which were generated by distinct neural sources (Makeig et al., 2004; Onton et al., 2006). ICA, therefore, enabled the parsing of neural activity uniquely related to the Pe from that related to the execution of the speeded awareness response and other co-incident neuro-cognitive phenomena extraneous to error awareness. After isolating the Pe in this way, we demonstrate via a combination of within-subjects single-trial analyses, ROC classification analysis and between-subjects correlations that this component is closely tied to the latency of the awareness response, suggesting it provides an index of the emergence of error awareness.

STUDY 1

MATERIALS AND METHODS

Participants

Twenty-nine participants took part in this study. Two were excluded due to poor accuracy on the task (<30% correctly withheld No-go trials). A further participant with no observable Pe component was also excluded from all analyses. This left a final sample of 26 participants (14 female), with a mean age of 23.1 years (SD = 5.4). All participants were right-handed, had normal or corrected-to-normal vision, no history of psychiatric illness or head injury, and reported no history of color-blindness. All participants were asked to refrain from ingesting any caffeine on the day of testing. They provided written informed consent before testing began, and all procedures were approved by the Trinity College Dublin ethics committee and in accordance with the Declaration of Helsinki. Participants received a gratuity of €20 to cover any expenses incurred on the day of testing.

Study design

We employed the EAT, (see **Figure 1**) developed by Hester and colleagues (Hester et al., 2005, 2012). The EAT is a Go/No-go response inhibition paradigm in which a serial sequence of color words is presented to participants, with congruency between the font color of the word and its semantic content manipulated across trials. In the iteration of the paradigm used here, subjects were required to respond as quickly as possible with a single "A" button press in situations where the semantic content of the word and the color in which it was presented were incongruent (Go trial), and to withhold this response in two different circumstances: (1) when the word presented on the current trial was the same as that presented on the previous trial ("Repeat"

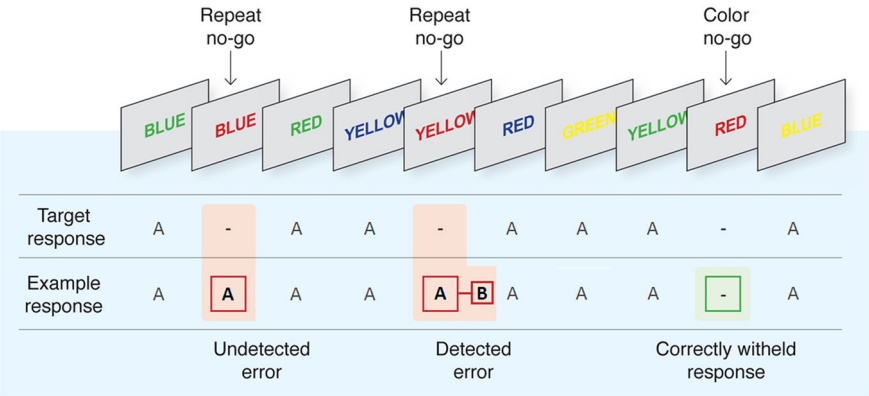


FIGURE 1 | The Error Awareness Task [EAT; Hester et al. (2005, 2012)]. Participants were required to make a speeded button press ("A") to all incongruent stimuli (word and color not matching) and to withhold from responding to congruent stimuli or when a word was repeated on consecutive trials. Participants were also required to press a different button ("B") as soon as possible following commission errors to indicate error awareness.

No-go), and (2) when the meaning of the word and its font color matched ("Congruent" No-go). In the event that participants failed to withhold to either type of No-go stimulus (a commission error; referred to hereafter as an "error press"), they were required to press a second "B" button *as quickly as possible* when they realized their error (referred to hereafter as an "awareness press"). The inclusion of this speeded awareness press represents a departure from previous iterations of the EAT (Hester et al., 2005, 2012; O'Connell et al., 2007; Shalgi et al., 2007, 2009) and other error awareness paradigms (e.g., Nieuwenhuis et al., 2001) which required participants to delay this response for a fixed time. This aspect of the study design was advantageous as it provided a trial-by-trial measure of the timing of error awareness. Participants were instructed to perform the task "as quickly and as accurately as possible."

Each participant was required to complete an automated training session which provided a standardized set of instructions and practice protocols in three separate steps. First, participants were asked to make speeded presses to a sequence of 10 standard Go trials. Second, participants were presented with a sequence of 18 Go trials interceded by both a Repeat No-go stimulus and a Congruent No-go stimulus to which they were asked to withhold responding and to signal any errors. In the event that any stimulus was responded to inappropriately (by withhold on a Go trial, error press on a No-go trial and/or lack of awareness press following this error) the participant was automatically provided with feedback outlining their error and forced to repeat that training protocol until perfect accuracy was achieved. Third, participants progressed to an extended practice session without performance feedback, which lasted approximately 4 min and allowed the experimenter to assess the extent to which the participant was capable of performing the task without further instruction. In the event of any obvious persisting problems in task performance during this practice session (low accuracy/no awareness press to errors), the entire training protocol was repeated. The large majority of participants were able to complete the practice session without any repetition.

All instructions and stimuli were presented using the "Presentation" software suite (NeuroBehavioral Systems, San Francisco, CA). Participants were instructed to use the thumb of their right hand for both "A" and "B" button responses (Microsoft "Sidewinder" Controller). Every participant was administered at least eight blocks of the EAT. Where possible (allowing for time constraints), we administered more blocks in order to maximize the number of error trials available for analysis. On average, participants completed of 9.5 (SD = 0.7) blocks (range 8–10). Each block consisted of 224 word presentations, 200 of which were Go stimuli and 24 of which were No-go stimuli (12 Repeat No-gos and 12 Congruent No-gos). All stimuli were presented for 400 ms, followed by an inter-stimulus interval of 1600 ms. The duration of each block was, therefore, approximately 7.5 min. Stimuli were presented in a pseudo-random order, with a minimum of three Go trials between any two No-go trials. Stimuli appeared 0.25° over a white fixation cross and on a gray background. Although the results are not reported here, pupil diameter was recorded throughout task performance (Eyelink 1000, SR Research). Participants used a table-mounted head-rest which fixed their distance from the computer monitor at 80 cm for the entire duration of the task. They were also instructed to maintain fixation at the fixation cross during task performance in order to minimize eye movements.

The EAT contains three main trial-types: correct "Go trials," where an incongruent Go stimulus was followed by a correct "A" button press; "unaware errors," where either type of No-go stimulus (Repeat or Congruent) was followed by a failure to withhold an "A" button press (i.e., an error press) and no subsequent awareness press; and, most importantly, "aware errors," where an initial error press to a No-go stimulus was followed by an awareness press before the onset of the next stimulus. Trials where the awareness press occurred after the onset of the next stimulus were counted as an aware error when calculating participants' average behavioral measures of error awareness, but were omitted from all ERP analyses. All mean values below are quoted ± SD. In order to maximize trial numbers in our analyses, and because

there was no significant difference in average Aware RT between Congruent No-go trials (564.6 ± 93 ms) and Repeat No-go trials (543.4 ± 76 ms; $p = 0.11$), we did not distinguish between the two trial types in any analyses.

EEG acquisition and processing

Continuous EEG was acquired using an ActiveTwo system (BioSemi, Netherlands) from 64 scalp electrodes, configured to the standard 10/20 setup and digitized at 512 Hz. Vertical and horizontal eye movements were recorded using two vertical electro-occulogram (EOG) electrodes placed above and below the left eye and two horizontal EOG electrodes placed at the outer canthus of each eye, respectively. EEG data were processed using Matlab via the EEGLAB toolbox (Delorme and Makeig, 2004). Continuous EEG data were re-referenced offline to the average reference, high-pass filtered to 0.5 Hz, low-pass filtered up to 95 Hz and notch-filtered at 50 Hz. Noisy EEG channels were interpolated using spherical spline interpolation. Response-locked data were epoched from 400 ms before to 1600 ms after correct Go responses and error presses, and were baseline-corrected relative to the interval −400 to −200 ms. Waveforms locked to the speeded awareness press were also calculated by epoching data on aware error trials from 500 ms before to 100 ms after this event, using the same pre-error press baseline period. Preliminary artifact rejection was performed using an epoch rejection algorithm in the Fully Automated Statistical Thresholding for EEG artifact Rejection (FASTER) toolbox (Nolan et al., 2010), which identifies artifactual epochs based on their deviation from the mean of each channel, their variance and their amplitude range. Any epoch which exceeded a threshold of ±3 standard deviations on any of these measures was discarded. Data from all electrodes were then subjected to temporal ICA using infomax (Bell and Sejnowski, 1995) and implemented in EEGLAB (Delorme and Makeig, 2004). ICs dominated by EOG artifacts or other noise transients were identified by the FASTER automatic ICA component rejection feature (Nolan et al., 2010), again applying a threshold of ±3 standard deviations to several IC measures including the median IC gradient, spectral slope and correlation with EOG activity. ICs found to violate any of these criteria were subtracted from the EEG.

Inspection of the grand-average response-locked ERP on aware error trials revealed a large positivity maximal at CPz/Pz at a latency of approximately 400 ms post-error, consistent with the classic Pe (Falkenstein et al., 2000; Overbeek et al., 2005). This component was not present on unaware error trials. The ICA decompositions were subsequently employed to parse neural activity on aware error trials related to the Pe and error awareness from that related specifically to the introduction of the speeded awareness press. The initial ICA decomposition was conducted on response-locked data including correct Go trials and all commission errors, but only aware error trials and the corresponding IC activation time-courses therein were used for the selection of Pe ICs, since the Pe was only present in this condition. Previous research across different neuroimaging modalities has implicated multiple candidate brain regions in error awareness and in generating the Pe (van Veen and Carter, 2002; Hermann et al., 2004; Hester et al., 2005; van Boxtel et al., 2005; Klein et al., 2007; O'Connell et al., 2007). It is, therefore, unlikely that ICA decomposition would delineate one distinct IC capturing all the variance in this component. For this reason, we did not limit our search for Pe ICs to one IC per participant. Pe ICs were selected manually for each participant (see Onton et al., 2006 for manual IC selection applied to the P300 component). For an IC to warrant selection, it was required to have a positive peak within the average latency range of the Pe (200–600 ms) when back-projected to original EEG space, and a scalp weighting which was comparable to the average Pe topography for that participant within that latency range (see Results for description of an automated IC selection procedure which produced very similar results). As a further constraint on the selection process, we used the DIPFIT 2 plugin for EEGLAB to localize the equivalent dipole locations of all IC scalp maps. Any IC which was localized to left motor or pre-motor cortex (contralateral to the responding hand) was excluded from consideration. A mean of 2.96 ± 1.0 ICs were selected per participant for further analysis, and their activation profiles summed to estimate variation in the EEG uniquely related to the Pe component for that participant (hereafter referred to as "Pe_C"). **Figure 2A** shows the average Pe_C topography and time-course across all participants when back-projected to original EEG-space, and **Figure 3C** highlights individual Pe_C topographies and time-courses for a sample of four individual participants.

Measures and analysis

The task yielded several behavioral measures of interest. Accuracy was calculated as the percentage of correct withholds over all No-go trials. Error Awareness was calculated as the percentage of Aware Errors out of all commission errors made. Error RT was defined as the latency between stimulus onset and error press, and calculated for both aware errors and unaware errors. Lastly, Aware RT was defined as the latency of the speeded awareness response relative to the preceding error press. For both RT measures, trials which were characterized by outlier values of ± 3 standard deviations from the mean were excluded from all ERP analyses.

In all cases the Pe was measured by averaging the activation profile of a cluster of centro-parietal electrodes (CPz, CP1, CP2, Pz, P1, P2), within predefined latency windows which varied between analyses. For the initial comparison of average Pe_C waveforms locked separately to the error press and the awareness press, Pe_C amplitude was defined as the maximum amplitude 200–600 ms post-error press in the former, and the maximum amplitude in the 400 ms window preceding the awareness press in the latter. Peak amplitude measures were employed here to facilitate the direct comparison of component amplitude between waveform-types; a *mean* amplitude measure was not appropriate because any chosen latency window over which to average would not have been directly comparable across waveforms which are locked to events with independent temporal distributions. All further analyses focused exclusively on waveforms locked to the error press. For analyses examining such waveforms *averaged across trials*, component amplitude was defined as the mean voltage from 200 ms to 600 ms post-error. By contrast, single-trial measurements were taken from a latency window that extended

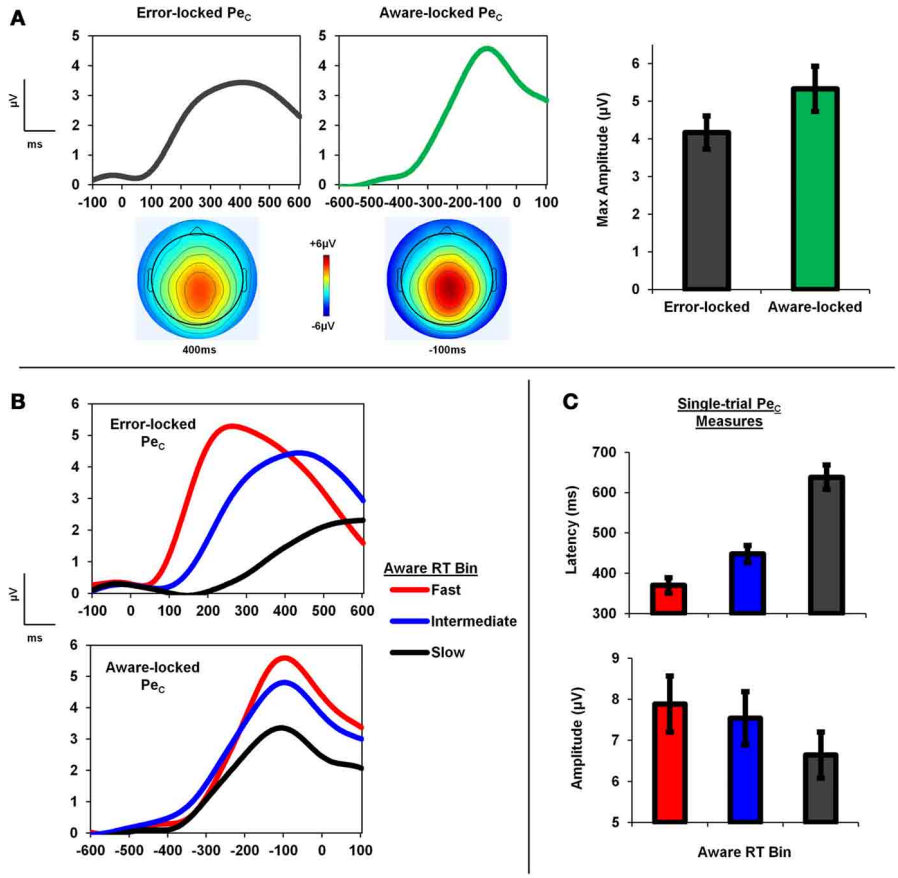


FIGURE 2 | Pe$_C$ is locked to awareness, not error commission. (A) Group average topography and time-course of the Pe$_C$ when locked to both the error press and subsequent awareness press on aware error trials. Note the identical topographies, but greater component amplitude when locked to the timing of the awareness response. **(B)** Pe$_C$ waveforms binned according to Aware RT and locked to the initial error press and subsequent awareness press. **(C)** Single-trial Pe$_C$ latency increased as Aware RT slowed, and single-trial Pe$_C$ amplitude decreased.

from 200 ms post-error press to the slowest awareness press for each participant (mean = 1196 ± 174 ms after error press), and so varied across participants. Single-trial Pe$_C$ latency and amplitude were defined as the timing and amplitude, respectively, of the maximum voltage within this broad latency range on each trial. For all post-ICA analyses, a low-pass filter of 6 Hz was applied to the data in order to improve measure reliability (Spencer, 2004).

Primary analyses focused on the comparison of Pe$_C$ waveforms locked to both the error press and subsequent awareness press in order to determine the event to which the Pe$_C$ was most closely time-locked. We then quantified the strength of the relationship between the Pe$_C$ and the timing of awareness via within-subjects, trial-by-trial correlations of Pe$_C$ latency with Aware RT. We also employed receiver-operating-characteristic (ROC) curve analysis to determine the degree to which information contained in the Pe$_C$ could be used to accurately classify commission error trials as with or without awareness (see Results for details).

A key advantage of ICA decomposition is that it also allows for the complete removal of the variance in the EEG associated with each distinct IC. This is seen most often when ICA is employed for the identification and removal of eye-blink and other ocular artifacts (e.g., Viola et al., 2009), though the same principle applies to ICs which are uniquely related to specific cognitive processes. This enabled us to completely remove all variance associated with the Pe$_C$ from aware error ERP waveforms and compare the residual potentials to those evoked on unaware error trials. In order to avoid the possible confounding influence of differing trial numbers across conditions when directly comparing the aware and unaware traces, aware error trials included in this analysis (mean trial count after artifact rejection = 64.7 ± 17.4) were randomly matched for number of unaware error trials (mean trial count after artifact rejection = 32.8 ± 19.6). Measures were extracted for each of 100 iterations of this random trial-matching procedure, and statistical analysis was carried out on the average measures across all iterations. For this analysis, three participants were excluded because they committed less than 10 unaware errors over the entire testing session.

Statistical tests consisted of bivariate correlations (Pearson's r), paired-samples t-tests and repeated-measures analyses of variance (ANOVAs) where appropriate. Greenhouse–Geisser corrected degrees of freedom were used in cases of violated sphericity with corrected p-values reported.

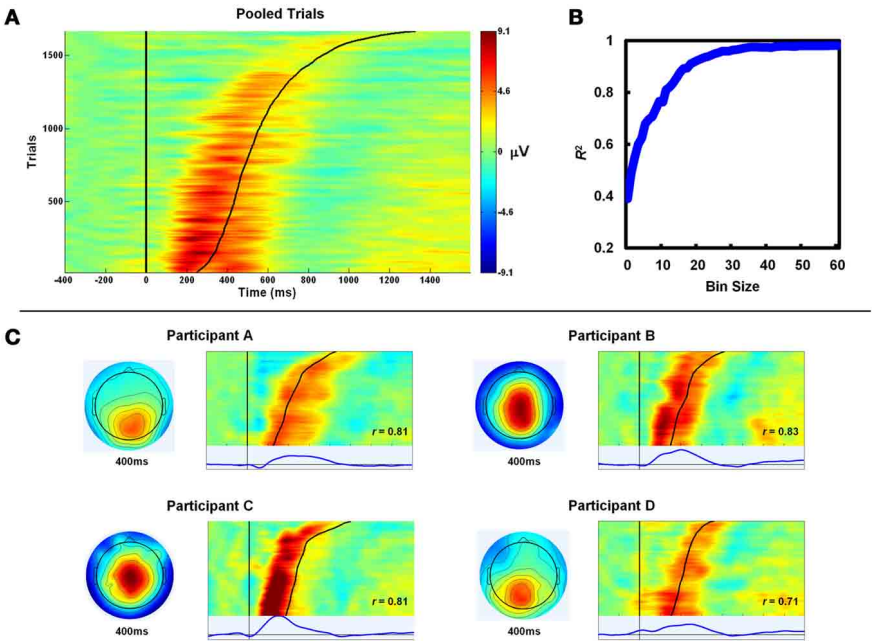


FIGURE 3 | Pe_C latency correlates with the timing of awareness. **(A)** Single-trial aware error Pe_C waveforms locked to the erroneous response, pooled across participants and sorted by Aware RT (black line). For visualization, trials were smoothed across bins of 30 trials with a Gaussian-weighted moving average. Note how Pe_C peak latency closely tracks Aware RT. **(B)** To minimize the effect of noise in the single-trial measures on the observed relationship, the pooled single-trials were again sorted by Aware RT and, through repeated iterations, averaged across bins of increasing size (1–60; see text for details). The variance in Aware RT explained by Pe latency (R^2) increased with bin size up to approximately 25 trials. **(C)** Topographies and trial sortings (including average component waveforms) for a subset of four participants. *R*-values at inset refer to correlation coefficients between single-trial Pe_C latency and Aware RT for that participant.

Table 1 | Study 1 behavioral measures.

	Mean (SD)
Accuracy (%)	56.1 (11.8)
Awareness (%)	71.1 (15.4)
Go RT (ms)	521.6 (116.6)
Error RT (aware errors; ms)	490.0 (107.3)
Error RT (unaware errors; ms)	569.0 (133.9)
Aware RT (ms)	611.0 (88.2)

RESULTS

Behavior

Relevant behavioral measures from Study 1 are highlighted in **Table 1**. Mean Error RT on aware error trials was significantly faster than mean correct Go RT ($t_{25} = 4.03$, $p < 0.001$), which in turn was faster than mean Error RT on unaware error trials ($t_{25} = 3.97$, $p = 0.001$). It is also noteworthy that there was considerably greater within-subject variability (SD) in Aware RT (236.3 ± 64.5 ms) compared to Go RT (143.6 ± 40.5 ms) and Error RT on both aware (119.5 ± 33.4 ms), and unaware error trials ($124/5 \pm 60.7$ ms; all $p < 0.001$). Trial-by-trial variance in Error RT did not relate to variance in Aware RT (mean within-subjects $r = -0.01 \pm 0.03$).

Pe_C latency correlates with the timing of error awareness

We first sought to determine whether the Pe_C was better related to the timing of initial error commission or to the timing of error signaling, and examined this question by comparing component amplitude in the average aware error waveforms locked to the initial error press and the subsequent awareness press. We reasoned that the weaker the temporal relationship between the Pe_C and a given event is, the lower the amplitude of the respective average waveform should be due to increased latency jitter. We found that peak Pe_C amplitude in the average waveforms locked to the awareness press ($5.33 \pm 3.06\,\mu V$) was significantly greater than peak amplitude as measured from the error-locked waveforms ($4.17 \pm 2.24\,\mu V$; $t_{25} = 4.48$, $p < 0.001$; **Figure 2A**).

To explore the relationship between the Pe_C and Aware RT, single-trial Pe_C latency and amplitude values on aware errors were sorted according to Aware RT, and divided into three bins consisting of Fast, Intermediate, and Slow Aware RT trials. A One-Way repeated-measures ANOVA revealed a significant main effect of bin on Pe_C latency [$F_{(1.5, 37.9)} = 104.64$, $p < 0.001$], driven by faster latencies in the Fast compared to the Slow Aware RT bins as revealed by planned *post-hoc* contrasts ($t_{25} = 11.77$, $p < 0.001$; **Figure 2C**). A separate ANOVA also revealed a significant main effect of Aware RT bin on Pe_C amplitude [$F_{(1.7, 43.4)} = 23.81$, $p < 0.001$], which was due to larger component amplitudes in the Fast compared to the Slow bins ($t_{25} = 5.68$, $p < 0.001$; see **Figure 2C**).

To better quantify the strength of the relationship between the Pe_C and the timing of the speeded awareness press we conducted within-subjects, trial-by-trial correlations of Pe_C latency with Aware RT. On average, mean single-trial Pe_C latency

(487 ± 109 ms) occurred significantly earlier than mean Aware RT measured across the same trials (573 ± 90 ms; $t_{25} = 4.36$, $p < 0.001$). As **Figure 3A** illustrates, however, there was a robust trial-by-trial correlation between both measures (mean $r = 0.60 \pm 0.21$; $t_{25} = 14.75$, $p < 0.001$). Hence, single-trial Pe_C latency accounted for 36% of the variance in Aware RT. To verify the effects of filtering on this result, we repeated the same analysis using a higher cut-off for the low-pass filter (30 Hz). This change produced a comparable but smaller correlation between single-trial Pe_C latency and Aware RT (mean $r = 0.49 \pm 0.19$), indicating the use of a 6 Hz low-pass filter helped to eliminate noise in the single-trial measures. As is frequently the case with RT measures, the distributions of Pe_C latency and Aware RT were both positively skewed and this may have partly driven the strength of the observed correlation between these variables. To rule out this possibility, their distributions were log-transformed and the correlation coefficients recalculated. Transformation had no effect on the strength of the relationship between these measures (mean r-value using log-transformed distributions: 0.59 ± 0.21).

For comparison, the single-trial, within-subjects correlation coefficients were also calculated using the original Pe waveforms, with all measures defined identically to those employed above. The results showed that the isolation of Pe ICs led to a stronger relationship between peak latency and Aware RT than was observed using the original Pe waveforms (mean r-value: 0.54 ± 0.23) although the difference did not reach statistical significance ($t_{25} = 1.67$, $p = 0.11$).

We also replicated our within-subjects correlational findings using Pe_Cs which were selected using completely objective criteria. For this analysis, all ICs for a given participant were back-projected to EEG space before calculating mean signal amplitude (averaged over the relevant Pe channels defined above) and scalp topography, across the relevant Pe latency window (200–600 ms post-error), for each IC. Spatial correlations between the resulting topographies of each IC and that participant's raw Pe topography, averaged across the same latency range, were subsequently calculated. Both values (spatial correlation and mean amplitude) for each IC were normalized relative to the mean and standard deviation across all ICs for that participant (converted into z-scores), and a combined score was derived for each IC by obtaining the average of the two z-scores. By basing automatic IC selection on this combined score, we placed equal weight on both criteria employed during the manual IC selection process: (1) topographic similarity with the original Pe component, and (2) a readily apparent positive peak within the latency range of the Pe. Using an arbitrary threshold score of $+1.5$ z, a mean of 2.5 (± 1.2) ICs were selected by this procedure, which was slightly lower than the number of ICs selected manually ($t_{25} = 2.13$, $p < 0.05$). A robust correlation between the single-trial latency of the automatically selected Pe_Cs and Aware RT was observed within-participants (mean r-value: 0.58 ± 0.24) which was slightly lower but statistically equivalent to the strength of the relationship observed between these measures after manual IC selection ($t_{25} = 1.25$, $p = 0.22$).

Despite the combination of ICA and low-pass filtering, measuring peak component latency in the single-trial can be inaccurate because of the inherent noise in the EEG at this level of analysis (e.g., Spencer, 2004). Noise may, therefore, have diminished the sensitivity of the within-subjects correlational analyses toward revealing the true strength of the relationship between Pe_C latency and Aware RT. In an effort to circumvent this issue, all single-trial aware error traces were pooled across subjects, sorted by Aware RT and, through repeated iterations, averaged across bins of increasing size (1–60). Separate linear regression analyses were conducted for each iteration with Pe_C latency as the predictor variable and Aware RT as the dependent variable. Here, Pe_C latency was defined as the latency of the maximum voltage of the average Pe_C waveform in each bin, between 200 and1600 ms post-error. Hypothetically, including more trials in this averaging process by increasing bin size should eliminate more of the noise in the single-trial waveforms, and hence should provide a truer representation of the strength of the Pe_C latency/Aware RT relationship. As can be seen in **Figure 3B**, employing larger bin sizes substantially increased the strength of the relationship between Pe_C latency and Aware RT, to a point where they shared greater than 95% of their variance (>25 trials per bin).

Pe_C predicts error awareness

We employed a receiver operating characteristic (ROC) curve analysis (see, for example, Quiroga et al., 2005; Einhauser et al., 2010; Steinhauser and Yeung, 2010) to quantify how well the Pe_C could predict error awareness on a trial-by-trial basis. For this analysis, all IC weights manually selected for the Pe_Cs were also back-projected to EEG-space for the unaware error condition. Classification accuracy was quantified as the area under the curve (AUC; also referred to as Az) of True Positive rate (percentage aware error trials correctly classified) plotted against False Positive rate (percentage unaware error trials incorrectly classified) across all levels of discrimination threshold: accuracy is 50% if classification is at chance based on the information contained in the Pe_C, and 100% if classification is perfect. Separate ROC analyses were conducted for each participant on mean component amplitude information in discrete time bins along the whole Pe_C time-course (window width of 20 ms, moving in 20 ms increments), comparing aware versus unaware error traces. Extracted measures were similar to those for earlier ERP analyses: maximum classification accuracy over the entire time-course (analogous to maximum component amplitude), and the latency at which this occurred relative to error press. Any cases with extracted classification accuracy or latency values exceeding ± 3 SD from the mean were excluded.

Average maximum classification accuracy across participants was $75.2 \pm 7.2\%$. The temporal evolution of awareness-predictive activity closely matched the time-course of the average Pe_C waveform on aware error trials (see **Figure 4A**). Further, average latency of maximum classification accuracy was 445.4 ± 131.2 ms post-error, which is statistically equivalent to the peak latency of the Pe_C when derived from the average component waveforms on aware error trials (432.1 ± 99.8 ms; $t_{21} < 1$, $p = 0.6$).

We also examined the earliest latency post-error at which Pe_C amplitude significantly predicted error awareness. For this analysis, all aware and unaware error trials were pooled across participants and permutation testing was employed to calculate a bootstrapped significance threshold for each discrete ROC time bin along the Pe_C time-course. For each time bin, component

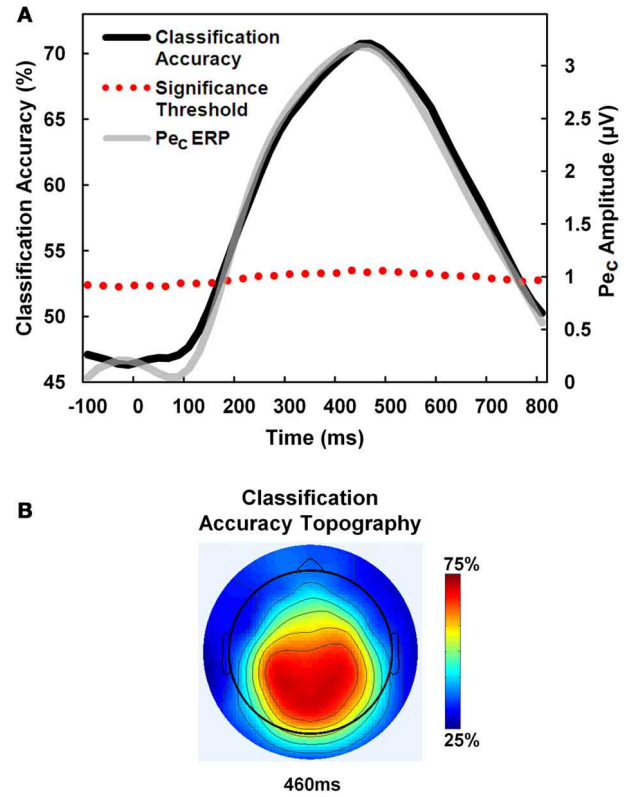


FIGURE 4 | Pe_C amplitude predicts error awareness. A receiver-operating-characteristic (ROC) curve classifier analysis was employed to predict error awareness using the Pe_C. **(A)** Classification accuracy (black line) using pooled single-trials across participants rose above a bootstrapped significance threshold (dotted red line; see text for details) just 180 ms after error commission, and closely traced the time-course of the average Pe_C on aware errors (gray line). **(B)** Classification accuracy was maximal over centro-parietal electrodes.

amplitude values from all trials were randomly assigned to aware and unaware conditions according to the relative proportion of each trial type in the overall pool, and a value for classification accuracy was calculated for that bin. This was repeated 1000 times to estimate a permuted distribution of accuracy values for each time bin representing the null hypothesis that accuracy was not greater than chance. If the observed classification accuracy for a time bin derived from the *unshuffled* trials was greater than 1.96 standard deviations above the mean of the permuted distribution, then component amplitude at that time point was said to significantly predict error awareness at $p < 0.05$. Using this method, we found that the earliest latency at which the Pe_C could reliably predict awareness was 180 ms post-error (see **Figure 4A**).

Between-subjects correlations

Between-subjects bivariate correlations were also carried out to test for relationships between the task-related behavioral measures of error awareness (mean Aware RT; Error Awareness) and our hypothesized neural indices of the emergence of awareness (mean single-trial Pe_C latency and amplitude; see **Figure 5**). The single-trial measures were used here because they provide a better reflection of the characteristics and dynamics of each participant's Pe_C; the average waveforms, by contrast, lose substantial amounts of information pertaining to component amplitude and latency, which may significantly distort or obscure any observed correlations. Outlier participants specific to each correlation were classified as those with studentized deleted residuals above or below 3, and excluded from that analysis.

Mean single-trial Pe_C latency was positively correlated with mean Aware RT ($r = 0.53$, $p < 0.01$), and negatively correlated with participants' Error Awareness ($r = -0.53$, $p < 0.01$). This indicates that participants whose Pe_Cs peaked relatively early were also faster at indicating awareness that they had committed an error, and were generally aware of a greater number of the errors they committed. In addition, mean single-trial Pe_C amplitude showed a strong negative correlation with mean Aware RT ($r = -0.61$, $p < 0.001$), and was weakly positively correlated with Error Awareness ($r = 0.36$, $p = 0.075$). Hence, participants with high-amplitude Pe_Cs were faster at indicating error awareness, and were aware of a greater proportion of the errors they committed.

Pe_C removal from aware errors

Working under the hypothesis that the Pe reflects the emergence of error awareness, we reasoned that if selected Pe_Cs truly contained all of the variation in the EEG related to the Pe component and were subsequently removed from the original aware error waveforms, there should be no difference between the average response-locked ERPs elicited by errors made with and without awareness. The amplitude of the original aware error ERP within the latency range of the Pe (mean amplitude: $4.01 \pm 2.3\,\mu V$) was substantially reduced when all variance associated with the manually selected Pe_Cs was removed ($1.42 \pm 1.1\,\mu V$; $t_{22} = 8.64$, $p < 0.001$) but a small, statistically significant difference in the amplitude of the aware and unaware ($0.82 \pm 1.4\,\mu V$) error waveforms was still observed ($t_{22} = 2.38$, $p < 0.05$; **Figure 6**). Although this suggests that IC selection did not capture the Pe component in its entirety, the difference in amplitude between these waveforms was not substantial ($0.6\,\mu V$).

STUDY 2

To completely rule out the possibility that the observed Pe_C latency/Aware RT correlation in Study 1 was driven by the requirement of an overt, speeded motor response to indicate awareness, a second study that included two EAT conditions was conducted. In condition 1, participants completed the same version of the task reported in Study 1; in condition 2 the same participants completed the EAT without the requirement to signal their errors. One concern when selecting relevant ICs in Study 1 related to the fact that Pe-related activity extends to central scalp locations, and hence may not have been dissociated via ICA from any central motor-related activity. Study 2, by contrast, allowed for the derivation of a set of IC weights in condition 2 that were completely uncontaminated by neural activity related to motor-preparation or task-switching and could subsequently be applied to the EEG data from condition 1 for estimation of the Pe_C.

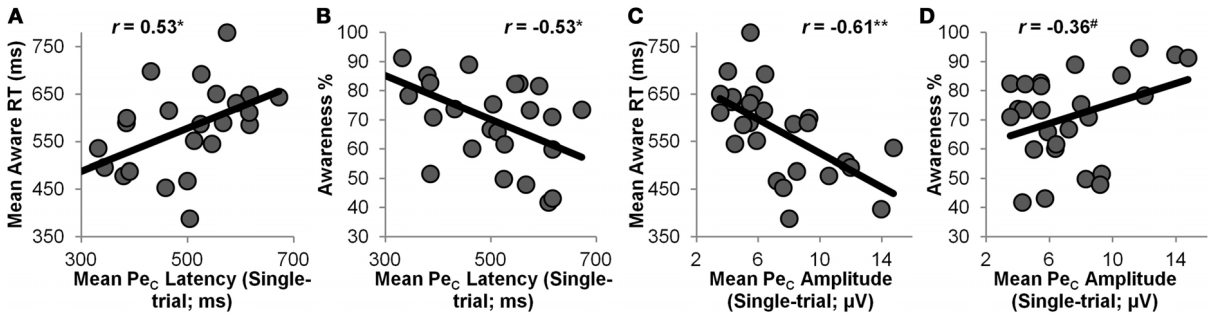


FIGURE 5 | Between-subjects correlations. Mean single-trial Pe_C latency was positively correlated with mean Aware RT **(A)** and negatively correlated with % error awareness over the entire task **(B)**; mean single-trial Pe_C amplitude showed the opposite trends, correlating negatively with Aware RT **(C)** and positively with % awareness **(D)**. * = $p < 0.01$, ** = $p < 0.001$, # = $p < 0.1$.

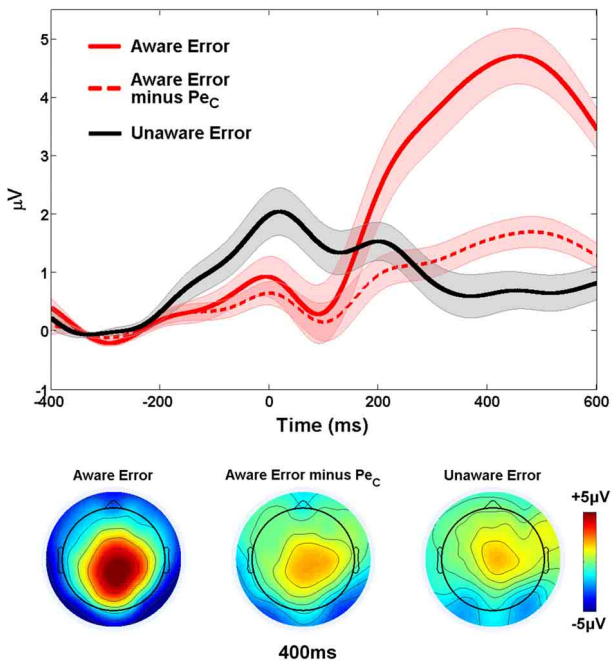


FIGURE 6 | Validation of Study 1 Pe_Cs. The difference in waveform amplitude between aware and unaware errors within the relevant Pe_C latency range was substantially reduced when all variance related to the Pe_C was removed from aware error traces. Shaded regions around waveforms represent error bars (SEM).

MATERIALS AND METHODS

A separate cohort of 16 participants took part in Study 2. One was excluded due to an insufficient number of aware error trials for analysis (<10), and an additional participant was excluded because they did not have an observable Pe component. This resulted in a final sample of 14 participants (four male) with a mean age of 23.6 ± 3.9 years, who met the same inclusion criteria as used in Study 1.

Task design and setup was identical to Study 1, except here all participants completed five blocks of the primary task ("Report" condition) and five blocks in which no awareness press was required even when they were aware of committing an error ("No-Report" condition; order of completion counter-balanced across participants). It was, therefore, impossible to differentiate aware from unaware errors in the No-Report condition, though we can infer, based on the results of the present study and other research employing the EAT (e.g., O'Connell et al., 2007; Shalgi et al., 2009), that the majority of errors in this condition were aware errors. Study 2 therefore, yielded two main trial-types of interest: aware errors, derived from the Report condition and defined as in Study 1 (mean trial count = 27.9 ± 10.9 after artifact rejection); and No-Report errors, defined as any error of commission in the No-Report condition (mean = 35.8 ± 11.0). There were too few unaware errors in the Report condition for reliable analysis (mean = 11.9± 9.5).

EEG acquisition and pre-processing followed the procedures employed in Study 1, with the exception that only response-locked data from correct Go trials and No-Report errors were decomposed via ICA. IC weights from this decomposition were then applied to the response-locked aware error data from the Report condition, and Pe ICs were manually selected based on the resulting IC activation time-courses and scalp topographies in the manner outlined previously. A mean of 3.0 ± 1.4 ICs were selected per participant for further analysis. **Figure 7A** shows the average time-course and topography of the resulting Pe_Cs, across participants, when back-projected to original EEG-space. Single-trial and average component waveform measures were defined in an identical manner to Study 1.

Primary analyses of the Study 2 data were restricted to aware error trials and focused on replicating the within-subjects correlation effects between Pe_C latency and Aware RT found in Study 1. Correlation coefficients were calculated for these analyses using original and log-transformed distributions (see Study 1 Materials and methods). The limited trial count for aware errors in Study 2 prevented the binning of trials by Aware RT as in Study 1. Further analyses of the Study 2 data focused on determining whether or not we successfully isolated relevant Pe ICs by comparing average waveforms from aware error and No-Report error trials both with and without inclusion of the manually selected Pe_Cs (see Study 2 Results for details). These analyses employed the iterative randomized trial-matching procedure described previously, here matching No-Report errors for number of aware errors.

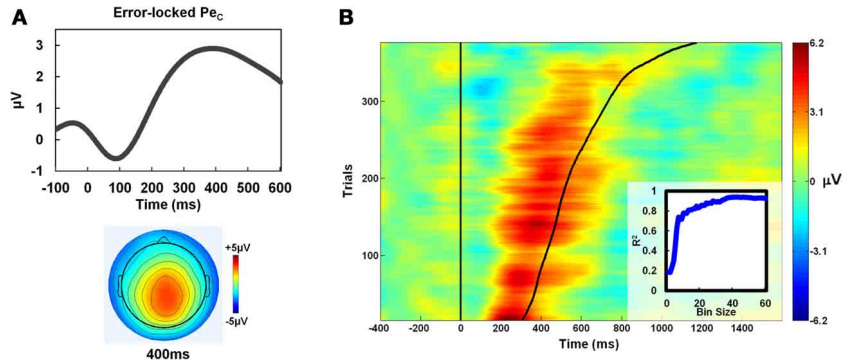


FIGURE 7 | Replication of Pe_C/Aware RT correlation. (A) Group average topography and time-course of the error-locked Pe_C in Study 2, where Pe_Cs were selected using ICA weights derived from a control condition which did not require an awareness press, were comparable to those observed in Study 1 (see **Figure 2**). **(B)** The relationship between Pe_C latency and Aware RT reported in Study 1 (see **Figure 3**) was also apparent in Study 2. This relationship was similarly strengthened as bin size increased using the iterative trial smoothing procedure (shown at inset; compare to **Figure 3B**).

Table 2 | Study 2 behavioral measures; mean (SD).

	Report	No-Report
Accuracy (%)	67.1 (13.3)	67.3 (9.4)
Awareness (%)	76.7 (12.2)	–
Go RT (ms)	521.5 (110.7)	523.8 (96.3)
Error RT (aware errors; ms)	483.0 (85.1)	522.2 (95.8)
Error RT (unaware errors; ms)	600.2 (181.3)	
Aware RT (ms)	640.4 (141.2)	–

RESULTS

Relevant behavioral measures from Study 2 (**Table 2**) were comparable to those observed in Study 1.

As illustrated by **Figure 7B**, the relationship between Pe_C latency and Aware RT reported in Study 1 was also apparent in Study 2. Within-subjects correlations between these measures showed a robust relationship (mean $r = 0.54 \pm 0.20$) that was comparable to that observed in Study 1. This was also the case when the same correlation coefficients were calculated from log-transformed distributions (mean $r = 0.56 \pm 0.19$). Hence, deriving IC weights from the No-Report condition and applying these to the Report condition for Pe IC selection did not affect the strength of the relationship between Pe_C latency and the timing of the awareness response. The iterative trial smoothing procedure employed on the pooled aware error traces in Study 1 also yielded similar results in Study 2: employing larger bin sizes for averaging substantially increased the strength of the Pe_C latency/Aware RT relationship, to the point where they shared greater than 90% of their variance (**Figure 7B**; inset).

We also replicated our within-subjects correlational findings from Study 2 using the automated IC selection protocol outlined previously. A mean of 2.9 (±1.4) ICs were selected by this procedure, which was equivalent to the number of ICs selected manually for Study 2 ($t_{13} < 1$, $p = 0.80$). A robust correlation between the single-trial latency of the automatically selected Pe_Cs and Aware RT was observed within-participants (mean r-value: 0.53 ± 0.23) which, as in Study 1, was slightly lower but statistically equivalent to the strength of the relationship observed between these measures after manual IC selection ($t_{13} < 1$, $p = 0.81$).

Because an unknown number of No-Report error trials are actually unaware errors which do not elicit a Pe component, it was expected that mean amplitude of the average Pe ERP on aware error trials should be greater than the mean amplitude of average No-Report error waveforms. However, we again reasoned that if all Pe ICs were successfully selected and subsequently removed from these original aware error and No-Report error waveforms there should be no difference between the mean amplitudes extracted from the two trial-types, since the variance related to error awareness which was initially driving the amplitude differences should now be eliminated from the data. To address this question, a 2×2 repeated-measures ANOVA tested for ERP amplitude differences with factors of trial-type (aware error vs. No-Report error) and ERP-type (with-Pe_C vs. without-Pe_C). This analysis (see **Figure 8**) revealed main effects of trial-type [$F_{(1, 13)} = 9.09$, $p = 0.01$; driven by greater mean amplitudes on aware error trials compared to no-report error trials] and ERP-type [$F_{(1, 13)} = 10.16$, $p < 0.01$; driven by greater mean amplitudes in ERPs with-Pe_C compared to without-Pe_C], and a significant trial-type X ERP-type interaction [$F_{(1, 13)} = 10.34$, $p < 0.01$]. *Post-hoc* comparisons revealed this interaction to be driven by greater amplitudes on aware error trials (mean amplitude $= 4.17 \pm 2.57\,\mu V$) as opposed to no-report error trials ($2.22 \pm 2.59\,\mu V$) when ERPs incorporated variance related to the Pe_C ($t_{13} = 4.11$, $p = 0.001$), while there was no difference between trial-types when Pe_Cs were subtracted from the ERPs ($1.78 \pm 1.50\,\mu V$ vs. $1.07 \pm 1.14\,\mu V$; $t_{13} = 1.45$, $p = 0.2$). Hence, the selected Pe ICs were highly relevant to error awareness.

Lastly, we tested whether or not the Pe_Cs were contaminated by neural activity related to motor execution. In order to isolate neural activity specific to the awareness press, aware error minus No-Report error difference waveforms (low-pass filtered to 20 Hz) were derived from average error-locked ERPs both with and without Pe_Cs. Lateralized neural activity from 200 ms before to 100 ms after the mean Aware RT for each participant (mean $= 614.4 \pm 132.3$ ms) was examined by comparing mean ERP

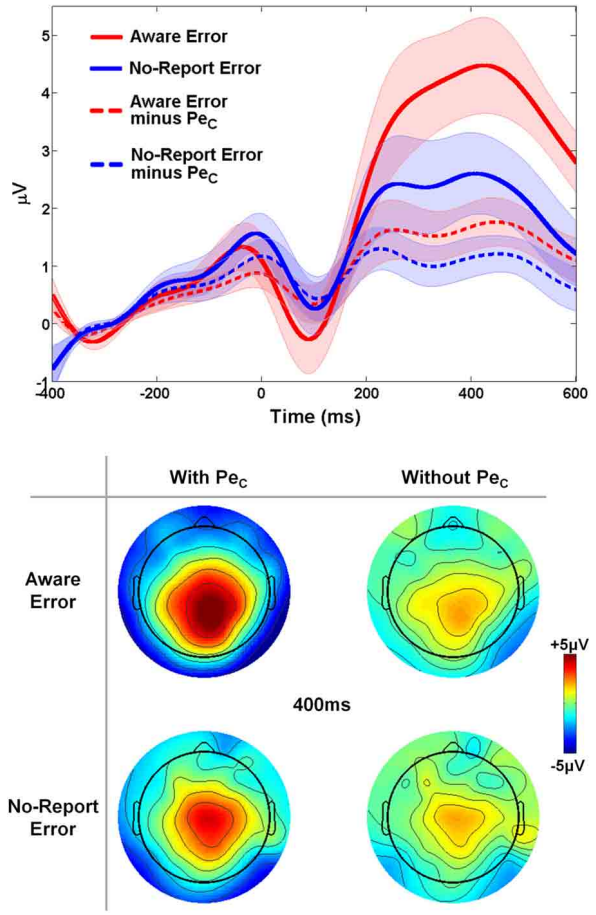


FIGURE 8 | Validation of Study 2 Pe$_C$s. The difference in Pe$_C$ amplitude between aware and no-report errors was eliminated when all variance related to the Pe$_C$ was removed from both traces. Shaded regions around waveforms represent error bars (SEM).

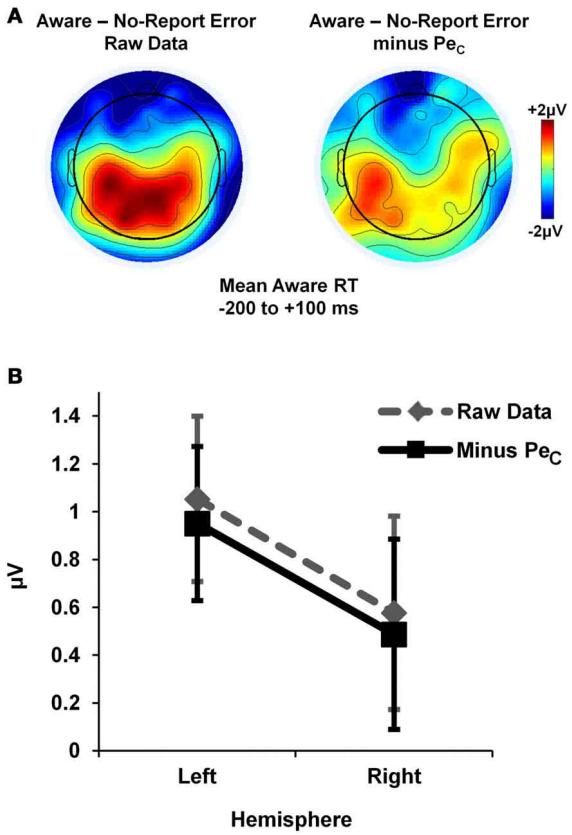


FIGURE 9 | Study 2 Pe$_C$s did not include lateralised motor activity. Motor-evoked activity unique to the awareness press was isolated by deriving error-locked difference waveforms for aware minus no-report errors. Average topographies within a broad Aware RT latency range suggested the presence of lateralized activity **(A)**, which was not affected by the removal of all variance associated with the Pe$_C$ from the difference waveforms **(B)**.

amplitudes across this window at a selection of electrodes corresponding to left (C3, C5, CP5) and right (C4, C6, CP6) motor regions. Amplitude differences were tested via 2 × 2 repeated-measures ANOVA with factors of lateralization (left vs. right hemisphere) and ERP-type (with-Pe$_C$ vs. without-Pe$_C$), which revealed no significant main effect of ERP-type [$F_{(1, 13)} = 1.89$, $p = 0.19$], no main effect of lateralization [$F_{(1, 13)} < 1$, $p = 0.36$; though trends suggested more positive amplitudes over left compared to right hemisphere; see **Figure 9**] and, critically, no lateralization X ERP-type interaction [$F_{(1, 13)} < 1$, $p = 0.96$]. Hence, the presence of lateralized activity over motor regions was not contingent on the inclusion or removal of the Pe$_C$, suggesting Pe$_C$s were uncontaminated by motor-related neural activity.

DISCUSSION

In the present study, participants performed a modified version of the EAT Go/No-go response inhibition task (Hester et al., 2005, 2012) in which error awareness was indicated via a speeded manual response separate from that used for Go stimuli. This aspect of study design provided an opportunity to test the hypothesis that the Pe component reflects the emergence of error awareness by examining the relationship between the Pe and the timing of the awareness response, but the potential for motor activity to confound such an analysis presented a significant methodological challenge. After employing ICA to isolate activity uniquely related to the Pe (the "Pe$_C$"), we demonstrated for the first time that Pe latency closely tracked the latency at which participants indicated awareness. This effect was replicated in a second study, in which Pe ICs were derived from a control condition which did not require an explicit awareness response. The intimate relationship between the Pe and error awareness was further illustrated by findings that the mean amplitude and latency of this component were related to behavioral measures of error awareness between-subjects, and that Pe waveforms could be used to reliably predict error awareness at the single-trial level. Our results highlight the utility of ICA as a methodological control and the benefit of employing a speeded awareness press as a measure of the timing of error awareness, while also having implications for interpreting the functional significance of the Pe.

Isolating the Pe

Despite our use of dipole source modeling to constrain the process, the selection of ICs based on information contained in their topographies and time-courses is somewhat subjective (Onton et al., 2006). The question of whether or not we successfully isolated relevant Pe ICs and avoided those related to extraneous neural signals is therefore an important one. Comparison of error trials with and without awareness suggested that, even after Pe_C removal from aware error waveforms, a small but significant amount of residual variance remained in the EEG which differentiated the two trial-types within the latency range of the Pe. ICA effectively decomposes activity in the EEG into orthogonal components which have discrete underlying neural generators. Given the highly heterogeneous collection of candidate Pe generators and the generally conservative nature of our IC selection process, it is perhaps unsurprising that the selected Pe_Cs did not account for all of the variation in the EEG which was related to the Pe. Nevertheless, we are confident that our IC selection yielded Pe_Cs that reflect the dynamics of the original Pe component and are highly relevant to error awareness. A strong relationship was observed between Pe_C latency and Aware RT which was slightly stronger than that observed in the raw data (a result replicated using completely objective IC selection criteria), and we found that Pe_C amplitude in the single-trial could be used to accurately classify error trials as with or without awareness. Pe_C latency and amplitude were also significantly related to error awareness at the between-subjects level: participants whose Pe_C peaked earlier and was of greater amplitude were faster at indicating error awareness and were also aware of a greater percentage of the errors they committed. These findings all suggest that the Pe_C in the present study represents a neural signal which is tightly coupled to error awareness. In our second study, we demonstrated that any lateralized motor-related activity specifically introduced by the speeded awareness press was still present in the EEG when Pe_Cs were subtracted from the data, and hence did not affect the observed relationship between the Pe_C and the timing of error awareness.

These findings highlight the utility of ICA as a methodological tool which can be used to parse neural activity related to distinct cognitive processes from other stimulus- or response-evoked signals. In the domain of error awareness, we have demonstrated that this tool can be employed effectively to isolate the Pe component from the potential confounds associated with a speeded awareness press, and hence rendered this important behavioral measure readily accessible to future ERP studies.

Pe functional significance

By building a speeded awareness response into our experimental design, we have shown that the latency of the Pe component is closely tied to the moment at which the participant first signals awareness of their actions. Our results have implications for interpreting the functional significance of the Pe. Previous research has often struggled to disambiguate whether the Pe reflects processes that lead to the emergence of error awareness, or is instead one of the consequences of error awareness (Overbeek et al., 2005; but, see Steinhauser and Yeung, 2010). Two novel findings of the present study suggest that the Pe reflects the real-time emergence of error awareness and not one of its sequelae. Firstly, the evolution of the Pe_C comfortably preceded any overt indication of awareness: On average the Pe_C peaked 80–90 ms before the average awareness response. We also found that single-trial Pe_C amplitude was a significant predictor of the presence or absence of error awareness from a latency of up to 400 ms before the average timing of the speeded awareness response (and just 180 ms after initial error commission). Hence, information contained early in the time-course of the Pe_C was predictive of whether or not error awareness would ultimately be achieved. These findings are consistent with the hypothesis that the Pe component reflects conscious recognition that an error has occurred (Nieuwenhuis et al., 2001; Endrass et al., 2005), but also suggest that the emergence of awareness, as reflected in the Pe, is a dynamic and cumulative process which begins at an early stage after error commission. The latter analysis also represents an important replication and extension of the previous finding that single-trial Pe amplitude, as derived by logistic regression classifier analysis as opposed to the present ICA-based approach, was predictive of individual error signaling (Steinhauser and Yeung, 2010).

Recent research (Steinhauser et al., 2008; Steinhauser and Yeung, 2010) has conceptualized the emergence of error awareness as a "decision process," in which awareness is only achieved after sufficient evidence of initial error commission has been accumulated to pass an independent decision threshold (the response criterion). Response criterion was explicitly manipulated in one recent ERP study (Steinhauser and Yeung, 2010) and the findings supported the possibility that Pe amplitude indexes the strength of accumulated evidence, which in turn informs decisions about the accuracy of the preceding response. A potential corollary of this model which was not addressed in that paper is that Pe peak latency marks the time at which the criterion has been met, and hence should be closely tied to the timing of error awareness (assuming a reasonably consistent delay between the decision threshold being passed and categorical decision *output* being generated). Our findings of a strong relationship between Pe_C latency and Aware RT are entirely consistent with this hypothesis, and complement those of Steinhauser and Yeung (2010) in highlighting the importance of both amplitude *and* latency characteristics of the Pe in predicting error awareness.

An interesting question that arises from our study is whether Pe peak latency marks the culmination of the accumulation process or whether evidence accumulation continues after the response criterion has been passed. Although we cannot adjudicate conclusively between these alternatives, two aspects of our data would appear to support the former possibility. First, we found that peak ROC classification accuracy was coincident with Pe_C peak latency and declined immediately thereafter. Second, the time-lag between Pe_C peak latency and the subsequent awareness press (80–90 ms on average) is consistent with empirical estimates in non-human primates (Roitman and Shadlen, 2002; Gold and Shadlen, 2007) of the time it takes for the crossing of a perceptual decision threshold to be transmuted into an overt behavioral response, albeit in a different response modality.

One consequence of the assumption that the proposed evidence accumulation process does indeed cease upon crossing the response criterion is that peak Pe *amplitude* should reflect

the level of the response criterion itself. Interpreting our Pe_C amplitude results in this regard is more complicated. A fundamental characteristic of computational frameworks which model decision processes is that a lower criterion should equate to faster RTs, and this has been found for error signaling latency using a connectionist model of error detection (Steinhauser et al., 2008). Our results may appear to the contrary insofar as Pe_C amplitude was diminished for trials on which Aware RT was comparatively *slow*; hence, trials which might theoretically be characterized by a high response criterion were instead marked by particularly low Pe_C amplitudes. This within-subjects amplitude effect was also reflected in our between-subjects correlations, where participants with the largest Pe components (and, perhaps, the highest response criteria) were actually fastest at indicating error awareness. However, models of perceptual decision-making also posit the onset latency of evidence accumulation and the quality of the decision evidence itself to be other critical parameters affecting the RT distribution (Ratcliff and Smith, 2010). The extent to which these factors may interact with response criterion in driving variability in the latency of error signaling is unknown, and it is difficult to draw concrete inferences from our Pe_C amplitude effects without measuring them in some way. Further, it should be noted that the neural mechanisms for accumulating evidence about sensory events versus internal monitoring processes may not be the same (cf. Steinhauser and Yeung, 2010), and so the same principles may not apply to both phenomena. Generally, our findings serve to highlight some of the complexities inherent in interpreting the emergence of error awareness in a decision-making framework, and warrant further exploration in studies specifically designed to systematically manipulate relevant parameters of the decision process.

Utility of speeded awareness signaling

Although speeded error signaling has been used previously in behavioral and computational modeling studies of error detection (Rabbitt et al., 1978; Rabbitt and Vyas, 1981; Rabbitt, 1990, 2002; Steinhauser et al., 2008), only one ERP study to our knowledge has employed a measure analogous to our speeded awareness press (Ullsperger and von Cramon, 2006). This study did not explore the relationship between the latency of this measure and the Pe in significant detail. A number of important issues are raised by the use of a speeded awareness press in the present research which should be considered when studying the electrophysiology of error awareness, and suggest that this measure is preferable to a delayed awareness response.

A high majority of ERP studies on error awareness have measured the Pe component as the amplitude of the average waveform locked to the initial erroneous response. By this method, many studies have reported significant differences in Pe morphology across a range of clinical groups (e.g., Brazil et al., 2009; O'Connell et al., 2009a; Olvet et al., 2010; Larson et al., 2011; Luijten et al., 2011; Perez et al., in press; Peterburs et al., 2012) and genotypes (Frank et al., 2007; Althaus et al., 2010; Biehl et al., 2011), and via pharmacological manipulation (Bartholow et al., 2012). In many cases these differences have been reported in the absence of an explicit indication of error awareness, and the present results highlight an important limitation of characterizing the Pe in this way. We have convincingly demonstrated via within-subjects, trial-by-trial correlational analyses that the Pe is better-related to the timing of error awareness, as indexed by Aware RT, than to the timing of initial error commission. Hence, the amplitude of the average error-locked Pe waveform will be partly determined by the amount of variability in the timing of error awareness. As a consequence, previously observed amplitude effects on the Pe may in fact be partly or even exclusively due to differences in the timing of awareness, and not to failures of awareness *per se*. A truer representation of component amplitude from the average waveform will necessarily be gleaned when trials are locked to the timing of error awareness, and not to the initial erroneous response as has traditionally been the case. This subtle distinction may have important implications for the interpretation of awareness deficits in clinical populations.

Behaviorally, we replicated previous studies (Endrass et al., 2005; O'Connell et al., 2007) in showing that aware errors were characterized by significantly faster initial RTs than correct Go trials, which in turn were faster than responses on unaware error trials. These differences have been interpreted as reflecting the possibility that aware errors are predominantly driven by failures of response inhibition, whereas unaware errors are precipitated by lapses of attention (O'Connell et al., 2007, 2009b; Shalgi et al., 2007). However, our Aware RT measure was characterized by substantial intra-subject variability, and its distribution was highly positively skewed. These characteristics indicate that a proportion of aware error trials are marked by very late error awareness, which may potentially be indicative of an attentional lapse as opposed to inhibitory failure. Aware RT represents a potentially useful measure for future studies to disentangle possible sub-types of aware errors, which may be characterized by distinct neural signatures and antecedent conditions and could not be distinguished using a delayed indication of awareness. Generally, these findings also point to the emergence of error awareness as being a highly variable process, and more complex than its usual characterization in the ERP literature as a binary, "present-or-absent" phenomenon would suggest.

The question of what actually drives the substantial variability in the timing of awareness signaling has not been addressed in the present study. Recent research has demonstrated that fluctuations in baseline attentional state in the inter-target interval can predict upcoming performance trends (Eichele et al., 2008; O'Connell et al., 2009c; MacDonald et al., 2011; Murphy et al., 2011). It may be the case that these markers are also related to fluctuations in the timing of error awareness. Further, a critical issue in performance monitoring research lies in trying to link the unique contributions of both the implicit and conscious recognition processes underlying error detection. One potential question for future research relates to the extent to which early, pre-conscious error detection mechanisms, as indexed by the ERN, drive the subsequent timing of error awareness, as indexed in our study by Pe_C latency and Aware RT. Although the ERN is not typically modulated by error awareness (though see Wessel et al., 2011), one recent study found that ERN amplitude was modestly correlated with Pe amplitude on a trial-by-trial basis (Hughes and Yeung, 2011). This suggests that the two may be at least partially related. In addition, medial-frontal

theta power (4–7 Hz), which has been proposed to reflect the recruitment of cognitive control (Cavanagh et al., 2009; Cohen et al., 2009; Cohen and Cavanagh, 2011) and to drive the ERN (Luu et al., 2004), has recently been shown to dynamically adjust response criterion according to changing levels of response conflict (Cavanagh et al., 2011). As previously described, response criterion is an important determinant of error signaling latency in connectionist computational models of error detection (Steinhauser et al., 2008), and it is feasible that fluctuations in the recruitment of control, as indexed by medial-frontal theta power, may be one mechanism by which variation in the timing of awareness arises. Further research in these areas is warranted.

CONCLUSIONS

The present study employed a novel variant of the EAT, which included a speeded manual response to indicate error awareness, in order to explore the relationship between the Pe component and the precise timing of error awareness. After isolating the Pe from other stimulus- and response-evoked signals using ICA, we demonstrated via within-subjects single-trial analyses that this component's peak latency is tightly related to the latency of the awareness response, as opposed to the timing of initial error commission. Further, component amplitude and latency correlated with behavioral indices of error awareness at the between-subjects level, and single-trial component amplitude was shown to reliably classify trials as with or without awareness up to 400 ms before the average timing of the awareness response. These results highlight the intimate association between the Pe and the emergence of error awareness, and also the utility of employing a speeded awareness press as an index of the timing of awareness.

ACKNOWLEDGMENTS

This research was supported by an Irish Research Council for Science, Engineering, and Technology (IRCSET) "Embark Initiative" grant awarded to Peter R. Murphy, an ARC Fellowship to Robert Hester (FT110100088), and an IRCSET "Empower" Fellowship to Redmond G. O'Connell. The authors also acknowledge funding support via the HEA PRTLI Cycle 3 program of the EU Structural Funds and the Irish Government's National Development Plan 2002–2006. We thank Hugh Nolan and Simon Kelly for their assistance with aspects of data analysis and Misha Byrne for assistance with figure design.

REFERENCES

Althaus, M., Groen, Y., Wijers, A. A., Minderaa, R. B., Kema, I. P., Dijck, J. D., and Hoekstra, P. J. (2010). Variants of the SLC6A3 (DAT1) polymorphism affect performance monitoring-related cortical evoked potentials that are associated with ADHD. *Biol. Psychol.* 85, 19–32.

Arieli, A., Sterkin, A., Grinvald, A., and Aertsen, A. (1996). Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. *Science* 273, 1868–1871.

Bartholow, B. D., Henry, E. A., Lust, S. A., Saults, J. S., and Wood, P. K. (2012). Alcohol effects on performance monitoring and adjustment: affect modulation and impairment of evaluative cognitive control. *J. Abnorm. Psychol.* 121, 173–186.

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

Biehl, S. C., Dresler, T., Reif, A., Scheuerpflug, P., Deckert, J., and Herrmann, M. J. (2011). Dopamine transporter (DAT1) and dopamine receptor D4 (DRD4) genotypes differentially impact on electrophysiological correlates of error processing. *PLoS One* 6:e28396. doi: 10.1371/journal.pone.0028396

Bland, A. R., Mushtaq, F., and Smith, D. V. (2011). Exploiting trial-to-trial variability in multimodal experiments. *Front. Hum. Neurosci.* 5:80. doi: 10.3389/fnhum.2011.00080

Botvinick, M. M., Cohen, J. D., and Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends Cogn. Sci.* 8, 539–546.

Brazil, I. A., de Bruijn, E. R., Bulten, B. H., von Borries, A. K., van Lankveld, J. J., Buitelaar, J. K., and Verkes, R. J. (2009). Early and late components of error monitoring in violent offenders with psychopathy. *Biol. Psychiatry* 65, 137–143.

Cavanagh, J. F., Cohen, M. X., and Allen, J. J. (2009). Prelude to and resolution of an error: EEG phase synchrony reveals cognitive control dynamics during action monitoring. *J. Neurosci.* 29, 98–105.

Cavanagh, J. F., Wiecki, T. V., Cohen, M. X., Figueroa, C. M., Samanta, J., Sherman, S. J., and Frank, M. J. (2011). Subthalamic nucleus stimulation reverses mediofrontal influence over decision threshold. *Nat. Neurosci.* 14, 1462–1467.

Cohen, M. X., and Cavanagh, J. F. (2011). Single-trial regression elucidates the role of prefrontal theta oscillations in response conflict. *Front. Psychol.* 2:30. doi: 10.3389/fpsyg.2011.00030

Cohen, M. X., van Gaal, S., Ridderinkhof, K. R., and Lamme, V. A. (2009). Unconscious errors enhance prefrontal-occipital oscillatory synchrony. *Front. Hum. Neurosci.* 3:54. doi: 10.3389/neuro.09.054.2009

Colebatch, J. G. (2007). Bereitschaftspotential and movement-related potentials: origin, significance, and application in disorders of human movement. *Mov. Disord.* 22, 601–610.

Debener, S., Ullsperger, M., Siegel, M., and Engel, A. K. (2006). Single-trial EEG-fMRI reveals the dynamics of cognitive function. *Trends Cogn. Sci.* 10, 558–563.

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

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

Eichele, H., Juvodden, H. T., Ullsperger, M., and Eichele, T. (2010). Maladaptation 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 Ulsperger, M. (2008). Prediction of human errors by maladaptive changes in event-related brain networks. *Proc. Natl. Acad. Sci. U.S.A.* 105, 6173–6178.

Einhauser, W., Koch, C., and Carter, O. L. (2010). Pupil dilation betrays the timing of decisions. *Front. Hum. Neurosci.* 4:18. doi: 10.3389/fnhum.2010.00018

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

Falkenstein, M., Hohnsbein, J., and Hoormann, J. (1995). "Event-related potential correlates of errors in reaction tasks," in *Perspectives of Event-related Potentials Research*, eds G. Karmos, M. Molnar, V. Scépe, I. Czigler, and J. E. Desmedt (Amsterdam: Elsevier), 287–296.

Falkenstein, M., Hohnsbein, J., Hoorman, J., and Blanke, L. (1990). "Effects of errors in choice reaction tasks on the ERP under focused and divided attention," in *Psychophysiological Brain Research*, eds C. H. M. Brunia, A. W. K. Gaillard, and A. Kok. (Tilberg, Netherlands: Tilberg University Press), 192–195.

Falkenstein, M., Hohnsbein, J., Hoormann, J., and Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components. II. 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.

Frank, M. J., D'Lauro, C., and Curran, T. (2007). Cross-task individual differences in error processing: neural, electrophysiological, and genetic components. *Cogn. Affect. Behav. Neurosci.* 7, 297–308.

Gehring, W. J., Goss, B., Coles, M. G., Meyer, D. E., and Donchin, E. (1993). A neural system for

error detection and compensation. *Psychol. Sci.* 4, 385–390.

Gold, J. I., and Shadlen, M. N. (2007). The neural basis of decision making. *Annu. Rev. Neurosci.* 30, 535–574.

Hermann, M., Rommler, J., Ehlis, A., Heidrich, A., and Fallgatter, A. (2004). Source localisation (LORETA) of the error-related negativity (ERN) and positivity (Pe). *Cogn. Brain Res.* 20, 294–299.

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., Nandam, L. S., O'Connell, R. G., Wagner, J., Strudwick, M., Nathan, P. J., Mattingley, J. B., and Bellgrove, M. A. (2012). Neurochemical enhancement of error awareness. *J. Neurosci.* 32, 2619–2627.

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., Niuwenhuis, S., Mars, R. B., and Coles, M. G. H. (2004). "Anterior cingulate cortex, selection for action, and error processing," in *Cognitive Neuroscience of Attention*, ed M. I. Posner (New York, NY: Guilford), 219–231.

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

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.

Larson, M. J., Fair, J. E., Farrer, T. J., and Perlstein, W. M. (2011). Predictors of performance monitoring abilities following traumatic brain injury: the influence of negative affect and cognitive sequelae. *Int. J. Psychophysiol.* 82, 61–68.

Larson, M. J., and Perlstein, W. M. (2009). Awareness of deficits and error processing after traumatic brain injury. *Neuroreport* 20, 1486–1490.

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

Luijten, M., van Meel, C. S., and Franken, I. H. (2011). Diminished error processing in smokers during smoking cue exposure. *Pharmacol. Biochem. Behav.* 97, 514–520.

Luu, P., Tucker, D. M., and Makeig, S. (2004). Frontal midline theta and the error-related negativity: neuropsychological mechanisms of action regulation. *Clin. Neurophysiol.* 115, 1821–1835.

MacDonald, J. S., Mathan, S., and Yeung, N. (2011). Trial-by-trial variations in subjective attentional state are reflected in ongoing prestimulus EEG alpha oscillations. *Front. Psychol.* 2:82. doi: 10.3389/fpsyg.2011.00082

Makeig, S., Debener, S., Onton, J., and Delorme, A. (2004). Mining event-related brain dynamics. *Trends Cogn. Sci.* 8, 204–210.

Mintz, A. R., Addington, J., and Addington, D. (2004). Insight in early psychosis: a 1-year follow-up. *Schizophr. Res.* 67, 213–217.

Murphy, P. R., Robertson, I. H., Balsters, J. H., and O'Connell, R. G. (2011). Pupillometry and P3 index the locus coeruleus-noradrenergic arousal function in humans. *Psychophysiology* 48, 1532–1543.

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.

Nolan, H., Whelan, R., and Reilly, R. B. (2010). FASTER: fully automated Statistical thresholding for EEG artifact rejection. *J. Neurosci. Methods* 192, 152–162.

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. (2009a). 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.

O'Connell, R. G., Dockree, P. M., Bellgrove, M. A., Turin, A., Ward, S., Foxe, J. J., and Robertson, I. H. (2009b). Two types of action error: electrophysiological evidence for separable inhibitory and sustained attention neural mechanisms producing error on Go/No-go tasks. *J. Cogn. Neurosci.* 21, 93–104.

O'Connell, R. G., Dockree, P. M., Robertson, I. H., Bellgrove, M. A., Foxe, J. J., and Kelly, S. P. (2009c). Uncovering the neural signature of lapsing attention: electrophysiological signals predict errors up to 20s before they occur. *J. Neurosci.* 29, 8604–8611.

O'Keeffe, F. M., Dockree, P. M., and Robertson, I. H. (2004). Poor insight in traumatic brain injury mediated by impaired error processing?: evidence from electrodermal activity. *Brain Res. Cogn. Brain Res.* 22, 101–112.

Olvet, D. M., Klein, D. N., and Hajcak, G. (2010). Depression symptom severity and error-related brain activity. *Psychiatry Res.* 179, 30–37.

Onton, J., Westerfield, M., Townsend, J., and Makeig, S. (2006). Imaging human EEG dynamics using independent component analysis. *Neurosci. Biobehav. Rev.* 30, 808–822.

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

Perez, V. B., Ford, J. M., Roach, B. J., Woods, S. W., McGlashan, T. H., Srihari, V. H., Loewy, R. L., Vinogradov, S., and Mathalon, D. H. (in press). Error monitoring dysfunction across the illness course of schizophrenia. *J. Abnorm. Psychol.* doi: 10.1037/a0025487. [E-pub ahead of print].

Peterburs, J., Gajda, K., Koch, B., Schwarz, M., Hoffmann, K. P., Daum, I., and Bellebaum, C. (2012). Cerebellar lesions alter performance monitoring on the antisaccade task – an event-related potentials study. *Neuropsychologia* 50, 379–389.

Quiroga, R. Q., Reddy, L., Kreiman, G., Koch, C., and Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature* 435, 1102–1107.

Rabbitt, P. (1966a). Error-correction without external error signals. *Nature* 212, 438.

Rabbitt, P. (1966b). Errors and error correction in choice-response tasks. *J. Exp. Psychol.* 71, 264–272.

Rabbitt, P. (1990). Age, IQ and awareness, and recall of errors. *Ergonomics* 33, 1291–1305.

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

Rabbitt, P., Cumming, G., and Vyas, S. (1978). Some errors of perceptual analysis in visual search can be detected and corrected. *Q. J. Exp. Psychol.* 30, 319–332.

Rabbitt, P., and Vyas, S. (1981). Processing a display even after you make a response to it: how perceptual errors can be corrected. *Q. J. Exp. Psychol. A.* 33, 223–239.

Ratcliff, R., and Smith, P. L. (2010). Perceptual discrimination in static and dynamic noise: the temporal relation between perceptual encoding and decision making. *J. Exp. Psychol. Gen.* 139, 70–94.

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.

Roitman, J. D., and Shadlen, M. N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J. Neurosci.* 22, 9475–9489.

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.

Shalgi, S., O'Connell, R. G., Deouell, L. Y., and Robertson, I. H. (2007). Absent minded but accurate: delaying responses increases accuracy but decreases error awareness. *Exp. Brain Res.* 182, 119–124.

Spencer, K. M. (2004). "Averaging, detection, and classification of single-trial ERPs," in *Event-related Potentials: A Methods Handbook*, ed T. C. Handy (Cambridge, MA: MIT Press), 209–228.

Steinhauser, M., Maier, M., and Hubner, R. (2008). Modeling behavioral measures of error detection in choice tasks: response monitoring versus conflict monitoring. *J. Exp. Psychol. Hum. Percept. Perform.* 34, 158–176.

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

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. (2006). How does error correction differ from error signaling? An event-related potential study. *Brain Res.* 1105, 102–109.

van Boxtel, G., van der Molen, M., and Jennings, J. (2005). Differential involvement of the anterior cingulate cortex in performance monitoring during a

stop-signal task. *J. Psychophysiol.* 19, 1–10.

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.

Viola, F. C., Thorne, J., Edmonds, B., Schneider, T., Eichele, T., and Debener, S. (2009). Semi-automatic identification of independent components representing EEG artifact. *Clin. Neurophysiol.* 120, 868–877.

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.

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.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 31 January 2012; accepted: 12 March 2012; published online: 28 March 2012.

Citation: Murphy PR, Robertson IH, Allen D, Hester R and O'Connell RG (2012) An electrophysiological signal that precisely tracks the emergence of error awareness. Front. Hum. Neurosci. ***6****:65. doi: 10.3389/fnhum.2012.00065*

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 22 June 2012
doi: 10.3389/fnhum.2012.00178

Temporospatial dissociation of Pe subcomponents for perceived and unperceived errors

Tanja Endrass*, Julia Klawohn, Julia Preuss and Norbert Kathmann

Institute for Psychology, Humboldt University Berlin, Berlin, Germany

Edited by:
Claudia Danielmeier, Radboud University Nijmegen, Netherlands

Reviewed by:
K. Richard Ridderinkhof, University of Amsterdam, Netherlands
Wim Notebaert, Ghent University, Belgium

****Correspondence:***
Tanja Endrass, Institute for Psychology, Humboldt University Berlin, Rudower Chaussee 18, 12489 Berlin, Germany.
e-mail: endrass@gmail.com

Previous research on performance monitoring revealed that errors are followed by an initial fronto-central negative deflection (error-related negativity, ERN or Ne) and a subsequent centro-parietal positivity (error positivity, Pe). It has been shown that error awareness mainly influences the Pe, whereas the ERN seems unaffected by conscious awareness of an error. The aim of the present study was to investigate the relation of ERN and Pe to error awareness in a visual size discrimination task in which errors are not elicited by impulsive responding but by perceptual difficulty. Further, we applied a temporospatial principal component analysis (PCA) to examine whether the temporospatial subcomponents of the Pe would differentially relate to error awareness. Event-related potential (ERP) results were in accordance with earlier studies: a significant error awareness effect was found for the Pe, but not for the ERN. Interestingly, a modulation with error perception on correct trials was found: correct responses considered as incorrect had larger correct-related negativity (CRN) and lager Pe amplitudes than correct responses considered as correct. The PCA yielded two relevant spatial factors accounting for the Pe (latency 300 ms). A temporospatial factor characterized by a centro-parietal positivity varied significantly with error awareness. Of the two temporospatial factors corresponding to ERN and CRN, one factor with central topography varied with response correctness and subjective error perception on correct responses. The PCA results indicate that the error awareness effect is specifically related to the centro-parietal subcomponent of the Pe. Since this component has also been shown to be related to the importance of an error, the present variation with error awareness indicates that this component is sensitive to the salience of an error and that salience secondarily may trigger error awareness.

Keywords: error awareness, error-related negativity, error positivity, principal component analysis, PCA

INTRODUCTION

Performance monitoring is an essential prerequisite for adaptive behavior and implements adjustment processes, such as error detection and subsequent post-error slowing. Over the past years numerous psychophysiological and neuroimaging studies investigated the neural basis of performance monitoring and error processing. Event-related potential (ERP) studies identified a fronto-central negativity that emerges shortly after the execution of incorrect responses, the error-related negativity (ERN, Gehring et al., 1993) or error negativity (Ne, Falkenstein et al., 1991). Sometimes a smaller negative deflection is also observed for correct responses, the correct-related negativity (CRN, Ford, 1999; Vidal et al., 2000). The ERN is followed by the error positivity (Pe), a centro-parietal positive deflection that peaks between 200 and 400 ms after response onset (Falkenstein et al., 1991, 2000; Overbeek et al., 2005). These components are considered to indicate error-related brain activity but it is not fully clear whether they reflect functionally dissociate aspects of error processing. While source localization studies suggest that the ERN is generated in the anterior cingulate cortex (ACC), or more precisely, Hyphenation: gene-rated in the posterior medial frontal cortex (pMFC, Dehaene et al., 1994; Van Veen and Carter, 2002; Debener et al., 2005), the source of the Pe is more difficult to determine, and heterogeneous results were obtained (Herrmann et al., 2004; O'Connell et al., 2007; Vocat et al., 2008).

Although it is broadly agreed that the ERN and Pe are linked to error commission, their functional significance still remains to be clarified. Currently, the ERN is regarded as a negative reinforcement learning signal (Holroyd and Coles, 2002), or as an index for response conflict or error likelihood (Botvinick et al., 2001; Yeung et al., 2004; Brown and Braver, 2007, 2008). These models assume that the ERN or its underlying ACC activity plays a key role in the recruitment of cognitive control in response to erroneous actions in order to prevent future errors (Ridderinkhof et al., 2004; Ullsperger et al., 2004). Since the Pe most consistently varies with motivational salience and subjective error perception, it is considered to reflect evaluative aspects of error processing (Falkenstein et al., 2000; Overbeek et al., 2005).

The sensitivity to error awareness is perhaps the most important distinction between both error-related ERP components. Results were rather consistent in that reduced Pe components were reported for unaware/unperceived compared to aware/perceived errors (Nieuwenhuis et al., 2001; Endrass et al., 2005, 2007; O'Connell et al., 2007; Shalgi et al., 2009; Dhar et al., 2011; Wessel et al., 2011). Whereas the Pe seems clearly sensitive to error awareness, the majority of studies found no effect

on the ERN (Nieuwenhuis et al., 2001; Endrass et al., 2005, 2007; O'Connell et al., 2007; Shalgi et al., 2009; Dhar et al., 2011). However, there is some evidence that an error awareness effect may also be present for the ERN (Maier et al., 2008; Hewig et al., 2011; Wessel et al., 2011). Interestingly, a recent study provided evidence that ERN and Pe reflect dissociate aspects of error processing in a choice selection task. In that task errors were compared between a masking and a conflict condition. While the ERN was substantially reduced in the masking condition, the Pe varied with error awareness (Hughes and Yeung, 2011). Neuroimaging studies examining error awareness reported that the activity of the anterior insula was significantly modulated by conscious error perception and suggested this region to indirectly contribute to the emergence of a Pe (Hester et al., 2005; Klein et al., 2007). The same region was shown to be sensitive to interoceptive awareness (Critchley et al., 2004, 2005), and is reliably associated with autonomous activation (Ullsperger et al., 2010). Therefore, it was suggested that error awareness emerges when information for error commission aggregates within the salience network, and its presence is to some extent indicated by the Pe (Ullsperger et al., 2010).

Additional evidence that the Pe may be related to saliency of incorrect response can be derived from the fact that topography and time course of the Pe displays similarities to the P300 or P3b (Falkenstein et al., 1999; Leuthold and Sommer, 1999; Hajcak et al., 2003; Overbeek et al., 2005; Ridderinkhof et al., 2009). The P3b is a stimulus-locked positive brain potential that is elicited in response to rare and motivationally significant events (Picton, 1992; Nieuwenhuis et al., 2005). This component has been considered to reflect the response of the locus coeruleus-norepinephrine system to the outcome of internal decision-making processes (Nieuwenhuis et al., 2005). Errors are rare and salient events that trigger a cascade of central nervous and autonomous changes and are considered as events that are motivationally significant and thus elicit a P3b (e.g., Hajcak et al., 2003; O'Connell et al., 2007; Wessel et al., 2011). Hence, the Pe might be understood as a P3b to error commission. Using principal components analysis (PCA) it has been shown that the Pe is composed of two different subcomponents: a fronto-central component that shares spatial distribution with the ERN and a centro-parietal component that is similar to the P300 (Arbel and Donchin, 2009). While emphasis on accuracy did not cause significant variation of the fronto-central component, the centro-parietal component was enlarged for this condition. These results highlight the association between the centro-parietal component of the Pe and error salience.

The aim of the current study is to investigate error awareness effects on ERN and Pe in a choice selection task. In this task errors are elicited due to perceptual difficulty instead of failures to withhold a response. Specifically, we were interested in the structure of the underlying processes of the Pe and conducted a temporospatial PCA (Dien, 2010a; Endrass et al., 2012). The aim was to disentangle overlapping electrophysiological activity captured in the Pe, and to examine the sensitivity of the subcomponents of ERN and Pe to error awareness. In the ERP analysis we expected to replicate earlier findings showing that error awareness selectively affects the Pe, but not the ERN. In addition, we examined more closely the effect of subjective error perception in correct responses. With the PCA analysis we intended to replicate the previously identified subcomponents of the Pe (Arbel and Donchin, 2009) and to examine whether these subcomponents would be differentially sensitive to error awareness. Since the centro-parietal subcomponent of the Pe was associated with error salience, we expected only this component to be sensitive to error awareness.

MATERIALS AND METHODS

PARTICIPANTS

Seventeen healthy undergraduate students (nine female, 20–30 years, mean ± SD, 23.4 years ± 3.1) of the Humboldt-University Berlin voluntarily took part in the present experiment[1]. The participants either received monetary remuneration or class credit points for their participation. All of them were in good health, with no history of psychiatric or neurological disease, and had normal or corrected to normal vision. In accordance to the ethical guidelines of the Declaration of Helsinki, written informed consent was enquired before the experimental procedures started.

TASK AND EXPERIMENTAL PROCEDURE

A visual size discrimination task was employed with a display showing two dots to compare. Participants were seated at a distance of 70 cm in front of a 19-in computer screen. Each stimulus comprised a standard (visual angle of 2.4°) and a comparison dot that varied in size with task difficulty (2.5, 2.6, and 2.7° in the difficult, intermediate and easy condition, respectively). Both dots were presented in white color against dark background left and right to a fixation cross (distance 1.4°).

Each trial started with a fixation cross displayed for 600–1000 ms, followed by the stimulus array presented for 500 ms. After that, the screen turned blank for 1000 ms. Participants were asked to select the larger dot by a left or right response button press. Instruction equally emphasized the importance of speed and accuracy. Subsequently, a display requesting for accuracy ratings was shown for 1000 ms. Participants were instructed to evaluate whether the previous response was correct, incorrect, or they were unsure. The experiment comprised a total number of 832 trials administered in four blocks. The target stimuli were presented equally frequent at both sides of the screen and subjects were informed that always one of the dots was larger than the other. The hard and intermediate conditions encompassed 208 trials each and 416 trials were presented for the easy condition, all distributed evenly over the four blocks, and displayed in pseudo-randomized order. A larger number of easy trials were presented to obtain a reasonable number of error trials in that condition since the goal was to compare errors between difficulty levels. The completion of the whole experimental task

[1]The data presented here were also included in a previous study from our group that focused on differences in post-response potentials following errors and correct responses and addressed the question whether ERN and CRN reflect different aspects of performance monitoring (Endrass et al., 2012). Specifically, the previous analysis compared correct and incorrect responses in easy, intermediate, and difficult condition. The results presented here were not reported in the previous report and data analysis followed a different rational.

lasted about 60 min including a practice block and short breaks between blocks.

ELECTROENCEPHALOGRAPHIC RECORDING AND ANALYSIS

The electroencephalogram (EEG) was recorded from 63 electrodes, with 61 equidistant electrode positions including Cz as recording reference placed in an electrode cap (EasyCap, Herrsching, Germany) and two external electrodes located below the left and right eye. The ground electrode was placed below T1. It was ensured that electrode impedances were below 5 kΩ. EEG was continuously recorded at a sampling rate of 500 Hz. Amplifiers (BrainAmp BrainProducts, Gilching Germany) used a high-pass filter with a time constant of 10 s (0.0159 Hz) and a low-pass filter set at 250 Hz.

Responses were obtained from two force-sensitive response devices. Force was continuously recorded in two separate recording channels together with the EEG signals. Devices were calibrated to 200 μV equaling 1N. Participants were instructed to rest their index fingers on the devices, and to press them in order to respond. For data analysis, first, the mean activity of a 200 ms pre-stimulus baseline interval was subtracted. Then, an algorithm searched for amplitude maxima exceeding predefined thresholds for partial responses (minimum of 0.25 N equaling 50 μV) and full responses (minimum of 0.5 N). Given an above-threshold activity, response onsets were marked at the initial force onset as indicated by an amplitude change of 20 μV within 20 ms. This method allowed to detect both partial responses as well as full responses and to distinguish them for further analyses. In the present study, only purely correct responses which were not preceded by any partial incorrect responses within the same trial were included into analyses. Accordingly, error trials were only taken into account if they were not preceded by partial correct reactions within the trial.

For ERP analysis, EEG data were filtered off-line with a low-pass filter set at 40 Hz (12 db/Oct) as well as with a 50 Hz notch filter, and re-referenced to average reference. Correction for eye-movements and blink artifacts was applied using the multiple source eye correction method (Berg and Scherg, 1994) implemented in BESA5 (Brain Electrical Source Analysis, MEGIS Software GmbH, Gräfelfing, Germany). Epochs of 1200 ms starting 200 ms before the first response in each trial were obtained from continuous EEG data. Epochs were baseline corrected using the 200 ms pre-response window. Segments containing amplitude changes exceeding 200 μV, or voltage steps of more than 100 μV between consecutive data points were rejected from further analysis. For each participant, four averages were computed: correct responses rated as correct (perceived correct) or incorrect (unperceived correct) as well as erroneous responses rated as correct (unperceived errors) or incorrect (perceived errors). Although a behavioral pilot study (10 participants) indicated that it would be possible to obtain a sufficient number of perceived and unperceived errors in all three difficulty levels, only very few participants ($N = 8$) had a minimum number of six errors in each condition. Therefore, the factor task difficulty could not be taken into account for this analysis. For visual presentation, grand averages were filtered with a 15 Hz low-pass filter.

For ERP analysis, ERN and CRN amplitudes were quantified as mean amplitudes between 60 and 140 ms post-response at fronto-central electrode sites (Fz, FCz, and Cz). The Pe was measured as mean amplitudes at the electrodes Cz, CPz, and Pz between 300 and 500 ms after response onset. ERP amplitudes were statistically analyzed with repeated measurement ANOVAs with the factors Response Type (correct vs. incorrect), Response Rating (response perceived as correct vs. incorrect) and Electrode Site. Greenhouse-Geisser correction was applied when appropriate. All statistical analyses for the present study were conducted with IBM SPSS Statistics (Version 19.0, Chicago).

A covariance-based two-step temporospatial PCA was computed on individual response-locked ERP averages using the ERP PCA Toolkit 2.06 (Dien, 2010a,b). In accordance with Dien et al., 2005, a covariance matrix and Kaiser normalisation was applied. The temporospatial PCA extracts linear combinations of data that distinguish patterns of electrocortical activity across all time points and recording sites (see also Dien and Frishkoff, 2005). The temporospatial sequence of analyses was chosen since this was found to be most effective in simulation studies (Dien, 2010b). First, the temporal PCA was computed using the individual averages of each participant over all 63 electrodes, for correct and incorrect responses in the two response rating conditions (perceived as correct vs. incorrect). Each dataset consisted of 600 time points (–200 to 1000 ms). A scree plot was used to limit extracted factors in number, resulting in the promax rotation that yielded 19 temporal factors. Then, in order to analyze their spatial distribution, separate spatial PCA (infomax rotation) was applied to each temporal factor. In total, the temporospatial PCA yielded 76 factor combinations (four spatial factors extracted for each of the 19 temporal factors). Only those temporospatial factors that uniquely accounted for more than 1% of the total variance in the data were included in further analyses (Kayser and Tenke, 2005; Foti et al., 2009, 2011). Note that the amount of explained variance by one factor is related to the total variance in the data, i.e., all time points and all electrodes. Factor scores of these factors were plotted as "virtual ERPs" and averaged for both response types and rating conditions. The temporal factors corresponding to ERN/CRN and Pe, as our ERP components of interest, were selected by temporal characteristics of the PCA waveforms (Dien et al., 2005, 2010; Foti et al., 2011). The resulting factor scores were submitted to statistical analysis using repeated measurement ANOVA with the factors Response Type and Response Rating.

Error and correct awareness were determined as the percentages of errors and correct responses that were adequately perceived as incorrect or correct, respectively. Behavioral data were analyzed by repeated-measurement ANOVAs. Awareness was analyzed with the factors Response Type (error vs. correct) and Difficulty (easy, intermediate vs. difficult). Analysis of reaction time data involved the factors Response Type and Response Rating. Post-error adjustment effects were analyzed in terms of subsequent reaction time (post-error slowing) and response correctness. For post-error slowing, correct reaction times following perceived and unperceived correct and incorrect responses were analyzed by an ANOVA with the factors Preceding Response Type and Preceding Response Rating of the preceding response. To examine whether there was a relative

increase of response correctness after erroneous responses (post-error correctness), the percentage of correct responses following perceived and unperceived errors and correct responses was assessed.

RESULTS

BEHAVIORAL DATA

Main behavioral results are presented in **Table 1**. On average, participants committed 16.1% (±SD 5.2) errors and 80.2% (±SD 6.6) of all responses were correct. The percentage of correct perceptions, i.e., the amount of correctly classified correct and incorrect responses, varied with Response Type, [$F_{(1, 16)} = 142.59$, $p < 0.001$, $\eta^2 = 0.90$], and Difficulty, [$F_{(2, 32)} = 17.27$, $p < 0.001$, $\varepsilon = 0.88$, $\eta^2 = 0.52$]. The amount of correct perceptions was significantly higher for correct than incorrect responses (mean ± SD, 93.2% ± 3.3 vs. 51.5% ± 13.6) and for easy and intermediate compared to difficult trials (mean ± SD, 73.6% ± 8.7, 76.7% ± 8.2, vs. 66.8% ± 6.6), $p < 0.008$. Further, these effects resulted in a significant interaction of both factors, [$F_{(2, 32)} = 3.83$, $p = 0.044$, $\varepsilon = 0.79$, $\eta^2 = 0.19$]. While relative error perception was lower for easy and difficult compared to intermediate trials (mean ± SD, 50.9% ± 17.7, 44.5% ± 14.0, vs. 58.9% ± 15.3, easy, difficult and intermediate condition, respectively), $p = 0.025$ and $p = 0.007$, easy and difficult conditions did not differ, $p = 0.335$. For correct responses, more correct perceptions were found for easy and intermediate compared to difficult trials, $p < 0.001$ and $p = 0.002$ (mean ± SD, 96.3% ± 2.1, 94.5% ± 4.1, vs. 89.0% ± 6.2, easy, intermediate, and difficult condition, respectively)[2].

The analysis of reaction time data revealed a main effect of Response Type, [$F_{(1, 16)} = 6.40$, $p = 0.022$, $\eta^2 = 0.29$], that was qualified by an interaction of Response Type and Response Rating, [$F_{(1, 16)} = 12.29$, $p = 0.003$, $\eta^2 = 0.43$]. Overall, error reaction times were faster than correct reaction times. Mean reaction times were faster for perceived errors compared to unperceived errors, $p = 0.047$ (mean ± SD, 336 ms ± 50, 351 ms ± 60), and for perceived correct compared to unperceived correct responses, $p = 0.001$ (mean ± SD, 346 ms ± 42, 364 ms ± 49).

Concerning post-error slowing, reaction times of correct responses following perceived and unperceived errors as well as those following perceived and unperceived correct responses were analyzed (mean values in **Table 1**). Only the main effect of Preceding Response Rating was significant, [$F_{(1, 16)} = 17.06$, $p = 0.001$, $\eta^2 = 0.52$], but neither the main effect of Preceding Response Type nor the interaction were significant, $F < 1$. *Post-hoc* comparisons indicate that reaction times were slower in trials following reactions that were subjectively perceived as incorrect (mean ± SD, 358 ± 40) compared to trials following reactions perceived as correct (mean ± SD, 346 ± 40), irrespective of the actual correctness of the preceding response.

In contrast, the analysis of post-error correctness only revealed a significant main effect of Preceding Response Type, [$F_{(1, 16)} = 4.81$, $p = 0.043$, $\eta^2 = 0.23$]. Overall, accuracy was higher following incorrect responses (mean ± SD, 82.7% ± 7.3) than following correct responses (mean ± SD, 79.3% ± 7.0). The main effect of Preceding Response Rating and the interaction were not significant, $F < 2.4$.

Table 1 | Behavioral results and amplitudes of response-locked negativities (ERN/Ne and CRN) at electrode FCz in the three difficulty conditions (means and standard deviations).

	Task difficulty					
	Easy		Intermediate		Difficult	
	M	SD	*M*	SD	*M*	SD
Error rates (%)	14.9	6.1	18.1	7.2	31.3	5.9
Error RT (ms)	309	67	321	60	341	60
Correct RT (ms)	338	46	340	47	345	50
Error awareness (%)	47.3	19.1	49.6	16.3	37.6	15.4
Correct awareness (%)	94.3	3.5	91.8	6.2	81.9	10.1
Error correction (%)	32.8	17.2	29.9	16.6	16.6	11.1
ERN/Ne amplitude (μV)	−4.11	2.81	−4.22	2.60	−3.21	1.91
CRN amplitude (μV)	−1.91	2.24	−2.17	2.17	−2.22	1.89

Note: RT, reaction time.

ERP RESULTS

ERPs and topographies for both response types (correct and incorrect responses) and rating conditions (perceived as correct and incorrect) are displayed in **Figure 1**. ERN and CRN are apparent at fronto-central electrode sites peaking about 100 ms following response onset. Pe is evident at central electrodes in the time range between 200 and 600 ms after response onset.

Statistical analysis of ERN/Ne and CRN amplitudes yielded significant main effects of Electrode, [$F_{(2, 32)} = 19.43$, $p < 0.001$, $\varepsilon = 0.59$, $\eta^2 = 0.55$], and Response Type, [$F_{(1, 16)} = 11.92$, $p = 0.003$, $\eta^2 = 0.43$]. These main effects were further specified by a significant interaction of Electrode × Response Type, [$F_{(2, 32)} = 9.80$, $p = 0.002$, $\varepsilon = 0.73$, $\eta^2 = 0.38$]. While no main effect of Response Rating or interaction of Response Rating and Response Type was found, $F < 2.14$, $p > 0.163$, the three-way interaction reached statistical trend level, [$F_{(2, 32)} = 2.80$, $p = 0.105$, $\varepsilon = 0.60$, $\eta^2 = 0.149$]. Follow-up comparisons of the Electrode × Response Type interaction revealed significantly larger ERN than CRN amplitudes at FCz and Cz, but not at Fz, $p = 0.004$, $p = 0.001$, and $p = 0.149$, respectively. CRN amplitudes were largest at Fz compared to FCz and Cz, $p = 0.041$ and $p < 0.001$, respectively. ERN amplitudes were larger at FCz compared to Fz and Cz, $p < 0.001$, but did significantly differ only between FCz and Cz. *Post-hoc* analysis of the three-way interaction revealed that perceived and unperceived errors did not differ at any electrode site, $p > 0.48$. In contrast, unperceived correct responses elicited larger CRN amplitudes than perceived correct responses at FCz, $p = 0.056$.

[2]Since error awareness varied between difficulty conditions and a larger number of trials were presented in the easy condition, the absolute numbers of easy, intermediate, and difficult trials were compared between perceived and unperceived errors. An interaction between Difficulty and Response Rating was found at trend level, [$F_{(2, 16)} = 3.55$, $p = 0.058$, $\eta^2 = 0.18$]. While the number of perceived and unperceived errors did not differ in the easy (31.0 vs. 26.3) and difficult condition (24.9 vs. 29.5), significantly more perceived than unperceived error occurred in the intermediate condition (19.6 vs. 13.5), $p = 0.025$. Hence, the number of errors did not differ between easy and difficult condition for both rating types.

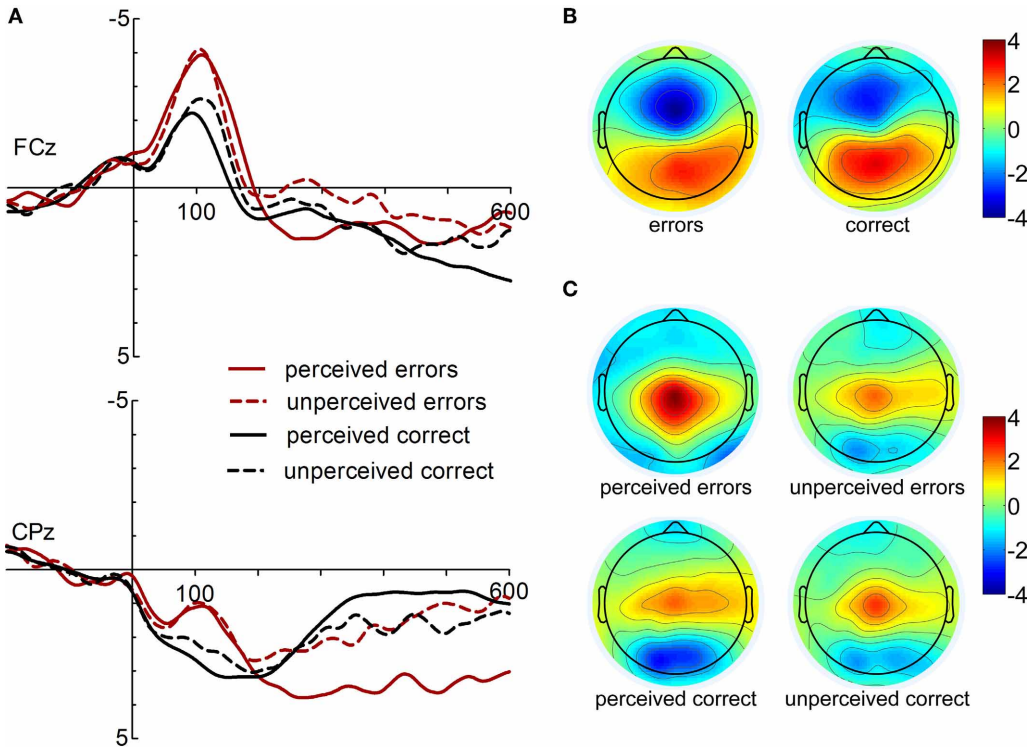


FIGURE 1 | (A) Response-locked ERPs following correct and incorrect response at electrodes FCz and CPz (left). **(B)** Topographies of the response-related negativities (ERN and CRN, 100 ms following response onset). **(C)** Topographies of the post-response positivity (Pe, 300 ms).

Supplementary paired *t*-tests were used to examine the Response Rating effect for ERN and CRN amplitudes at several fronto-central and midline electrode sites (FCz, FC1, FC2, Fz, F1, F2, Cz, CPz, and Pz). The comparison of perceived and unperceived errors yielded no significant difference at any electrode location, all comparisons $t < 1$, $p > 0.50$. Yet, larger CRN amplitudes were found for unperceived correct responses compared to perceived correct responses at FCz, $t(16) = 2.06$, $p = 0.056$, FC1, $t(16) = 4.71$, $p < 0.001$, F1, $t(16) = 1.91$, $p = 0.075$, CPz, $t(16) = 2.50$, $p = 0.024$, and Pz, $t(16) = 2.88$, $p = 0.011$.

The analysis of Pe amplitudes revealed significant main effects of the factors Electrode, $[F_{(2,\ 32)} = 14.83,\ p = 0.001,\ \varepsilon = 0.56,\ \eta^2 = 0.48]$, Response Type, $[F_{(1,\ 16)} = 22.47,\ p < 0.001,\ \eta^2 = 0.58]$, and Response Rating, $[F_{(1,\ 16)} = 30.32,\ p < 0.001,\ \eta^2 = 0.66]$. Further, the following interactions were significant: Response Type × Response Rating, $[F_{(1,\ 16)} = 11.08,\ p = 0.004,\ \eta^2 = 0.41]$, and Electrode × Response Type, $[F_{(2,\ 32)} = 10.20,\ p = 0.004,\ \varepsilon = 0.56,\ \eta^2 = 0.40]$. Overall, Pe amplitudes were more positive for incorrect compared to correct responses. Thus, the effect of Response Rating is more pronounced for errors than for correct responses. Larger Pe amplitudes were found for perceived compared to unperceived errors (mean difference ± SD, errors: 1.82 μV ± 1.26), $p < 0.001$, and for unperceived compared to perceived correct responses (mean difference ± SD: 0.75 μV ± 1.07), $p = 0.011$. In fact, amplitudes did not differ between unperceived errors and unperceived correct responses. While the Pe is significantly more positive at Cz and CPz compared to Pz, $p = 0.004$ and $p < 0.001$, the Pe is larger for incorrect than correct responses at CPz and Pz, but not at Cz, $p < 0.001$, $p < 0.001$, and $p = 0.124$, respectively[3].

PCA RESULTS

The application of the temporospatial PCA revealed 19 temporal factors and four spatial factors for each temporal factor,

[3]In addition, ERN and Pe amplitudes were analyzed with the factors Electrode, Response Rating and Difficulty (easy vs. difficult) in a subsample of 12 participants who committed at least six errors in each condition. For ERN amplitudes significant main effects of Electrode, $[F_{(2,\ 22)} = 5.12,\ p = 0.015,\ \varepsilon = 0.58,\ \eta^2 = 0.32]$, and Difficulty, $[F_{(1,\ 11)} = 24.40, p < 0.001,\ \eta^2 = 0.69]$, were found. ERN amplitudes were smaller in the difficult compared to the easy condition. A statistical trend was found for the interaction of Electrode and Difficulty, $[F_{(2,\ 22)} = 3.68,\ p = 0.077,\ \varepsilon = 0.52,\ \eta^2 = 0.25]$. While the difficulty effect was significant at FCz and Cz ($p < 0.001$), it only reached trend level at Fz ($p = 0.09$). Importantly, a main effect of Response Rating or an interaction of Response Rating and Difficulty were not found, $[F_{(1,\ 11)} < 1]$. In addition, separate comparisons between perceived and unperceived errors in the easy and difficult condition did not reach significance ($p > 0.3$). The analysis of the Pe showed significant main effects of Electrode, $[F_{(2,\ 22)} = 13.33,\ p = 0.001,\ \varepsilon = 0.59,\ \eta^2 = 0.40]$, and Response Rating, $[F_{(1,\ 11)} = 25.87,\ p < 0.001,\ \eta^2 = 0.56]$. Enhanced Pe amplitudes were elicited by perceived compared to unperceived errors. The main effect of Difficulty, $[F_{(1,\ 11)} = 1.41,\ p = 0.25]$, or the interaction with Response Rating, $[F_{(1,\ 11)} = 0.09, p = 0.76]$, were not significant. Separate comparisons indicate that the Pe enhancement for perceived errors was present in the easy ($p = 0.028$) and the difficult condition ($p = 0.001$).

resulting in 76 factor combinations. The factors corresponding to the ERP components of interest were identified by their temporal characteristics, one peaking at 300 ms post-response and presumably reflecting the error positivity (TF2) and the other peaking at 104 ms corresponding to response-related negativities (TF3). Of the eight resulting temporospatial factors, two corresponding to the Pe (TF2SF1, TF2SF2) and two corresponding to the ERN/CRN (TF3SF1, TF3SF2) accounted for at least 1% of the variance and were analyzed further.

In **Figure 2**, time course and topographies of the two spatiotemporal factors corresponding to the ERN and CRN are displayed. The first factor (TF3SF1, **Figure 2A**) had a central topography and uniquely accounted for 2.28 % of variance in the solution whereas the second factor (TF3SF2, **Figure 2B**) displayed a frontal negativity and a parietal positivity and accounted for 1.82% of variance. The central factor (TF3SF1) was more negative for errors than correct responses as revealed by a significant main effect of Response Type, $[F_{(1,\ 16)} = 25.62,\ p < 0.001,\ \eta^2 = 0.62]$. Although the main effect of Response Rating and the interaction with Response Type only reached a statistical trend, $[F_{(1,\ 16)} = 3.46,\ p = 0.081,\ \eta^2 = 0.18]$ and $[F_{(1,\ 16)} = 3.97,\ p = 0.064,\ \eta^2 = 0.20]$, follow-up comparisons for the interaction were conducted. While factor scores for perceived and unperceived errors did not differ, $p = 0.75$, a significant awareness effect was found for correct responses, $p = 0.003$. Factor scores of the frontal factor (TF3SF2) revealed a significant main effect of Response Rating, $[F_{(1,\ 16)} = 11.96,\ p = 0.003,\ \eta^2 = 0.43]$. Overall, factor scores were more pronounced for responses that were subjectively perceived as correct. A statistical trend was found for the main effect Response Type, $[F_{(1,\ 16)} = 3.10,\ p = 0.098,\ \eta^2 = 0.162]$, indicating also more pronounced factor scores for correct than for incorrect responses. The interaction was not significant, $F < 1.2$. However, while factor scores were more pronounced for perceived than for unperceived correct responses, $p = 0.014$, the difference between perceived and unperceived errors was not significant, $p = 0.17$.

The first factor corresponding to the Pe (TF2SF1, **Figure 3A**), had a centro-parietal distribution and uniquely accounted for 2.27 % of the variance in the solution whereas the second factor (TF2SF2, **Figure 3B**) showed a fronto-central distribution

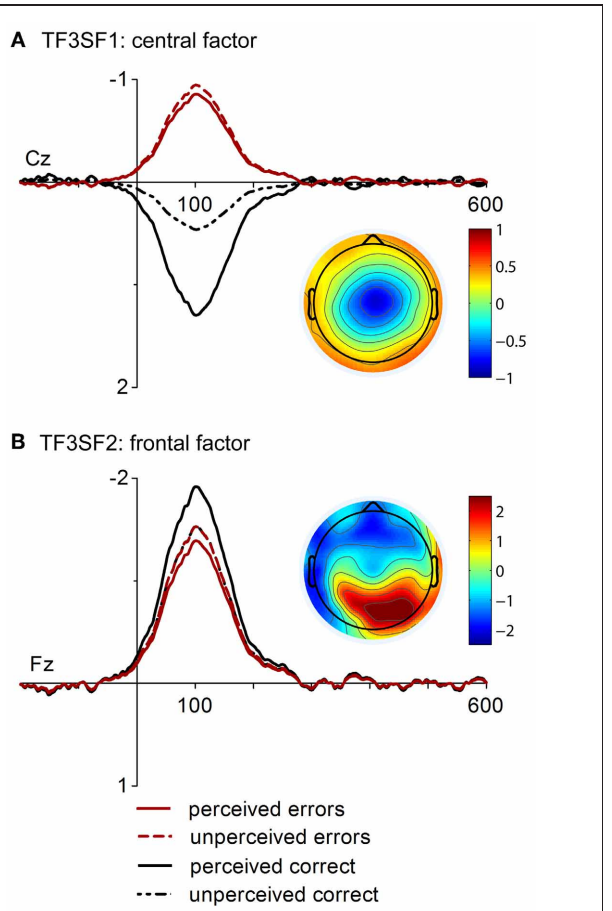


FIGURE 2 | "Virtual ERPs" and topographic maps (at 100 ms) of the spatial factors contributing to ERN and CRN. The waveforms are presented at electrodes Cz and Fz where factor scores were at maximum. Waveforms and topographies of the central factor (**A**. TF3SF1) and the frontal factor (**B**. TF3SF2) are depicted.

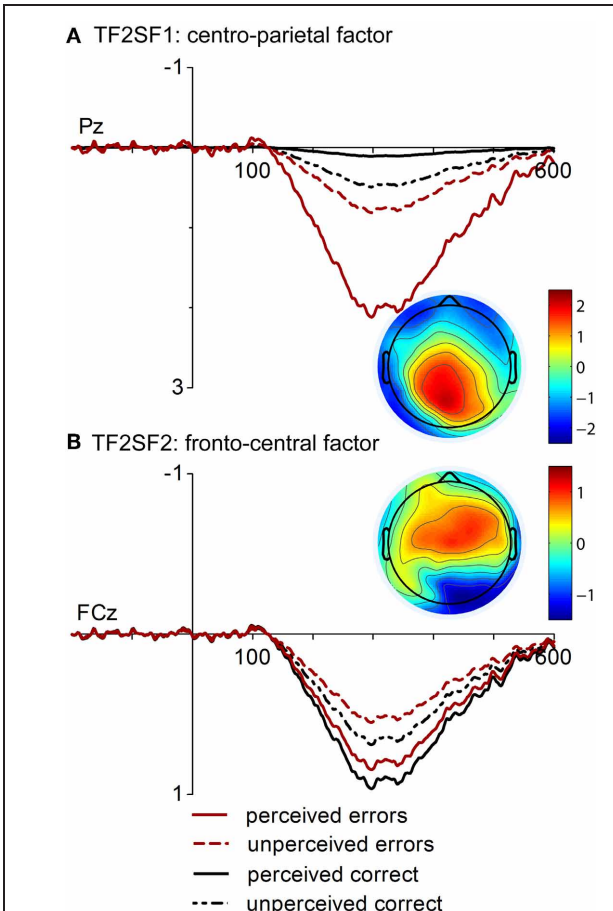


FIGURE 3 | "Virtual ERPs" and topographic maps (at 300 ms) of the spatial factors contributing to Pe. The waveforms are presented at electrodes Pz and FCz where factor scores were at maximum. Waveforms and topographies of the centro-parietal factor (**A**. TF2SF1) and of the fronto-central factor at FCz (**B**. TF2SF2) are depicted.

and accounted for 1.49 % of variance. The analysis of the centro-parietal factor (TF2SF1) showed significant main effects of Response Type, [$F_{(1, 16)} = 14.78$, $p = 0.001$, $\eta^2 = 0.48$], and Response Rating, [$F_{(1, 16)} = 13.24$, $p = 0.002$, $\eta^2 = 0.45$]. These main effects were further specified by their significant interaction, [$F_{(1, 16)} = 6.13$, $p = 0.025$, $\eta^2 = 0.28$]. Follow-up comparisons revealed that factor scores were more pronounced for both perceived and unperceived errors. However, the effect of Response Rating was significant for incorrect, but not for correct responses, $p = 0.001$ vs. $p = 0.183$. Factor scores of ERPs elicited by unperceived errors did not differ from those of unperceived correct responses, $p = 0.432$. Finally, the analysis of the fronto-central factor of the Pe (TF2SF2) yielded no significant main effects but a significant interaction of Response Type and Response Rating, [$F_{(1, 16)} = 6.05$, $p = 0.026$, $\eta^2 = 0.27$]. While an effect of Response Rating was found neither for incorrect nor for correct responses, a statistical significant difference between perceived correct responses and unperceived errors was found, $p = 0.045$.

DISCUSSION

The present study investigated error awareness effects in a choice selection task in which errors are caused by perceptual difficulty instead of impulsive responding like in tasks used in most of the error awareness studies so far. The question was whether ERN and Pe amplitudes would be sensitive to the awareness of response correctness. Importantly, the main objective was to disentangle overlapping electrophysiological activity captured in the Pe by means of temporospatial PCA. Thereby, the sensitivity of the Pe subcomponents to error awareness was examined. To this end, a visual discrimination task with three difficulty levels and subsequent accuracy judgments was applied to elicit perceived and unperceived errors as well as perceived and unperceived correct responses. Although it was not possible to examine awareness effects for each difficulty level separately, we obtained a sufficient number of perceived and unperceived incorrect and correct responses when task difficulty was not considered.

ERP results based on classical amplitude analysis indicate that the ERN following errors was significantly larger than the negativity following correct responses. There was no difference in ERN size between perceived and unperceived errors, which is consistent with earlier studies (Nieuwenhuis et al., 2001; Endrass et al., 2005, 2007; O'Connell et al., 2007; Shalgi et al., 2009; Dhar et al., 2011; Hughes and Yeung, 2011), but inconsistent with more recent reports showing an amplitude reduction for unperceived errors (Maier et al., 2008; Hewig et al., 2011; Wessel et al., 2011). In fact, the current study revealed numerically identical ERN amplitudes for both error types. With regard to correct responses, CRNs tended to be larger for correct responses falsely judged as errors compared to perceived correct responses (Scheffers and Coles, 2000). Although the current data do not support a variation of ERNs with perceived incorrectness, the increase of the CRN with subjective error perception supports the view that the early response-related negativity might be important for the emergence of error awareness. While no distinction between error types was found the increase of the CRN with error perception indicates that the ERN might qualify as a necessary but not as a sufficient precondition for error awareness.

The PCA revealed two relevant temporospatial factors in the time range of response-related negativities: the first was characterized by a central negativity and the second by a frontal negativity and a parietal positivity. This factor solution was already shown in an earlier analysis of the current dataset which examined the effect of task difficulty and response correctness (Endrass et al., 2012). In that study, the central factor varied with response correctness and difficulty, while the frontal factor was not sensitive to task manipulations. In the current study the central factor also varied with response correctness and a trend for an interaction with perceived accuracy was found. This interaction indicated that factor scores varied with error perception on correct trials, but not on incorrect ones. Factor scores were larger for correct responses that were falsely considered as incorrect. These results indicate that the central factor is not only sensitive to the distinction between errors and correct responses but is also modulated by false error perception on correct responses. Thus, the modulation of the central factor is in accordance with ERP findings and might represent the underlying activity that modulates response-related negativity amplitudes. Furthermore, the variation with response correctness and its sensitivity to error perception on correct trials further supports the idea that this factor represents specific aspects of error processing. This conclusion cannot be drawn for the frontal factor that varied only with perceived accuracy and was most pronounced for perceived correct responses. Since an influence of response correctness was also absent in our previous analysis, this factor was interpreted to represent an outcome-independent monitoring process contributing to both ERP components (Endrass et al., 2012). Certainly, more studies are needed to thoroughly understand the function of this component.

In contrast to response-related negativities, the error positivity, determined as conventional amplitude measure, showed a distinct variation with error awareness (Nieuwenhuis et al., 2001; Endrass et al., 2005, 2007; O'Connell et al., 2007; Shalgi et al., 2009; Dhar et al., 2011; Hughes and Yeung, 2011; Wessel et al., 2011). The current study supports these findings and shows that a variation of the Pe can also be found in a task where errors are caused by perceptual difficulty instead of failed response inhibition towards an imperative stimulus (Hughes and Yeung, 2011). Importantly, the Pe varied not only between perceived and unperceived errors but was also more positive for unperceived compared to perceived correct responses. Therefore, the Pe may represent a gradual measure for error awareness with its amplitude being most pronounced for perceived errors and least pronounced for perceived correct responses. This view is compatible with the conclusion drawn by Hewig and colleagues who found a similar modulation of the Pe amplitude by subjective correctness (Hewig et al., 2011). Regarding the idea that multiple changes in the salience network accumulate to the conscious perception of an error (Ullsperger et al., 2010), it seems that the modulation of the error positivity with subjective error perception on correct trials activates the salience network which leads to false error detection on correct trails. Although the current study found a variation of the Pe with false error perception on correct trials, the Pe in that condition was smaller compared to perceived errors. This indicates that subjective error perception

may not depend on the presence of a pronounced Pe. Therefore, the current data are consistent with the assumption that the Pe, like the ERN, represents an internal error signal which is related to error processing and potentially to subjective error awareness, but it is not an all-or-nothing process that triggers error perception. Instead, error awareness might arise from multiple inputs at various stages during error processing (Steinhauser and Yeung, 2010; Ullsperger et al., 2010).

PCA analysis disentangled the error positivity into two temporospatial PCA factors as in the previous analysis of the present data (Endrass et al., 2012). Like in the previous analysis, a modulation of the centro-parietal factor with response correctness was found. Interestingly, the current study revealed an effect of subjective awareness on this same factor. The largest factor scores were found for perceived errors and smallest for perceived correct responses, while unperceived errors and unperceived correct responses did not differ and layed in-between. This modulation is fairly consistent with the results at ERP level suggesting that this factor represents the underlying process that triggers error awareness effects in the Pe. A centro-parietal factor was also identified in a previous PCA study. Therefore, the current study supports the assumption that the Pe represents a P300-like component to error commission (Arbel and Donchin, 2009). This component varied with error significance (Arbel and Donchin, 2009) and with error awareness in the current study. The modulation of the centro-parietal component subsequent to error commission may reflect the importance of an error and thus the saliency of that event. Consequently, the current PCA results support the idea that the Pe reflects the activity of a salience network that leads to error detection.

Post-error slowing, i.e., response time slowing following incorrect responses depended on subjective accuracy ratings and was more pronounced following perceived than unperceived errors. Therefore, it was argued that post-error slowing depended on the process reflected by the Pe and is also elicited by perceived errors (Nieuwenhuis et al., 2001; Endrass et al., 2007; Klein et al., 2007; Wessel et al., 2011). However, the current study not only revealed post-response slowing following perceived errors but also following unperceived correct responses. Therefore the slowing of response times in the subsequent trial occurred only when the previous response was perceived as incorrect independently of the objective correctness of that response (Hewig et al., 2011). Thus, it seems that post-error slowing occurred without a pronounced Pe in case of false error perception on correct trials. An alternative account considers post-error slowing as automatic behavioral adjustment by implementing compensatory cognitive control in the medio-frontal cortex (Botvinick et al., 2001; Gehring and Fencsik, 2001; Debener et al., 2005). However, the current data are also incompatible with that interpretation since unperceived errors elicited an ERN but were not followed by post-error slowing. A recent study found that post-response slowing occurred as a consequence to infrequent events irrespective of response correctness (Notebaert et al., 2009). Thus, it was considered as a slowing of task-relevant processing caused by an orienting response towards unexpected events. Although the current data cannot help to disentangle whether post-response slowing is an adaptive mechanism to prevent future errors or a distortion of stimulus processing leading to slower responses, it seems that this process is related to perceived incorrectness of a response rather than an automatic adjustment process outside of conscious error perception.

Whereas subsequent response times were influenced by a subjective error perception, post-error correctness was only influenced by effective error commission. Therefore, it might be assumed that the ERN which was present for perceived and unperceived errors is more closely related to post-error correctness. This relation appears independent of subjective error awareness. Thus, the current data may suggest that ERN and Pe reflect two partially independent error monitoring mechanisms leading to different adjustments at the behavioral level. One internal mechanism reflected in the ERN that detects errors and initiates subsequent adjustment of cognitive control improving correctness of a subsequent action. This adjustment might be accomplished outside of awareness but nevertheless related to automatic error processing, like increased theta coupling between the medial and lateral prefrontal cortex (Cavanagh et al., 2009), autonomic changes (Wessel et al., 2011), or the later error positivity. The second mechanism might be reflected in the Pe or, more specifically, its centro-parietal subcomponent. It is conceived as a P300-like component to motivationally salient events, i.e. to perceived incorrect responses. Subjective error perception though appears not the only modulator of the Pe since it was reduced in case of false error perception on correct trials. Thus, subjective error awareness may depend on gradual inputs during error processing like changes ERN, Pe, or autonomic response.

Possible limitations of the current study should be noted. First, error perception rates were quite low. Only 51.5% of the errors were recognized by the participants. However, similar rates were found in other experiments, like antisaccade tasks (Nieuwenhuis et al., 2001; Endrass et al., 2007). In addition, only 6.8% of the correct responses were falsely perceived as errors which leads to an acceptable discrimination accuracy ($d' = 1.52$) suggesting that participants were able to judge response accuracy. Second, we were not able to separately analyze error awareness and task difficulty. Although both perceived and unperceived errors and correct responses were present at all difficulty levels, the amount of misclassifications, i.e., incorrect perceptions of the responses, varied with perceptual difficulty, especially for correct responses. Because more trials were presented in the easy condition the numbers of committed easy and difficult errors within perceived and unperceived errors did not differ. In addition, the analysis of a subsample of participants who had more than six perceived and unperceived errors in the easy and the difficult condition showed that an effect of error awareness was present for the Pe in both difficulty levels, but no variation of the ERN was found. Hence, the current findings of a difference between perceived and unperceived errors appear not to be caused by unequal amounts of easy and difficult trials. Nevertheless, the question whether the error awareness effect on Pe amplitudes is independent of task difficulty needs further examination.

To summarize, the current study replicates previous findings and demonstrates that subjective error awareness selectively modulates Pe but not ERN amplitudes. The temporospatial PCA identified two underlying factors for both ERP components.

The factors in the time range of ERN were characterized by a central and a frontal topography. While the central factor putatively reflecting error processing was modulated by objective response correctness, it was unrelated to subjective error perception. The frontal factor which appears to reflect outcome-independent monitoring (Endrass et al., 2012) varied with subjective awareness irrespective of actual response correctness. The PCA factors underlying the Pe had a fronto-central and a centro-parietal distribution. Only the latter factor was related to error awareness: a selective enhancement was found for perceived errors compared to all other response types. Taken together these results support the functional distinctions of subcomponents of both the ERN and the Pe. The PCA results suggest that Pe modulations by experimental conditions are mostly due to variations of the centro-parietal subcomponent. Therefore, the Pe may represent a P300-like response reflecting motivational significance of an error and saliency processing.

ACKNOWLEDGMENTS

This work was supported by the German Research Foundation (EN 906/1-1).

REFERENCES

Arbel, Y., and Donchin, E. (2009). Parsing the componential structure of post-error ERPs: a principal component analysis of ERPs following errors. *Psychophysiology* 46, 1179–1189.

Berg, P., and Scherg, M. (1994). A multiple source approach to the correction of eye artifacts. *Electroencephalogr. Clin. Neurophysiol.* 90, 229–241.

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. (2007). Risk prediction and aversion by anterior cingulate cortex. *Cogn. Affect. Behav. Neurosci.* 7, 266–277.

Brown, J. W., and Braver, T. S. (2008). A computational model of risk, conflict, and individual difference effects in the anterior cingulate cortex. *Brain Res.* 1202, 99–108.

Cavanagh, J. F., Cohen, M. X., and Allen, J. J. (2009). Prelude to and resolution of an error: EEG phase synchrony reveals cognitive control dynamics during action monitoring. *J. Neurosci.* 29, 98–105.

Critchley, H. D., Tang, J., Glaser, D., Butterworth, B., and Dolan, R. J. (2005). Anterior cingulate activity during error and autonomic response. *Neuroimage* 27, 885–895.

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

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

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

Dien, J. (2010a). The ERP PCA toolkit: an open source program for advanced statistical analysis of event-related potential data. *J. Neurosci. Methods* 187, 138–145.

Dien, J. (2010b). Evaluating two-step PCA of ERP data with Geomin, Infomax, Oblimin, Promax, and Varimax rotations. *Psychophysiology* 47, 170–183.

Dien, J., Beal, D. J., and Berg, P. (2005). Optimizing principal components analysis of event-related potentials: matrix type, factor loading weighting, extraction, and rotations. *Clin. Neurophysiol.* 116, 1808–1825.

Dien, J., and Frishkoff, G. (2005). "Principal components analysis of event-related potential datasets," in *Event-Related Potentials: A Methods Handbook,* ed T. C. Handy (Cambridge, MA: The MIT Press), 189–208.

Dien, J., Michelson, C. A., and Franklin, M. S. (2010). Separating the visual sentence N400 effect from the P400 sequential expectancy effect: cognitive and neuroanatomical implications. *Brain Res.* 1355, 126–140.

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

Endrass, T., Klawohn, J., Gruetzmann, R., Ischebeck, M., and Kathmann, N. (2012). Response-related negativities following correct and incorrect responses: evidence from a temporospatial principal component analysis. *Psychophysiology* 49, 733–743.

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.

Falkenstein, M., Hohnsbein, J., Hoormann, J., and Blanke, L. (1991). Effects of cross-modal divided attention on late ERP components: II. 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.

Falkenstein, M., Hoormann, J., and Hohnsbein, J. (1999). ERP components in Go Nogo tasks and their relation to inhibition. *Acta Psychol.* 101, 267–291.

Ford, J. M. (1999). Schizophrenia: the broken P300 and beyond. *Psychophysiology* 36, 667–682.

Foti, D., Hajcak, G., and Dien, J. (2009). Differentiating neural responses to emotional pictures: evidence from temporal-spatial PCA. *Psychophysiology* 46, 521–530.

Foti, D., Weinberg, A., Dien, J., and Hajcak, G. (2011). Event-related potential activity in the basal ganglia differentiates rewards from nonrewards: temporospatial principal components analysis and source localization of the feedback negativity. *Hum. Brain Mapp.* 32, 2207–2216.

Gehring, W. J., and Fencsik, D. E. (2001). Functions of the medial frontal cortex in the processing of conflict and errors. *J. Neurosci.* 21, 9430–9437.

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.

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.

Herrmann, M. J., Rommler, J., Ehlis, A. C., Heidrich, A., and Fallgatter, A. J. (2004). Source localization (LORETA) of the error-related-negativity (ERN/Ne) and positivity (Pe). *Cogn. Brain Res.* 20, 294–299.

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.

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.

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

Kayser, J., and Tenke, C. E. (2005). Trusting in or breaking with convention: towards a renaissance of principal components analysis in electrophysiology. *Clin. Neurophysiol.* 116, 1747–1753.

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.

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

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.

Nieuwenhuis, S., Aston-Jones, G., and Cohen, J. D. (2005). Decision making, the P3, and the locus coeruleus-norepinephrine system. *Psychol. Bull.* 131, 510–532.

Nieuwenhuis, S., Ridderinkhof, K. R., Blow, J., Band, G. P. H., 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.

Notebaert, W., Houtman, F., Opstal, F. V., Gevers, W., Fias, W., and Verguts, T. (2009). Post-error slowing: an orienting account. *Cognition* 111, 275–279.

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.

Picton, T. W. (1992). The P300 wave of the human event-related potential. *J. Clin. Neurophysiol.* 9, 456–479.

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.

Scheffers, M. K., and Coles, M. G. H. (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.

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., Volz, K. G., and Von Cramon, D. Y. (2004). A common neural system signaling the need for behavioral changes. *Trends Cogn. Sci.* 8, 445–446; author reply 446–447.

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.

Vocat, R., Pourtois, G., and Vuilleumier, P. (2008). Unavoidable errors: a spatio-temporal analysis of time-course and neural sources of evoked potentials associated with error processing in a speeded task. *Neuropsychologia* 46, 2545–2555.

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.

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.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 17 January 2012; accepted: 31 May 2012; published online: 22 June 2012.

Citation: Endrass T, Klawohn J, Preuss J and Kathmann N (2012) Temporospatial dissociation of Pe subcomponents for perceived and unperceived errors. Front. Hum. Neurosci. ***6****:178. doi: 10.3389/fnhum.2012.00178*

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 19 June 2012
doi: 10.3389/fnhum.2012.00177

Error-related anterior cingulate cortex activity and the prediction of conscious error awareness

Catherine Orr* and Robert Hester

Department of Psychological Sciences, University of Melbourne, Melbourne, VIC, Australia

Edited by:
Tilmann A. Klein, Max Planck Institute for Human Cognitive and Brain Sciences, Germany

Reviewed by:
Michael X. Cohen, University of Amsterdam, Netherlands
Simon van Gaal, University of Amsterdam, Netherlands

****Correspondence:***
Catherine Orr, Department of Psychological Sciences, University of Melbourne, Melbourne, VIC 3010, Australia.
e-mail: corr@unimelb.edu.au

Research examining the neural mechanisms associated with error awareness has consistently identified dorsal anterior cingulate cortex (ACC) activity as necessary but not predictive of conscious error detection. Two recent studies (Steinhauser and Yeung, 2010; Wessel et al., 2011) have found a contrary pattern of greater dorsal ACC (dACC) activity [in the form of the error-related negativity (ERN)] during detected errors, but suggested that the greater activity may instead reflect task influences (e.g., response conflict, error probability) and or individual variability (e.g., statistical power). We re-analyzed fMRI BOLD data from 56 healthy participants who had previously been administered the Error Awareness Task (EAT), a motor Go/No-go response inhibition task in which subjects make errors of commission of which they are aware (Aware errors), or unaware (Unaware errors). Consistent with previous data, the activity in a number of cortical regions was predictive of error awareness, including bilateral inferior parietal and insula cortices, however, in contrast to previous studies, including our own smaller sample studies using the same task, error-related dACC activity was significantly greater during aware errors when compared to unaware errors. While the significantly faster RT for aware errors (compared to unaware) was consistent with the hypothesis of higher response conflict increasing ACC activity, we could find no relationship between dACC activity and the error RT difference. The data suggests that error awareness is associated with error-related dACC activity but that the role of this activity is probably best understood in relation to the activity in other regions. Activity in the dACC may be important to conscious error detection, but it remains unclear what task and individual factors influence error awareness.

Keywords: error awareness, anterior cingulate cortex, performance monitoring

INTRODUCTION

Goal directed behavior requires the ability to recognize appropriate responses and to flexibly adjust behavior in response to an error. Even in the absence of explicit feedback, people demonstrate characteristic reactions following an error and often spontaneously correct their response (Nieuwenhuis et al., 2001; O'Connell et al., 2009). These behaviors suggest the activity of a performance monitoring system, which evaluates actions and allows adaptive adjustments in attention and cognitive control mechanisms to reduce the likelihood of repeating an error (Ullsperger and von Cramon, 2001). Cognitive neuroimaging research has consistently implicated the prefrontal cortex (PFC) and the anterior cingulate cortex (ACC) in error processing (Garavan et al., 2003; Ridderinkhof et al., 2004). Moreover, hypoactivity in this network has been associated with deficits in error-related processing and subsequent behavioral adjustments observed in populations with schizophrenia (Morris et al., 2006), Attention Deficit Hyperactivity Disorder (ADHD; Burgio-Murphy et al., 2007), Alzheimer's disease (Mathalon et al., 2003), and substance use disorders (Franken et al., 2007).

It has been suggested that error processing can proceed, at least in part, outside of conscious awareness. We are not aware of every error we commit and research explicitly assessing error awareness has shown that even spontaneously corrected errors can go unreported (Endrass et al., 2005, 2007; Wessel et al., 2011). Although some post-error adaptations can occur outside of awareness, Ullsperger and colleagues (2010) argued that conscious awareness is critical to the affective significance of an error. It is likely that the affective evaluation of one's performance influences overall motivation and longer-term strategic approach to the task. Reduced awareness of errors has been associated with a lack of insight into maladaptive behaviors in drug addiction (Hester et al., 2007), ADHD (O'Connell et al., 2009), and psychopathy (Brazil et al., 2009). It is important, therefore, to understand the conditions under which errors reach consciousness and the neural correlates of error awareness.

Electrophysiological research has consistently associated a positively deflecting event-related potential (ERP) observed 100–200 ms following an error (the error positivity: Pe) with awareness of the error (Endrass et al., 2005; Overbeek et al., 2005). Typically, the Pe is reduced or absent for those errors of which the participant remains subjectively unaware. It has been suggested that the Pe may be a context specific P3, a component associated with attentional orienting to stimuli of motivational significance (Ridderinkhof et al., 2009). The P3/Pe is believed to arise from activity in the prefrontal and parietal cortices (Soltani and Knight,

2000), and is recorded as an average over a centro-parietal electrode site. Functional neuroimaging research corroborates these suggestions in implicating a network of frontal and parietal regions in error awareness, notably the insulae (Hester et al., 2005, 2009a; Klein et al., 2007), bilateral inferior parietal (Hester et al., 2005, 2009a), and bilateral mid frontal (Hester et al., 2005, 2009a) cortices.

Performance monitoring is also associated with a negatively deflecting ERP that occurs immediately following the erroneous response, the error-related negativity (ERN) (Falkenstein et al., 1991), which is believed to reflect a stage of error processing that is functionally distinct from that reflected in the Pe (Overbeek et al., 2005). For example, although the ERN is significantly larger following an error than a correct response (Dehaene et al., 1994), ERN magnitude does not reliably discriminate aware from unaware errors (Endrass et al., 2005; O'Connell et al., 2007; but cf. Scheffers and Coles, 2000 and refer Wessel in this issue for a comprehensive review). It has been suggested that the ERN represents encoding of an element of the task environment or of task performance associated with errors; specifically, it has been proposed that the ERN reflects response conflict (Botvinick et al., 2001; Yeung et al., 2004), thwarted reward expectations (Holroyd et al., 2004), or detected changes in error likelihood (Brown and Braver, 2005).

The ERN has been localized to the ACC (Halgren et al., 2002; Debener et al., 2005; Wang et al., 2005) and fMRI research is largely consistent with the ERP findings in showing that, although BOLD activity in dorsal ACC (dACC) is greater for errors than correct responses, this region is not sensitive to error awareness (Hester et al., 2005, 2009a; Klein et al., 2007). The role of the ERN/dACC in error processing is unclear, but elevated levels of dACC activity have been associated with error-related autonomic arousal (Critchley et al., 2005) and with more conservative responding following an error, measured as post-error slowing (PES) of reaction time (RT) (Garavan et al., 2002; Kerns et al., 2004). Autonomic arousal and PES have been shown to be reduced or absent following unaware errors (Nieuwenhuis et al., 2001; Endrass et al., 2007; O'Connell et al., 2009). It is "surprising" (Klein et al., 2007, p. 1779), therefore, that those studies that have explicitly examined the role of dACC in error awareness have consistently reported equivalent levels of dACC activity for aware and unaware errors in healthy controls.

In contrast to the predominance of past research, two recent studies (Steinhauser and Yeung, 2010; Wessel et al., 2011) have reported sensitivity in dACC to error awareness in the form of greater ERN magnitude for aware than unaware errors. While it appears that these recent findings contradict the extensive literature on the role of the ERN/dACC, in this issue Wessel reviews past research and argues that the ERN/dACC is critically involved in error awareness. Wessel discusses findings of error sensitivity in the ERN (Maier et al., 2008) and the relationship between the ACC and awareness in other paradigms (Dehaene et al., 2003; Mayr, 2004) and suggests that methodological considerations may have obscured the role of ERN in error awareness in past research. Wessel argues that those studies that limit the time participants have to acknowledge an error, or in which signaling an error requires an additional "error awareness" response may induce a conservative response bias in which some liminal errors go unreported (cf. Nieuwenhuis et al., 2001; Woodman, 2010; Wessel et al., 2011). Alternatively, it was suggested that the inconsistent ERN/error awareness relationship might be the result of low statistical power in the reported studies. After reviewing past research, Wessel and colleagues (2011) concluded that it is common for studies to demonstrate numerically larger ERN for aware than for unaware errors (e.g., Endrass et al., 2007; Hughes and Yeung, 2011), but that this effect falls short of statistical significance. The authors proposed that small sample sizes and low numbers of unreported errors in past research mean that conclusions that the ERN is not sensitive to error awareness may have been drawn from research with insufficient statistical power.

Review of the fMRI research suggests that these studies are also vulnerable to the criticism that they may have lacked sufficient statistical power to determine the presence or absence of a relationship between the dACC and error awareness. Due to the practicalities of collecting fMRI data, sample sizes are comparatively small and experimental sessions are typically of a length that limits the number of unreported errors. As with the electrophysiological findings relating to the ERN, fMRI studies have consistently reported numerically larger BOLD signal change in the dACC for aware than unaware errors (Hester et al., 2005, 2009a; Klein et al., 2007). The fact that this difference has failed to reach statistical significance may be due, as was argued to be the case for ERP research (Wessel et al., 2011), to a lack of statistical power. Alternatively, a recent investigation using multimodal neuroimaging techniques (Agam et al., 2011) suggested that the ERN may originate in the posterior cingulate cortex (PCC), raising the possibility of a functional dissociation between the ERN and activity in the dACC. In light of these suggestions, the relationship between the dACC, erroneous responses, and error awareness warrants reinvestigation.

The aim of the present study was to determine whether more powerful statistical analyses would demonstrate a relationship between the dACC and error awareness. The present study combined the samples of three previous studies (Hester et al., 2005, 2009a, 2012) to form a group of 56 healthy controls. The studies employed the Error Awareness Task (EAT: Hester et al., 2005), a motor Go/No-go task modified to provide a measure of error awareness and to optimize unaware errors. The behavioral performance and cortical activation levels of the composite sample were analyzed to reassess the neural mechanisms associated with error awareness and error-related behavior.

MATERIALS AND METHODS

SUBJECTS AND TASK DESIGN

Fifty-six right-handed participants (seven female, mean age 27, range: 20–41), reporting no history of neurological or psychological impairment, completed a version of the EAT (Hester et al., 2005). The experimental procedure was approved by the ethics committees of the relevant institutions (University of Melbourne, Trinity College Dublin, and University of Queensland), and participants provided written, informed consent at the beginning of the experimental session. The EAT is a motor Go/No-go response inhibition task, modified to allow participants to acknowledge errors of commission of which they are aware (refer **Figure 1**).

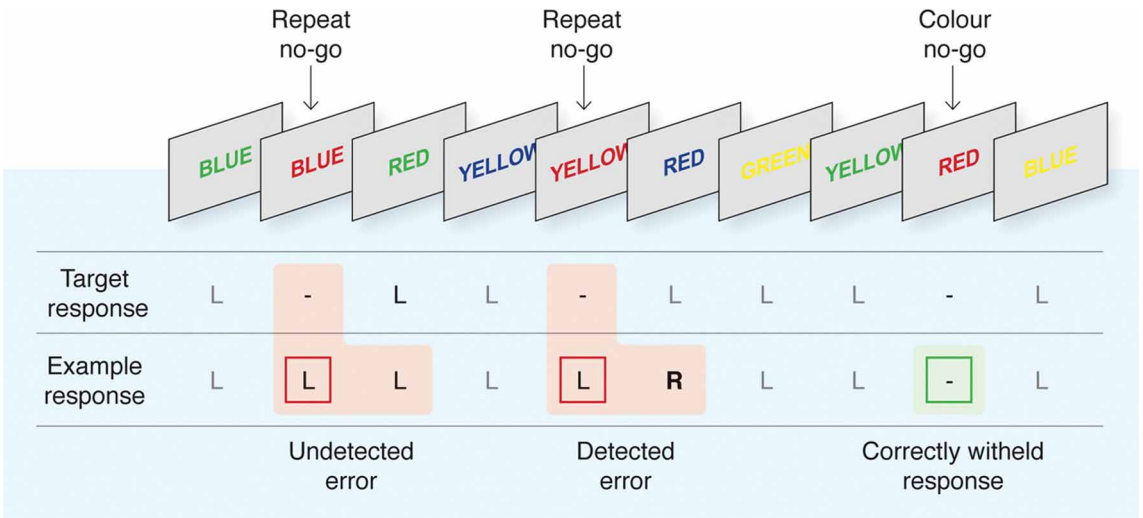


FIGURE 1 | The Error Awareness Task (EAT). The EAT presents a serial stream of single color words in colored fonts. Participants were trained to respond to each of the words with a single "Go trial" button press, and to withhold this response when either of two different circumstances arose. The first was if the same word was presented on two consecutive trials (Repeat No-go), and the second was if the word and font of the word matched (Color No-go). To indicate "error awareness" participants were trained to forego the regular go trial button response and instead to execute the alternative "error awareness" response following any commission error. Past studies have demonstrated that error-related BOLD signal is uninfluenced by the awareness response itself (Hester et al., 2005). Although levels of awareness undoubtedly vary on a continuum, we made a qualitative distinction between "Aware" and "Unaware" errors to facilitate our event-related fMRI analysis. Figure reproduced from Hester and colleagues (2012).

Color names, printed in colored font, were presented in a serial stream. The three samples completed versions of the task that differed slightly in the stimulus presentation and inter-stimulus interval duration, 900/600 ms (Hester et al., 2005, 2009a) or 800/700 ms (Hester et al., 2012). Participants were instructed to respond to the presentation of each stimulus with a button press (a "Go" response) unless the stimulus was a "lure," in which case they were to withhold their response ("No-go" response). Lures could take two forms: a "Repeat lure" in which the same word was presented on two consecutive trials; and a "Color lure" defined by the congruence of the color name and the font color. Color lures were defined by incongruence between the color word and the font in the two earlier studies (Hester et al., 2005, 2009a); and by congruence in the third study (Hester et al., 2012). Adopting competing inhibition rules exploits the different strengths of the stimulus-response relationships such that the overlearned behavior of reading a word would make the Repeat rule more salient than the Color rule. Previous research has suggested that this may cause the Color rule to be suppressed, producing more Color errors than Repeat errors and potentially affecting participants' awareness of the errors (Hester et al., 2005, 2009a). Participants were instructed to execute an alternative "error awareness" response on the "Go" trial following an error.

The 2005 and 2009 studies presented five blocks of 225 trials during fMRI data acquisition with lure trials pseudo randomly distributed across the 1125 experimental trials. The 2012 study presented six blocks of 225 trials. There were minor differences to the ratio of No-go: Go trials across the three experiments, with 128 lures presented in the earlier study (Hester et al., 2005) and 125 and 150, respectively, in the two later studies (Hester et al., 2009a, 2012). On average, a lure was presented every 8.95 trials, corresponding to an average inter-lure interval of 13.42 s. While some other variations in the design of these tasks existed, only the commission errors made during the aforementioned design were considered, with the assumption made that the event-related analysis would minimize the influence of unrelated task variance.

SCANNING PARAMETERS AND DATA ANALYSES

Scanning for the original study (Hester et al., 2005) was conducted using contiguous 5 mm sagittal slices covering the entire brain from a 1.5 T Siemens Vision scanner using a single shot, T2*-weighted echo-planar pulse sequence (TE = 50 ms; TR = 2000 ms; FOV = 256 mm; 64 × 64 matrix). High resolution T1-weighted structural MPRAGE images (FOV = 256 mm, thickness = 1.0 mm isotropic with no gap) were acquired prior to functional imaging to allow subsequent activation localization and for spatial normalization. Stimuli were delivered using an IFIS-SA stimulus-delivery system (MRI Devices Corp., Waukesha, Wisconsin), which was equipped with a head-coil-mounted 640 × 480 LCD panel. This shielded LCD screen is mounted on the head-coil, directly in the subjects' line of vision. Foam padding was used to limit head movements within the coil.

Scanning for the second study (Hester et al., 2009a) was conducted using a 3T Philips Achieva scanner in which foam padding was used to restrict head movements. Contiguous 3.5 mm sagittal slices covering the entire brain were collected using a gradient-echo echo-planar imaging (EPI) sequence (TE = 35 ms; TR = 2000 ms; FOV = 224 mm). High resolution T1-weighted structural MPRAGE images (FOV = 256 mm, isotropic 0.9 mm voxels) were acquired following functional imaging to allow subsequent activation localization and spatial normalization. Stimuli

were back-projected onto a screen at the subject's feet and were viewed with the aid of prism glasses attached to the inside of the radio-frequency head-coil.

Scanning for the third study (Hester et al., 2012) using a whole-body 1.5 Tesla Siemens scanner with a gradient-echo EPI sequence. EPI images were acquired using a gradient-echo pulse sequence and sequential slice acquisition ($T_R = 2000$ ms, $T_E = 30$ ms, flip angle = 90°, 29 contiguous slices of 3 mm thickness, 10% gap, in-plane resolution of 3.6 × 3.6 pixels in a FOV of 384 mm). Activation data were registered to high resolution T1-weighted isotropic (1 mm^3) structural MPRAGE images. Only the data from the placebo condition of this study was used for this reanalysis.

Behavioral data from each participant were used to categorize the trial events into successful responses (stops), aware errors, and unaware errors. All analyses were conducted using AFNI software (http://afni.nimh.nih.gov/afni/) (Cox, 1996). Following image reconstruction, the time-series data were motion-corrected using 3D volume registration (least-squares alignment of three translational and three rotational parameters). Activation outside the brain was removed using edge detection techniques.

Separate haemodynamic impulse response functions (IRFs) at 2 s temporal resolution were calculated using deconvolution techniques for aware errors, unaware errors, and stop events. Response functions for all regressor events were initiated at image acquisition onsets because the presentation of all epochs-of-interest was timed to coincide with the beginning of the 2 s TR-cycle. A non-linear regression program determined the best-fitting gamma-variate function for these IRFs as previously described (Murphy and Garavan, 2005). The area under the curve of the gamma-variate function was expressed as a percentage of the area under the baseline. The baseline in this design is indicative of task-related go trial processing that remains after the variance related to the other types of events have been removed.

The percentage area (event-related activation) map voxels were re-sampled at 1 mm^3 resolution, then spatially normalized to standard MNI space (MNI 152 template), and spatially blurred with a 3 mm isotropic root mean squared Gaussian kernel. Group activation maps for errors were determined with one-sample *t*-tests against the null hypothesis of zero event-related activation changes (i.e., no change relative to baseline). Significant voxels within group maps passed a voxelwise statistical threshold ($t = 5.23$, $p \leq 0.00001$) and were required to be part of a larger 84 μl cluster of contiguous significant voxels. By using a combination of probability thresholding and cluster thresholding, the aim is to maximize the power of the statistical test while holding the likelihood of false-positives to a minimum. To determine the cluster threshold we use a program called 3dClustSim. The program is provided with the number of voxels in the group map, the spatial correlation of voxels (must be contiguous on three sides), and the voxelwise threshold. The program then runs a series of Monte Carlo simulations (10000 iterations for our study) to determine the frequency of clusters of varying sizes produced by chance. From this frequency distribution, we then select the cluster size (84 μl given our parameters) that occurs less than 1% of the time by chance, to give a threshold of $p = 0.01$ (corrected).

The activation clusters from whole-brain analyses of aware and unaware errors were used to create an OR map for the purposes of an ROI analysis. An OR map includes the voxels of activation indicated as significant from either of the constituent maps, which are presented in Appendix. The mean activation for clusters in the combined error map was then calculated for the purposes of an ROI analysis, deriving mean activation levels for aware and unaware errors, that were compared using repeated measures *t*-tests, corrected via a modified Bonferroni procedure for multiple comparisons (Keppel, 1991).

RESULTS

Performance indices for the EAT are summarized in **Table 1**. Participants correctly withheld their responses on 52.8% of the No-go trials, with significantly more successful inhibitions for Color than for Repeat lures (62.6 vs. 42.9%: $t_{(55)} = 9.76$, $p < 0.001$). Participants reported awareness of 82.5% of their errors (range 15–99%), with a significantly greater proportion of Color errors endorsed than Repeat errors (85.5 vs. 77.2%: $t_{(55)} = 2.907$, $p < 0.006$). Participants' awareness of errors was not related to their inhibition accuracy ($r = -0.0115$, ns). The speed of commission error responses was significantly related to awareness of the error [$F_{(2,\ 110)} = 7.222$, $p < 0.002$, $\eta_p^2 = 0.116$]. *Post-hoc* comparisons demonstrated that this effect was due to significantly faster responses for aware errors of commission than for either unaware errors or for correct "Go" responses [$t_{(55)} = 3.403$, $p < 0.002$ and $t_{(55)} = 2.680$, $p < 0.01$, respectively].

Behavioral adjustments following No-go trials were assessed by calculating the difference between RT for the trials that immediately preceded and succeeded the lure. Participants demonstrated significant speeding of responses following aware errors [$t_{(55)} = 10.519$, $p < 0.001$]. This finding is inconsistent with other reports, in which participants slow their rate of responding following an error (Rabbitt, 1966) and this PES is exacerbated by error awareness (Nieuwenhuis et al., 2001). Reductions in RT following aware errors have consistently been reported using

Table 1 | Behavioral performance: inhibition accuracy and reaction time on the EAT.

Category	Mean (SD)
INHIBITION ACCURACY (%)	
Color lure	42.9 (23.6)
Repeat lure	62.6 (21.9)
ERROR AWARENESS (%)	
Color lure	85.5 (17.7)
Repeat lure	77.2 (19.9)
REACTION TIME (ms)	
Go trial	488.2 (14.93)
Aware error	479.1 (16.99)
Unaware error	535.7 (22.88)
POST-LURE ADJUSTMENT (POST-LURE RT—PRE-LURE RT)	
Correct inhibition	−118.1 (100.90)
Aware error	−124.0 (88.24)
Unaware error	27.3 (81.14)

the EAT and it is possible that the post-error speeding observed in this paradigm is due to participants learning that lures were widely spaced. This suggestion is consistent with the significant decrease in RT for the Go trial following a correct inhibition [$t_{(29)} = 118.124$, $p < 0.001$], but is confounded by the fact that following aware errors, participants will execute the "awareness" response without the need to process the post-lure stimulus. PES was observed following errors of which the participant remained unaware [$t_{(55)} = 2.517$, $p < 0.02$]. In order to verify that this effect was not due to strategic slowing of RT with increasing temporal distance from the last perceived lure, a surrogate data-set was generated by selecting Go trials that occurred at the same point following a consciously perceived lure as the unaware error trial. When comparing RT for those trials that immediately preceded and succeeded the surrogate trials, no significant difference was observed [$t_{(38)} = 0.53$, ns], suggesting that post-unaware error slowing was not due to changes in response strategy across trials.

The event-related functional analysis of aware and unaware errors identified eight clusters that showed significantly greater increases in activity associated with aware errors than unaware errors (**Table 2**). Consistent with previous reports, we observed error awareness sensitivity in clusters in the bilateral insula cortices (**Figure 2C**) and the IPL (**Figure 2B**). We also revealed activity in midline structures (dACC, pre-SMA, and SMA; **Figure 2A**) that discriminated aware from unaware errors[1]. When repeating the analysis on a subset of participants who endorsed an equivalent number of color and repeat lure errors [$t_{(31)} = 2.01$, ns], activity in all eight of the clusters identified in the full sample continued to show greater levels of BOLD activity associated with aware than with unaware errors.

Although dACC activity was sensitive to error awareness, there was no relationship between dACC activity and behavioral measures typically associated with error awareness. The speed with which errors of commission were executed was not associated with the degree of increase in dACC activity for either aware ($r = -0.04$, ns) or unaware ($r = -0.047$, ns) errors. Nor were post-error RT adjustments correlated with the change in dACC activity for either aware ($r = -0.064$, ns) or unaware ($r = -0.007$, ns) errors. The magnitude of the discrepancy between BOLD activity in the dACC associated with aware errors and that associated with unaware errors was not related to the speed with which erroneous responses were made or the post-error adjustments in RT.

For aware errors, the relative speed of the erroneous response was related to activity in the SMA ($r = 0.309$, $p < 0.03$), such that slower aware error RT (relative to average Go trial RT) was associated with higher levels of activity in this region. Following an aware error, the post-error adjustment in RT (post-error RT-pre-lure RT) was negatively correlated with activity in the R-IPL ($r = -0.277$, $p < 0.04$) and with both clusters centered on the L-IPL ($r = -0.312$, $p < 0.02$[2] and $r = -0.330$, $p < 0.02$[3]), such that greater levels of activity in these regions predicted greater speeding of RT on the post-error trial. Unaware error RT and post-unaware-error RT were not correlated with BOLD activity in any of the functionally defined ROI. For none of the ROI did the magnitude of the discrepancy between BOLD activity associated with aware errors or that associated with unaware errors correlate with the speed with which erroneous responses were made or the post-error adjustment in RT.

DISCUSSION

The current study was designed to reassess the relationship between dACC activity and error awareness in light of recent reports that an electrophysiological measure of dACC activity (ERN) discriminated aware from unaware errors (Steinhauser and Yeung, 2010; Wessel et al., 2011). In addition to regions previously implicated in error awareness when assessed using the EAT, the insula and the bilateral IPL (Hester et al., 2005, 2009a), we demonstrated that aware errors were associated with significantly greater BOLD signal change in the dACC and supplementary motor cortex (SMC) than unaware errors. Although

[1]Note that, although the center of mass of activity in these midline structures is reported as falling in the left hemisphere, the lateral extent of these clusters was such that the bilateral dACC, pre-SMA, and SMA can be considered to discriminate between aware and unaware errors.

[2]Center of Mass MNI co-ordinates: $x = -49$, $y = -34$, $z = 50$.
[3]Center of Mass MNI co-ordinates: $x = -51$, $y = -47$, $z = 40$.

Table 2 | Regions that showed significantly greater BOLD signal change for aware than unaware errors.

Structure	*p*	Vol. μl	Brodmann Area	HS	Center of mass (MNI Coordinates)		
					x	*y*	*z*
Inferior parietal	<0.00005	1332	40	L	−49	−34	50
Inferior parietal	<0.0007	2758	40	R	49	−49	40
Inferior parietal/SMG	<0.0004	1135	40	L	−51	−47	40
dACC	<0.0006	1510	32	L	−1	19	39
Insula	<0.03	813	13	L	−36	14	2
Insula	<0.004	1167	13	R	44	13	4
MidCingulate/pre-SMA	<0.01	216	6	L	−3	−14	52
SMA	<0.0004	107	6	L	0	−3	61

Positive values for x, y, and z coordinates denote, respectively, locations that are right, posterior, and superior relative to the anterior commissure. Significance test results indicate areas of increased activation associated with aware relative to unaware errors assessed using t-test.

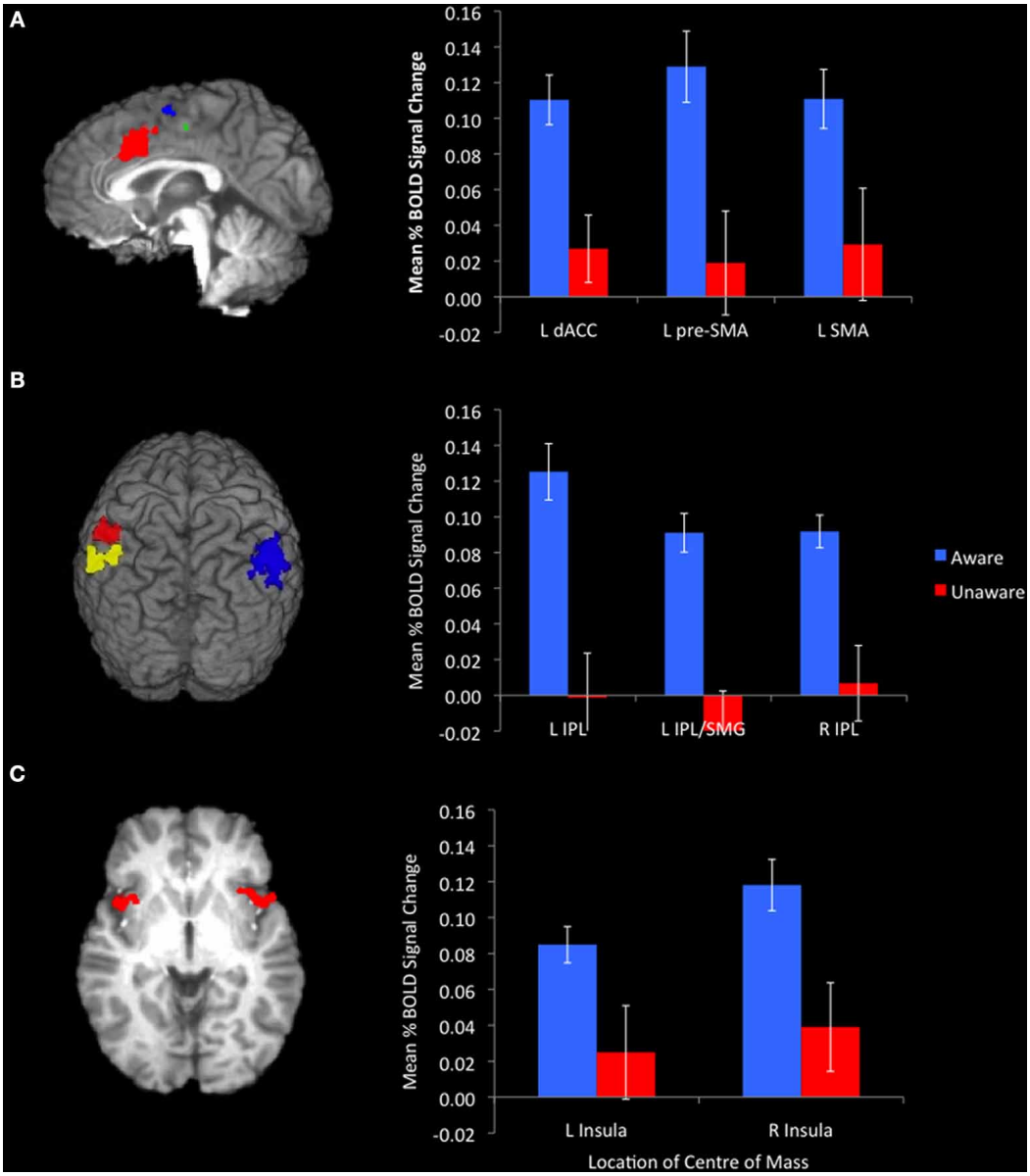


FIGURE 2 | Awareness of errors was associated with significantly greater activation in midline structures. [**(A)** sagittal slice at $x = -1$. Red cluster centered on the dACC, blue cluster centered on the SMA, green cluster centered on the pre-SMA)], the bilateral IPL [**(B)** red cluster centered on the L-IPL, yellow cluster centered on the L-IPL/SMG, blue cluster centered on the R-IPL], and insulae [**(C)** axial slice at $z = 0$]. Bar graphs represent mean %BOLD signal change relative to baseline for aware and unaware errors.

the dACC showed error sensitivity, there was no direct correlation between error-related activity in this region and behaviors typically associated with error awareness (error RT and post-error RT). For this reason, and because this awareness sensitivity was only observed when reanalyzing a composite sample of 56 participants, we suggest that the relationship between dACC, error awareness, and error-related behaviors is indirect and best understood in relation to the activity in other brain regions.

An indirect relationship between dACC and error-related behaviors is consistent with the predictions of a model recently described by Holroyd and Yeung (2011). The authors suggested that the dACC is involved in the selection and evaluation of appropriate sequences of actions directed to attaining a particular goal (Botvinick et al., 2009), rather than the constituent actions themselves. For example, error-related activity in the dACC may index competition between multiple response strategies or the negative reinforcement of an unsuccessful strategy. If activity in the dACC is related to long term goal attainment, rather than the value of individual responses, it is possible that it reflects an aggregation of performance relevant information such as motivational significance, response selection parameters, and the

efficient balance between speed and accuracy. In the context of error processing, activity in the dACC may need to surpass a certain threshold (Yeung et al., 2004) for awareness and post-error adjustments to occur. In this way, activity in the dACC would covary with error commission and contribute to the likelihood that an error will reach awareness, but may not be sufficient to produce awareness in isolation.

Increased activity in the insula cortex has consistently been associated with error awareness (Hester et al., 2005, 2009a; Klein et al., 2007), an effect that is likely due to the central role of the insula in interoceptive awareness (Critchley et al., 2004; Craig, 2009). The insula is implicated in the regulation and awareness of autonomic responses (Critchley et al., 2000; Craig, 2002) and error-related activity in the autonomic nervous system (ANS) is greater following an aware than an unaware error (O'Connell et al., 2007, 2009). There is evidence that dACC provides top-down input to systems regulating autonomic activity (Critchley et al., 2001, 2002) and that the resulting visceral sensation is perceived as a "feeling," indexed by insula activity (Critchley, 2005). Craig (2009) described this relationship as the ACC signal providing a measure of the motivational significance of an event, which is interpreted in the insula in the context of information about social and cognitive conditions to produce awareness of an event and the actor's subjective experience of it. This model is consistent with the suggestion that error-related activity in the dACC may feedforward into regions directly responsible for error awareness (Wessel et al., 2011).

The relationship between the dACC, ANS response, and the insula provides a possible means by which activity in the dACC could mediate error awareness. It is unclear, however, what is communicated by the dACC signal. One influential hypothesis argues that dACC activity indexes conflict or dissonance, in the form of a mismatch between two possible response mappings (Coles et al., 2001) or as conflict between the executed response and the correct response (Yeung et al., 2004). In any task, a proportion of errors may occur due to pre-emptive responding; it has been argued that these errors will be higher in conflict than those errors that occur due to loss of attentional set or impaired stimulus perception as the still evolving representation of the correct response will conflict with the executed response (Yeung et al., 2004). These errors will, obviously, be associated with comparatively fast RT but may also be more likely to be reported as participants become aware of both the executed response and the correct response (Scheffers and Coles, 2000; Shalgi et al., 2007; Steinhauser and Yeung, 2010). If ACC activity indexes high conflict trials, we might predict a relationship between activity in dACC, error awareness, and RT. Consistent with this suggestion, we observed faster RT for aware errors than for either unaware errors or correct "Go" responses. dACC activity, however, was not correlated with RT for either aware or unaware errors. As RT is an indirect measure of response conflict, and the EAT is not designed to discriminate between high conflict and low conflict errors, it is possible that this task is not sensitive to the relationship between dACC and conflict (cf. Hughes and Yeung, 2011). We believe, however, that evaluation of a response and awareness of its appropriateness most likely occurs after the response has been executed. In this way, error RT may affect awareness but would not correlate directly with activity in regions believed to be associated with error processing (insula, IPL, dACC).

In the current study, we reported a correlation between activity in the SMA and aware error RT such that higher levels of activity in the SMA were associated with slower aware-error RT. SMC (comprising the SMA and the pre-SMA) is reliably implicated in successful response inhibition (Aron et al., 2007; Simmonds et al., 2008; Sharp et al., 2010). As such, elevated activity in this region associated with aware errors may reflect an insufficient or belated inhibitory response. Garavan and colleagues (2002) reported that clusters of BOLD activity in the SMC were implicated in both successful response inhibition and in commission errors in a Go/No-go task. By combining the fMRI analysis with EEG, the authors demonstrated that commission errors were related to delayed, rather than absent or insufficient, inhibition-related activity. In the context of the EAT, as participants become aware of the No-go signal and the drive to inhibit increases, RT will slow but not enough to facilitate inhibition if the prepotent "Go" response was initiated prematurely. This slowing of error-RT is not necessarily a conscious adjustment that depends on awareness of an imminent error, but it seems probable that awareness of the No-go signal is more likely to produce an attempt to inhibit the prepotent response and, should this be unsuccessful, awareness of the error. It has been suggested that the SMC and the dACC serve complimentary roles in performance monitoring: signaling that we are at imminent risk of committing an error or that an error has occurred, respectively (Garavan et al., 2003). It is possible that measures of response conflict indexed by the SMC contribute to the dACC signal by providing an indication of the likelihood of an error and, in this way, dACC activity would covary with response conflict but not correlate directly with error-RT. Alternatively, activity in the SMA and dACC may be driven by activity in another source.

Consistent with previous research using the EAT (Hester et al., 2005, 2009a), regions in the bilateral IPL discriminated aware from unaware errors. Along with the ACC and the insula, the IPL forms part of the frontoparietal control system described by Vincent and colleagues (2008). Seeley and colleagues (2007) argued that within this network the dACC and the insula are associated with the salience of an event and regions in the parietal cortex, including the bilateral IPL, act on events identified as salient. It has been suggested the role of the IPL in cognitive control is to maintain sustained attention on task goals (Singh-Curry and Husain, 2009), task parameters (Dosenbach et al., 2006), and response contingencies (Bunge et al., 2003). It is likely that an error constitutes a salient event, and subsequent application of cognitive control may cause participants to re-orient their attention to the task at hand (Coull et al., 1996; Singh-Curry and Husain, 2009). In the current study, activity in the IPL was correlated with the post-error adjustment in RT following an aware error such that elevated levels of BOLD activity predicted faster responses on the post-error trial. In the context of the EAT, a salient error may cause an increase in attention to the task parameters or stimulus-response contingencies, thus affecting the speed with which the error awareness response is made on the subsequent trial.

It must be noted that, although activity in the IPL has previously been associated with error awareness when assessed by the EAT, this activity did not discriminate aware from unaware errors in the anti-saccade task (AST: Klein et al., 2007). It is possible that Klein and colleagues did not observe sensitivity to errors in the IPL due to insufficient statistical power. Alternatively, the EAT may induce a task specific effect due to the response parameters of the Go/No-go task. The EAT requires participants to acknowledge an error with the "error awareness" response rather than the prepotent "Go" response; it is possible that activity in the IPL, pre-SMA, and SMA reflects the intention, preparation, and initiation of the alternative response (Fried et al., 1991; Desmurget and Sirigu, 2009; Desmurget et al., 2009). We believe this to be unlikely, as it would require the preparation of the post-error response to commence during the execution of the erroneous response. Moreover, a previously published study using the EAT (Hester et al., 2005) included a condition that assessed the effect of the additional "awareness" response on BOLD activity in the absence of an error. Only one region in the left middle temporal gyrus showed significant activation for both these odd-ball trials and aware errors, suggesting that the response demands imposed by the task do not account for additional activity observed in the IPL and SMC for aware errors. Unfortunately, this condition was not included in the subsequent studies using the EAT (Hester et al., 2009a, 2012) so we were not able to verify this suggestion with more powerful analysis. In the future, the contribution of the error awareness response to the dACC activity in this paradigm could be assessed by temporally dissociating the awareness response from the commission error itself.

The EAT is not well suited to analyzing post-error behaviors; as discussed above, post-error RT for aware errors is confounded by the requirements of the error awareness response. In error awareness studies using the AST, PES is typically only observed following aware errors (Nieuwenhuis et al., 2001; Wessel et al., 2011) and has been correlated with error-related BOLD activity in the pre-SMA (Klein et al., 2007). Nonetheless, the observation of a small but statistically significant degree of slowing following errors of which the participant was unaware is reliably reported in studies using the EAT and is consistent with reports from two recent studies in which errors were induced by presenting some stimuli below the threshold required for conscious awareness (Pavone et al., 2009; Cohen et al., 2009). Both studies reported PES following aware and unaware errors, but the degree of slowing was greater following errors of which the participants were aware. In the current study, PES following unaware errors was not correlated with the error-related BOLD activity in any of the functionally defined ROI. Using EEG and Granger causality analysis, Cohen and colleagues (2009) revealed increased top-down control over task-relevant sensory regions measured as synchrony between the medial frontal cortex (MFC) and occipital cortex (OCC) following commission errors. The strength of the synchrony between MFC and task-relevant regions was observed to increase following both aware and unaware errors, but was significantly stronger following aware errors. The authors argued that increased synchrony may be the means by which MFC enacts cognitive control by enhancing processing in task-relevant regions, and that this process does not depend on conscious awareness. The BOLD response measured by fMRI does not represent or correlate with all patterns of neural activity, so it may be that increased synchrony leading to PES following unaware errors is not detected in the univariate BOLD response. It is interesting to note that Cohen and colleagues reported no correlation between the degree of PES and the strength of MFC-OCC synchrony; it is possible that increased synchrony is sufficient to induce the less pronounced PES observed following unaware errors but that awareness of the error can induce deliberate remedial strategies over and above these mechanisms, which modulate the nature or degree of post-error behaviors. This suggestion is consistent with the results of Klein and colleagues (2007) in which PES was correlated with error-related activity in the pre-SMA, but only for those errors of which the participant was aware.

It is generally accepted that error awareness is adaptive. Klein and colleagues (2007) reported that the error rate following an aware error was significantly lower than that following an unaware error, suggesting that error awareness improves subsequent performance. In the context of a Go/No-go task, Hester and colleagues (2009b) demonstrated that error-related activity in the dACC, insula, and IPL predicted successful inhibition on the following No-go trial, occurring up to 20 s later. Although Hester and colleagues (2009b) did not explicitly assess awareness, when considered with the results of the current study, it suggests that error awareness and subsequent inhibition success are subserved by common neural mechanisms. It is surprising, therefore, that we observed no correlation between a participant's level of error awareness and their overall inhibition performance. The absence of a relationship between error awareness and inhibition performance in the EAT has previously been reported by O'Connell and colleagues (2009). It is possible that the application of two inhibition contingencies in the EAT disrupts the relationship between error awareness and subsequent performance, reflecting the role of dACC activity as a reinforcement learning signal (Holroyd and Coles, 2002; Holroyd and Yeung, 2011). This suggestion highlights the need for future studies to assess whether error-related neural activity produces a general, transient increase in the conservatism of responding and a subsequent increase in accuracy, or a context specific, learned adaptation. It would be particularly informative to explore the role of error awareness in these dynamics.

Although models of performance monitoring have consistently implicated the dACC in error processing, the relationship between activity in this region and error awareness has not been reported in previously published fMRI investigations (Hester et al., 2005, 2009a; Klein et al., 2007) and not reliably observed in previously published studies using ERP (Wessel et al., 2011). The increased statistical power afforded by the composite sample used in the current study has shown that error-related BOLD activity in the dACC discriminated aware from unaware errors, consistent with recent reports of error awareness sensitivity in the ERN (Steinhauser and Yeung, 2010; Wessel et al., 2011). Although the results of the present study are consistent with the longstanding belief of a functional relationship between the ERN and BOLD activity in the dACC, Agam and colleagues (2011) have recently suggested that the ERN

is generated in the PCC. We did not observe error-related BOLD activity in the PCC, so are unable to comment on Agam and colleagues' suggestion that the ERN/PCC response is an error-specific signal, which is then communicated to the dACC to facilitate behavioral adjustments. The BOLD response is not a direct measure of underlying neural activity and it is possible that the ERN is the result of synchronization of neural populations or the disinhibition of cortical neurons (Agam et al., 2011), phenomena that may not produce significant modulation of the hemodynamic response. It will be necessary to apply multimodal neuroimaging techniques to various error awareness paradigms in order to characterize the role of task and performance factors in the hemodynamic and electrophysiological markers of performance monitoring. Of particular interest will be exploring the influence of the ways in which error awareness is assessed. Those experiments that present some stimuli outside of conscious awareness provide a means of assessing the neural correlates of error awareness without the need of an additional error awareness response that might be vulnerable to failures of memory, task switch costs, or response bias (Woodman, 2010). It remains to be determined whether those errors that occur due to insufficient stimulus representation are as motivationally significant as those that occur due to cognitive failures and whether this mediates error-related neural activity.

The current study has demonstrated a network of regions associated with the motoric, cognitive, and motivational components of performance monitoring that discriminated aware from unaware errors. Analyzing the performance of 56 participants revealed error awareness sensitivity in the dACC, but the EAT does not allow us to determine which performance or contextual factors influenced the response of this region and the emergence of error awareness. Our findings are consistent with the suggestion that dACC activity is indirectly related to error awareness and may feed forward into regions directly responsible for consciousness (Wessel et al., 2011) and remedial post-error behavioral adjustments. Error-related activity in the dACC is probably best understood as part of a network in which the constituent regions are differentially sensitive to the demands of the task and the performance context, and that error awareness and error-related behaviors rely on the pattern of activity in this network.

ACKNOWLEDGMENTS

The authors would like to thank Drs. Mark Bellgrove, Hugh Garavan, Kevin Murphy, and John Foxe for their assistance with the data collection for this study and Mr. Mykalos Byrne for assistance with figures. This work was supported by grants from the Australian Research Council (DP1092852) and National Health and Medical Research Council (1008044) to Robert Hester.

REFERENCES

Agam, Y., Hamalainen, M. S., Lee, A. K., Dyckman, K. A., Friedman, J. S., Isom, M., Makris, N., and Manoach, D. S. (2011). Multimodal neuroimaging dissociates hemodynamic and electrophysiological correlates of error processing. *Proc. Natl. Acad. Sci. U.S.A.* 108, 17556–17561.

Aron, A. R., Behrens, T. E., Smith, S., Frank, M. J., and Poldrack, R. A. (2007). Triangulating a cognitive control network using diffusion-weighted magnetic resonance imaging (MRI) and functional MRI. *J. Neurosci.* 27, 3743–3752.

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.

Botvinick, M. M., Niv, Y., and Barto, A. C. (2009). Hierarchically organized behaviour and its neural foundations: a reinforcement learning perspective. *Cognition* 113, 262–280.

Brazil, I. A., de Bruijn, E. R. A., Bulten, B. H., von Borries, A. K. L., van Lankveld, J. J. D. M., Buitelaar, J. K., and Verkes, R. J. (2009). Early and late components of error monitoring in violent offenders with psychopathy. *Biol. Psychiatry* 65, 137–143.

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

Bunge, S. A., Kahn, I., Wallis, J. D., Miller, E. K., and Wagner, A. D. (2003). Neural circuits subserving the maintenance and retrieval of abstract rules. *J. Neurophysiol.* 90, 3419–3428.

Burgio-Murphy, A., Klorman, R., Shaywitz, S. E., Fletcher, J. M., Marchione, K. E., Holahan, J., Stuebing, K. K., Thatcher, J. E., and Shaywitz, B. A. (2007). Error-related event-related potentials in children with attention-deficit hyperactivity disorder, oppositional defiant disorder, reading disorder, and math disorder. *Biol. Psychol.* 75, 75–86.

Cohen, M. X., van Gaal, S., Ridderinkhof, K. R., and Lamme, V. A. F. (2009). Unconscious errors enhance prefrontal-occipital oscillatory synchrony. *Front. Hum. Neurosci.* 3:54. doi: 10.3389/neuro.09.054.2009

Coles, M. G. H., 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.

Coull, J. T., Frith, C. D., Frackowiak, R. S. J., and Grasby, P. M. (1996). A fronto-parietal network for rapid visual information processing: a PET study of sustained attention. *Neuropsychologia* 34, 1085–1095.

Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* 29, 162–173.

Craig, A. D. (2002). How do you feel? Interoception: the sense of the physiological condition of the body. *Nat. Rev. Neurosci.* 3, 655–666.

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

Critchley, H. D. (2005). Neural mechanisms of autonomic, affective, and cognitive integration. *J. Comp. Neurol.* 493, 154–166.

Critchley, H. D., Mathias, C. J., and Dolan, R. J. (2002). Fear conditioning in humans: the influence of awareness and autonomic arousal on functional neuroanatomy. *Neuron* 33, 653–663.

Critchley, H. D., Melmed, R. N., Featherstone, E., Mathias, C. J., and Dolan, R. J. (2001). Brain activity during biofeedback relaxation: a functional neuroimaging investigation. *Brain* 124, 1003–1012.

Critchley, H. D., Tang, J., Glaser, D., Butterworth, B., and Dolan, R. J. (2005). Anterior cingulate activity during error and autonomic response. *Neuroimage* 27, 885–895.

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

Critchley, H. D., Elliott, R., Mathias, C. J., and Dolan, R. J. (2000). Neural activity relating to generation and representation of galvanic skin conductance responses: a functional magnetic resonance imaging study. *J. Neurosci.* 20, 3033–3040.

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

Dehaene, S., Sergent, C., and Changeux, J. P. (2003). A neuronal network model linking subjective reports and objective physiological data during conscious perception. *Proc. Natl. Acad. Sci. U.S.A.* 100, 8520–8525.

Desmurget, M., Reilly, K. T., Richard, N., Szathmari, A., Mottolese, C., and Sirigu, A. (2009). Movement intention after parietal cortex stimulation in humans. *Science* 324, 811–813.

Desmurget, M., and Sirigu, A. (2009). A parietal-premotor network for

movement intention and motor awareness. *Trends Cogn. Sci.* 13, 411–419.

Dosenbach, N. U. F., Visscher, K. M., Palmer, E. D., Miezin, F. M., Wenger, K. K., Kang, H. C., Burgund, E. D., Grimes, A. L., Schlagger, B. L., and Petersen, S. E. (2006). A core system for the implementation of task sets. *Neuron* 50, 799–812.

Endrass, T., Reuter, B., and Kathmann, N. (2005). Effects of error awareness on error related brain potentials in an antisaccade task. *J. Psychophysiol.* 19, 114.

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.

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

Franken, I. H. A., Rassin, E., and Muris, P. (2007). The assessment of anhedonia in clinical and non-clinical populations: further validation of the Snaith-Hamilton Pleasure Scale (SHAPS). *J. Affect. Disord.* 99, 83–89.

Fried, I., Katz, A., Mccarthy, G., Sass, K. J., Williamson, P., Spencer, S. S., and Spencer, D. D. (1991). Functional organization of human supplementary motor cortex studied by electrical-stimulation. *J. Neurosci.* 11, 3656–3666.

Garavan, H., Ross, T. J., Kaufman, J., and Stein, E. A. (2003). A midline dissociation between error-processing and response-conflict monitoring. *Neuroimage* 20, 1132–1139.

Garavan, H., Ross, T. J., Murphy, K., Roche, R. A. P., and Stein, E. A. (2002). Dissociable executive functions in the dynamic control of behavior: Inhibition, error detection, and correction. *Neuroimage* 17, 1820–1829.

Halgren, E., Boujon, C., Clarke, J., and Wang, C. (2002). Rapid distributed fronto-parieto-occipital processing stages during working memory in humans. *Cereb. Cortex* 12, 710–728.

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., Madeley, J., Murphy, K., and Mattingley, J. B. (2009a). Learning from errors: error-related neural activity predicts improvements in future inhibitory control performance. *J. Neurosci.* 29, 7158–7165.

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

Hester, R., Nandam, S., O'Connell, R., Wagner, J., Strudwick, M., Nathan, P., Mattingley, J., and Bellgrove, M. (2012). Neurochemical enhancement of conscious error awareness. *J. Neurosci.* 32, 2619–2627.

Hester, R., Simoes-Franklin, C., and Garavan, H. (2007). Post-error behavior in active cocaine users: poor awareness of errors in the presence of intact performance adjustments. *Neuropsychopharmacology* 32, 1974–1984.

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

Holroyd, C. B., Nieuwenhuis, S., Yeung, N., Nystrom, L., Mars, R. B., Coles, M. G. H., and Cohen, J. D. (2004). Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals. *Nat. Neurosci.* 7, 497–498.

Holroyd, C. B., and Yeung, N. (2011). "An integrative theory of anterior cingulate cortex function: Option selection in hierarchical reinforcement learning," in *The Neural Basis of Motivational and Cognitive Control,* eds R. B. Mars, J. Sallet, M. F. S. Rushworth, and N. Yeung (Cambridge, MA: MIT Press), 333–349.

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

Keppel, G. (1991). *Design and Analysis: A Researcher's Handbook.* New Jersey, Englewood Cliffs NJ: Prentice Hall.

Kerns, J. G., Cohen, J. D., MacDonald, A. W., Cho, R. Y., Stenger, V. A., and Carter, C. S. (2004). Anterior Cingulate conflict monitoring and adjustments in control. *Science* 303, 1023–1026.

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.

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.

Mathalon, D. H., Bennett, A., Askari, N., Gray, E. M., Rosenbloom, M. J., and Ford, J. M. (2003). Response-monitoring dysfunction in aging and Alzheimer's disease: an event-related potential study. *Neurobiol. Aging* 24, 675–685.

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

Morris, S. E., Yee, C. M., and Nuechterlein, K. H. (2006). Electrophysiological analysis of error monitoring in schizophrenia. *J. Abnorm. Psychol.* 115, 239–250.

Murphy, K., and Garavan, H. (2005). Deriving the optimal number of events for an event-related fMRI study based on the spatial extent of activation. *Neuroimage* 27, 771–777.

Nieuwenhuis, S., Ridderinkhof, K. R., Blow, J., Band, G. P. H., 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.

Pavone, E. F., Marzi, C. A., and Girelli, M. (2009). Does subliminal visual perception have an error-monitoring system? *Eur. J. Neurosci.* 30, 1424–1431.

Rabbitt, P. M. A. (1966). Errors and error correction in choice-response tasks. *J. Exp. Psychol.* 71, 264–272.

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 Nieuwenhuiss, S. (2004). The role of the medial frontal cortex in cognitive control. *Science* 306, 443–447.

Scheffers, M. K., and Coles, M. G. H. (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.

Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., Reiss, A. L., and Greicius, M. D. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *J. Neurosci.* 27, 2349–2356.

Shalgi, S., O'Connell, R. G., Deouell, L. Y., and Robertson, I. H. (2007). Absent minded but accurate: delaying responses increases accuracy but decreases error awareness. *Exp. Brain Res.* 182, 119–124.

Sharp, D. J., Bonnelle, V., De Boissezon, X., Beckmann, C. F., James, S. G., Patel, M. C., and Mehta, M. A. (2010). Distinct frontal systems for response inhibition, attentional capture, and error processing. *Proc. Natl. Acad. Sci. U.S.A.* 107, 6106–6111.

Simmonds, D. J., Pekar, J. J., and Mostofsky, S. H. (2008). Meta-analysis of Go/No-go tasks demonstrating that fMRI activation associated with response inhibition is task-dependent. *Neuropsychologia* 46, 224–232.

Singh-Curry, V., and Husain, M. (2009). The functional role of the inferior parietal lobe in the dorsal and ventral stream dichotomy. *Neuropsychologia* 47, 1434–1448.

Soltani, M., and Knight, R. T. (2000). Neural origins of the P300. *Crit. Rev. Neurobiol.* 14, 199–224.

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

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., 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.

Vincent, J. L., Kahn, I., Snyder, A. Z., Raichle, M. E., and Buckner,

R. L. (2008). Evidence for a frontoparietal control system revealed by intrinsic functional connectivity. *J. Neurophysiol.* 100, 3328–3342.

Wang, C. M., Ulbert, I., Schomer, D. L., Marinkovic, K., and Halgren, E. (2005). Responses of human anterior cingulate cortex microdomains to error detection, conflict monitoring, stimulus-response mapping, familiarity, and orienting. *J. Neurosci.* 25, 604–613.

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.

Woodman, G. F. (2010). Masked targets trigger event-related potentials index 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.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 31 January 2012; accepted: 31 May 2012; published online: 19 June 2012.

Citation: Orr C and Hester R (2012) Error-related anterior cingulate cortex activity and the prediction of conscious error awareness. Front. Hum. Neurosci. ***6****:177. doi: 10.3389/fnhum.2012.00177*

APPENDIX

In order to generate the constituent maps for error-related activity associated with aware and unaware errors, group activation maps were determined with one-sample *t*-tests against the null hypothesis of zero event-related activation changes (i.e., no change relative to baseline). Significant voxels within group maps passed a voxelwise statistical threshold ($p \leq 0.0001$) and were required to be part of a larger 84 μl cluster of contiguous significant voxels.

The event-related functional analysis of aware errors identified nine clusters that showed significantly greater activity associated with aware errors than with correct go-trial responses (**Table A1**).

The event-related functional analysis of unaware errors identified eight clusters that showed significantly greater activity associated with unaware errors than with correct go-trial responses (**Table A2**).

Table A1 | Regions that showed significantly greater BOLD signal change for aware errors than for correct go-trial responses.

Structure	Volume μl	Brodmann area	HS	Center of mass (MNI Coordinates)		
				x	*y*	*z*
Inferior parietal	7776	40	L	−48	−40	46
Inferior parietal	6529	40	R	47	−49	40
dACC	4878	32	L	0	9	43
Insula	3020	13	R	44	13	4
Insula	2169	13	L	−41	13	3
Middle frontal gyrus	138	46	R	44	39	26
Extrastriate cortex	118	7	L	−9	−77	31
Precentral gyrus/SMA	116	4/6	L	−28	−27	70
Middle frontal gyrus	84	9	R	39	12	46

Positive values for x, y, and z coordinates denote, respectively, locations that are right, posterior, and superior relative to the anterior commissure. Significance test results indicate areas of increased activation associated with aware relative to unaware errors assessed using t-test.

Table A2 | Regions that showed significantly greater BOLD signal change for unaware errors than for correct go-trial responses.

Structure	Volume μl	Brodmann area	HS	Center of mass (MNI Coordinates)		
				x	*y*	*z*
Inferior occipital gyrus	260	18	R	33	−87	−11
Middle occipital gyrus	217	19	L	−38	−78	−2
Lingual gyrus	164	18	R	17	−81	−14
Middle occipital gyrus	127	19	R	34	−76	−15
Superior temporal gyrus	125	38	R	39	1	−17
Cerebellum (Declive)	98		L	−18	−67	−27
Superior temporal gyrus	96	41	R	52	−24	7
Caudate	88		R	13	22	7

Positive values for x, y, and z coordinates denote, respectively, locations that are right, posterior, and superior relative to the anterior commissure. Significance test results indicate areas of increased activation associated with aware relative to unaware errors assessed using t-test.

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 09 April 2012
doi: 10.3389/fnhum.2012.00077

Acute tryptophan depletion attenuates brain-heart coupling following external feedback

Erik M. Mueller[1]*, Elisabeth A. Evers[2], Jan Wacker[1] and Freddy van der Veen[3]

[1] *Department of Psychology, Philipps-Universität Marburg, Marburg, Germany*
[2] *Department of Neuropsychology and Psychopharmacology, Maastricht University, Maastricht, Netherlands*
[3] *Institute of Psychology, Erasmus University Rotterdam, Rotterdam, Netherlands*

Edited by:
Tilmann A. Klein, Max Planck Institute for Human Cognitive and Brain Sciences, Germany

Reviewed by:
Redmond O'Connell, Trinity College Dublin, Ireland
Sören Enge, Technische Universität Dresden, Germany

****Correspondence:***
Erik M. Mueller, Department of Psychology, Philpps-Universität Marburg, Gutenbergstrasse 18, 35032 Marburg, Germany.
e-mail: erik.mueller@staff.uni-marburg.de

External and internal performance feedback triggers neural and visceral modulations such as reactions in the medial prefrontal cortex and insulae or changes of heart period (HP). The functional coupling of neural and cardiac responses following feedback (cortico-cardiac connectivity) is not well understood. While linear time-lagged within-subjects correlations of single-trial EEG and HP (cardio-electroencephalographic covariance tracing, CECT) indicate a robust negative coupling of EEG magnitude 300 ms after presentation of an external feedback stimulus with subsequent alterations of heart period (the so-called N300H phenomenon), the neurotransmitter systems underlying feedback-evoked cortico-cardiac connectivity are largely unknown. Because it has been shown that acute tryptophan depletion (ATD), attenuating brain serotonin (5-HT), decreases cardiac but not neural correlates of feedback processing, we hypothesized that 5-HT may be involved in feedback-evoked cortico-cardiac connectivity. In a placebo-controlled double-blind cross-over design, 12 healthy male participants received a tryptophan-free amino-acid drink at one session (TRP−) and a balanced amino-acid control-drink (TRP+) on another and twice performed a time-estimation task with feedback presented after each trial. N300H magnitude and plasma tryptophan levels were assessed. Results indicated a robust N300H after TRP+, which was significantly attenuated following TRP−. Moreover, plasma tryptophan levels during TRP+ were correlated with N300H amplitude such that individuals with lower tryptophan levels showed reduced N300H. Together, these findings indicate that 5-HT is important for feedback-induced covariation of cortical and cardiac activity. Because individual differences in anxiety have previously been linked to 5-HT, cortico-cardiac coupling and feedback processing, the present findings may be particularly relevant for futures studies on the relationship between 5-HT and anxiety.

Keywords: feedback processing, FRN, tryptophan depletion, serotonin, vagal, heart rate

INTRODUCTION

We rapidly process internal and external feedback signals in order to optimally interact with our environment. Neural signatures of error or negative feedback processing (e.g., event-related potentials, fMRI activation) are often accompanied by changes of behavior (Debener et al., 2005; Eichele et al., 2010; Mueller et al., 2011): the commission of an error in a speeded reaction time task can lead to subsequent slowing of reaction times; the presentation of negative feedback following a particular decision can alter decision-making in the future. The adaptiveness of such behavioral changes with regard to task performance is still under investigation (Notebaert et al., 2009). Negative feedback in real life (e.g., a loose stone when climbing a rock) may often indicate a sudden need for intensive and effortful compensatory actions (quickly grasping another stone), associated with a sudden requirement for energy. Consistently, it has also been demonstrated that error and feedback signals trigger reactions in the periphery such as changes in heart rate (Crone et al., 2003; Hajcak et al., 2003; Wessel et al., 2011). Because these peripheral reactions are sensitive to feedback valence, even in the case of abstract feedback stimuli, which can hardly be adequately represented at the subcortical level, it is likely that cortical structures are involved in the modulation of cardiac responses at some point.

Consistent with this proposal Mueller et al. (2010) recently demonstrated a robust covariation between cortical (e.g., EEG) and subsequent cardiac chronotropic (i.e., heart period) modulations. Using a probabilistic gambling task (Sato et al., 2005), it was shown that a feedback stimulus indicating whether participants won or lost a small amount of money elicited among others (a) a large frontomedial amplitude which peaked at 300–400 ms and (b) a relative acceleration of heart rate which occurred around 3–4 s later. Most importantly, using a novel within-subject single-trial correlation method termed Cardio-Electroencephalographic Covariance Tracing (CECT) it was shown, that regardless of whether a win or loss was signaled, the frontomedial single-trial EEG-amplitude at 300–400 ms correlated with subsequent cardiac acceleration in the same trial. That is, a relatively larger positivity 300 ms after feedback presentation predicted relatively

more cardiac acceleration 4 s later. Because this phenomenon is reflected in a negative within-subject correlation between EEG amplitude at 300 ms and subsequent heart interbeat interval (IBI) changes we termed it N300H. The N300H was robust across individuals in the initial study (Mueller et al., 2010) and was recently replicated in another sample (Panitz et al., 2011). In both studies, the N300H did not differ between positive and negative feedback valence. Using a low resolution localizing approach Panitz et al. (2011) associated the N300H with increased insula current source density 300 ms after feedback presentation—an observation converging with the importance of the insular cortices for autonomic regulation (Gianaros et al., 2004).

The finding that feedback-locked EEG with a latency of 300 ms relates to heart period modulations converges with prior work on the P300 event-related potential component. It has long been speculated that P300 and evoked changes in autonomic activity are linked to each other, for example as different correlates of the orienting response which may function to facilitate information processing and/or action preparation (Graham and Clifton, 1966; Donchin, 1981; Donchin et al., 1984; Rushby et al., 2005; Polich, 2007; Nieuwenhuis et al., 2011). Both, P300 and evoked changes of heart period are modulated by motivational significance (Duncan-Johnson and Donchin, 1977; Fowles et al., 1982; Sato et al., 2005), which could vary across feedback presentations and thereby induce cortico-cardiac covariation as indicated by N300H. Finally, because one putative generator of the P300 is the insula (Linden, 2005), there appears to be structural overlap between sources linked to P300 and to N300H.

The neurotransmitter systems involved in this feedback-related cortico-cardiac connectivity phenomenon are currently unknown. Given the plethora of brain structures involved in feedback processing and the neurogenic modulation of peripheral responses to external stimuli (Benarroch, 1997; Ullsperger and von Cramon, 2003; Critchley et al., 2005; Wager et al., 2009) several transmitter systems including dopaminergic, noradrenergic, and serotoninergic processes may be of relevance. While dopamine may play an important role for modulating cortical and behavioral action monitoring processes (Holroyd and Coles, 2002; Ullsperger, 2010; Mueller et al., 2011), in the present paper we focus on the serotonergic system, which may be of particular relevance for transmitting cortically processed feedback information to the viscera.

In vitro studies have shown that several 5-HT receptor subtypes are expressed at vagal nuclei binding sites and at the nucleus tractus solitarius (Manaker and Verderame, 1990; Thor et al., 1992), which modulates heart rate indirectly (e.g., by its projections to the nucleus ambiguous). Moreover, it has been demonstrated in a variety of animal studies, that central 5-HT is involved in control of cardiac vagal output (Jordan, 2005). Because feedback-related alterations in heart period show a relatively short latency (e.g., 1–3 s; Crone et al., 2003; 3–4 s: Mueller et al., 2010), it can be assumed that quick vagal rather than slow sympathetic modulations (Berntson et al., 1997) underlie feedback-evoked cardiac responses. Therefore, by affecting cardiac vagal output, 5-HT may modulate the transmission of feedback-related information from the brain to the viscera.

In line with this assumption, van der Veen et al. (2008) demonstrated, that a manipulation of the 5-HT system modulated peripheral (i.e., heart rate) but not central (i.e., EEG) responses to a negative feedback stimulus. Specifically, participants were tested at two sessions. In a double-blind crossover design participants received either a drink that did or a drink that did not contain the important 5-HT precursor tryptophan, thereby presumably increasing or lowering central 5-HT synthesis and release (Bell et al., 2001). At both sessions, participants performed a time-estimation task (Miltner et al., 1997), in which performance feedback was given after each trial. The authors analyzed the feedback-related negativity as a cortical indicator for feedback processing and heart IBI for the subsequent heart beats as a cardiac marker for feedback reactivity. Following the drink that did contain tryptophan (TRP+), it was shown, that feedback induced an increase in FRN amplitude for negative vs. positive feedback and a (valence independent) decrease of heart period for the three following heartbeats. Following the drink that did not contain tryptophan (TRP−), negative feedback induced an increase in FRN amplitude that was comparable to the TRP+ session. However, the relative cardiac deceleration after negative feedback was significantly smaller after TRP− as compared to TRP+.

One explanation of these findings is, that a cortical feedback-related process, which is not captured by the FRN but associated with the cardiac response is attenuated by 5-HT. In addition, the findings by van der Veen et al. (2008) could indicate, that 5-HT influences how cortically processed feedback-information is subsequently transmitted to the periphery. Under the assumption, that the N300H captures such cortico-cardiac transmission following feedback presentation (Mueller et al., 2010), it could thus be hypothesized that TRP−, by lowering central 5-HT, reduces feedback-related N300H values compared to TRP+. This hypothesis was tested in the present study by reanalyzing the data of van der Veen et al. with the CECT-approach.

METHODS

PARTICIPANTS

Data of 12 healthy male participants (mean age: 22.5 years; SD = 5 years) with complete EEG and ECG data for both sessions (TRP+ and TRP−) was available from the study previously published by van der Veen et al. (2008).

STIMULI

Participants performed a time estimation task based on a paradigm developed by Miltner et al. (1997) and the details of the task can be found elsewhere (van der Veen et al., 2008). Participants had to estimate a 1 s interval starting at the onset of a visual cue by pushing a button and received feedback 1 s after the given response. An exclamation mark ("!") was presented if the estimation was correct, a "+" was presented when the estimation was too long and a "−" was presented when the estimation was too short. The feedback stimulus was presented for 1 s and was followed by a 2 s inter-trial interval. Percentage of positive and negative feedback was kept at about 50% by varying the interval in which the estimation was labeled as correct.

PROCEDURE

Participants were tested in a randomized double-blind cross-over design and details are described elsewhere (van der Veen et al., 2008). Participants were tested on 2 separate days. On 1 day (TRP−) they ingested the tryptophan depleted mixture which contained 200 ml of tap water plus 75 g of the amino-acids L-alanine, L-glycine, L-histidine, L-isoleucine, L-leucine, L-lysine, L-phenylalanine, L-proline, L-serine, L-threonine, L-tyrosine, L-valine, L-arginine, L-cysteine, and L-methionine. On the other day (TRP+) they ingested the balanced amino-acid mixture which contained the same ingredients plus 3.0 g tryptophan. Participants were tested 5 h after ingestion of the mixture. Of each participant, a blood sample was taken right before drinking the amino acid mixture and right before the psychophysiological assessment, i.e., about 5 h after drinking the amino-acid mixture. As described in van der Veen et al. (2008), the raw plasma concentration of free tryptophan was determined from these samples.

EEG AND ECG

As described in detail elsewhere (van der Veen et al., 2008), ECG was derived from pre-cordial leads and was sampled at 1000 Hz. R spikes were detected online with an accuracy of 1 ms, manually checked and corrected if necessary. The ECG was converted into a (stepwise) IBI-time course with a sampling rate of 500 Hz where the value at each sample reflects the distance between the preceding and succeeding r-spike in ms. EEG was recorded at sensors F3, Fz, F4, C3, Cz, C4, Pz, and at one ocular channel using a sampling rate of 500 Hz. Epochs with non-systematic artifacts in either EEG or IBI were manually removed while Independent Component Analysis was used to correct for ocular artifacts in the EEG.

CECTs

To compute CECTs, the continuous EEG was first segmented into epochs from 0 to 500 ms relative to the feedback stimulus and baseline corrected for the preceding 500 ms. EEG epochs were then binned into 50 bins of 10 ms each (corresponding to a sampling rate of 100 Hz) and mean voltage magnitudes were determined for each bin. With this data, for each participant and each channel an EEG matrix for subsequent CECT analyses was constructed which consisted of about 200–240 rows (depending on the number of artifact-free trials) and 50 columns (corresponding to the 50 bins per epoch). In analogy to the EEG, the continuous IBI-trace was segmented into epochs from 0 to 2500 ms relative to the feedback stimulus and baseline corrected for the preceding 250 ms. IBI epochs were binned into 10 bins of 250 ms each (4 Hz) and mean IBI values were determined for each bin. For each participant an IBI matrix for CECT analyses was constructed that consisted of about 200–240 rows (trials) and 10 columns (bins). CECT matrices were computed for each participant and channel, separately by correlating each column of the corresponding EEG matrix with each column of the IBI matrix. Thus, for every CECT matrix 10×50 within-subject correlations were computed (correlations over trials). These were now structured in a 10×50 correlation matrix in which each row represents one particular IBI bin and each column represents one particular EEG bin. To illustrate the interpretation of CECT matrices, a correlation displayed in column 1, row 2, reflects how much the EEG magnitude from 0 to 10 ms (EEG bin 1) after a feedback stimulus correlates with the change in IBI from 250 to 500 ms after a feedback stimulus (IBI bin 2).

Grand average CECTs (as displayed in **Figure 1**) were computed by averaging the CECT matrices across participants (separately for each channel). To first test for an effect of acute tryptophan depletion (ATD) a mean CECT-value (averaged across all time bins) was compared between TRP+ and TRP− session using a paired-samples t-test. To then test whether an N300H component could be replicated in the present study, CECT-values were tested against zero (across participants, separately for each session) by conducting one-sample t-tests for each cell of the CECT-matrix at channel Cz ($p < 0.05$, two-sided). With this analysis an N300H like component was evident in the TRP+ session, which extended from 150 to 300 ms in the EEG time domain and from 250 to 1750 ms in the IBI time domain. To assess the between-subject correlation of N300H and tryptophan level we extracted individual peak CECT-values in this time window (see **Figure 1**). To test for effects of feedback valence, separate CECTs were computed for positive and negative feedback and mean CECT-values in the previously specified time-window were tested against zero and compared between feedback types.

RESULTS

THE EFFECT OF ATD ON PLASMA TRYPTOPHAN LEVEL

As expected and as previously reported for the entire sample, the oral intake of the tryptophan-deficient amino-acid mixture reduced the levels of free tryptophan concentrations in blood plasma from $M = 53$ (SD $= 6$) at T_0 (i.e., before the drink was taken) to $M = 17$ (SD $= 8$) at T_5 (i.e., 5 h after the drink was taken just before the task began) ($t_{(11)} = 12.28$, $p < 0.001$). In contrast, the ingestion of the nutritionally balanced mixture at the TRP+ session yielded an increase of plasma tryptophan levels from T_0($M = 50$; SD $= 14$) to T_5 ($M = 105$; SD $= 41$) ($t_{(11)} = 5.99$, $p < 0.001$).

BEHAVIORAL PERFORMANCE

On average, participants had 112 correct responses, 40 under-estimations and 88 over-estimations of the 1 s epoch (the window was continuously adapted to provide about equal amounts of positive and negative feedback). There was no effect of ATD on the number of correct responses, under-estimations or over-estimations (ps > 0.5).

EVENT-RELATED POTENTIAL AND EVEN-RELATED CARDIAC RESPONSE

Feedback-evoked a P300-like deflection, which extended from 200 to 500 ms (peak at 360 ms, channel Cz) and a triphasic cardiac response (0–1000 ms: deceleration; 1000–2000 ms: acceleration; 2000–2500 ms: deceleration). ATD did not influence the P300 amplitude ($p > 0.5$), but reduced the cardiac response to feedback for bins 3, 4, and 5 (i.e., 500–1250 ms; $p < 0.05$). Effects of feedback-valence were not analyzed in the present report (see van der Veen et al., 2008, for effects of feedback-valence).

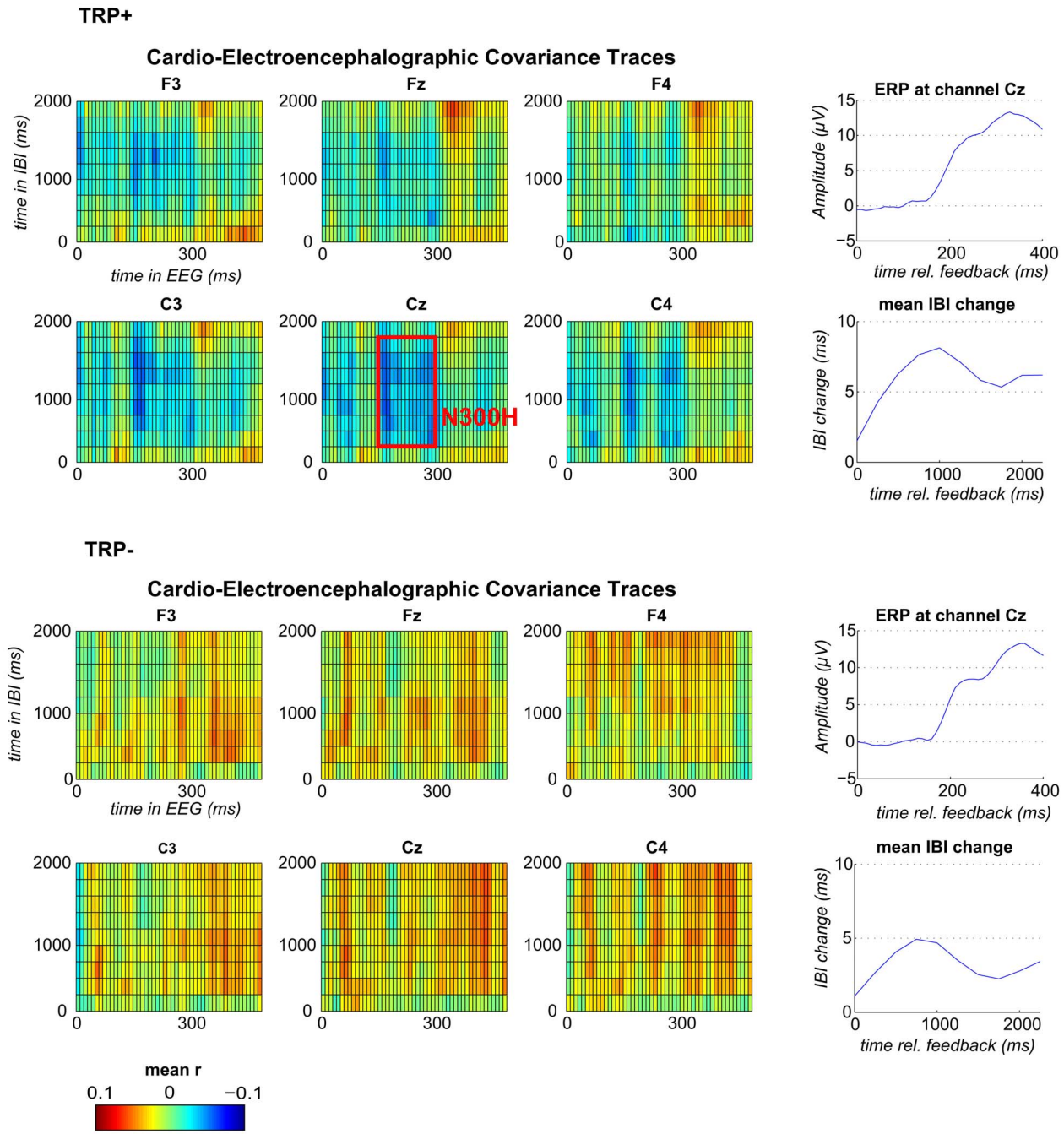


FIGURE 1 | Feedback-locked CECTs, IBI-traces and event-related potentials (ERP) at channel Cz following TRP+ (top) and TRP− (bottom). Blue indicates negative and red positive within-subject correlations averaged across subjects (range from $r = -0.1$ to $r = +0.1$). The red box illustrates the time window, which was used for correlation analyses and for analyses of feedback valence.

CARDIO-ELECTROENCEPHALOGRAPHIC COVARIANCE TRACING

Effects of session on the overall CECT

A paired samples t-test comparing mean CECT-values (whole CECT epoch from 0 to 500 ms in the EEG and 0 to 2500 ms in the IBI time domain) between TRP+ and TRP− session indicated a significant effect of session for the overall CECT ($t_{(11)} = 3.26$, $p < 0.01$).

Replication of N300H

Because ATD affected the overall CECT, separate CECTs for each session were tested for N300H. As shown in **Figures 1** and **2** CECT-analyses for the TRP+ session revealed an N300H-like component that was negative and significantly different from zero ($p < 0.05$, two-sided, 19 adjacent time-bins). The peak average within-subject correlation within this cluster was at 280 ms in the

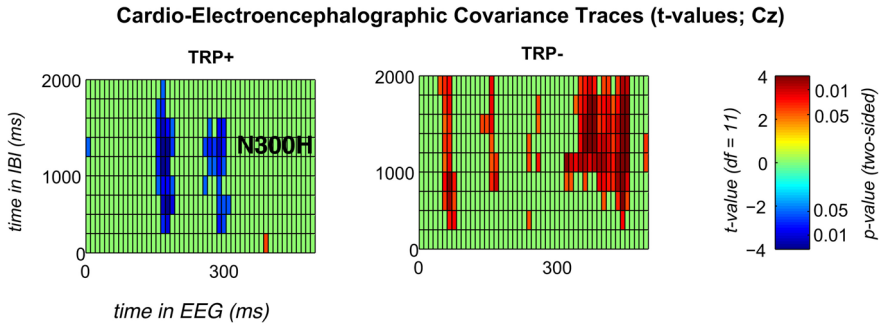


FIGURE 2 | Significant *t*-values for feedback-locked CECTs at channel Cz following placebo (left) or acute tryptophan depletion (right). *T*-values are thresholded with $p < 0.05$ (two sided). Any non-significant *t*-values are plotted in green color.

EEG and at 1250–1500 ms in the IBI time domain ($t_{(11)} = -3.6$, $p < 0.005$). There was also a slightly earlier portion of the EEG (150–180 ms) that showed a significant negative covariation with subsequent IBI values. It should be noted that the small sample size (and consequently limited statistical power) precluded application of a conservative Bonferroni-correction for 10×50 tests (resulting in $\alpha' = 0.0001$), which might be considered appropriate in the light of the limited a priori knowledge about the timing of N300H in the present task (Mueller et al., 2010). The statistical power of such Bonferroni-corrected *t*-tests is < 0.09, which were (unsurprisingly) non-significant.

Of relevance, in the TRP− session an unexpected, temporally less specific *positive* association between feedback-evoked EEG and IBI values reached significance (*p*s < 0.05). As shown in **Figures 1** and **2**, a negative N300H like component could not be observed following ATD (all *p*s > 0.05).

Effects of feedback valence and ATD on N300H

Mean N300H values in the time window from 150 to 300 ms in the EEG time domain and from 250 to 1750 ms in the EEG time domain were significantly different from zero for positive ($t_{(11)} = 2.23$, $p < 0.05$) and negative ($t_{(11)} = 2.50$, $p < 0.031$) feedback in the TRP+ but not TRP− condition (*p*s > 0.09, two-sided). A TRP × Feedback-Valence ANOVA yielded a main effect for TRP ($F_{(1, 12)} = 12.6$, $p < 0.005$), confirming the findings for the valence-independent CECTs reported above. There was also a trend for a TRP × Valence interaction $F_{(1,12)} = 3.5$, $p < 0.09$, indicating that N300H to negative feedback tended to be less affected by ATD than N300H to positive feedback. Consistent with Mueller et al. (2010), there was no main effect of feedback valence on N300H.

Correlations between plasma tryptophan level and N300H

In addition to these group effects, we examined the association between plasma tryptophan level and N300H aiming to probe whether interindividual variations of tryptophan level are also related to N300H. The tryptophan level was measured at two occasions (T_0 and T_5) on each day (TRP+ and TRP−). Overall, the T_0 tryptophan level at the TRP+ day was correlated with the T_5 tryptophan level at TRP+ day ($r = 0.74$, $p < 0.005$, one-sided) and T_0 tryptophan level at TRP− day ($r = 0.54$, $p < 0.05$, one-sided), indicating that interindividual differences in the tryptophan level were reliably measured and relatively stable over time, although the average tryptophan level was elevated by the nutritionally balanced drink as reported above. Importantly, the tryptophan level at TRP+ day (averaged across T_0 and T_5) was significantly correlated with N300H at the TRP+ day ($r = -0.68$, $p < 0.05$, **Figure 3**). Consistent with the experimental findings, the direction of these correlations indicates that a reduced level of plasma tryptophan predicted lower (i.e., less negative) N300H. As reported above at T_5 of the TRP− day, plasma tryptophan levels were severely reduced. These reduced tryptophan levels were not significantly associated with the (attenuated) N300H in that session ($p > 0.4$, **Figure 3**), possibly due to blunted between-subject variation after TRP− (SD = 8 μMol/l) vs. TRP+ (SD = 41 μMol/l). Together, these correlative findings indicate that baseline variations of plasma tryptophan levels are also related to N300H.

CONTROL ANALYSIS

To rule out the alternative explanation that ATD reduced cortico-cardiac within-subject correlations by affecting trial-to-trial

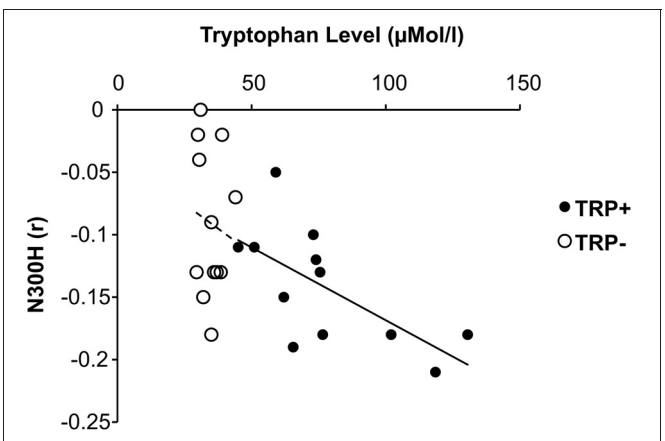


FIGURE 3 | Relationship between plasma tryptophan level and cortico-cardiac connectivity as indicated by peak N300H amplitude under placebo and acute tryptohan depletion (ATD).

variation of cardiac responses we tested whether there was a difference in the cross-trial variance of feedback-locked IBI (from 250 to 1750 ms) between TRP+ and TRP−. However, this was not the case ($p > 0.5$).

DISCUSSION

The goal of the present study was to test whether a manipulation of central 5-HT would influence cortico-cardiac connectivity evoked by feedback stimuli. Central 5-HT was manipulated through administration of drinks that either contained no tryptophan (TRP−), serving to deplete this precursor of 5-HT synthesis, or elevated tryptophan levels (TRP+) in two separate sessions (double-blind cross-over design). Five hours later, when plasma levels of tryptophan were reduced in the TRP− vs. TRP+ session, participants conducted a time estimation task in which feedback was given after each trial. To measure cortico-cardiac connectivity, CECT-analyses were conducted, in which feedback-evoked-changes in single-trial EEG magnitudes are systematically correlated with time-lagged feedback-evoked changes in heart period. Importantly, overall CECTs were significantly different between the TRP+ and TRP− session. In the TRP+ session, we replicated a previously reported phenomenon (N300H), indicating a significant coupling between centromedial EEG magnitude about 300 ms after feedback presentation and subsequent accelerations of heart rate (i.e., smaller IBIs). Of particular relevance, N300H was absent following TRP−, indicating that a pharmacological downregulation of 5-HT synthesis attenuated the negative covariation of feedback-evoked EEG and heart period changes or even induced a positive covariation (as indicated by the red spots in **Figures 1** and **2**). Moreover, interindividual differences in plasma tryptophan levels at the TRP+ session were correlated with N300H values measured in the same session. Together these findings provide converging evidence that tryptophan plays a role in the covariation of cortical and cardiac feedback-related activity as indicated by N300H. One possible interpretation of these preliminary findings is that 5-HT is crucial for transmitting feedback-related information from the brain to the viscera.

As outlined in **Figure 4** cortico-cardiac coupling following feedback presumably involves several structures, which may communicate in parallel and bidirectionally. The cortical processing of feedback includes the anterior (mid−) cingulate cortex and the insulae (Ullsperger and von Cramon, 2003), two highly relevant regions for central autonomic control (Benarroch, 1997; Gianaros et al., 2004; Critchley, 2005). Via projections to key regions in midbrain and pons these feedback-relevant structures can indirectly modulate input to the nucleus tractus solitarus in the dorsal medulla and the nucleus ambiguous [path (a) and (b)], where cell bodies of the myelinated cardiac vagal preganglionic neurons [path (c)] are located (Benarroch, 1997; Jordan, 2005). Accordingly, feedback-related information can be transmitted relatively quickly from the prefrontral cortex (where stimulus valence of rather abstract stimuli can be assessed) to the heart (where modulations may prepare subsequent alterations of behavior).

In the present study CECTs indicated a covariation of feedback-evoked single-trial EEG at 300 ms and changes in IBI from 250 to 1750 ms in the TRP+ session (N300H). Because the EEG in the TRP+ session thus correlated with an earlier proportion of IBI than in the Mueller et al. (2010) study (3000–4000 ms) and because a different paradigm was used in the two studies, it is possible that this cortico-cardiac covariation reflects a different phenomenon than N300H in the Mueller et al. (2010) study. It could be hypothesized, that tasks with short intertrial intervals and/or tasks where feedback crucially depends on participants' performance (like in the current study) trigger earlier neurogenic accelerations of heart period (as observed in the event-related cardiac responses of the current vs. Mueller et al., 2010 study) and thus earlier cortico-cardiac correlation patterns. However,

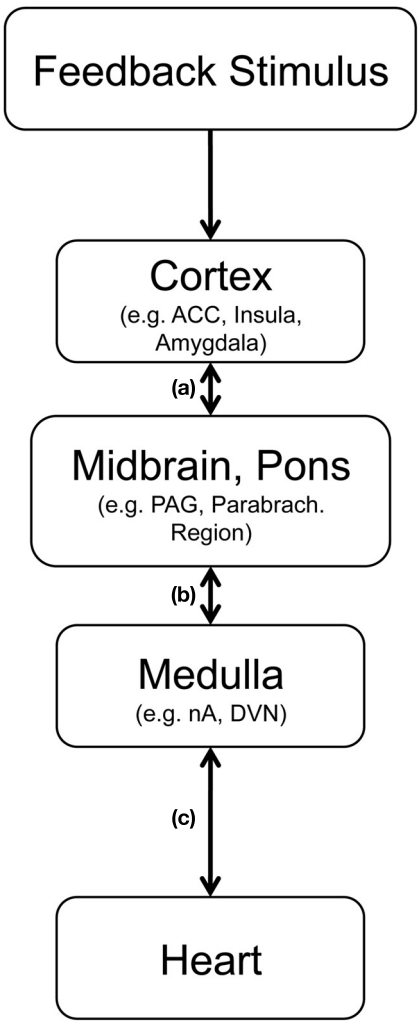


FIGURE 4 | Hypothetical pathway for abstract feedback stimuli to trigger vagally mediated cardiac reactions, (adapted from Jordan, 2005; Thayer and Lane, 2007). After sensory processing feedback properties are analyzed in cortical structures including the anterior cingulate cortex (ACC), the insula and the amygdala. Relevant information is transmitted to the periaqueductal gray in the midbrain and to parabrachial region in the pons **(a)**. From there medullar regions including the nucleus tractus solitarius, the nucleus ambiguous (nA) and dorsal vagal nucleus (DVN) can be activated **(b)**. Myelinated cardiac preganglionic vagal neurons, which have a high density of 5-HT receptors, have their cell-bodies in the nA and modulate heart rate by projecting to the sinoatrial node **(c)**.

there are several similarities between the N300H in the current study and N300H in the Mueller et al study, such as latency in the EEG domain (300 ms), peak localization (centromedial electrodes) and polarity of the correlation (negative). Nevertheless, future studies are needed to clarify whether the effects of ATD reported here generalize to N300H evoked by other tasks.

CECTs provide a method to measure cortico-cardiac coupling with higher temporal precision than neuroimaging-based approaches. However, when interpreting the latencies of N300H in the present study, it should be considered, that IBI values only change with every heartbeat, and that an IBI value at a given time (e.g., at 250 ms) is determined by the latency of the subsequent heart beat (which may occur several 100 ms later). Thus, the temporal resolution in the IBI domain is somewhat blurred. Nevertheless, under the assumption that N300H is driven by the combined paths (a), (b), and (c), the present findings indicate that it took less than 1500 ms (i.e., 1750 ms minus 300 ms) for the feedback signal to be transmitted from the cortex to the heart (or alternatively, to be simultaneously transmitted from a third region to the cortex and the heart). Thus, the cortical response 300 ms after a given feedback stimulus predicts how much the heart accelerates about one to two beats later.

As outlined in the introduction the N300H and the P300 event-related potential show overlap with regard to latency and scalp topography. However, it should be emphasized that they are not the same phenomenon. The P300 reflects the *averaged* summed electrophysiological activity of several brain structures that become active around 300ms after stimulus presentation (Soltani and Knight, 2000). The scalp P300, therefore, reflects a mixture (Makeig et al., 2002; Polich, 2007), which has previously been linked to a variety of phenomena including autonomous reactions, stimulus probability, motivational significance, attention, and task performance (Duncan-Johnson and Donchin, 1977; Isreal et al., 1980; Donchin et al., 1984; Li et al., 2009; Nieuwenhuis et al., 2011). In contrast, the N300H only reflects a highly specific portion of *variance* of EEG 300 ms after a feedback stimulus: that portion that is shared with variance of beat-to-beat intervals hundreds of ms later. Accordingly, N300H shows a distinct topography and latency (Mueller et al., 2010) and may be more differentially sensitive to manipulations of neurotransmitter systems than the overall P300 event-related potential.

In the present study, we were predominantly interested in the effect of 5-HT on cortico-cardiac connectivity in feedback processing. Following TRP− vs. TRP+, we found relatively lower N300H, which explains the previously reported finding in the same data, namely that ATD had no effect on cortical feedback-related signatures but attenuated the cardiac concomitants of feedback processing. These findings are consistent with a predominant role of 5-HT for path (c), which has been demonstrated in studies with cats and rats before (Jordan, 2005). However, given the widespread distribution of 5-HT in the human brain (Cools et al., 2008), future studies are necessary to further specify the particular path(s) that involves 5-HT for transmitting feedback-related information from the brain to the heart. Obviously, 5-HT may not be the only monoaminergic transmitter-system involved. Dopamine influences how errors (i.e., internal performance feedback) are processed at cortical sites such as the anterior (mid-) cingulate cortex and the effect of dopamine at these sites may be further modulated by 5-HT (Mueller et al., 2011). Moreover, widespread noradrenaline release triggered at the Locus Ceruleus following relevant stimuli such as feedback may lead to parallel modulations of P300-like amplitudes [i.e., upward arrow (a)] and visceral reactions [downward arrow (b)] and thereby induce or amplify a N300H like covariation of P300 and cardiac acceleration (Nieuwenhuis et al., 2011). As a consequence we suggest that the monoaminergic mechanisms underlying N300H are investigated in future studies by also using catecholaminergic challenge tests and/or molecular genetic assessments. In this regard, two particularly promising candidate genes are the widely studied COMT Val158Met polymorphism and a recently reported polymorphism in the promoter region of the norepinephrine transporter gene (Kim et al., 2006), both of which have previously been linked to P300 amplitude (Gallinat et al., 2003) and latency (Enge et al., 2011), respectively.

In the TRP+ session, the tryptophan level predicted the amplitude of N300H. Individuals with higher plasma tryptophan levels, possibly associated with higher levels of 5-HT, showed a stronger covariation of cortical and cardiac activity after feedback stimuli. Accordingly, individual differences in cortico-cardiac connectivity may be biologically linked to 5-HT. Interestingly, both, 5-HT and cortico-cardiac connectivity may be of high relevance for individual differences in trait anxiety. For example, the short allele of the 5-HT transporter polymorphism (5-HTTLPR) predicts high levels of extracellular 5-HT and is associated with elevated risk for high trait anxiety and negative emotionality (Lesch et al., 1996; Sen et al., 2004), particularly in individuals who were previously exposed to adverse life events (Karg et al., 2011). Individuals who show stronger visceral reactions to cortically processed (feedback) information may also experience increased subjective anxiety more often, given that interoception of peripheral reactions (i.e., awareness of increased heart rate) may increase subjective anxiety (Paulus and Stein, 2006). Together, this raises the question, whether genetically driven differences in 5-HT activity may be linked to anxiety by modulating how potentially threatening information such as (negative) external feedback is transmitted from the brain to the viscera. Future studies examining serotoninergic polymorphisms, trait anxiety and N300H like phenomena using large samples will be needed to address this hypothesis, which emerges from the present findings. Because gender is known to influence serotoninergic neurotransmission (Jovanovic et al., 2008), negative emotionality and vagal reactivity (Thayer et al., 1998), such studies should either focus on one gender (as in the present study) or—if sample sizes are sufficiently large—explicitly include gender as an independent variable.

It is worth noting, that the effects of ATD on evoked cardiac response are significant for the initial deceleratory (0–1250 ms) but not for the acceleratory phase, while the N300H, which is also affected by ATD, suggests an association between EEG and heart period from 250 to 1750 ms. At the moment, detailed mechanistic explanations for this dissociation between evoked IBI and CECTs would be speculative. However, it should be noted, that cardiac responses are influenced by a variety of more or less opposing processes, which may occur in parallel. Accordingly, the

observed cardiac response pattern (deceleration—acceleration—deceleration) may reflect the overall net effect of several (cortical and non-cortical) processes while the CECT only captures heart period fluctuations that covary with feedback-evoked cortical activity. The present N300H findings indicate that increased (positive) EEG around 300 ms covaries with a *relative* cardiac acceleration from 250 to 1750 ms after TRP+ and that this covariation is lower or even reversed after TRP−. At which latency and to which degree this relative cardiac acceleration contributes to the net evoked IBI is a different question, which should be addressed in future studies.

Consistent with the assumption of parallel partially opposing brain-heart processes there appeared to be a positive correlation between feedback-evoked EEG and IBI in the TRP− session. Unlike the N300H, this association did not show a precise temporal or spatial localization, although the largest correlation cluster at Cz appeared from 300 to 400 ms. One could speculate, that while earlier feedback-evoked centromedial brain activity (i.e., 200–300 ms) is linked to cardiac acceleration, for example to prepare the organism for subsequent actions, more widespread brain activity with a later onset is linked to cardiac deceleration, possibly associated with information intake. If 5-HT differently contributes to these cortico-cardiac processes the polarity of the net CECT could be modulated by ATD as in the present study. However, it should be emphasized that this cortico-cardiac covariation in the TRP− session was unexpected with regard to polarity, spatial, and temporal localization. The current interpretations thus remain speculative until this pattern is replicated in a larger independent sample.

Two limitations deserve attention. First of all, the present study investigated a relatively small sample. Thus, although an N300H-like phenomenon was detected after TRP+, this component was not significantly different from zero after rigorously controlling for $n = 500$ computed t-tests and applying a conservative significance level of $\alpha = 0.0001$. To investigate the effect of TRP+ vs. TRP− however, the use of a cross-over design, allowed us to test substantially smaller samples than would be needed to achieve a comparable power with a between-subjects design. Moreover, the detected effects were relatively large and reached statistical significance at an uncorrected level ($p < 0.005$) despite the small sample. However, future studies using experimental manipulations of 5-HT with larger samples and other methods (e.g., using selective 5-HT reuptake inhibitors or molecular genetics) in combination with the CECT approach could further strengthen the present findings. As a second limitation, CECTs that were computed separately for positive and negative feedback stimuli may not have had enough trials for reliable within subject correlations. Because single-trial EEG and single-trial IBI show high levels of noise and non-specific fluctuations, respectively, individual CECT-correlations are typically relatively small ($r < 0.2$) and need a large number of trials to be reliable. Due to the small sample size and the relatively small number of feedback-specific trials, effects of feedback valence or interactions of feedback valence and ATD may not have had enough statistical power to reach significance in the present study. Although the lack of a main effect of feedback valence on N300H is consistent with Mueller et al. (2010), interactions of feedback valence and 5-HT cannot be ruled out with the present findings. Particularly, the role of 5-HT in cortico-cardiac processing of negative feedback could be relevant in the abovementioned context of anxiety, which could be more closely investigated in future studies.

Despite these limitations, we showed for the first time, that the covariation of cortical and cardiac activity following external feedback presentation in humans is affected by plasma tryptophan availability, which determines the rate of central 5-HT synthesis. Our findings thereby provide indirect evidence that 5-HT is of relevance for the interaction between the brain and the heart following feedback presentation.

ACKNOWLEDGMENTS

This study was supported by a TOP grant (912-02-050) from ZonMW-NWO to Dr. Evers and a DFG grant DFG WA 2593/2-2 to Dr. Wacker.

REFERENCES

Bell, C., Abrams, J., and Nutt, D. (2001). Tryptophan depletion and its implications for psychiatry. *Br. J. Psychiatry* 178, 399–405.

Benarroch, E. E. (1997). *Central Autonomic Network: Functional Organization and Clinical Correlations.* Armonk, NY: Futura Pub. Co.

Berntson, G. G., Bigger, J. T. Jr., Eckberg, D. L., Grossman, P., Kaufmann, P. G., Malik, M., Nagaraja, H. N., Porges, S. W., Saul, J. P., Stone, P. H., and van der Molen, M. W. (1997). Heart rate variability: origins, methods, and interpretive caveats. *Psychophysiology* 34, 623–648.

Cools, R., Roberts, A. C., and Robbins, T. W. (2008). Serotoninergic regulation of emotional and behavioural control processes. *Trends Cogn. Sci.* 12, 31–40.

Critchley, H. D. (2005). Neural mechanisms of autonomic, affective, and cognitive integration. *J. Comp. Neurol.* 493, 154–166.

Critchley, H. D., Tang, J., Glaser, D., Butterworth, B., and Dolan, R. J. (2005). Anterior cingulate activity during error and autonomic response. *Neuroimage* 27, 885–895.

Crone, E. A., van der Veen, F. M., van der Molen, M. W., Somsen, R. J., van Beek, B., and Jennings, J. R. (2003). Cardiac concomitants of feedback processing. *Biol. Psychol.* 64, 143–156.

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.

Donchin, E. (1981). Presidential address, 1980. Surprise! . . . Surprise? *Psychophysiology* 18, 493–513.

Donchin, E., Heffley, E., Hillyard, S. A., Loveless, N., Maltzman, I., Ohman, A., Rosler, F., Ruchkin, D., and Siddle, D. (1984). Cognition and event-related potentials. II. The orienting reflex and P300. *Ann. N.Y. Acad. Sci.* 425, 39–57.

Duncan-Johnson, C. C., and Donchin, E. (1977). On quantifying surprise: the variation of event-related potentials with subjective probability. *Psychophysiology* 14, 456–467.

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

Enge, S., Fleischhauer, M., Lesch, K. P., Reif, A., and Strobel, A. (2011). Serotonergic modulation in executive functioning: linking genetic variations to working memory performance. *Neuropsychologia* 49, 3776–3785.

Fowles, D. C., Fisher, A. E., and Tranel, D. T. (1982). The heart beats to reward: the effect of monetary incentive on heart rate. *Psychophysiology* 19, 506–513.

Gallinat, J., Bajbouj, M., Sander, T., Schlattmann, P., Xu, K., Ferro, E. F., Goldman, D., and Winterer, G. (2003). Association of the G1947A COMT (Val(108/158)Met) gene polymorphism with prefrontal P300 during information

processing. *Biol. Psychiatry* 54, 40–48.

Gianaros, P. J., van der Veen, F. M., and Jennings, J. R. (2004). Regional cerebral blood flow correlates with heart period and high-frequency heart period variability during working-memory tasks: implications for the cortical and subcortical regulation of cardiac autonomic activity. *Psychophysiology* 41, 521–530.

Graham, F. K., and Clifton, R. K. (1966). Heart-rate change as a component of the orienting response. *Psychol. Bull.* 65, 305–320.

Isreal, J. B., Chesney, G. L., Wickens, C. D., and Donchin, E. (1980). P300 and tracking difficulty: evidence for multiple resources in dual-task performance. *Psychophysiology* 17, 259–273.

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.

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.

Jordan, D. (2005). Vagal control of the heart: central serotonergic (5-HT) mechanisms. *Exp. Physiol.* 90, 175–181.

Jovanovic, H., Lundberg, J., Karlsson, P., Cerin, A., Saijo, T., Varrone, A., Halldin, C., and Nordstrom, A. L. (2008). Sex differences in the serotonin 1A receptor and serotonin transporter binding in the human brain measured by PET. *Neuroimage* 39, 1408–1419.

Karg, K., Burmeister, M., Shedden, K., and Sen, S. (2011). The serotonin transporter promoter variant (5-HTTLPR), stress, and depression meta-analysis revisited: evidence of genetic moderation. *Arch. Gen. Psychiatry* 68, 444–454.

Kim, C. H., Hahn, M. K., Joung, Y., Anderson, S. L., Steele, A. H., Mazei-Robinson, M. S., Gizer, I., Teicher, M. H., Cohen, B. M., Robertson, D., Waldman, I. D., Blakely, R. D., and Kim, K. S. (2006). A polymorphism in the norepinephrine transporter gene alters promoter activity and is associated with attention-deficit hyperactivity disorder. *Proc. Natl. Acad. Sci. U.S.A.* 103, 19164–19169.

Lesch, K. P., Bengel, D., Heils, A., Sabol, S. Z., Greenberg, B. D., Petri, S., Benjamin, J., Muller, C. R., Hamer, D. H., and Murphy, D. L. (1996). Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. *Science* 274, 1527–1531.

Li, R., Keil, A., and Principe, J. C. (2009). Single-trial P300 estimation with a spatiotemporal filtering method. *J. Neurosci. Methods* 177, 488–496.

Linden, D. E. (2005). The P300: where in the brain is it produced and what does it tell us? *Neuroscientist* 11, 563–576.

Makeig, S., Westerfield, M., Jung, T. P., Enghoff, S., Townsend, J., Courchesne, E., and Sejnowski, T. J. (2002). Dynamic brain sources of visual evoked responses. *Science* 295, 690–694.

Manaker, S., and Verderame, H. M. (1990). Organization of serotonin 1A and 1B receptors in the nucleus of the solitary tract. *J. Comp. Neurol.* 301, 535–553.

Miltner, W. H. R., Braun, C. H., and Coles, M. G. H. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a "generic" neural system for error detection. *J. Cogn. Neurosci.* 9, 788–798.

Mueller, E. M., Makeig, S., Stemmler, G., Hennig, J., and Wacker, J. (2011). Dopamine effects on human Error processing depend on catechol-*O*-methyltransferase VAL158MET genotype. *J. Neurosci.* 31, 15818–15825.

Mueller, E. M., Stemmler, G., and Wacker, J. (2010). Single-trial electroencephalogram predicts cardiac acceleration: a time-lagged P-correlation approach for studying neurovisceral connectivity. *Neuroscience* 166, 491–500.

Nieuwenhuis, S., De Geus, E. J., and Aston-Jones, G. (2011). The anatomical and functional relationship between the P3 and autonomic components of the orienting response. *Psychophysiology* 48, 162–175.

Notebaert, W., Houtman, F., Opstal, F. V., Gevers, W., Fias, W., and Verguts, T. (2009). Post-error slowing: an orienting account. *Cognition* 111, 275–279.

Panitz, C., Wacker, J., Stemmler, G., and Mueller, E. M. (2011). "Temporal coupling of cortical and cardiac activity," in *Psychologie und Gehirn*, eds D. Hagemann and P. Kirsch (Heidelberg: Universität Heidelberg), 199.

Paulus, M. P., and Stein, M. B. (2006). An insular view of anxiety. *Biol. Psychiatry* 60, 383–387.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118, 2128–2148.

Rushby, J. A., Barry, R. J., and Doherty, R. J. (2005). Separation of the components of the late positive complex in an ERP dishabituation paradigm. *Clin. Neurophysiol.* 116, 2363–2380.

Sato, A., Yasuda, A., Ohira, H., Miyawaki, K., Nishikawa, M., Kumano, H., and Kuboki, T. (2005). Effects of value and reward magnitude on feedback negativity and P300. *Neuroreport* 16, 407–411.

Sen, S., Burmeister, M., and Ghosh, D. (2004). Meta-analysis of the association between a serotonin transporter promoter polymorphism (5-HTTLPR) and anxiety-related personality traits. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 127B, 85–89.

Soltani, M., and Knight, R. T. (2000). Neural origins of the P300. *Crit. Rev. Neurobiol.* 14, 199–224.

Thayer, J. F., and Lane, R. D. (2007). The role of vagal function in the risk for cardiovascular disease and mortality. *Biol. Psychol.* 74, 224–242.

Thayer, J. F., Smith, M., Rossy, L. A., Sollers, J. J., and Friedman, B. H. (1998). Heart period variability and depressive symptoms: gender differences. *Biol. Psychiatry* 44, 304–306.

Thor, K. B., Blitz-Siebert, A., and Helke, C. J. (1992). Autoradiographic localization of 5HT1 binding sites in autonomic areas of the rat dorsomedial medulla oblongata. *Synapse* 10, 217–227.

Ullsperger, M. (2010). Genetic association studies of performance monitoring and learning from feedback: the role of dopamine and serotonin. *Neurosci. Biobehav. Rev.* 34, 649–659.

Ullsperger, M., and Von Cramon, D. Y. (2003). Error monitoring using external feedback: specific roles of the habenular complex, the reward system, and the cingulate motor area revealed by functional magnetic resonance imaging. *J. Neurosci.* 23, 4308–4314.

van der Veen, F. M., Mies, G. W., van der Molen, M. W., and Evers, E. A. (2008). Acute tryptophan depletion in healthy males attenuates phasic cardiac slowing but does not affect electro-cortical response to negative feedback. *Psychopharmacology (Berl.)* 199, 255–263.

Wager, T. D., Waugh, C. E., Lindquist, M., Noll, D. C., Fredrickson, B. L., and Taylor, S. F. (2009). Brain mediators of cardiovascular responses to social threat Part I: reciprocal dorsal and ventral sub-regions of the medial prefrontal cortex and heart-rate reactivity. *Neuroimage* 47, 821–835.

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.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 27 January 2012; accepted: 20 March 2012; published online: 06 April 2012.

Citation: Mueller EM, Evers EA, Wacker J and van der Veen F (2012) Acute tryptophan depletion attenuates brain-heart coupling following external feedback. Front. Hum. Neurosci. 6:77. doi: 10.3389/fnhum.2012.00077

frontiers in
HUMAN NEUROSCIENCE

REVIEW ARTICLE
published: 04 February 2013
doi: 10.3389/fnhum.2013.00014

Error awareness and the insula: links to neurological and psychiatric diseases

Tilmann A. Klein[1,2]*, Markus Ullsperger[3,4,5] and Claudia Danielmeier[3]

[1] *Department of Neurology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany*
[2] *Day Care Clinic for Cognitive Neurology, University Clinic Leipzig, Leipzig, Germany*
[3] *Donders Institute for Brain, Cognition, and Behaviour, Radboud University Nijmegen, Nijmegen, Netherlands*
[4] *Institute of Psychology, Otto-von-Guericke University Magdeburg, Magdeburg, Germany*
[5] *Center for Behavioral Brain Sciences, Magdeburg, Germany*

Edited by:
Hauke R. Heekeren, Freie Universität Berlin, Germany

Reviewed by:
Redmond O'Connell, Trinity College Dublin, Ireland
Alexandros Goulas, Maastricht University, Netherlands

****Correspondence:***
Tilmann A. Klein, Department of Neurology, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstrasse 1a, 04103 Leipzig, Germany.
e-mail: tklein@cbs.mpg.de

Becoming aware of errors that one has committed might be crucial for strategic behavioral and neuronal adjustments to avoid similar errors in the future. This review addresses conscious error perception ("error awareness") in healthy subjects as well as the relationship between error awareness and neurological and psychiatric diseases. We first discuss the main findings on error awareness in healthy subjects. A brain region, that appears consistently involved in error awareness processes, is the insula, which also provides a link to the clinical conditions reviewed here. Then we focus on a neurological condition whose core element is an impaired awareness for neurological consequences of a disease: anosognosia for hemiplegia (AHP). The insular cortex has been implicated in both error awareness and AHP, with anterior insular regions being involved in conscious error processing and more posterior areas being related to AHP. In addition to cytoarchitectonic and connectivity data, this reflects a functional and structural gradient within the insula from anterior to posterior. Furthermore, studies dealing with error awareness and lack of insight in a number of psychiatric diseases are reported. Especially in schizophrenia, attention-deficit hyperactivity disorder, (ADHD) and autism spectrum disorders (ASD) the performance monitoring system seems impaired, thus conscious error perception might be altered.

Keywords: insula, error awareness, anosognosia, lack of insight, conscious error perception, error-related negativity (ERN), error positivity (Pe)

INTRODUCTION

For daily life it is important that we become aware of the consequences of our actions, of failures and limitations that force us to change our behavior and strategies. In clinical settings, reduced conscious perception of errors has been associated with poor insight in consequences of neurological conditions (O'Keeffe et al., 2004). Whereas it is still unclear whether conscious perception of errors is a necessary prerequisite for all kinds of post-error adjustment (cf. Danielmeier and Ullsperger, 2011), in situations when several people work together it certainly is, because only after conscious detection and appreciation of an error it can be communicated to others and appropriate measures can be taken. This review deals with brain areas that have been shown to play a role in conscious error detection (or "error awareness") in functional magnetic resonance imaging (fMRI) or patient studies. Additionally, electrophysiological studies addressing error awareness and their functional and clinical relevance will be discussed.

Relevant brain areas in the context of error awareness are the posterior medial frontal cortex (pMFC), the thalamus and, as we want to argue in the course of this review, most important, the anterior insula. The insula seems to be crucial for error awareness, because fMRI studies revealed that the insula is consistently activated for consciously perceived errors compared to unperceived errors (Klein et al., 2007a; Hester et al., 2009). Recently, the insula has been suggested to be of relevance for interoception (Craig, 2009, 2011). On the one hand, interoception might contribute to conscious error detection processes, because errors elicit a number of autonomic responses, e.g., changes in heart rate (Wessel et al., 2011) and skin conductance responses (O'Connell et al., 2007), that could potentially be detected by the (anterior) insula. On the other hand, lesions in more posterior regions of the insula have been associated with anosognosia for hemiplegia (AHP, Karnath et al., 2005). AHP describes the unawareness of motor deficits that are related to hemiplegia. Vocat and Vuilleumier (2010) proposed that anosognosia is a multi-componential disorder affecting bodily awareness (amongst other things), or in other words, affecting interoception. Thus, both error awareness and AHP might be linked through interoception or the proper integration of interoception and exteroception. The potential relationship between error awareness and AHP has already been discussed by Vocat and Vuilleumier (2010). Since error awareness processes have been located in the inferior anterior part of the insula, and AHP can be observed after lesions in more posterior parts of the insula, we propose that there is a functional gradient in the insula from anterior to posterior that reflects different aspects of interoception. A similar gradient has also been observed in cytoarchitectonics

and structural as well as functional connectivity analyses of the insular cortex. In this review we want to argue that the insular cortex, due to its cytoarchitectonic layout and its functional as well as structural connectivity, is perfectly suited to play a key role in error awareness. The processing of interoceptive information might deliver information that supports error awareness. The recently proposed role of the insula as a relay station regulating interactions between brain networks involved in external attention and interoceptive cognition (Menon and Uddin, 2010) fits well with the proposed role of the insula in error awareness. Interoceptive information supports error awareness, which in turn might lead to an orienting reaction to the now salient external event.

In the following, we will start with a brief overview over the research on error awareness and its electrophysiological correlates. Then, we will report the neuroanatomical and neurochemical basis of error awareness, with a special focus on the insular cortex. The insular focus and the concept of interoception will lead to a brief discussion on AHP. To complete the picture on error awareness, we selectively report findings on those psychiatric disorders where (a) structural or functional changes in the insula have been reported (among other changes in various brain areas), and (b) electrophysiological studies on error processing exist that suggest an impairment in error awareness.

Error awareness describes the ability to consciously perceive one's own mistakes. A mistake is the failure to achieve the intended goal of an action. Current views suggest that error awareness can be explained by an accumulating evidence account (Ullsperger et al., 2010; Vocat and Vuilleumier, 2010; Wessel et al., 2011; Wessel, 2012). This account describes the accumulation of evidence for an error from very different sources, e.g., pMFC activity, proprioceptive and other sensory input that deviates from expectation, and/or changes in the autonomic nervous system. Thus, each event is evaluated as to whether it indicates or predicts an action outcome that is different (worse) than intended. For example, a deviation of the motor efference copy and/or the proprioceptive and sensory feedback from predictions made in forward models of the action (Desmurget and Grafton, 2000) can indicate that the entire action is going to fail. Later, the observation of the outcome itself deviating from the goal provides additional evidence for the mistake. Moreover, when two alternative response tendencies compete, the resulting response conflict has been suggested to provide evidence for the error (Yeung et al., 2004). These pieces of evidence, which by themselves can be expressed as deviations from predictions (prediction errors), accumulate during and after the action. Evidence accumulation can start as early as the action is initiated, but the point of awareness can be temporally detached from the actual error (e.g., in underdetermined responding, error awareness can only occur after external feedback). Vocat and Vuilleumier (2010) suggest a comparable mechanism, for explicit awareness of motor impairments, i.e., the integration of information from different channels.

It should be noted that the evidence accumulation account outlined above is compatible with predictive coding accounts of awareness and motor action control (e.g., Friston et al., 2010; Seth et al., 2011). Whether error awareness itself is a product of another higher-level predictive-coding mechanism that, for example, compares the predicted task performance with the accumulating prediction error evidence remains to be investigated.

Reduced error awareness can occur under normal as well as pathological conditions. One major determinant may be the type of error that is committed. Depending on the complexity of the task, the level of processing and the information available, different error types can be detected with different reliability (Reason, 1990). During action slips and lapses that occur during skill-based, routine behavior usually all information to detect the error is available such that almost all errors are consciously perceived. For example, in speeded choice reaction time tasks, such as the Eriksen Flanker task, where subjects have to respond to a centrally presented target stimulus and ignore (conflicting) stimuli next to the target, usually 90% or more errors are detected by healthy participants (Ullsperger and Von Cramon, 2006; Seifert et al., 2011). In contrast, mistakes of planning or judgment during rule-based or knowledge-based behavior are less easy to detect (Reason, 1990). Particularly, if errors result from failures of interpretation and comprehension of the current task situation, they are often performed with high confidence and are therefore often missed. In underdetermined, overwhelmingly complex situations, participants have a low confidence in their responses, but without feedback they are unable to determine whether their response was correct or erroneous. Errors can also result from insufficient perceptual information, for instance, when stimuli are degraded or masked. In this case, the necessary sensory information for performance monitoring processes is missing, so that errors cannot be noticed. If errors result from general decreases of arousal and a disengagement from the task (Eichele et al., 2008), their likelihood to be consciously perceived can be expected to decrease. This may be particularly true for errors that occur after sleep deprivation (Scheffers et al., 1999; Chee et al., 2008), but this hypothesis still needs to be tested. Indeed Shalgi et al. (2007) were able to show that greater task monotony (presumably via reduced arousal) reduces the number of errors that are consciously perceived. Finally errors can result from failures in the processing of the perceptual properties of the stimulus (see also section "Experimental Paradigms to Investigate Error Awareness").

Usually, error awareness has been studied by asking participants whether they noticed having made a mistake, since it has been unclear whether error awareness can be quantified reliably in a more direct and objective way, i.e., without asking participants after every trial. However, recent studies suggest that the amplitude of the error positivity (Pe) might be a good quantitative correlate of error awareness (Murphy et al., 2012; see below), particularly when quantified in single trials and/or time-locked to the error-signaling response (see below), since the Pe seems to be linked to the time when the subject presses the error-signaling button.

Often, participants are asked to signal any encountered error by pressing an "error signaling button" (Rabbitt, 1968). This procedure may, however, induce some response bias, because for responses considered correct no motor response is needed. Furthermore, short inter-trial intervals may prevent participants from signaling errors despite being aware of them. A number

of studies therefore explicitly asked participants after each trial, whether they deemed the preceding behavior correct or incorrect (Endrass et al., 2007; Klein et al., 2007a; Logan and Crump, 2010; Wessel et al., 2011).

EXPERIMENTAL PARADIGMS TO INVESTIGATE ERROR AWARENESS

Three kinds of tasks have been used to study error awareness. As discussed in Ullsperger et al. (2010), they appear to interfere with the accumulation of error evidence at different stages, thereby resulting in a sufficient number of errors that remain unconscious. (1) When the detection of stimuli is rendered increasingly difficult, for example by degrading visibility (Scheffers and Coles, 2000) or metacontrast masking (Maier et al., 2008; Cohen et al., 2009; Steinhauser and Yeung, 2010), participants not only make more errors, they are also less certain about their performance and miss a number of mistakes. (2) Oculomotor tasks, such as the antisaccade task, have been very successful in inducing unperceived errors (Nieuwenhuis et al., 2001; Endrass et al., 2007; Klein et al., 2007a; Wessel et al., 2011). It appears that error evidence from proprioception and sensory (visual) input is rather weak for short and immediately corrected prosaccades, such that they are easily overlooked (Ullsperger et al., 2010). (3) In complex task sets consisting of a number of competing and constantly to-be-monitored rules, some rule representation may be dominant and others only weakly represented. Errors related to one rule may then remain undetected more frequently. This principle has been successfully applied in a number of studies using a Go/NoGo task with two different NoGo conditions (Hester et al., 2005; O'Connell et al., 2007). The typical error awareness task in these studies consisted of color words printed in congruent or incongruent ink (as in a Stroop task). The majority of stimuli were congruent words, serving as signal for a Go response. In contrast, when incongruent stimuli appeared (rule 1) or a color word was repeated in two successive trials (rule 2), subjects had to withhold their response (NoGo). Continuously monitoring both congruency and repetitions appears to be difficult and leads to many NoGo errors that subjects are not aware of.

ELECTROPHYSIOLOGICAL CORRELATES OF ERROR AWARENESS

Performance monitoring is associated with a number of neural correlates that appear to be differentially modulated by conscious error perception. Based on early findings in antisaccade tasks (Nieuwenhuis et al., 2001; Klein et al., 2007a) and the Go/NoGo "error awareness task" (O'Connell et al., 2007) it was assumed for a long time that the error-related negativity (ERN) (Falkenstein et al., 1990; Gehring et al., 1993), a frontocentral event-related potential occurring shortly after erroneous button presses in speeded choice reaction time tasks, was present on all error trials and unaffected by conscious error perception. In contrast, the later and more posterior Pe (Falkenstein et al., 1990) was present only when errors were perceived consciously (Nieuwenhuis et al., 2001; Endrass et al., 2007). Similarly, neuroimaging studies seemed to suggest that the pMFC, the putative generator of the ERN, was active on both reported and unreported errors, whereas the anterior insula was specifically modulated by error awareness (Ullsperger et al., 2010).

However, a recent study using an antisaccade task (Wessel et al., 2011) as well as studies using degraded or masked stimuli (Scheffers and Coles, 2000; Steinhauser and Yeung, 2010) showed that the ERN may co-vary with error awareness as well. Smaller ERN amplitudes are associated with a lower likelihood to consciously perceive the error. Shalgi and Deouell (2012) were able to show that the amplitude of the ERN is related to error awareness and that it co-varies with the individual confidence with which an answer was made (higher amplitude in aware errors for confident subjects). In line with this, more recent fMRI studies reported greater pMFC activity in aware compared to unaware errors (Hester et al., 2009, 2012; Orr and Hester, 2012; see also "Posterior Medial Frontal Cortex"). Current views suggest that the ERN (and feedback-related negativity, FRN) reflects the processing of single pieces of objective evidence for an error (or other events requiring adaptation). For example, when stimulus-induced evidence is low, the ERN amplitude is low (Scheffers and Coles, 2000). In a flanker task study with response feedback, in most trials feedback is redundant and not associated with an additional negativity (De Bruijn et al., 2004; Gentsch et al., 2009). When, for any reason on some trials efference copy or perceptual information available at the time of the response was reduced (behaviorally reflected in prolonged remedial action times), not only the ERN was reduced in amplitude but also an FRN appeared in the same trial (Gentsch et al., 2009). Thus, the additional feedback information was used to disambiguate the situation. In such trials, two small pieces of evidence for an error occurred in short succession and were both reflected in medial frontal negativities, namely the (reduced) ERN and (increased) FRN. This is compatible with the view that error evidence accumulates with new incoming information related to action outcome. When sufficient evidence has accumulated, this may be the basis of error awareness. In contrast to the ERN, the Pe reflects the subjective (accumulated) evidence associated with conscious awareness (cf. Wessel, 2012). A recent study suggests that the Pe amplitude and latency correlates with the subject's indication of error awareness and predicts reliably whether an error would be consciously perceived or not (Murphy et al., 2012). Thus, the Pe appears to be a good measure of error awareness. Murphy et al. (2012), however, suggest investigating the Pe locked to the error-signaling response and not time-locked to the response. This should make clear that a reduced amplitude is really due to diminished awareness and not to for example a higher variability in the latency of error awareness.

FUNCTIONS OF THE INSULAR CORTEX

Several reviews about the functional neuroanatomy of the insula have been published recently (Kurth et al., 2010; Menon and Uddin, 2010; Cauda et al., 2011; Kelly et al., 2012). Therefore, we only want to give a brief overview over functions that have been associated with this brain area (see **Figure 1**). In line with the cytoarchitechtonic gradient in the insula (Mesulam and Mufson, 1982a; see below)—from agranular cortex in the (inferior) anterior part to dysgranular cortex in the middle part to granular cortex in the posterior part—Cauda et al. (2011) reported two

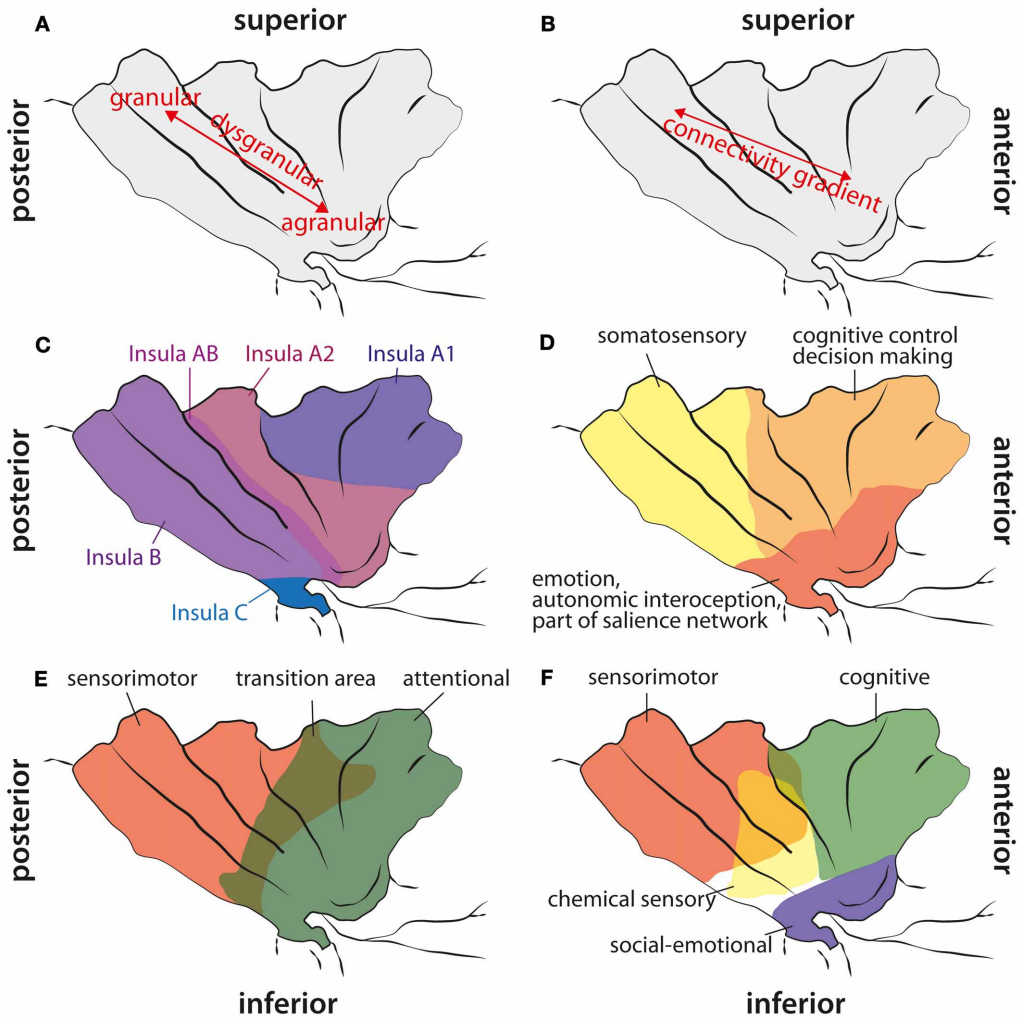


FIGURE 1 | Cytoarchitectonic, structural connectivity, and functional maps of the human insula. (A) Cytoarchitectonic gradient from agranular cortex in the anterior inferior insula via dysgranular cortex to granular cortex in the posterior part of the insula. **(B)** Structural connectivity gradient in the insula according to Cerliani et al. (2012). Unlike in other brain areas (e.g., premotor cortex), they did not find any clear border between insula regions based on the structural connectivity profile; instead, they reported a gradual change in connectivity patterns from anterior to posterior insular areas. **(C)** Cytoarchitectonic map adapted from Von Economo and Koskinas (1925). They did not find any agranular areas within the insula (except from a fronto-insular region anterior to what is shown here), but a less granulated area "Insula A1" and stronger granulated areas more posterior. Note, that they explicitly report a transition area "Insula AB" between anterior and posterior insular regions. **(D)** Functional areas of the insula according to Deen et al. (2011). **(E)** Functional differentiation of the insula according to Cauda et al. (2011). Note, that they report a transition area between the anterior and the posterior part. **(F)** Functional areas in the insula according to Kurth et al. (2010). They reported four different areas, also with a clear overlap in the middle aspects of the insula.

overlapping functional networks, an attention-related network anterior, and a sensorimotor network posterior, with a large overlap of both networks in mid-insula areas. By means of a meta-analysis of functional neuroimaging data, Kurth et al. (2010) found four distinct functional regions within the insula. They described the inferior anterior part of the insula in terms of social-emotional processes, the superior anterior part in relation to cognitive processes, a chemical sensory area in the middle part and a sensorimotor area in the posterior part, with considerable overlap between functional areas especially in the middle part of the insula. Based on resting state data, Kelly et al. (2012) reported up to nine different functional clusters within the insula, also with considerable overlap between these clusters. In agreement with other studies, they found cognitive and attentional processes to be located in more anterior parts, emotional aspects in inferior parts, and sensorimotor and interoceptive processes in posterior parts. Additionally, Mutschler et al. (2009) reported consistent activation of the inferior anterior insula in relation to peripheral physiological changes. As reviewed already by Augustine (1996), the insula is engaged in a wide variety of functions, such as visceral sensory and motor processes, vestibular processes, limbic integration, motor association, and language-related auditory processing. In the last decade, the role of the insula in interoception has been emphasized, as well as its role in emotional

and interoceptive awareness or awareness in general (Critchley et al., 2004; Craig, 2011; Simmons et al., 2012). Recently, it has been suggested that the right fronto-insular cortex plays a crucial role in switching activity between different functional networks, especially the default mode and an executive network (Sridharan et al., 2008), or that the anterior insula is involved in detecting novel salient stimuli in different modalities (Menon and Uddin, 2010). This last hypothesis is in agreement with the suggestion that the anterior insular cortex (AIC) is part of a salience network, consisting of the AIC, the anterior cingulate cortex, the amygdala, and the inferior frontal gyrus (IFG) (Seth et al., 2011). The notion that the AIC belongs to a salience network fits well with observations that the AIC plays a crucial role in error awareness (e.g., Klein et al., 2007a; see below), because consciously perceived errors are obviously salient events, whereas unnoticed errors are not. Furthermore, there are strong intra-insular connections (Augustine, 1996; Kurth et al., 2010), suggesting that posterior parts might feed information into the salience network located in AIC. An interruption of this process due to lesions within the insula might result in a mismatch in bodily or sensorimotor perceptions. Especially the awareness for limb functioning and the sense for limb ownership seem to require intact insular functions. As pointed out by Karnath et al. (Karnath et al., 2005; Baier and Karnath, 2008; Karnath and Baier, 2010), especially the right posterior insular was repeatedly found in lesion analysis studies with stroke patients to be a central element in the process of interoceptive awareness necessary for intact sense of limb functioning and limb ownership. Berti et al. (2005) also report that, besides lesions in motor and premotor areas, lesions to prefrontal areas like BA 46 and the insula are differentially involved in AHP as well (but less frequent). More recently, however, Vocat et al. (2010) reported lesions to the anterior insula as being crucial for AHP especially during the hyperacute (three days post insult) phase.

NEUROANATOMICAL BASIS OF ERROR AWARENESS

A few brain areas have been associated with conscious error perception. Most studies suggest that the anterior insula is crucial for error awareness. Besides the insula, the pMFC (comprising the pre-supplementary motor area and an area that is equivalent to the ACC in monkeys, i.e., the anterior mid-cingulate cortex, aMCC; cf. Vogt, 2005), and the thalamus might be important for error awareness. In the majority of studies reporting insula activations, the aMCC is co-activated with the insula (cf. Craig, 2009). In the following, insula anatomy and connectivity will be described briefly. Then, we will report studies providing evidence that the anterior insula, the thalamus, and the pMFC are crucial for conscious error perception.

INSULA: STRUCTURE AND CONNECTIVITY

The anterior inferior part of the human insula consists of agranular cortex. Specific cytoarchitectonic areas of the insula are preferentially connected to cytoarchitectonically similar areas in other parts of the brain (Mesulam and Mufson, 1982b; Mufson and Mesulam, 1982), that is, agranular areas are predominantly connected to other agranular cortical areas, e.g., the anterior cingulate cortex. Cytoarchitectonically, there is a gradient from agranular cortex, located in the anterior insula, to dysgranular cortex located in the middle part of the insula and to granular cortex in the posterior insula (see **Figure 1**). However, Von Economo and Koskinas (1925) doubted that there are agranular areas in the insula (except of a fronto-insular transition area at the anterior border of the insula), but their data also suggest the presence of a cytoarchitectonic gradient from dysgranular cortex in the superior anterior insula to granular cortex in posterior insula regions.

The von Economo neurons (VENs) have been found in both the anterior cingulate cortex and the frontal insular cortex in humans and great apes (Von Economo, 1926; Allman et al., 2010; Seeley et al., 2012), and recently also in macaque monkeys (Evrard et al., 2012). They appear to be projection neurons and most likely project to the frontal pole, other frontal, and insular areas, the septum, and the amygdala (Allman et al., 2010). Allman et al. (2010) found that the protein, which is encoded by the DISC1 gene (disrupted in schizophrenia), is preferentially expressed in VENs, thereby relating these neurons to a genetic basis of schizophrenia.

In macaque monkeys the insular cortex is characterized by widespread anatomical connections (for an overview see Cerliani et al., 2012), among them projections to autonomic nuclei in the brainstem and several thalamic nuclei. Cerliani et al. (2012) investigated white matter connections of the insula cortex in humans by using diffusion-weighted imaging (DWI). This type of connectivity data can be used to parcellate brain regions according to their connectivity profile. Usually, at the border between two functionally different regions, a clear change in the connectivity profile can be observed. However, Cerliani et al. (2012) reported that this connectivity-based parcellation did not yield reliable, clearly distinguishable subregions for the insula, since they did not find any of these abrupt changes in the connectivity profile of the insula. The authors instead argue that their connectivity data suggest a gradient in connectivity profiles from the anterior to posterior insula, which show a large overlap in their connectivity profile without any distinct borders. According to their results, the anterior (agranular) part of the insula is mainly connected to the orbitofrontal cortex, the amygdala, and the rostral part of the IFG, whereas the posterior (granular) part of the insula is connected to parietal and posterior temporal areas, caudal parts of the IFG, and the lateral premotor cortex. The dysgranular insular cortex in between shows some overlap in the connectivity pattern with both the anterior and posterior insula. While the anterior dysgranular areas are more similar to the connectivity pattern of the agranular insula, the posterior dysgranular areas are more similar to the connectivity pattern of the granular insula. Thus, in line with the cytoarchitectonic gradient from agranular to granular cortex, there also is a connectivity gradient from anterior to posterior in the insula without any distinct borders that could potentially have been defined based on abrupt changes in connectivity (Cerliani et al., 2012). However, one limitation of this study is that only brain areas, that are part of the probabilistic cytoarchitectonic map from Juelich (cf. Cerliani et al., 2012 for a complete overview of used maps), have been included as potential target areas. Therefore, some brain areas, such as the aMCC, that are connected to the insula in rhesus monkeys (Morecraft

et al., 2012), but which are not yet included in these maps, could not be found as projection targets of the insula. The large diversity of insula connections makes this brain area well suited for the integration of external signals with interoception (Seth et al., 2011).

Cauda et al. (2011) used resting state fMRI measures to identify functional networks of the insula. They found an anterior and a posterior network in the insula and a transition area in between, thus, corroborating the insular gradient in the cytoarchitectonic data and the DWI study by Cerliani et al. (2012). The anterior insula was functionally connected to the rostral ACC, middle and inferior lateral frontal cortex and temporoparietal areas. The posterior insula was part of a network consisting of the superior posterior cingulate cortex, motor areas (including pre-SMA and premotor cortex), somatosensory areas, temporal cortex, and parts of the occipital lobe. The authors described the anterior insular network as "attention network" and the posterior insular network as "sensorimotor network."

Deen et al. (2011) divided the insula in 3 subregions based on functional connectivity measures: an anterior inferior part, an anterior superior part, and a posterior part. The inferior part of the anterior insula was most strongly connected to the pregenual ACC, while the superior part of the anterior insula was mainly connected to the aMCC. The posterior insula was functionally connected to posterior MCC. The results by Deen et al. (2011) suggest that areas of the insula are systematically connected to the medial frontal cortex (MFC) with more anterior insular regions being connected to more anterior MFC regions and more posterior insular regions to more posterior MFC areas. Besides the functional connections between insula and MFC, extensive connectivity with other brain areas were reported (Deen et al., 2011): the inferior anterior insula was connected to opercular cortex, the posterior IFG, and the superior temporal sulcus. The superior anterior insula showed functional connections with visual areas, the medial thalamus, opercular and posterior orbitofrontal cortex, pre-supplementary motor cortex, precentral sulcus, supramarginal gyrus, and the anterior IFG. The posterior insula was connected with motor (including SMA) and somatosensory areas, opercular cortex, pre-SMA, and the medial thalamus. All insular subregions were interconnected, providing a structural basis for the communication between different parts of the insula, i.e., between somatosensory and attention-related areas.

Co-activations of brain regions, and thus potential functional networks, can also be demonstrated in fMRI meta-analysis as well as in spatial independent component analysis (ICA) of fMRI data. A meta-analysis of performance monitoring showed co-activation of anterior insula, aMCC, and thalamic regions (Klein et al., 2007a; Ullsperger et al., 2010). Similarly, the posterior MFC, thalamus, and anterior insula were repeatedly covered by the same independent components, suggesting a highly similar signal time course in these regions (Dosenbach et al., 2007; Eichele et al., 2008; Danielmeier et al., 2011).

In conclusion, the insular cortex is involved in at least 2–3 functional networks. Both macaque cytoarchitectonics and human connectivity studies (Cauda et al., 2011; Deen et al., 2011) suggest that there might be no distinct subdivisions in the insula, but instead, that there is an anterior-to-posterior gradient in both cytoarchitectonics and connectivity. Craig (2011) suggested a functional gradient within the insula, with posterior insula regions reflecting the objective stimulation strength (e.g., of painful stimuli), and anterior regions reflecting subjective feelings related to this stimulus. This suggestion is in line with interoceptive processes that have been associated with the insula.

THE INSULA AND AWARENESS DEFICITS

Especially the anterior inferior insula seems to be involved in error awareness. In an antisaccade task, the anterior insula was the only brain area distinguishing between consciously perceived and unperceived errors (Klein et al., 2007a). In an fMRI study employing the error awareness Go/NoGo task described above, it has been shown that consciously perceived errors go along with larger BOLD responses in the right insula and in left inferior parietal regions (Hester et al., 2009).

According to Kurth et al. (2010), the anterior insula is functionally related to attentional and cognitive processes and autonomic responses. A meta-analysis (Mutschler et al., 2009) has associated the anterior inferior insula with autonomic responses, such as changes in heart rate or skin conductance rate. Furthermore, this part of the insula is often co-activated with the amygdala. Intraoperative stimulations of the insula in epileptic patients led to changes in cardiac responses (Oppenheimer et al., 1992). Wessel et al. (2011) have recently described a link between error awareness and cardiac responses. Following perceived errors, a stronger heart beat deceleration was observed compared to unperceived errors. Craig (2009) suggested that the right AIC activity correlates with subjective feelings of body states, e.g., pain or awareness of heartbeats. Similarly, Paulus et al. (2009) suggested that the functional role of the insula is to detect discrepancies between the predicted body state and the actual body state. Recently, Seth et al. (2011) proposed a model of awareness (or "presence") in general, i.e., not restricted to error awareness. They suggested that the insula is crucial for the integration of interoceptive and exteroceptive signals, and the anterior insula is assumed to be a "comparator or error module" (Seth et al., 2011). This is likely to also apply to error awareness, as already discussed in the Introduction. However, at this point it is unclear whether the autonomic response is cause, result or correlate of error awareness.

As mentioned above, the posterior region of the insula is connected to the SMA and premotor areas (Cerliani et al., 2012). Lesions in the right posterior insula can lead to AHP (Karnath et al., 2005), which can be defined as selective disorder of awareness for motor deficits (Spinazzola et al., 2008). This supports the notion by Craig (2009) that the insular cortex in general is related to awareness. While the anterior insula has been associated with error awareness, the middle and posterior insular cortex seem to be associated with the awareness of motor and somatosensory processes (Karnath et al., 2005; Spinazzola et al., 2008). Thus, depending on the exact lesion location within the insula, one might observe different, domain-specific awareness deficits. This gradient in awareness deficits from anterior to posterior insular areas might reflect the underlying connectivity and cytoarchitectonics gradient within the insula.

THALAMUS AND AWARENESS DEFICITS

Some studies suggest that thalamic lesions can also impair error awareness and lead to anosognosia (De Witte et al., 2011; Peterburs et al., 2011; Seifert et al., 2011). In a review by De Witte et al. (2011), two patients were described who suffered from anosognosia after bilateral thalamic lesions. In a study by Seifert et al. (2011), patients with chronic thalamic lesions were asked to participate in a flanker task and signal their errors with a button press. While the age matched control group signaled 85% of their errors on average, the patient group indicated only 39% of their erroneous responses. This suggests that in patients, suffering from thalamic lesions, a majority of errors is not perceived consciously. This result has been replicated in an antisaccade task (Peterburs et al., 2011). Patients with thalamic lesions signaled their errors significantly less often than the healthy control group. Thus, there is preliminary evidence that error awareness is reduced following thalamic lesions. However, in both studies an error signaling procedure was used. This procedure has some disadvantages, e.g., patients who generally respond slower than healthy individuals might prefer to avoid additional button presses in between trials. Thus, they might notice their errors but miss to indicate them (cf. Wessel, 2012). However, significantly reduced Pe amplitudes in these patients provide additional evidence for impaired error awareness (Seifert et al., 2011). A further, previously unpublished analysis of the data by Seifert and colleagues broken down by lesioned thalamic subregions revealed that the Pe is absent in patients with lesions focused on the mediodorsal nucleus and only marginally reduced in amplitude in patients with focal lesions in the ventral anterior and ventrolateral anterior nuclei. Interestingly, the ERN showed the opposite pattern of impairment. Thus, it appears that the basal-ganglia-thalamocingulate circuit is involved in ERN generation, whereas the more arousal-related circuitry of the mediodorsal nucleus plays a role in error awareness and generation of the Pe. Given the limited sample size, further studies with thalamic patients are necessary that involve a procedure where participants are required to evaluate the accuracy of their response after every trial (e.g., as described in Klein et al., 2007a; Wessel et al., 2011).

POSTERIOR MEDIAL FRONTAL CORTEX

There are mixed results with respect to the role of the pMFC in error awareness. While earlier studies did not find any difference in pMFC activity between perceived and unperceived errors (Hester et al., 2005; Klein et al., 2007a), recent studies did find a difference in medial frontal areas (Hester et al., 2009, 2012; Orr and Hester, 2012). Furthermore, recent electrophysiological studies on error awareness found larger ERN amplitudes in consciously perceived errors compared to unperceived errors (Scheffers and Coles, 2000; Steinhauser and Yeung, 2010; Wessel et al., 2011; Shalgi and Deouell, 2012; for a review see Wessel, 2012). While error correction appears to be affected by lesions of the pMFC (Swick and Turken, 2004; Modirrousta and Fellows, 2008), studies directly addressing error awareness in patients with pMFC lesions are lacking.

In sum, most error awareness studies identify the anterior inferior insula as crucial neuronal correlate of conscious error perception, but there is also preliminary evidence that the pMFC and thalamic regions are important structures for error awareness.

DRUGS AFFECTING CONSCIOUS ERROR PERCEPTION

It has been shown that the use of certain drugs attenuates the response of the aMCC to errors or diminishes the ERN. This has been demonstrated for cocaine, opioids, and alcohol (Ridderinkhof et al., 2002; Kaufman et al., 2003; Forman et al., 2004). Furthermore, the dopaminergic D2 receptor antagonist haloperidol reduces the ERN response to errors in flanker tasks (Zirnheld et al., 2004; De Bruijn et al., 2006), and there is evidence that smaller ERN amplitudes go along with reduced error awareness (see above). Moreover, subjects with lower D2 receptor density showed attenuated pMFC responses to negative feedback (Klein et al., 2007b). Therefore, the question arises whether the use of certain drugs also affects error awareness and, more specifically, whether dopamine (DA) plays a crucial role in conscious error detection. In the following, we will review those studies that investigated error awareness under pharmacological challenges.

Hester et al. (2007) showed that cocaine use can lead to reduced error awareness. They investigated a group of active cocaine users with the Go-NoGo error awareness task described above. Cocaine is assumed to exert its influence by blocking the re-uptake of DA, norepinephrine (NE) and serotonin and thereby increasing extracellular DA levels in those brain areas with afferents from the mesolimbic DA system (cf. Kuhar et al., 1991; Jocham et al., 2007). Thus, a long-term effect of cocaine use seems to be that the DA receptor density decreases (Volkow et al., 1990). This could explain an attenuated aMCC response in cocaine users (Kaufman et al., 2003; Klein et al., 2007b). However, error awareness in cocaine users was not decreased in general, but specific for certain error types. In the error awareness task employed by Hester et al. (2007), errors were committed when participants failed to withhold their response either to incongruent stimuli (first NoGo condition) or to the repetition of the same stimulus as in the trial before (repeat trials, second NoGo condition). Interestingly, cocaine users showed reduced error awareness only in repeat trial errors, but conscious error perception in incongruent trials was comparable to that in the control group. Therefore, one cannot unequivocally conclude that cocaine use leads to reduced error awareness. An alternative explanation could be that cocaine use might cause slight working memory or attentional impairments, and therefore, only repeat trial errors were noticed less often. Moreover, Garavan et al. (2008) showed that cocaine does affect insular activity. After i.v. cocaine administration, participants showed *enhanced* insular activity in response to performance errors. Note that an acute cocaine administration might evoke different effects than long-term cocaine usage.

In a later fMRI study using the same task, Hester et al. (2009) showed that error awareness is reduced in chronic cannabis users. In this study, reduced error awareness was associated with attenuated aMCC activity in cannabis users. There was also a relation between insula activity and error awareness: insula activity was negatively correlated with the amounts of cannabis used, and

higher cannabis craving was correlated with less error awareness. Insula activations are reliably found in craving paradigms (Garavan, 2010). The craving for abused drugs seems to be linked to hypocretin (orexin) transmission in the insula, which in turn might modulate DA release in connected brain areas (Kenny, 2011). In mice, hypocretin has been associated with modulations in wakefulness, and it has been suggested that the hypocretin neural network might initiate arousal responses (Adamantidis et al., 2007), which would be a plausible adjustment after errors. It has been suggested before that errors might elicit an orienting response (Notebaert et al., 2009; Wessel et al., 2012), which is associated with increased arousal. Thus, hypocretin effects in the insula might be two-fold: on the one hand it could potentially increase arousal, and on the other hand, it could influence DA and NE release.

A recent study showed an enhancing effect of methylphenidate (MPH) on error awareness (Hester et al., 2012). Healthy participants perceived more errors consciously when they were under the influence of MPH than in the placebo condition. MPH has also proven to be effective in the treatment of cognitive deficits that can be observed after traumatic brain injury (Willmott and Ponsford, 2009), which have been associated with reduced awareness (O'Keeffe et al., 2004). Since MPH is a DA and NE reuptake inhibitor, it can be seen as indirect DA agonist (cf. Hester et al., 2012). Thus, this study provides further evidence that error awareness could be related to DA levels. Furthermore, a study by Frank et al. (2007) showed an effect of the catechol-o-methyltransferase (COMT) genotype on the Pe, which is systematically linked to error awareness. The val/met polymorphism of the COMT gene has been associated with prefrontal DA levels (Egan et al., 2001; Bilder et al., 2004).

Although the number of studies investigating neurochemical aspects of conscious error perception is very limited, there is converging evidence that the catecholamines DA and NE are highly relevant neurotransmitters associated with error awareness. Most direct evidence for a relation between these neurotransmitters and error awareness has been collected with psychostimulants that increase extracellular DA and NE. The role of hypocretin needs further investigation, but it seems to modulate DA, NE, and serotonin release as well. Given its proposed role in the orienting reflex and the generation of the P300 (and Pe) potentials (Nieuwenhuis et al., 2005), NE is likely to be involved in error awareness. Based on the current knowledge on the role of DA and NE in performance monitoring and attention (Aston-Jones and Cohen, 2005; Jocham and Ullsperger, 2009; Cools, 2011) one can assume that these neurotransmitters play a role in strengthening the error signal and enhance subsequent central neural and autonomic activity changes that contribute to conscious error perception. Differentiating the contributions of DA and NE to error awareness is an important goal for future research.

INSULA INVOLVEMENT IN ANOSOGNOSIA FOR HEMIPLEGIA: A LINK TO ERROR AWARENESS?

As already proposed by Vocat and Vuilleumier (2010), AHP and error unawareness might share some neuroanatomical correlates. Poor insight into the consequences of a neurological disease is related to poor treatment outcome and reduced treatment compliance. Sometimes anosognosia is also accompanied by a disturbance of the sense of agency and the sense of limb ownership (Karnath and Baier, 2010). In general, anosognosia can be observed following both right and left hemispheric brain damage with some predominance of appearance following right hemispheric insult. In the acute phase of a neurological disease, anosognosia is observed quite often: In a meta-analysis of 27 studies a median of 26% of patients following right hemispheric and a median of 10% of patients following left hemispheric stroke showed signs of anosognosia (Jehkonen et al., 2006).

The insular cortex has often been associated with deficit awareness (Karnath et al., 2005; Prigatano, 2009; Craig, 2010). Other relevant brain areas in AHP are the prefrontal cortex, the inferior parietal lobe, the angular gyrus, the supramarginal gyrus, and the anterior temporal lobe (Prigatano and Shacter, 1991). Especially during the first days of acute illness, damage to the posterior insular is predictive for developing anosognosia. Vocat and colleagues showed that in patients with *sustained* unawareness frontal and parietal brain areas were also affected (Vocat et al., 2010).

There might be two subcomponents of error processing: an early component that is not dependent on any kind of proprioceptive feedback but solely based on the efference copy of the executed action, and a second component that is more about the evaluation of the error and potential adjustments to avoid future errors of a similar kind (see Vocat and Vuilleumier, 2010 for a similar account of deficit awareness). Similarly, Prigatano (2010) claims that self-awareness is necessary for performance monitoring. This self-referential information (provided by interoceptive awareness, Craig, 2010) has to be integrated with external information (supplied by exteroceptive awareness, Craig, 2010) in order to come up with an accurate view of the situation or the action and its outcome, respectively. It has been suggested that this integration takes place in the insular cortex (Craig, 2011). When the representation of internal or external information is corrupted (what might be the case following brain lesions), deficient decisions or profound problems in awareness of the outcome of a decision/action might be the consequence.

Although several brain areas have been discussed to play a role in AHP (for reviews on AHP in general see e.g., Vuilleumier, 2004; Vocat and Vuilleumier, 2010), the insular cortex seems to play a prominent role in deficit awareness. The complex connectivity patterns of anterior and posterior insular cortex might suggest that awareness in general is the product of a network of brain regions all providing different kinds of information which finally allow awareness for internal and external information (e.g., Vocat and Vuilleumier, 2010). However, studies directly linking symptoms of AHP to electrophysiological (Pe) or functional (fMRI) correlates of error awareness are missing so far.

PSYCHIATRIC ILLNESS, THE INSULAR CORTEX, AND ERROR AWARENESS

Psychiatric patients sometimes show a high degree of lack of insight into their psychiatric condition. Because lack of insight might be related to deficient monitoring processes and reduced

self-awareness, we review several studies that investigated error monitoring (mostly electrophysiological correlates of error monitoring or error awareness) in psychiatric patients. Lack of insight is, for example, a frequent observation in patients suffering from schizophrenia. Other psychiatric diseases like attention-deficit hyperactivity disorder (ADHD) and autism spectrum disorders (ASD) might also lead to patients' insensitivity to negative action outcomes, thereby promoting reduced error awareness. Since the insular cortex seems to play a crucial role in error awareness and AHP, an explicit focus will be put on the potential role of this area in disease symptomatology. This does, of course, not imply that a potential insular pathology alone accounts for the psychiatric disease under discussion.

CLINICAL SYMPTOMS

Misattribution of thoughts and events to external sources as a consequence of altered monitoring processes is one key symptom of schizophrenia (Frith, 1987; Ullsperger, 2006). At least 50% of schizophrenic patients are without awareness for their disorder (Pia and Tamietto, 2006). The question, whether or not error awareness is compromised in schizophrenia was directly addressed by Mathalon et al. (2002) using a task in which no error signaling response was required. They compared patients suffering from schizophrenia with healthy controls while both groups worked on a performance-monitoring task with concurrent EEG recordings. Compared to healthy controls, schizophrenic subjects (especially those with paranoid schizophrenia) showed smaller ERN and larger CRN (correct-related negativity) amplitudes, but no differences in the Pe amplitude and subsequent post-error slowing (PES) (Mathalon et al., 2002). The authors concluded that perception of an error is sufficient but not necessary for producing the ERN, but that it is necessary for producing the Pe. Deficits in self-monitoring as indexed by altered performance monitoring might underlie some of the positive symptoms observed in schizophrenia as Mathalon et al. (2002) suggest. Corollary discharge dysfunction in schizophrenia (Ford et al., 2001) could explain not only the sensory integration deficits but also misattributions of action outcomes and agency and therefore result in error awareness deficits.

Although schizophrenia affects various brain areas, we focus here on studies reporting changes in the insula. Cytoarchitectonic alterations in the inferior insular and enthorinal cortex were found by Jakob and Beckmann (1986) in a subsample of post-mortem brains of schizophrenia patients. Glahn et al. (2008) extended these findings with a meta-analysis by showing that gray matter of schizophrenic patients is reduced in a number of brain areas one of which being the bilateral insular cortex. Volume reduction of the left insular cortex in schizophrenia has been shown several times (Crespo-Facorro et al., 2000; Kasai et al., 2003; Kim et al., 2003). Using surface based morphometry in 57 schizophrenic patients Palaniyappan et al. (2011) demonstrated an inverse relationship between right posterior insula structure and the degree of insight in schizophrenia. The authors concluded that the right posterior insula might be the basis for insight as it allows for interoceptive awareness and self-appraisal of emotional states within a functional network that comprises also distant brain areas. In an extensive review of the existing literature Wylie and Tregellas (2010) summarized the role of the insular cortex in schizophrenia symptomatology. They concluded that damage to the insula or damage to a greater network comprising insular cortex could underlie many of the sensory-integration deficits observed in schizophrenia (Wylie and Tregellas, 2010). In a recent paper Williamson and Allman (2012) review the role of different functional brain networks in schizophrenia. Incorporating studies using voxel-based morphometry (VBM), DWI, and resting state functional MRI they report different brain networks as potentially relevant for schizophrenia. Besides (mal-) functioning of a network related to directed effort (comprising superior anterior and posterior cingulate cortex, auditory cortex, and the hippocampus) also its (disturbed) interactions with a brain network related to representations of thoughts, feelings, and actions (with the frontal and temporal pole and the frontoinsular cortex) might be involved in the pathophysiology of schizophrenia.

ADHD is associated with abnormalities in response to performance errors (O'Connell et al., 2009). PES, which might be indirectly linked to error awareness, was found to be reduced in a large sample of ADHD children (Schachar et al., 2004).

O'Connell et al. (2009) investigated electrophysiological correlates of performance monitoring in adult ADHD subjects. While making more errors, the ADHD patients showed an attenuated early and late Pe following erroneous responses. Furthermore, these subjects showed reduced electrodermal reactivity to an error, thus suggesting that errors had less emotional relevance to them. Wiersema et al. (2009) corroborated the findings by O'Connell et al. in ADHD subjects by showing reduced Pe amplitudes which were correlated with ADHD symptoms, normal sized ERN and no differences in behavioral adaptation following an error. They concluded that early automatic error detection processes are not affected in ADHD, but that there are illness related differences in later evaluative processes.

Several other studies investigated error processing in children with ADHD. It has been shown that ADHD children committed twice as many errors as healthy controls and did not show post-error behavioral adaptations, like PES (Schachar et al., 2004; Wiersema et al., 2005). As in ADHD adults, ADHD children also showed reduced Pe amplitudes as compared to healthy controls (Wiersema et al., 2005; Zhang et al., 2009).

Patients suffering from ASD sometimes show perseverative behavior, which might be interpreted as a consequence of impaired performance monitoring. These patients might be less sensitive to the course or the outcome of their actions thereby having an increased risk to repeat behavior over and over again. A role for the insular cortex might be assumed in this disorder: Ebisch et al. (2011) showed in a group of 12–20-year-old adolescents with high-functioning ASD in resting state fMRI that functional connectivity between right anterior and bilateral posterior insula with different other brain regions (posterior: inferior and superior somatosensory cortices; anterior: amygdala) was diminished as compared to results from a control group. Di Martino et al. (2009) were able to show an increased likelihood for hypoactivation of right anterior insula in studies analyzing brain activity in social tasks with patients suffering from ASD. Santos et al. (2011) investigated the frontoinsular cortex

(layer V) of children suffering from autism using a stereologic approach. They found a significantly higher ratio of VENs to pyramidal neurons in the frontoinsular cortex of autistic children (for a discussion of VEN in insular cortex see "Insula: Structure and Connectivity"). Furthermore, they found a trend for an increase of the total number of VENs in frontoinsular cortex in autistic patients compared to their respective controls. The authors interpret these findings in terms of a potential neuronal overpopulation which might finally lead to increased interoception sometimes being described as part of the autistic syndrome.

Sokhadze et al. (2010) confronted autistic children with a task suitable for investigating error processing (no error-signaling response required). They demonstrated that these children committed significantly more errors, showed a smaller ERN with a larger latency, a trend toward a reduction in Pe amplitude compared to controls and no signs of PES; they rather showed post-error-speeding. The authors discuss that these alterations in performance monitoring might lead to reduced error awareness thereby allowing no successful adaptation of behavior following an error. Furthermore, Vlamings et al. (2008) were able to show that children diagnosed with ASD showed a smaller ERN and Pe and reduced PES.

DISCUSSION: ERROR AWARENESS AND PSYCHIATRIC ILLNESS

Direct studies of error awareness in psychiatric patients are rare. There is indirect evidence for altered error awareness in schizophrenia, ADHD, and autism based on ERN and Pe amplitudes. In schizophrenia patients, the ERN seems to be diminished, but the Pe seems to be unaffected. Given that reduced ERN amplitudes have been associated with impaired error awareness (Wessel, 2012), but perhaps only when accompanied by reduced Pe amplitudes, it remains unclear whether error awareness is attenuated in schizophrenia. In contrast, in ADHD the Pe is reduced, whereas the ERN seems to be unaffected. Based on the results by Murphy et al. (2012), showing that Pe latency and amplitude predicts error awareness, this suggests that error awareness is compromised in ADHD. In ASD both the ERN and Pe are diminished, suggesting that error awareness is reduced in this disease. However, the conclusions on error awareness in psychiatric disorders are based on a very small number of studies, and further studies investigating error awareness directly with proper error awareness tasks are needed. Since the insular cortex has often been associated with error awareness, we reported evidence for changes in cytoarchitectonics, cortical volume, and functional connectivity in the insula of schizophrenia or ASD patients. Furthermore, a recent review by Menon and Uddin (2010) suggested a prominent role for the anterior insula acting together with the aMCC as a key player within a so-called salience network which identifies behaviorally relevant internal and external stimuli. The anterior insula is thereby not only thought to detect salient stimuli but also to initiate switches between other major brain networks (i.e., default mode network and central-executive network), modulate autonomic activity (via interplay with the posterior insula) and finally having direct access to the motor system via the aMCC. The authors conclude that alterations in the integrity of this salience network might underlie the symptomatology of different psychiatric disorders. Taken together it is not only the structural and functional integrity of the insular cortex per se that seems necessary for mental health but also intact structural and functional connectivity between insular cortex and other brain areas is needed for intact cognitive functioning. Whether structural or functional changes of insular cortex are also directly responsible for the alterations in electrophysiological indices of error awareness observed in patients suffering from either disease remains, however, speculative.

CONCLUSIONS

Deficits in performance monitoring in general, and error awareness in particular, might result from different pathological changes in the brain. The anterior insula has been discussed as part of an attentional network, and activity in this part of the insula is related to error awareness, whereas more posterior insula areas represent sensorimotor processes. AHP has been described as deficit in the re-representation of sensorimotor processes or as disorder of awareness for motor deficits and can be observed after posterior insular lesions. The anterior and posterior parts of the insula are highly interconnected. Thus, the insular cortex could be a structural link between error awareness and awareness of deficits or changes due to neurological or psychiatric diseases. Craig (2011) suggested that the integration of interoceptive and exteroceptive information takes place in the insula. Similarly, Kurth et al. (2010, p. 519) described the integration of "[. . .]different qualities into a coherent experience of the world[. . .]" as one potential role of the insula. This integration might be disturbed in anosognosia and insight deficits. Although evidence is rather indirect yet, error awareness seems to be attenuated in schizophrenia, ADHD, and autism. Further studies are needed with respect to the underlying

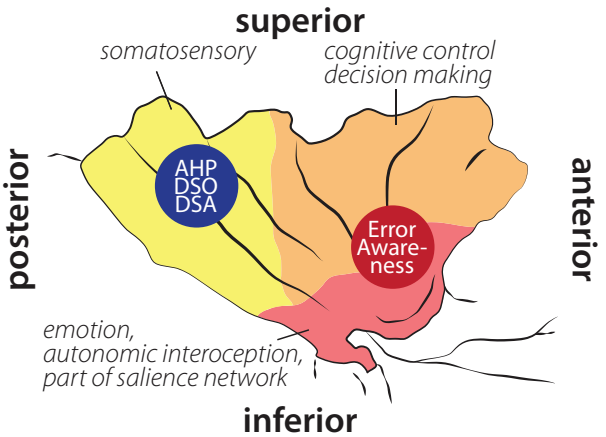


FIGURE 2 | Schematic illustration of insular cortex involvement in error awareness, anosognosia for hemiplegia (AHP), disturbed sense of ownership (DSO) and disturbed sense of agency (DSA) overlaid on schematic drawing of functional areas within the insula according to Deen et al. (2011). Localization of AHP, DSO, and DSA based on Karnath et al. (2005); Baier and Karnath (2008); Karnath and Baier (2010). Localization of error awareness based on Klein et al. (2007a); Hester et al. (2009, 2012); Orr and Hester (2012).

neurotransmitter systems involved in error awareness, but preliminary evidence indicates a prominent role of the dopaminergic system.

In sum, the insula appears to receive and process information on surprising and unwanted states. The anterior insula is involved in (potential) problems with action performance, such as errors, unexpected outcomes (Wessel et al., 2012), or increased necessity of effort (Croxson et al., 2009; Prevost et al., 2010). The posterior insula seems more involved in integrating somatosensory input (see **Figure 2**). Hemiplegia results in unusual and erroneous somatosensory and proprioceptive feedback. An integrating feature of insular activity is that, if evidence for salient action course or salient body perception is high, it becomes active, is involved in the orienting response, and consequently in awareness. The finding that in AHP posterior insular cortex is affected supports this view—the salience of missing or distorted feedback from the hemiplegic limbs is not detected and processed appropriately. It remains to be tested, however, whether anterior insular lesions impair error awareness.

ACKNOWLEDGMENTS

We thank Heike Schmidt-Duderstedt for graphical support. Supported by the federal state of Saxony-Anhalt and the "European Regional Developement Fund" (ERDF 2007–2013), Vorhaben: Center for Behavioral Brain Sciences (CBBS) FKZ: 1211080005.

REFERENCES

Adamantidis, A. R., Zhang, F., Aravanis, A. M., Deisseroth, K., and De Lecea, L. (2007). Neural substrates of awakening probed with optogenetic control of hypocretin neurons. *Nature* 450, 420–424.

Allman, J. M., Tetreault, N. A., Hakeem, A. Y., Manaye, K. F., Semendeferi, K., Erwin, J. M., et al. (2010). The von Economo neurons in frontoinsular and anterior cingulate cortex in great apes and humans. *Brain Struct. Funct.* 214, 495–517.

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.

Augustine, J. R. (1996). Circuitry and functional aspects of the insular lobe in primates including humans. *Brain Res. Brain Res. Rev.* 22, 229–244.

Baier, B., and Karnath, H.-O. (2008). Tight link between our sense of limb ownership and self-awareness of actions. *Stroke* 39, 486–488.

Berti, A., Bottini, G., Gandola, M., Pia, L., Smania, N., Stracciari, A., et al. (2005). Shared cortical anatomy for motor awareness and motor control. *Science* 309, 488–491.

Bilder, R. M., Volavka, J., Lachman, H. M., and Grace, A. A. (2004). The catechol-O-methyltransferase polymorphism: relations to the tonic-phasic dopamine hypothesis and neuropsychiatric phenotypes. *Neuropsychopharmacology* 29, 1943–1961.

Cauda, F., D'Agata, F., Sacco, K., Duca, S., Geminiani, G., and Vercelli, A. (2011). Functional connectivity of the insula in the resting brain. *Neuroimage* 55, 8–23.

Cerliani, L., Thomas, R. M., Jbabdi, S., Siero, J. C., Nanetti, L., Crippa, A., et al. (2012). Probabilistic tractography recovers a rostrocaudal trajectory of connectivity variability in the human insular cortex. *Hum. Brain Mapp.* 33, 2005–2034.

Chee, M. W., Tan, J. C., Zheng, H., Parimal, S., Weissman, D. H., Zagorodnov, V., et al. (2008). Lapsing during sleep deprivation is associated with distributed changes in brain activation. *J. Neurosci.* 28, 5519–5528.

Cohen, M. X., Van Gaal, S., Ridderinkhof, K. R., and Lamme, V. A. (2009). Unconscious errors enhance prefrontal-occipital oscillatory synchrony. *Front. Hum. Neurosci.* 3:54. doi: 10.3389/neuro.09.054.2009

Cools, R. (2011). Dopaminergic control of the striatum for high-level cognition. *Curr. Opin. Neurobiol.* 21, 402–407.

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

Craig, A. D. (2010). "The insular cortex and subjective awareness," in *The Study of Anosognosia*, ed G. P. Prigatano (New York, NY: Oxford University Press), 63–88.

Craig, A. D. (2011). Significance of the insula for the evolution of human awareness of feelings from the body. *Ann. N.Y. Acad. Sci.* 1225, 72–82.

Crespo-Facorro, B., Kim, J., Andreasen, N. C., O'Leary, D. S., Bockholt, H. J., and Magnotta, V. (2000). Insular cortex abnormalities in schizophrenia: a structural magnetic resonance imaging study of first-episode patients. *Schizophr. Res.* 46, 35–43.

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

Croxson, P. L., Walton, M. E., O'Reilly, J. X., Behrens, T. E., Rushworth, M. F. (2009). Effort-based cost-benefit valuation and the human brain. *J. Neurosci.* 29, 4531–4541.

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

De Bruijn, E. R., Sabbe, B. G., Hulstijn, W., Ruigt, G. S., and Verkes, R. J. (2006). Effects of antipsychotic and antidepressant drugs on action monitoring in healthy volunteers. *Brain Res.* 1105, 122–129.

De Bruijn, E. R. A., Mars, R. B., and Hulstijn, W. (2004). "It wasn't me …or was it? How false feedback affects performance," in *Errors, Conflicts, and the Brain. Current Opinions on Performance Monitoring*, eds M. Ullsperger and M. Falkenstein (Leipzig: MPI for Human Cognitive and Brain Sciences), 118–124.

Deen, B., Pitskel, N. B., and Pelphrey, K. A. (2011). Three systems of insular functional connectivity identified with cluster analysis. *Cereb. Cortex* 21, 1498–1506.

Desmurget, M., and Grafton, S. (2000). Forward modeling allows feedback control for fast reaching movements. *Trends Cogn. Sci.* 4, 423–431.

De Witte, L., Brouns, R., Kavadias, D., Engelborghs, S., De Deyn, P. P., and Marien, P. (2011). Cognitive, affective and behavioural disturbances following vascular thalamic lesions: a review. *Cortex* 47, 273–319.

Di Martino, A., Ross, K., Uddin, L. Q., Sklar, A. B., Castellanos, F. X., Milham, M. P. (2009). Functional brain correlates of social and nonsocial processes in autism spectrum disorders: an activation likelihood estimation meta-analysis. *Biol. Psychiatry* 65, 63–74.

Dosenbach, N. U., Fair, D. A., Miezin, F. M., Cohen, A. L., Wenger, K. K., Dosenbach, R. A., et al. (2007). Distinct brain networks for adaptive and stable task control in humans. *Proc. Natl. Acad. Sci. U.S.A.* 104, 11073–11078.

Ebisch, S. J., Gallese, V., Willems, R. M., Mantini, D., Groen, W. B., Romani, G. L., et al. (2011). Altered intrinsic functional connectivity of anterior and posterior insula regions in high-functioning participants with autism spectrum disorder. *Hum. Brain Mapp.* 32, 1013–1028.

Egan, M. F., Goldberg, T. E., Kolachana, B. S., Callicott, J. H., Mazzanti, C. M., Straub, R. E., et al. (2001). Effect of COMT Val108/158 Met genotype on frontal lobe function and risk for schizophrenia. *Proc. Natl. Acad. Sci. U.S.A.* 98, 6917–6922.

Eichele, T., Debener, S., Calhoun, V. D., Specht, K., Engel, A. K., Hugdahl, K., et al. (2008). Prediction of human errors by maladaptive changes in event-related brain networks. *Proc. Natl. Acad. Sci. U.S.A.* 105, 6173–6178.

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

Evrard, H. C., Forro, T., and Logothetis, N. K. (2012). Von Economo neurons in the anterior insula of the macaque monkey. *Neuron* 74, 482–489.

Falkenstein, M., Hohnsbein, J., Hoormann, J., and Blanke, L. (1990). "Effects of errors in choice reaction tasks on the ERP under focused and divided attention," in *Psychophysiological Brain Research*, eds C. H. M. Brunia, A. W. K. Gaillard, and A. Kok (Tilburg: Tilburg University Press), 192–195.

Ford, J. M., Mathalon, D. H., Heinks, T., Kalba, S., Faustman, W.

O., and Roth, W. T. (2001). Neurophysiological evidence of corollary discharge dysfunction in schizophrenia. *Am. J. Psychiatry* 158, 2069–2071.

Forman, S. D., Dougherty, G. G., Casey, B. J., Siegle, G. J., Braver, T. S., Barch, D. M., et al. (2004). Opiate addicts lack error-dependent activation of rostral anterior cingulate. *Biol. Psychiatry* 55, 531–537.

Frank, M. J., D'Lauro, C., and Curran, T. (2007). Cross-task individual differences in error processing: neural, electrophysiological, and genetic components. *Cogn. Affect. Behav. Neurosci.* 7, 297–308.

Friston, K. J., Daunizeau, J., Kilner, J., and Kiebel, S. J. (2010). Action and behavior: a free-energy formulation. *Biol. Cybern.* 102, 227–260.

Frith, C. D. (1987). The positive and negative symptoms of schizophrenia reflect impairments in the perception and initiation of action. *Psychol. Med.* 17, 631–648.

Garavan, H. (2010). Insula and drug cravings. *Brain Struct. Funct.* 214, 593–601.

Garavan, H., Kaufman, J. N., and Hester, R. (2008). Acute effects of cocaine on the neurobiology of cognitive control. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 3267–3276.

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.

Gentsch, A., Ullsperger, P., and Ullsperger, M. (2009). Dissociable medial frontal negativities from a common monitoring system for self- and externally caused failure of goal achievement. *Neuroimage* 47, 2023–2030.

Glahn, D. C., Laird, A. R., Ellison-Wright, I., Thelen, S. M., Robinson, J. L., Lancaster, J. L., et al. (2008). Meta-analysis of gray matter anomalies in schizophrenia: application of anatomic likelihood estimation and network analysis. *Biol. Psychiatry* 64, 774–781.

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., Nandam, L. S., O'Connell, R. G., Wagner, J., Strudwick, M., Nathan, P. J., et al. (2012). Neurochemical enhancement of conscious error awareness. *J. Neurosci.* 32, 2619–2627.

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

Hester, R., Simoes-Franklin, C., and Garavan, H. (2007). Post-error behavior in active cocaine users: poor awareness of errors in the presence of intact performance adjustments. *Neuropsychopharmacology* 32, 1974–1984.

Jakob, H., and Beckmann, H. (1986). Prenatal developmental disturbances in the limbic allocortex in schizophrenics. *J. Neural Transm.* 65, 303–326.

Jehkonen, M., Laihosalo, M., and Kettunen, J. (2006). Anosognosia after stroke: assessment, occurrence, subtypes and impact on functional outcome reviewed. *Acta Neurol. Scand.* 114, 293–306.

Jocham, G., Lauber, A. C., Muller, C. P., Huston, J. P., and De Souza Silva, M. A. (2007). Neurokinin 3 receptor activation potentiates the psychomotor and nucleus accumbens dopamine response to cocaine, but not its place conditioning effects. *Eur. J. Neurosci.* 25, 2457–2472.

Jocham, G., and Ullsperger, M. (2009). Neuropharmacology of performance monitoring. *Neurosci. Biobehav. Rev.* 33, 48–60.

Karnath, H. O., and Baier, B. (2010). Right insula for our sense of limb ownership and self-awareness of actions. *Brain Struct. Funct.* 214, 411–417.

Karnath, H. O., Baier, B., and Nagele, T. (2005). Awareness of the functioning of one's own limbs mediated by the insular cortex? *J. Neurosci.* 25, 7134–7138.

Kasai, K., Shenton, M. E., Salisbury, D. F., Onitsuka, T., Toner, S. K., Yurgelun-Todd, D., et al. (2003). Differences and similarities in insular and temporal pole MRI gray matter volume abnormalities in first-episode schizophrenia and affective psychosis. *Arch. Gen. Psychiatry* 60, 1069–1077.

Kaufman, J. N., Ross, T. J., Stein, E. A., and Garavan, H. (2003). Cingulate hypoactivity in cocaine users during a GO-NOGO task as revealed by event-related functional magnetic resonance imaging. *J. Neurosci.* 23, 7839–7843.

Kelly, C., Toro, R., Di Martino, A., Cox, C. L., Bellec, P., Castellanos, F. X., et al. (2012). A convergent functional architecture of the insula emerges across imaging modalities. *Neuroimage* 61, 1129–1142.

Kenny, P. J. (2011). Tobacco dependence, the insular cortex and the hypocretin connection. *Pharmacol. Biochem. Behav.* 97, 700–707.

Kim, J. J., Youn, T., Lee, J. M., Kim, I. Y., Kim, S. I., and Kwon, J. S. (2003). Morphometric abnormality of the insula in schizophrenia: a comparison with obsessive-compulsive disorder and normal control using MRI. *Schizophr. Res.* 60, 191–198.

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

Klein, T. A., Neumann, J., Reuter, M., Hennig, J., Von Cramon, D. Y., and Ullsperger, M. (2007b). Genetically determined differences in learning from errors. *Science* 318, 1642–1645.

Kuhar, M. J., Ritz, M. C., and Boja, J. W. (1991). The dopamine hypothesis of the reinforcing properties of cocaine. *Trends Neurosci.* 14, 299–302.

Kurth, F., Zilles, K., Fox, P. T., Laird, A. R., and Eickhoff, S. B. (2010). A link between the systems: functional differentiation and integration within the human insula revealed by meta-analysis. *Brain Struct. Funct.* 214, 519–534.

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

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.

Mathalon, D. H., Fedor, M., Faustman, W. O., Gray, M., Askari, N., and Ford, J. M. (2002). Response-monitoring dysfunction in schizophrenia: an event-related brain potential study. *J. Abnorm. Psychol.* 111, 22–41.

Menon, V., and Uddin, L. Q. (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Struct. Funct.* 214, 655–667.

Mesulam, M. M., and Mufson, E. J. (1982a). Insula of the old world monkey. I. Architectonics in the insulo-orbito-temporal component of the paralimbic brain. *J. Comp. Neurol.* 212, 1–22.

Mesulam, M. M., and Mufson, E. J. (1982b). Insula of the old world monkey. III: efferent cortical output and comments on function. *J. Comp. Neurol.* 212, 38–52.

Modirrousta, M., and Fellows, L. K. (2008). Dorsal medial prefrontal cortex plays a necessary role in rapid error prediction in humans. *J. Neurosci.* 28, 14000–14005.

Morecraft, R. J., Stilwell-Morecraft, K. S., Cipolloni, P. B., Ge, J., McNeal, D. W., and Pandya, D. N. (2012). Cytoarchitecture and cortical connections of the anterior cingulate and adjacent somatomotor fields in the rhesus monkey. *Brain Res. Bull.* 87, 457–497.

Mufson, E. J., and Mesulam, M. M. (1982). Insula of the old world monkey. II: afferent cortical input and comments on the claustrum. *J. Comp. Neurol.* 212, 23–37.

Murphy, P. R., Robertson, I. H., Allen, D., Hester, R., and O'Connell, R. G. (2012). An electrophysiological signal that precisely tracks the emergence of error awareness. *Front. Hum. Neurosci.* 6:65. doi: 10.3389/fnhum.2012.00065

Mutschler, I., Wieckhorst, B., Kowalevski, S., Derix, J., Wentlandt, J., Schulze-Bonhage, A., et al. (2009). Functional organization of the human anterior insular cortex. *Neurosci. Lett.* 457, 66–70.

Nieuwenhuis, S., Aston-Jones, G., and Cohen, J. D. (2005). Decision making, the P3, and the locus coeruleus-norepinephrine system. *Psychol. Bull.* 131, 510–532.

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.

Notebaert, W., Houtman, F., Opstal, F. V., Gevers, W., Fias, W., and Verguts, T. (2009). Post-error slowing: an orienting account. *Cognition* 111, 275–279.

O'Connell, R. G., Bellgrove, M. A., Dockree, P. M., Lau, A., Hester, R., Garavan, H., et al. (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., et al. (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.

O'Keeffe, F. M., Dockree, P. M., and Robertson, I. H. (2004). Poor insight in traumatic brain injury

mediated by impaired error processing? Evidence from electrodermal activity. *Brain Res. Cogn. Brain Res.* 22, 101–112.

Oppenheimer, S. M., Gelb, A., Girvin, J. P., and Hachinski, V. C. (1992). Cardiovascular effects of human insular cortex stimulation. *Neurology* 42, 1727–1732.

Orr, C., and Hester, R. (2012). Error-related anterior cingulate cortex activity and the prediction of conscious error awareness. *Front. Hum. Neurosci.* 6:177. doi: 10.3389/fnhum.2012.00177

Palaniyappan, L., Mallikarjun, P., Joseph, V., and Liddle, P. F. (2011). Appreciating symptoms and deficits in schizophrenia: right posterior insula and poor insight. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 35, 523–527.

Paulus, M. P., Tapert, S. F., and Schulteis, G. (2009). The role of interoception and alliesthesia in addiction. *Pharmacol. Biochem. Behav.* 94, 1–7.

Peterburs, J., Pergola, G., Koch, B., Schwarz, M., Hoffmann, K. P., Daum, I., et al. (2011). Altered error processing following vascular thalamic damage: evidence from an antisaccade task. *PLoS ONE* 6:e21517. doi: 10.1371/journal.pone.0021517

Pia, L., and Tamietto, M. (2006). Unawareness in schizophrenia: neuropsychological and neuroanatomical findings. *Psychiatry Clin. Neurosci.* 60, 531–537.

Prevost, C., Pessiglione, M., Metereau, E., Clery-Melin, M. L., Dreher, J. C. (2010). Separate valuation subsystems for delay and effort decision costs. *J. Neurosci.* 30, 14080–14090.

Prigatano, G. P. (2009). Anosognosia: clinical and ethical considerations. *Curr. Opin. Neurol.* 22, 606–611.

Prigatano, G. P. (2010). "Historical observations relevant to the study of anosognosia," in *The Study of Anosognosia*, ed G. P. Prigatano (New York, NY: Oxford University Press), 3–16.

Prigatano, G. P., and Shacter, D. L. (1991). *Awareness of Deficit after Brain Injury: Clinical and Theoretical Issues.* New York, NY: Oxford University Press.

Rabbitt, P. M. (1968). Three kinds of error-signalling responses in a serial choice task. *Q. J. Exp. Psychol.* 20, 179–188.

Reason, J. (1990). *Human Error.* Cambridge: Cambidge University Press.

Ridderinkhof, K. R., De Vlugt, Y., Bramlage, A., Spaan, M., Elton, M., Snel, J., et al. (2002). Alcohol consumption impairs detection of performance errors in mediofrontal cortex. *Science* 298, 2209–2211.

Santos, M., Uppal, N., Butti, C., Wicinski, B., Schmeidler, J., Giannakopoulos, P., et al. (2011) von Economo neurons in autism: a stereologic study of the frontoinsular cortex in children. *Brain Res.* 1380, 206–217.

Schachar, R. J., Chen, S., Logan, G. D., Ornstein, T. J., Crosbie, J., Ickowicz, A., et al. (2004). Evidence for an error monitoring deficit in attention deficit hyperactivity disorder. *J. Abnorm. Child Psychol.* 32, 285–293.

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.

Scheffers, M. K., Humphrey, D. G., Stanny, R. R., Kramer, A. F., and Coles, M. G. (1999). Error-related processing during a period of extended wakefulness. *Psychophysiology* 36, 149–157.

Seeley, W. W., Merkle, F. T., Gaus, S. E., Craig, A. D., Allman, J. M., Hof, P. R., et al. (2012). Distinctive neurons of the anterior cingulate and frontoinsular cortex: a historical perspective. *Cereb. Cortex* 22, 245–250.

Seifert, S., Von Cramon, D. Y., Imperati, D., Tittgemeyer, M., and Ullsperger, M. (2011). Thalamocingulate interactions in performance monitoring. *J. Neurosci.* 31, 3375–3383.

Seth, A. K., Suzuki, K., and Critchley, H. D. (2011). An interoceptive predictive coding model of conscious presence. *Front. Psychol.* 2:395. doi: 10.3389/fpsyg.2011.00395

Shalgi, S., and Deouell, L. Y. (2012). Is any awareness necessary for an Ne? *Front. Hum. Neurosci.* 6:124. doi: 10.3389/fnhum.2012.00124

Shalgi, S., O'Connell, R. G., Deouell, L. Y., and Robertson, I. H. (2007). Absent minded but accurate: delaying responses increases accuracy but decreases error awareness. *Exp. Brain Res.* 182, 119–124.

Simmons, W. K., Avery, J. A., Barcalow, J. C., Bodurka, J., Drevets, W. C., and Bellgowan, P. (2012). Keeping the body in mind: Insula functional organization and functional connectivity integrate interoceptive, exteroceptive, and emotional awareness. *Hum. Brain Mapp.* doi: 10.1002/hbm.22113. [Epub ahead of print].

Sokhadze, E., Baruth, J., El-Baz, A., Horrell, T., Sokhadze, G., Carroll, T., et al. (2010). Impaired error monitoring and correction function in autism. *J. Neurother.* 14, 79–95.

Spinazzola, L., Pia, L., Folegatti, A., Marchetti, C., and Berti, A. (2008). Modular structure of awareness for sensorimotor disorders: evidence from anosognosia for hemiplegia and anosognosia for hemianaesthesia. *Neuropsychologia* 46, 915–926.

Sridharan, D., Levitin, D. J., and Menon, V. (2008). A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *Proc. Natl. Acad. Sci. U.S.A.* 105, 12569–12574.

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

Swick, D., and Turken, A. U. (2004). "Focusing on the anterior cingulate cortex: effects of focal lesions on cognitive performance," in *Cognitive Neuroscience of Attention*, ed M. I. Posner (New York, NY: Guilford Press), 393–406.

Ullsperger, M. (2006). Performance monitoring in neurological and psychiatric patients. *Int. J. Psychophysiol.* 59, 59–69.

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. (2006). How does error correction differ from error signaling? An event-related potential study. *Brain Res.* 1105, 102–109.

Vlamings, P. H., Jonkman, L. M., Hoeksma, M. R., Van Engeland, H., and Kemner, C. (2008). Reduced error monitoring in children with autism spectrum disorder: an ERP study. *Eur. J. Neurosci.* 28, 399–406.

Vocat, R., Staub, F., Stroppini, T., and Vuilleumier, P. (2010). Anosognosia for hemiplegia: a clinical-anatomical prospective study. *Brain* 133, 3578–3597.

Vocat, R., and Vuilleumier, P. (2010). "Neuroanatomy of impaired body awareness in anosognosia and hysteria: a multicomponent account," in *The Study of Anosognosia*, ed G. P. Prigatano (Oxford: Oxford University Press), 359–406.

Vogt, B. A. (2005). Pain and emotion interactions in subregions of the cingulate gyrus. *Nat. Rev. Neurosci.* 6, 533–544.

Volkow, N. D., Fowler, J. S., Wolf, A. P., Schlyer, D., Shiue, C. Y., Alpert, R., et al. (1990). Effects of chronic cocaine abuse on postsynaptic dopamine receptors. *Am. J. Psychiatry* 147, 719–724.

Von Economo, C. (1926). Eine neue art spezialzellen des lobus cinguli und lobus insulae. *Zeitschrift für die gesamte Neurologie und Psychiatrie* 100, 706–712.

Von Economo, C., and Koskinas, G. N. (1925). *Die Cytoarchitektonik der Hirnrinde des erwachsenen Menschen.* Wien: Springer.

Vuilleumier, P. (2004). Anosognosia: the neurology of beliefs and uncertainties. *Cortex* 40, 9–17.

Wessel, J. R. (2012). Error awareness and the error-related negativity: evaluating the first decade of evidence. *Front. Hum. Neurosci.* 6:88. doi: 10.3389/fnhum.2012.00088

Wessel, J. R., Danielmeier, C., Morton, J. B., and Ullsperger, M. (2012). Surprise and error: common neuronal architecture for the processing of errors and novelty. *J. Neurosci.* 32, 7528–7537.

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.

Wiersema, J. R., Van Der Meere, J. J., and Roeyers, H. (2005). ERP correlates of impaired error monitoring in children with ADHD. *J. Neural Transm.* 112, 1417–1430.

Wiersema, J. R., Van Der Meere, J. J., and Roeyers, H. (2009). ERP correlates of error monitoring in adult ADHD. *J. Neural Transm.* 116, 371–379.

Williamson, P. C., and Allman, J. M. (2012). A framework for interpreting functional networks in schizophrenia. *Front. Hum. Neurosci.* 6:184. doi: 10.3389/fnhum.2012.00184

Willmott, C., and Ponsford, J. (2009). Efficacy of methylphenidate in the rehabilitation of attention following traumatic brain injury: a randomised, crossover, double blind, placebo controlled inpatient trial. *J. Neurol. Neurosurg. Psychiatry* 80, 552–557.

Wylie, K. P., and Tregellas, J. R. (2010). The role of the insula in schizophrenia. *Schizophr. Res.* 123, 93–104.

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

Zhang, J. S., Wang, Y., Cai, R. G., and Yan, C. H. (2009). The brain

regulation mechanism of error monitoring in impulsive children with ADHD—an analysis of error related potentials. *Neurosci. Lett.* 460, 11–15.

Zirnheld, P. J., Carroll, C. A., Kieffaber, P. D., O'Donnell, B. F., Shekhar, A., and Hetrick, W. P. (2004). Haloperidol impairs learning and error-related negativity in humans. *J. Cogn. Neurosci.* 16, 1098–1112.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 05 September 2012; accepted: 14 January 2013; published online: 04 February 2013.

Citation: Klein TA, Ullsperger M and Danielmeier C (2013) Error awareness and the insula: links to neurological and psychiatric diseases. Front. Hum. Neurosci. 7:14. doi: 10.3389/fnhum.2013.00014

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 29 February 2012
doi: 10.3389/fnhum.2012.00032

Increased orienting to unexpected action outcomes in schizophrenia

Elena Núñez Castellar[1], Femke Houtman[1], Wim Gevers[2], Manuel Morrens[3,4], Sara Vermeylen[3,5], Bernard Sabbe[3,5] and Wim Notebaert[1]*

[1] *Department of Experimental Psychology, Ghent University, Ghent, Belgium*
[2] *Unité de Recherché en Neurosciences Cognitives, Université Libre de Bruxelles, Brussels, Belgium*
[3] *Collaborative Antwerp Psychiatric Research Institute, University of Antwerp, Antwerp, Belgium*
[4] *Psychiatric Centre "BroedersAlexianen," Boechout, Belgium*
[5] *Psychiatric Centre "St. Norbertushuis," Duffel, Belgium*

Edited by:
Claudia Danielmeier, Radboud University, Netherlands

Reviewed by:
Jan R. Wessel, University of California, San Diego, USA
Frederik Van Der Veen, Erasmus University, Netherlands

****Correspondence:***
Wim Notebaert, Department of Experimental Psychology, Ghent University, 9000 Ghent, Belgium.
e-mail: wim.notebaert@ugent.be

Although some recent research has indicated reduced performance monitoring in patients with schizophrenia, the literature on this topic shows some remarkable inconsistencies. While most studies suggest diminished error signals following error responses, some studies reported normal post-error slowing, while others reported reduced post-error slowing. Here we review these studies and highlight the most important discrepancies. Furthermore, we argue that overall error rates are a mostly neglected issue that can at least partly explain these discrepancies. It has been reported previously that post-error slowing depends on the error rates. Participants or patients that make more errors are likely to show decreased post-error slowing. Therefore, when a group of patients is compared to a group of controls, it is extremely important to match error rates. For this purpose, we developed a procedure where we matched individuals' error rates. In a task where subjects had to press a response key corresponding to one of four colors we manipulated the difficulty on an individual basis by varying the discriminability between the colors. Schizophrenic patients and a group of controls were tested with this procedure showing that differences in accuracy disappear. Interestingly, we can see that in patients, the color values that were needed to reach similar levels of accuracy correlate with the Positive and Negative Syndrome Scale (PANSS) scale, with higher PANSS requiring more color. Most important, we found that schizophrenic patients have increased rather than decreased post-error slowing when the inter-trial interval (ITI) is short. This result can be interpreted within the framework of the orienting account, as it has been demonstrated previously that schizophrenic patients show increased distractibility.

Keywords: post-error slowing, outcome expectancy, schizophrenia, orienting account

INTRODUCTION

Impairments in monitoring and regulation of self-generated thoughts and behavior are considered as a hallmark in schizophrenia (Frith and Done, 1989; Leudar et al., 1994; Mlakar et al., 1994; Johns et al., 2001). The idea that these impairments could be related to deficits in the ability to monitor error responses and to make subsequent adjustments of behavior (Frith and Done, 1989; Mlakar et al., 1994) has received support during the last years. Studies performed with functional neuroimaging and electrophysiological methods have confirmed the existence of functional abnormality in error processing in schizophrenia. Specifically, functional neuroimaging studies have shown reduced error-related activation in the anterior cingulate cortex (ACC), a brain area considered to play a critical role in performance monitoring (Carter et al., 2001; Laurens et al., 2003; Kerns et al., 2005; Polli et al., 2008). Also, in line with these findings, event-related potential (ERP) studies have demonstrated that the amplitude of the error-related negativity (ERN), an ERP component generated in the ACC and observed as a negative deflection maximal at 50–150 ms following an erroneous response, is also disturbed (Kopp and Rist, 1999; Alain et al., 2002; Mathalon et al., 2002; Bates et al., 2004). Specifically, it has not only been shown that the ERN amplitude is smaller in patients with schizophrenia when compared to controls but also that the ERN on correct trials is abnormally large in patients when compared to healthy subjects (Kopp and Rist, 1999; Alain et al., 2002; Mathalon et al., 2002).

Although this evidence suggests a clear deficit in performance monitoring in schizophrenia, studies that investigated behavioral adjustments following errors do not allow for a clear-cut conclusion. For instance, deficits in immediate error corrections were described in early behavioral studies (Malenka et al., 1982, 1986) but most recent studies observed intact error corrections in patients (Kopp and Rist, 1994, 1999; Polli et al., 2006, 2008; but for an exception see Turken et al., 2003). Likewise, some studies investigating behavioral slowing after errors (post-error slowing) reported a reduction or absence of post-error slowing in patients (Carter et al., 2001; Alain et al., 2002; Kerns et al., 2005) while others have shown that schizophrenic patients just as healthy individuals show post-error slowing (Kopp and Rist, 1999; Mathalon et al., 2002; Laurens et al., 2003; Polli et al., 2006, 2008).

Most interesting, studies in which the neural response and behavioral error-related adjustments have been simultaneously investigated, all have observed diminished error-related activity but intact post-error slowing in patients with schizophrenia (Kopp and Rist, 1999; Mathalon et al., 2002; Laurens et al., 2003; Polli et al., 2008). These findings might suggest that error monitoring mechanisms and the mechanisms involved in implementing subsequent adjustments in behavior dissociate.

The aim of the present study is to further investigate behavioral changes following errors in patients compared to control subjects, when controlling for some methodological issues that have been largely ignored by previous research. We will specifically focus on post-error slowing, which is a behavioral adaptation effect that has been broadly investigated in healthy subjects (e.g., Rabbitt, 1966; Laming, 1968, 1979; Rabbitt and Rodgers, 1977; Gehring et al., 1993; Hajcak et al., 2003; Hajcak and Simons, 2008)

Recently, it has been suggested that the post-error slowing is related to the relative infrequency of errors which causes attentional capture that in turn delays processing of the following stimulus (Notebaert et al., 2009; Núñez Castellar et al., 2010). This idea is based on both behavioral and electrophysiological evidence. Behaviorally, it has been demonstrated that when the expectancy of error and correct responses is manipulated, the post-error slowing depends on the frequency of errors; when errors are more frequent than correct responses post-*correct* slowing instead of post-*error* slowing can be observed. This association has been confirmed by means of ERPs when investigating the neurophysiological correlates of the post-error and post-correct slowing. The results have shown that the P3, a component that has been associated with attentional orienting to unexpected events (see for a review Friedman et al., 2001) is correlated with post-error slowing. This suggests that post-error slowing is driven by attentional mechanisms elicited by the unexpected nature of the error (Notebaert et al., 2009), although this interpretation is still debated considering that several studies have reported the ERN amplitude in a single trial basis to be a predictor of the post-error slowing (Gehring et al., 1993; Debener et al., 2005)

Interestingly, schizophrenic patients show an increased vulnerability to distraction by novel stimuli (Grillon et al., 1990; Braff, 1993) and research investigating this deficit has provided convergent evidence suggesting that the reorienting of processing resources to salient novel stimuli is also disturbed (Grillon et al., 1990; Braff, 1993; Gray, 1995; Kapur, 2003). Functional neuroimaging studies recently confirmed that in at least some cerebral areas involved in the processing of salient stimuli, the hemodynamic response elicited by orienting to novel stimuli is greater in patients than in healthy participants (Laurens et al., 2005). Therefore, based on the orienting account for post-error slowing, increased rather than decreased slowing following unexpected outcomes can be predicted in patients when compared to controls. In the present study we investigated this prediction by using the same adaptive four-choice RT task designed to manipulate error rates as in previous studies (see Notebaert et al., 2009 and Núñez Castellar et al., 2010). Given that it has been shown that the slowing is strongly influenced by the error frequency, we leveled out performance in controls and patients by manipulating color discriminability. In this task we control the accuracy by making the color discrimination easier or harder depending on the subjects' performance. Participants performed the task in two conditions: 75%-correct responses (expectancy for correct) and 35%-correct responses (expectancy for error). We expect that patients with higher symptom severity will need more color (higher discriminability) than patients with low symptom severity. This prediction is made based on previous research showing that schizophrenic patients show sensory processing deficits (Slaghuis, 2004; Butler et al., 2005) and diminished performance when compared to controls in color vision tests (Shuwairi et al., 2002).

More important, however, on the basis of the orienting account we expect schizophrenic patients to show increased post-error slowing in the 75% correct condition and post-correct slowing in the 35% correct condition at short inter-trial intervals (ITIs). With longer ITIs, we expect that the effects of attention reorienting will be reduced. Remarkably, this prediction is opposing cognitive control theories stating that decreased ACC activity and decreased ERNs in schizophrenia should result in decreased post-error slowing.

METHOD

PARTICIPANTS

Eighteen healthy adults (eight males) and 18 patients with schizophrenia (14 males) participated in the experiment. All provided informed consent, had the approval of the local ethical committee, and participated in accord with the Declaration of Helsinki. Each had normal or corrected-to-normal vision. Four participants were excluded from analyses because they did not reach with the adaptive program the expected performance levels (performance more than 2, 5 standard deviations away from the group mean). Thus, data of 17 patients and 15 controls are here reported. The three excluded controls performed the task above the expected performance levels.

Controls were medication-free volunteers without a history of psychiatric or neurological illness. Patients were stable, partially remitted, medicated in- and outpatients recruited from St Norbertushuis Duffel psychiatric center. All patients met Diagnostic and Statistical Manual of Mental Disorders (4th ed.; DSM–IV; American Psychological Association, 1994) criteria for schizophrenia, as diagnosed by an institutional psychiatrist. Clinical status was characterized with the Positive and Negative Syndrome Scale (PANSS; Kay et al., 1987). All patients were mildly ill at the time of the testing (group mean PANSS total 65, 2 ± 22, 4). Thirteen patients received stable doses of atypical antipsychotics as their primary medication, two received typical antipsychotics, and two typical antipsychotic adjunctive to the atypical medication. Patients and controls did not differ significantly in age and gender (see **Table 1**).

TASK

An adaptive four-choice RT task was used to manipulate error rates. This task was in principle equal to that published in previous studies (Notebaert et al., 2009; Núñez Castellar et al., 2010) but the response deadline and the amount of trials presented were modified. Stimuli were small colored squares presented centrally

Table 1 | Demographic data by group and rating scale scores for patients.

	Healthy controls	Schizophrenia patients	*t*	*p*
Age	32.1 (8.1)	33.1 (6.5)	0.41	0.68
Sex	6M/9F	13M/4F	−2.06*	0.09
PANSS Positive		13.6 ± 6.5		
PANSS Negative		18.5 ± 6.6		
PANSS general		33.6 ± 10.4		
PANSS total		65.2 ± 22.4		

*Non-parametric Mann–Whitney comparison.

on a white background. The brightness of the colors was adjusted automatically in order to keep every participant's performance at a pre-specified level (35% or 75% accuracy). Colors are described according to the HSV color model with three parameters: hue (0–360), saturation (0–100), and value (0–100). The four colors that were used in the practice trials were red (20, 100, 80), yellow (60, 100, 80), green (120, 100, 80), and blue (240, 100, 80). Participants responded to each of the four colors with one of the four buttons on a response box, using their left and right middle and index fingers. Four different color-to-button mappings were used, and participants were randomly assigned to one of these mappings.

Each trial started with a central fixation cross (500 ms) before the stimulus presentation. Then the stimulus was presented and remained on the screen for a maximum of 500 ms or until a response button was pressed. The response was immediately followed by feedback ("J" for correct and "F" for incorrect, corresponding to the Dutch words "juist" and "fout"). Following the feedback presentation four different ITIs were randomly intermixed (150 ms, 250 ms, 500 ms, and 750 ms). For the data analysis 150 ms, 250 ms were considered as short ITIs and 500 ms and 750 ms as long ITIs. The experiment started with a practice block (40 trials) without a response deadline. Later in a second practice block (40 trials) a response deadline of 2000 ms was introduced simultaneously with a feedback signal: "T", for "too slow" ("te laat," in Dutch). Afterwards two experimental blocks with the same response deadline were presented (200 trials). Each block corresponded to the 35%- and 75%-accuracy rate manipulation. The order in which the blocks were presented was counterbalanced and within each block, participants received a short break. On every trial, the program calculated the accuracy of the last 20 trials and adjusted the color value by 1 value point when accuracy deviated from the specified level (75% or 35%). The color value increased when accuracy was too low and decreased when accuracy was too high to reach the pre-specified accuracy levels (see more details in Notebaert et al., 2009). The experiment lasted about 30 min.

DATA ANALYSES

We used the procedure described by Notebaert et al. (2009) for the data analysis: we excluded trials that occurred before a stable accuracy level was reached. Likewise, trials with RTs faster than 200 ms or slower than 2.000 ms (response deadline) and trials that were preceded by these trials were excluded. In total, 28.1% of the trials were excluded.

The results showed that the adaptation procedure worked as expected for the 75% and 35% correct conditions, with accuracies of 76 % (±4.2%) and 40% (±3.1%) for the patients group, and of 75 % (±1.4%) and 38% (±3.2) for the control group, respectively. A *t*-test showed that these accuracy rates did not differ between patients and controls in the 75% accuracy condition [$t(30) = 1.35, p = 0.18$] or in the 35% accuracy condition [$t(30) = 1.21, p = 0.23$]. The order in which the conditions were administered did not yield significant effects; therefore, we omitted this factor in subsequent analyses.

COLOR NEEDED TO REACH 35% AND 75% ACCURACY

We used the adaptive program in order to obtain equal accuracy levels for patients and controls since we speculated that without this program, patients would make more errors, which would possibly lead to reduced post-error slowing (Notebaert et al., 2009). In order to stress the importance of this manipulation we calculated the amount of color needed (the value) for each participant. Although there was no overall group difference in color needed in both conditions [35: $t(30) = 0.24, p = 0.81$; 75: $t(30) = 0.61, p = 0.54$], the amount of color needed correlated significantly with PANSS scores for the patients (see **Figure 1**). In order to rule out the possibility that results were driven by outliers, outlier analyses were carried out. Outliers were identified by calculated the cook's distance measures for all the correlations reported. Cases with values above the percentile 50 (*F*-distribution) were identified as outliers. After excluding the cases labeled as outliers from the analyses, the significant results reported in the **Figure 1** remained unchanged.

The significant correlation between the color needed for patients and the PANSS scores underlines the importance of the adaptive program since without the program patients with higher PANSS scores would have made more errors than patients with milder symptoms.

POST-ERROR SLOWING

Reaction times for correct and error trials in short and long ITIs for patients and controls in the current trial (n), in the correct trials ($n + 1$) and the post-error slowing are reported in the **Table 2**.

For the analysis of the post-error slowing effect, RTs of correct trials were subjected to an analysis of variance (ANOVA) with Previous trial accuracy (correct vs. error), Accuracy condition (35% correct vs. 75% correct) and ITI (short vs. long) as within-subjects factors and Group (control vs. patients) as between-subjects factor.

The results revealed a significant main effect of Group [$F(1,30) = 14.38, p < 0.001$], indicating increased reaction times for patients when compared to controls (857 ms vs. 700 ms; $p < 0.001$). The main effects of Condition [$F(1,30) = 2.84, p = 0.10$], and Previous trial accuracy [$F(1,30) = 3.7, p = 0.06$] were not significant. The main effect of ITI was also not significant [$F(1,30) = 1.3, p = 0.26$].

The interaction between Condition and Accuracy of the previous trial was significant, [$F(1,14) = 11.69, p < 0.01$], replicating

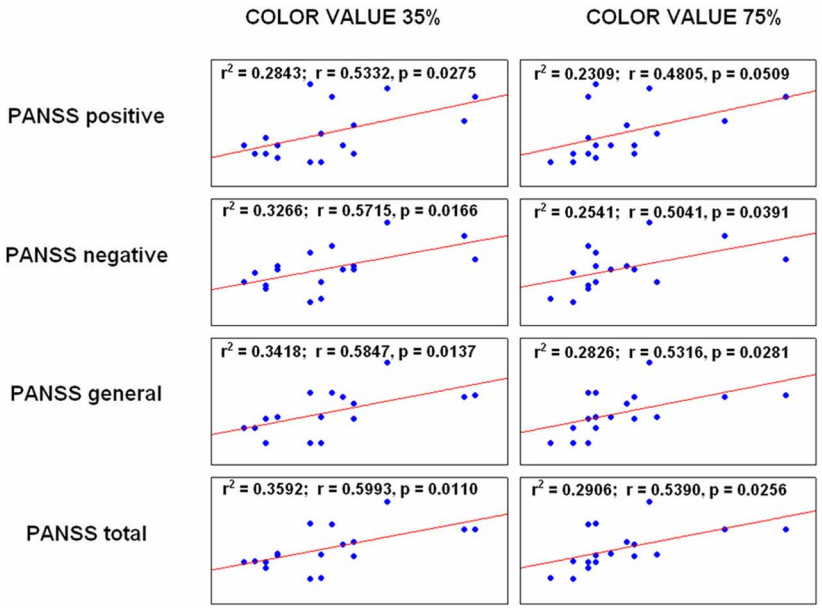


FIGURE 1 | Correlation between color value and PANSS scores.

Table 2 | Reaction times for correct and error trials in short and long ITIs for patients and controls in the current trial (*n*), in the correct trials (*n* + 1) and post-error slowing.

		n		*c*(*n* + 1)		Post-error slowing	
		35%	75%	35%	75%	35%	75%
PATIENTS							
ITI Long	Error	798	932	832	857	−37	−4
	Correct	852	861	869	861		
ITI Short	Error	801	941	791	922	−76	61
	Correct	830	872	868	862		
CONTROLS							
ITI Long	Error	624	764	655	744	−68	17
	Correct	692	730	723	728		
ITI Short	Error	629	770	638	706	−61	−2
	Correct	671	708	699	708		

the study of Notebaert and colleagues (2009). The interaction between Condition, ITI, and Group was significant [$F(1,30) = 4.31$, $p < 0.05$], but more important for the interpretation of the results, the four way interaction between Condition, Previous accuracy, Group, and ITI was significant as well [$F(1,30) = 4.46$, $p < 0.05$]. **Figure 2** shows that patients' slowing depends on the ITI while control participants' data are unaffected by ITI. ANOVAs performed separately for patients and control group confirm this pattern, showing that only for schizophrenic patients there is a significant interaction between Previous accuracy, Condition, and ITI [$F(1,16) = 6.67$, $p < 0.05$] with increased slowing following unexpected outcomes in short ITIs, while for the control group this is not the case [$F(1,14) = 0.27$, $p = 0.59$]. No other interactions were significant in the Four-Way ANOVA.

DISCUSSION

The aim of the present study was to further investigate error-related adjustments in schizophrenia. We tested the main prediction of a recent account that suggests the post-error slowing is driven by attentional mechanisms elicited by the unexpected nature of the error (Notebaert et al., 2009). For that purpose we used an adaptive four-choice reaction time task that allowed us to manipulate error rates. Every participant from the patient and control group performed a 35% and 75 % accuracy condition. Additionally the effect of short and long ITIs on the slowing was investigated.

Based on the orienting account we predicted that patients would show increased slowing at short ITIs considering their deficits in reorienting from the ongoing task and increased distraction by novel stimuli (Grillon et al., 1990; Braff, 1993).

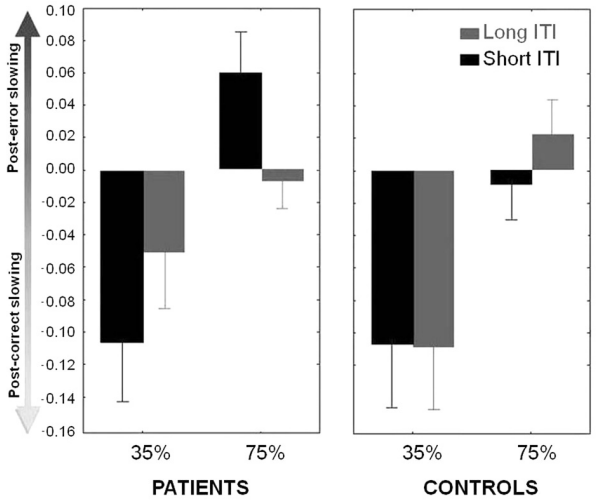


FIGURE 2 | Index of slowing (RTs after errors—RTs after correct responses/Individuals' mean RTs per condition). Positive difference values indicate post-error slowing while negative difference values indicate post-correct slowing. The graph shows increased slowing after unexpected correct and unexpected error responses in short ITIs for patients.

Consistent with the latter prediction, our results showed that in short ITIs, increased orienting to unexpected action outcomes in patients was observed. In other words, the size of the post-error and the post-correct slowing effects was significantly larger in patients than in controls in short ITIs. This finding is against one widely assumption of cognitive control theories, which suggest that the decreased ACC activity and decreased ERNs in schizophrenia that has consistently been reported in the extant literature (Kopp and Rist, 1994; Carter et al., 2001; Alain et al., 2002; Mathalon et al., 2002; Laurens et al., 2003; Bates et al., 2004; Kerns et al., 2005; Polli et al., 2008) should lead to decreased post-error slowing.

We speculate that since patients show deficits in performance monitoring as showed by studies demonstrating that the ERN amplitude is smaller in error trials at the same time that the ERN on correct trials is abnormally large (Kopp and Rist, 1999; Alain et al., 2002; Mathalon et al., 2002), it is likely that external feedback might be crucial for monitoring their performance. We hypothesize that attentional mechanisms might be more sensitive in patients than in controls to the external feedback producing a larger slowing. However, this is an issue that should be further investigated.

There is one aspect of the data that is unexpected, however. The control subjects do not show post-error slowing in both short and long ITIs in the 75% accuracy condition, while they do show post-correct slowing in the 35% condition. The absence of post-error slowing has also been reported in another study (Bates et al., 2004) where matched controls were compared to schizophrenic patients but most of the studies observed post-error slowing in controls (Kopp and Rist, 1999; Mathalon et al., 2002; Polli et al., 2008). On the one hand an important methodological difference between the present study and previous studies that have reported post-error slowing in the 75% accuracy condition (Notebaert et al., 2009; Núñez Castellar et al., 2010), is that the age-matched control subjects were on average 12 years older than the typical student population used in previous studies (e.g., Rabbitt and Rodgers, 1977; Jentzsch and Dudschig, 2009). Future studies should further investigate whether this is a factor that can influence the slowing in reaction times. On the other hand it is possible that in the 75% accuracy condition, errors are not surprising enough for the group of participants of the present study.

An important route for further research is to investigate error monitoring and behavioral adjustments following errors as a function of error type. Previous research has distinguished aware from unaware errors (e.g., Nieuwenhuis et al., 2001; Endrass et al., 2007), but one can also distinguish errors that were caused by perceptual factors (e.g., stimulus degradation), errors caused by short response deadlines forcing participants to give premature responses, errors caused by confusing mapping rules, errors caused by temporary confusion in a relatively simple mapping rule, errors caused by response conflict, and so on. One could argue that slowing after an error only makes sense when the error was caused by premature responding, as this would reduce the possibility of making yet another error. Consequently, within a framework where errors are considered to trigger adaptive behavior, one could explain the lack of post-error slowing in our control group by the fact that errors were not caused by premature responding. However, we like to add that in an identical experiment on a student population, where error RTs were not faster than correct RTs, we observed post-error slowing in this condition (Núñez Castellar et al., 2010). Nevertheless, we acknowledge that manipulating error type could be an important tool for increasing our understanding about error monitoring.

Two methodological aspects dissociate this study from previous ones (Kopp and Rist, 1999; Mathalon et al., 2002; Laurens et al., 2003; Polli et al., 2008). First, we leveled out performance in patients and matched their performance to controls. Overall accuracy levels play a key role in the modulation of post-error slowing in the sense that the more frequent the errors become, the smaller the size of the slowing gets (Notebaert et al., 2009). Importantly the correlation between average color needed for each patient and his or her PANSS scores for all subscales (see **Figure 1**) revealed that patients with higher PANSS would have made more errors when this was not controlled for. This is a relevant finding considering that in the present study it has been for the first time demonstrated that adaptive algorithms can successfully be used to control error rates in patients. This manipulation not only reduces drops-outs for a low number of error trials during data analysis, which is a common methodological problem faced by researchers investigating error-related processes, but also makes comparable the overall performance of patients with different levels of symptom severity. Consequently, we would like to plea for the use of similar adaptive programs when studying adaptive behavior in patients. Second, the manipulation of the ITIs showed to have an important effect in the slowing. When having a close look of the previous studies that have analyzed post-error slowing in schizophrenia, the ITIs fluctuated in a wide range going from 650 to 3000 ms (Bates et al., 2004; Laurens et al., 2005). Since here we have shown that short ITIs might enhance the size of the slowing, future studies should certainly consider this methodological aspect.

Taken together, our results showed that patients with schizophrenia who show increased distraction by novel stimuli (Grillon et al., 1990; Braff, 1993; Gray, 1995; Kapur, 2003; Laurens et al., 2005) also showed increased slowing at short ITIs when compared to controls. This finding has important implications not only at the theoretical level providing important evidence for the idea that the post-error slowing could rely on different mechanisms that the ones involved in error monitoring but also for future clinical studies. One possible explanation for the pattern of results in the patients' group can be found in the orienting account of the post-error slowing, which suggests that the slowing does not necessarily rely on error monitoring but is rather modulated by attentional mechanisms. Researchers and clinicians are advised to be careful in interpreting this behavioral effect as a marker for cognitive control (Notebaert et al., 2009) and to take into account relevant methodological aspects, like the duration of the ITIs and the error rates, when drawing conclusions from this behavioral measure.

ACKNOWLEDGMENTS

This work was supported by a doctoral grant from the Ghent University special research fund (BOF 07/DOS/011).

REFERENCES

Alain, C., McNeely, H. E., He, Y., Christensen, B. K., and West, R. (2002). Neurophysiological evidence of error-monitoring deficits in patients with schizophrenia. *Cereb. Cortex* 12, 840–846.

American Psychological Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders, Fourth ed.* Washington, DC: American Psychiatric Association.

Bates, A. T., Liddle, P. F., Kiehl, K. A., and Ngan, E. T. C. (2004). State dependent changes in error monitoring in schizophrenia. *J. Psychiatr. Res.* 38, 347–356.

Butler, P. D., Zemon, V., Schechter, I., Saperstein, A. M., Hoptman, M. J., Lim, K. O., Revheim, N., Silipo, G., and Javitt, D. C. (2005). Early-stage visual processing and cortical amplification deficits in schizophrenia. *Arch. Gen. Psychiatry* 62, 495–504.

Braff, D. L. (1993). Information-processing and attention dysfunctions in schizophrenia. *Schizophr. Bull.* 19, 233–259.

Carter, C. S., MacDonald, A. W., Ross, L. L., and Stenger, V. A. (2001). Anterior cingulate cortex activity and impaired self-monitoring of performance in patients with schizophrenia: an event-related fMRI study. *Am. J. Psychiatry* 158, 1423–1428.

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.

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.

Friedman, D., Cycowicz, Y. M., and Gaeta, H. (2001). The novelty P3: an event-related brain potential (ERP) sign of the brain's evaluation of novelty. *Neurosci. Biobehav. Rev.* 25, 355–373.

Frith, C. D., and Done, D. J. (1989). Experiences of alien control in schizophrenia reflect a disorder in the central monitoring of action. *Psychol. Med.* 19, 359–363.

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.

Gray, J. A. (1995). Dopamine release in the nucleus-accumbens – the perspective from aberrations of consciousness in schizophrenia. *Neuropsychologia* 33, 1143–1153.

Grillon, C., Courchesne, E., Ameli, R., Geyer, M. A., and Braff, D. L. (1990). Increased distractibility in schizophrenic-patients – electrophysiologic and behavioral evidence. *Arch. Gen. Psychiatry* 47, 171–179.

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.

Hajcak, G., and Simons, R. F. (2008). Oops!.. I did it again: an ERP and behavioral study of double-errors. *Brain Cogn.* 68, 15–21.

Jentzsch, I., and Dudschig, C. (2009). Why do we slow down after an error? Mechanisms underlying the effects of posterror slowing. *Q. J. Exp. Psychol.* 62, 209–218.

Johns, L. C., Rossell, S., Frith, C., Ahmad, F., Hemsley, D., Kuipers, E., and McGuire, P. K. (2001). Verbal self-monitoring and auditory verbal hallucinations in patients with schizophrenia. *Psychol. Med.* 31, 705–715.

Kapur, S. (2003). Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. *Am. J. Psychiatry* 160, 13–23.

Kay, S. R., Fiszbein, A., and Opler, L. A. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr. Bull.* 13, 261–276.

Kerns, J. G., Cohen, J. D., MacDonald, A. W., Johnson, M. K., Stenger, V. A., Aizenstein, H., and Carter, C. S. (2005). Decreased conflict- and error-related activity in the anterior cingulate cortex in subjects with schizophrenia. *Am. J. Psychiatry* 162, 1833–1839.

Kopp, B., and Rist, F. (1994). Error-correcting behavior in schizophrenic-patients. *Schizophr. Res.* 13, 11–22.

Kopp, B., and Rist, F. (1999). An event-related brain potential substrate of disturbed response monitoring in paranoid schizophrenic patients. *J. Abnorm. Psychol.* 108, 337–346.

Laming, D. R. J. (1968). *Information Theory of Choice Reaction Times*. New York, NY: Academic Press.

Laming, D. R. J. (1979). Autocorrelation of choice-reaction times. *Acta Psychol.* 43, 1381–1412.

Laurens, K. R., Kiehl, K. A., Ngan, E. T. C., and Liddle, P. F. (2005). Attention orienting dysfunction during salient novel stimulus processing in schizophrenia. *Schizophr. Res.* 75, 159–171.

Laurens, K. R., Ngan, E. T. C., Bates, A. T., Kiehl, K. A., and Liddle, P. F. (2003). Rostral anterior cingulate cortex dysfunction during error processing in schizophrenia. *Brain* 126, 610–622.

Leudar, I., Thomas, P., and Johnston, M. (1994). Self-monitoring in speech production – effects of verbal hallucinations and negative symptoms. *Psychol. Med.* 24, 749–761.

Malenka, R. C., Angel, R. W., Hampton, B., and Berger, P. A. (1982). Impaired central error-correcting behavior in schizophrenia. *Arch. Gen. Psychiatry* 39, 101–107.

Malenka, R. C., Angel, R. W., Thiemann, S., Weitz, C. J., and Berger, P. A. (1986). Central error-correcting behavior in schizophrenia and depression. *Biol. Psychiatry* 21, 263–273.

Mathalon, D. H., Fedor, M., Faustman, W. O., Gray, M., Askari, N., and Ford, J. M. (2002). Response-monitoring dysfunction in schizophrenia: an event-related brain potential study. *J. Abnorm. Psychol.* 111, 22–41.

Mlakar, J., Jensterle, J., and Frith, C. D. (1994). Central monitoring deficiency and schizophrenic symptoms. *Psychol. Med.* 24, 557–564.

Nieuwenhuis, S., Ridderinkhof, K. R., Blow, J., Band, G. P. H., 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.

Notebaert, W., Houtman, F., van Opstal, F., Gevers, W., Fias, W., and Verguts, T. (2009). Post-error slowing: an orienting account. *Cognition* 111, 275–279.

Núñez Castellar, E., Kühn, S., Fias, W., and Notebaert, W. (2010). Outcome expectancy and not accuracy determines posterror slowing: ERP support. *Cogn. Affect. Behav. Neurosci.* 10, 270–278.

Polli, F. E., Barton, J. J. S., Thakkar, K. N., Greve, D. N., Goff, D. C., Rauch, S. L., and Manoach, D. S. (2008). Reduced error-related activation in two anterior cingulate circuits is related to impaired performance in schizophrenia. *Brain* 131, 971–986.

Polli, F. E., Barton, J. J. S., Vangel, M., Goff, D. C., Iguchi, L., and Manoach, D. S. (2006). Schizophrenia patients show intact immediate error-related performance adjustments on an antisaccade task. *Schizophr. Res.* 82, 191–201.

Rabbitt, P. (1966). Errors and error correction in choice-response tasks. *J. Exp. Psychol. Gen.* 71, 264–272.

Rabbitt, P., and Rodgers, B. (1977). What does a man do after he makes an error? An analysis of response programming. *Q. J. Exp. Psychol.* 29, 727–743.

Shuwairi, S. M., Cronin-Golomb, A., McCarley, R. W., and O'Donnell, B. F. (2002). Color discrimination in schizophrenia. *Schizophr. Res.* 55, 197–204.

Slaghuis, W. L. (2004). Spatio-temporal luminance contrast sensitivity and visual backward masking in schizophrenia. *Exp. Brain Res.* 156, 196–211.

Turken, U., Vuilleumier, P., Mathalon, D. H., Swick, D., and Ford, J. M. (2003). Are impairments of action monitoring and executive control true dissociative dysfunctions in patients with schizophrenia? *Am. J. Psychiat.* 160, 1881–1883.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 17 October 2011; paper pending published: 28 October 2011; accepted: 11 February 2012; published online: 29 February 2012.

Citation: Núñez Castellar E, Houtman F, Gevers W, Morrens M, Vermeylen S, Sabbe B and Notebaert W (2012) Increased orienting to unexpected action outcomes in schizophrenia. Front. Hum. Neurosci. ***6****:32. doi: 10.3389/fnhum. 2012.00032*