Is there any electrophysiological evidence for subliminal error processing?

Shani Shalgi1* and Leon Y. Deouell2

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

The role of error awareness in executive control and modification of behavior is not fully understood. In line with many recent studies showing that conscious awareness is unnecessary for numerous high-level processes such as strategic adjustments and decision making, it was suggested that error detection can also take place unconsciously. The Error Negativity (Ne) component, long established as a robust error-related component that differentiates between correct responses and errors, was a fine candidate to test this notion: if an Ne is elicited also by errors which are not consciously detected, it would imply a subliminal process involved in error monitoring that does not necessarily lead to conscious awareness of the error. Indeed, for the past decade, the repeated finding of a similar Ne for errors which became aware and errors that did not achieve awareness, compared to the smaller negativity elicited by correct responses (Correct Response Negativity; CRN), has lent the Ne the prestigious status of an index of subliminal error processing. However, there were several notable exceptions to these findings. The study in the focus of this review (Shalgi and Deouell, 2012) sheds new light on both types of previous results. We found that error detection as reflected by the Ne is correlated with subjective awareness: when awareness (or more importantly lack thereof) is more strictly determined using the wagering paradigm, no Ne is elicited without awareness. This result effectively resolves the issue of why there are many conflicting findings regarding the Ne and error awareness. The average Ne amplitude appears to be influenced by individual criteria for error reporting and therefore, studies containing different mixtures of participants who are more confident of their own performance or less confident, or paradigms that either encourage or don’t encourage reporting low confidence errors will show different results. Based on this evidence, it is no longer possible to unquestioningly uphold the notion that the amplitude of the Ne is unrelated to subjective awareness, and therefore, that errors are detected without conscious awareness.

Keywords: error awareness, Ne/ERN, confidence, subliminal processing, wagering

INTRODUCTION

The human brain is wondrously efficient and complex, but by no means perfect, and errors are ubiquitous. Sometimes, we are acutely aware of having made an error, leading to the “oops” sensation, whereas at other times, errors go unnoticed. The detection of errors is crucial for successfully navigating through life and
Subliminal processing
Many, if not most, of our mental processes are inaccessible to conscious introspection, that is, we have no subjective experience of them. These subliminal processes, which are not available for explicit reporting, can be manifested in behavior (for example in post-error slowing) or by neuroimaging techniques (for example event-related potentials).

Error processing
The study of how an error is processed by the brain, under the assumption that errors are processed differentially than correct responses. It is debated whether errors are specifically registered as incorrect responses or are simply events linked to a large response conflict.

Error awareness
The study of how an error becomes recognized consciously. Unanswered questions in the field of error processing are: What brings about awareness of certain errors but not of others? Are errors which we are unaware of making, subliminally processed by the brain as errors at all? Or are they simply considered correct responses and not processed differentially?

Event Related Potentials (ERPs)
The electrophysiological neural response measured on the scalp which is directly related to a specific sensory, cognitive, or motor event (such as a certain stimulus or a response to a stimulus). Each ERP has a specific scalp distribution, latency and polarity (positive/negative) which defines it. ERPs are extracted by time-locking the raw EEG measurement to the triggering event, and usually averaging over several replications of the same event in order to increase the signal-to-noise ratio. Ideally, ERP components serve as physiological indices of specific cognitive functions.

effectively performing tasks. Error detection affords us the opportunity to modify behavior, which is essential for learning. Error detection may be particularly critical in high-risk tasks such as operating machinery or driving a car, where immediate error detection and correction can be vital to survival. What is not clear is the extent to which error detection occurs as a conscious or an unconscious process. For example, is conscious error detection a precondition for error correction or strategic adjustment of future behavior? Reduced awareness of errors has in fact been associated with a number of clinical conditions (ADHD, O’Connell et al., 2009; psychopathy, Brazil et al., 2009; drug addiction, Hester et al., 2009). Therefore, a more thorough understanding of the conditions under which errors reach consciousness and the neural correlates of error awareness may also have a significant clinical value.

The question of how consciousness is involved in error processing is interesting also because it is part of a much larger and profound question about the role of conscious awareness in our behavior. What function does consciousness serve? One way to answer this question is by elimination, that is, to ask what type of operations can be carried out in the absence of awareness. While it is generally accepted today that several perceptual, emotional and cognitive processes can run outside of awareness (e.g., Pessoa, 2005; Pessiglione et al., 2007, 2008; Bargh and Morsella, 2008; Sklar et al., 2012), it is also a common assumption that executive control operations do require conscious awareness for optimal performance (Mandler, 2002; Koch, 2004; Baars, 2005). This makes intuitive sense, as we usually become aware of those aspects in the internal or external environment that interfere with or interrupt routine action, or that defy our expectations and predictions, and errors, especially in simple tasks, are paradigmatically unexpected. Nevertheless, recent studies suggest that even cognitive control processes can be initiated unconsciously in some circumstances (see van Gaal and Lamme, 2012; van Gaal et al., 2012; Hassin, 2013 for reviews on this topic). Thus, the extent to which higher-level cognitive functions such as error processing and strategic adjustment of behavior can occur consciously remains unclear.

Throughout this review, we differentiate between two types of errors: Mistakes and Slips (see Reason, 1990). Slips are defined as errors that occur when the information needed to make the correct response is available to the decision making process (i.e., stimulus processing can be completed given sufficient time, and the task is understood). In contrast, Mistakes are errors made when the information or the algorithm needed to make the correct response is insufficient. For example, a typographical error of a familiar word is a Slip, while a wrong guess at the spelling of a novel word, or in solving a complex math problem, is a Mistake. In laboratory tasks, a Slip error could occur due to fast responding when the stimuli was seen but before it has been completely processed, or due to momentary lapse of attention to the task, and a Mistake could occur when the participant blinks at the time of the stimulus or did not fully understand the task. This implies that every Slip error can potentially be detected, whereas the same missing information that caused the Mistake would also hinder its detection. In our discussion of error processing in the brain, we refer only to Slip type errors, as processing can occur only in the presence of some information regarding the correct response1.

Electrophysiological indices of error processing
The scientific study of error monitoring began during the 1960s with a series of psychophysics studies. Rabbitt and co-workers’ pioneering work established that error monitoring relies on post-decision processing (e.g., Rabbitt, 1966, 2002; Rabbitt and Vyas, 1981). The in-depth study of the physiological characteristics of error processing began only two decades ago, with the discovery of two response-locked event related potentials (ERPs) that are enhanced following errors in contrast to correct responses. The first of these, termed the Error Negativity (Ne, Falkenstein et al., 1991; or ERN, Error-Related Negativity, Gehring et al., 1993) is a negative deflection in the ERP with a fronto-central distribution that begins around the time of the incorrect response2, and peaks roughly 50–100 ms thereafter. At the time of the Ne,

1Slips may go undetected because of failures to note the correct answer for a given trial, or because of failures to note a subtle erroneous response made (e.g., when the eye moves slightly in the wrong direction in an anti-saccade task, before a corrective movement is executed). A special type of errors that do not reach consciousness but are implicitly corrected, such as movement errors corrected by the cerebellum during sensory-motor adaptation will not be discussed here.

2While the Ne usually begins before the actual response time measured from a button press, it probably is more closely locked to the incorrect electromyographic (EMG) signal onset, which is not always measured in EEG experiments.
ERPs locked to correct responses typically show a significantly smaller negativity termed the Correct-Related Negativity (CRN; Ford, 1999; Vidal et al., 2000). It has recently been suggested using Independent Component Analysis (ICA) and source localization that the Ne and the CRN reflect the same underlying brain activity (Roger et al., 2010). The second error-related component is termed the Error Positivity (Pe, Falkenstein et al., 1991). The Pe is a more sustained component, which is absent after correct responses and is maximal at centro-parietal electrodes, peaking between 300 and 500 ms after the error. Both components appear independently of stimulus or response modality (Van ‘t Ent and Apkarian, 1999; Falkenstein et al., 2000; Shalgi et al., 2009) and across a number of paradigms designed to elicit errors, such as response inhibition tasks and choice reaction-time tasks. Based on converging evidence from different research methods, both components were localized to the Anterior Cingulate Cortex (ACC), albeit in different regions. The Ne is thought to originate from the dorsal/caudal ACC (ERP dipole localization: Dehaene et al., 1994; van Boxtel et al., 2005; O’Connell et al., 2007; LORETA: Herrmann et al., 2004; Trial-by-trial couplings of EEG and fMRI signals: Debener et al., 2005; Intracerebral recordings: Brazdil et al., 2002; though see Agam et al., 2011, for a more posterior location within the ACC). The Pe was localized to a comparatively more rostral region (van Veen and Carter, 2002; van Boxtel et al., 2005 and O’Connell et al., 2007, who also found contributions from the precuneus). Although some recent studies have questioned a direct relationship between the ACC and the Ne (see Orr and Hester, 2012), as a paralimbic area the ACC is uniquely positioned to integrate salient cognitive and emotional information with motor responses. It is active during the performance of complicated tasks, especially those involving conflict, and is thought to play a prominent role in the executive control of cognition (Botvinick et al., 2004; Ridderinkhof et al., 2004). The Ne and Pe also dissociate in their sensitivity to experimental manipulations, pharmacological substances and personality traits (for an overview, see Overbeek et al., 2005) and are therefore thought to play different roles in error processing. Due to its early latency after the error, the Ne has been linked to automatic, pre-conscious error monitoring, while the Pe has usually been associated with later conscious processes such as compensatory adjustments, emotional appraisal and mainly error awareness. However, the exact roles of both the Ne and Pe are still a matter of debate.

This review will focus solely on the role of the Ne with regard to error awareness.

**ERROR AWARENESS AND THE Ne**

The prevailing dogma for the past decade has been that the errors we make are registered by our brain even if we are unaware of them. Nieuwenhuis et al. (2001) showed that the Ne was equally seen after errors which subjects reported (henceforth “Aware Errors”), and after errors that were unnoticed (henceforth “Unaware Errors”). This result was striking because it implied that the whole complex and sophisticated sequence of processing—from stimulus recognition, decision making, up to the level of response selection and response monitoring—can be completed “below the surface.” That is, not only are stimuli and tasks processed unconsciously up to the semantic level (e.g., Costello et al., 2009; Mudrik et al., 2011; for a review see Kouider and Dehaene, 2007), but the brain also knows at a very early stage (perhaps even before the response was actually made) what should have been done regarding the stimuli and what was actually done. However, for some reason this information is not always relayed to consciousness. This starkly contradicts one of the most popular theories of the function of conscious deliberation, which is to deal with unexpected, non-routine situations (of which the error scenario is paradigmatic), designing intentional strategies for their handling, modifying behavior, and learning for future occasions. Although a prior study linked the Ne amplitude to subjective confidence of the response (Scheffers and Coles, 2000), the finding of equal amplitude Ne for aware and unaware errors was replicated several times in eleven different studies, including in our hands (Endrass et al., 2005, 2007, 2012; O’Connell et al., 2007, 2009; Pavone et al., 2009; Shalgi et al., 2009; Shalgi and Deouell, 2010; Dhar et al., 2011; Hughes and Yeung, 2011; for a review see Wessel, 2012), and led to the pervasive notion that error monitoring, as reflected by the Ne, can occur regardless of conscious awareness (e.g., Simons, 2009). The findings of these ERP studies were further supported by two event-related fMRI studies, which showed that activation of the ACC region associated with the generation of the Ne did not differ between Aware and Unaware Errors, while Aware Errors were associated with larger bilateral activation of prefrontal and parietal regions (Hester et al., 2005), or left anterior insula activity (Klein et al., 2007) compared to Unaware Errors. However, as more evidence accumulated, it became harder to ignore the data showing that error processing might
Wagering technique
A key problem of experiments that investigate conscious and subliminal processing is to objectively separate those stimuli or events which participants are consciously aware of and those which they are unaware of, due to the subjective nature of consciousness. The wagering technique is a method which has been proposed to solve this problem. In every trial of an experiment, after the response to primary task, the participant is asked to bet some money on whether their response is correct. If they are correct, they win the money; otherwise, they lose it. The basic idea is that people who know that they have information are usually willing to bet on it; that is, they are willing to “put their money where their mouth is.” In terms of assessing error awareness, the participants are first asked to judge whether their primary response was correct or incorrect, and then asked to bet on that judgment. In trials in which the participant made an error and judges it as correct, a high bet indicates that the participant was indeed unaware of their erroneous response, whereas a low bet would indicate some awareness of the error (of a trial that would otherwise be classified as an unnoticed error). The wagering method, by exploiting people’s desire to make money, is a more objective measure of awareness than asking participants to explicitly report on their confidence.

not occur without one being aware of the error. Six ERP studies (Sheffers and Coles, 2000; Maier et al., 2008; Steinhauser and Yeung, 2010; Woodman, 2010; Hewig et al., 2011; Wessel et al., 2011) showed that the Ne was significantly more negative for Aware Errors than Unaware Errors and correct responses, and three more studies which did not directly compare Aware and Unaware Errors were consistent with this notion (Praamstra et al., 2003; Pailing and Segalowitz, 2004; Selimbeyoglu et al., 2012).

A simple explanation which would reconcile the difference between the results of the two groups of studies might be that in those studies which found no difference between the Aware and Unaware Ne, many subjectively Aware Errors were not reported. In such cases, the group of trials designated “Unaware Errors” would be “contaminated” by many Aware errors with large Ne’s (see also Wessel, 2012). This would result in a reduction of the difference between the Aware and the Unaware Error ERPs (which are the average of trials in each group), becoming statistically insignificant. In fact, five of the studies which found no effect of awareness on the Ne, reported that its amplitude was numerically larger for Aware compared to Unaware Errors (Endrass et al., 2007; Shalgi et al., 2009; Shalgi and Deouell, 2010; Dhar et al., 2011; Hughes and Yeung, 2011; but see O’Connell et al., 2007, for the opposite result). However, there is no systematic difference between the experimental design of the experiments that did and did not find an effect of awareness on the Ne that would cause a response bias toward not signaling an error (see Wessel et al., 2011). In addition, error awareness rates in the different groups of studies were not noticeably different. Thus, the puzzle remained—can the Ne be taken as a signature of error detection which is independent of awareness of the error?

In his recent review of the relationship of the Ne with error awareness, which summarized thirteen studies, Wessel (2012) ruled out two more hypotheses as possible explanations for the presence or absence of error awareness effects on the Ne amplitude in different studies: (1) whether the stimuli were degraded or not in the given studies, and (2) whether errors were corrected or not (following Steinhauser and Yeung, 2010, who proposed that the amplitude of the Ne is an index of fast and automatic error correction). As an alternative, Wessel (2012) suggested that the studies that did not find a significant difference between the Ne for Aware and Unaware Errors but nevertheless, showed a numerical difference, may have suffered from low statistical power, increasing the probability of a type-2 error, i.e., the probability of not rejecting a null hypothesis, even though the alternative hypothesis is true. In light of this suggestion, Orr and Hester (2012) re-analyzed the fMRI data of three studies of error awareness, forming a composite sample of 56 subjects to increase statistical power. In contrast to their initial study (Hester et al., 2005, whose participants were included in this re-analysis), Orr and Hester found that the error-related dorsal ACC activity was significantly greater during Aware Errors compared to Unaware Errors, supporting the possibility that their initial null result was due to low statistical power. However, there are two important reasons to question the ability of the low-power hypothesis to adequately explain all the studies that failed to find an awareness effect on the Ne. First, numerical differences between Aware and Unaware Errors were found only in a few of the studies (in fact O’Connell et al., 2007, found a non-significant difference in the opposite direction). Second and most importantly, the finding of a similar Ne for Aware and Unaware errors cannot be viewed simply as a null result, because the imperative finding in those studies is in fact that Unaware Errors elicit a significant error-related component compared to correct responses (see Dhar et al., 2011, for an exception, and Endrass et al., 2005, who did not measure the Ne of correct responses).

The study in the focus of the current review (Shalgi and Deouell, 2012) proposed and empirically tested a solution to the above debate, which explains both types of results (difference/no difference between the Ne of Aware and Unaware Errors). In a choice-reaction time task, we asked our participants, on every trial, to judge whether they were accurate in their response or not (accuracy judgment). Immediately following this decision, we asked participants to bet money on their just-made accuracy judgment (wagering technique; Ullsperger et al., 2010). By examining the amount of their bet we effectively gained a measure of their subjective confidence in whether they were right or wrong in their accuracy judgment. We found that when participants made an error, yet were certain that they had not erred (indicated by their willingness to bet high that they had made a correct response), the Ne was significantly smaller than when subjects made an error, reported it, and were certain of it (betting high that they had indeed made an error). Moreover, when subjects erred and were certain that they did not, the Ne was comparable to
the CRN elicited by correct responses of which they were confident. Conversely, when subjects were correct but were uncertain of their performance (betting low on whether they were correct) or when they reported they erred but were uncertain of it, an intermediate CRN/Ne was elicited by all types of responses (Correct, Aware Errors, and Unaware Errors). These findings (schematized in Figure 1) suggest that the error detection as reflected by the Ne is after all dependent on subjective awareness of the error—when the level of awareness is more strictly determined using the wagering paradigm, no Ne is elicited without awareness. Thus, the Ne cannot be taken as evidence for error detection which is independent of error awareness.

Under the framework of signal detection theory (SDT; Green and Swets, 1966), both the "signal" (in our case, the error signal) and the background "noise" are associated with neuronal activity patterns that have a probabilistic distribution, and these distributions may overlap. This overlap creates ambiguous situations, requiring the subjects to make a decision (in the present case, to indicate that an error was or wasn’t made), even though the evidence is compatible with both options. SDT suggests that subjects set a threshold (criterion), and decide in favor of a "signal" only if the evidence exceeds this criterion. Under this framework, if subjects are conservative, they may refrain from declaring an error, even though there is a considerable evidence in favor of one (Figure 2A). Where does the wagering technique come into play within this scenario? Even if human subjects’ overt responses are based on their personal criterion (which is presumed to be stable within an experiment), they nonetheless make individual responses with different degrees of confidence. The process of confidence judgment is still debated (Clarke, 1959; Kunimoto et al., 2001; Pleskac and Busemeyer, 2010; Maniscalco and Lau, 2012), and the specific arguments are beyond the scope of this review. However, consistent with models of SDT addressing metacognitive judgments such as error detection (Clarke, 1959; Maniscalco and Lau, 2012; Rahnve et al., 2012), it seems reasonable to assume that, for a given subject with a set discrimination criterion, decisions made with high confidence occur when ambiguity is

3While the wagering technique has become popular in the field of communications, it has also been criticized (e.g., Clifford et al., 2008; Fleming and Dolan, 2010; Sandberg et al., 2010). Wagering related factors such as loss/risk aversion or subjective monetary strategies were brought up mainly to question the use of low bets as a signature of low confidence. However, in our use of wagering these concerns only serve to bolster the use of the high bets as a relatively pure measure of high confidence (for example, if subjects sometimes refrained from betting high due to loss aversion, then trials with the highest bet would undoubtedly be trials in which confidence in the accuracy judgment was strong enough to overcome loss aversion). Moreover, we did not directly compare between low and high bets.

Signal detection theory

Borrowing from the field of communications, SDT models our ability to discern between information-bearing patterns (the signal) and irrelevant patterns (noise). In the present case of error detection, the signal would be an error, and the "noise" refers to correct responses. Under the assumption that the noise signal has some probability distribution of possible patterns and the signal has another distribution, an observer needs to make a decision for every pattern, deciding whether it belongs to the noise distribution or to the signal distribution. Ambiguity arises when some patterns may come from either distribution, requiring a decision criterion (or threshold). The criterion specifies for which patterns the system will make a "signal" or "no signal" response, and depends on both the physical properties of the signal and the psychological state of the individual (e.g., experience, expectations, attention). For example, if a subject has a conservative response criterion for reporting errors, they may decide to report one only when they are absolutely sure that an error has been made (that is, when there is no ambiguity) and not report an error in any other case. Conversely, if one has a liberal criterion, they may report an error any time they feel their performance was less than perfect (that is, for all cases of ambiguity). Within this model, confidence may be determined by the distance from the decision threshold; the further away one is from the one's threshold, the more confident is the decision.
minimal. That is, when the signal in a given trial is far from the decision criterion (Figure 2B), or when the discriminability between the error and no error condition is greater (larger metacognitive $d'$)$^4$. Note that here the confidence measured is about the secondary decision (also known as the “Type 2 decision,” see Galvin et al., 2003) about the error signal (“did I make an error or not”), and not about the initial decision about the stimulus (“Type 1 decision”). Thus, by examining only trials with high bets, one can limit the analysis to trials where the subjects were fully aware of the error, or, more importantly, truly had no awareness of the error.

The results thus effectively resolve the issue of why there are conflicting findings regarding the Ne and error awareness, suggesting that there may be a range of subjective experiences of making an error, rather than an all-or-none phenomenon. The Ne amplitude thus appears to be influenced by the level of discriminability (in signal detection terms) as well as the criteria different individuals have for error reporting; studies containing different mixtures of participants who are more or less conservative in their error reporting will yield different results. Additionally, some paradigms may either encourage subjects to report errors even when they are less confident, while others may discourage reporting low-confidence errors, also affecting the results. Specifically, if something in the design of the experiment, the task instructions, or an experimenter bias, encourages the subjects to report errors only when they are highly confident, or if because of random selection the subjects in the experiment are relatively conservative, and no confidence measures are used, many error trials classified as Unaware Errors will in fact have some level of awareness, thus increasing the amplitude of the Ne in the putatively “unaware” category.

$^4$Indeed, our subjects varied considerably in their wagering performance; some subjects bet mainly using the highest bet and others mainly on the lowest amount, suggesting different $d'$ for detecting errors.
Recently, a study by Steinhauser and Yeung (2010) directly examined the effect of manipulation of the error reporting criteria on the amplitude of the Ne. The authors used a brightness discrimination task and asked subjects to judge whether they were right or wrong in every trial. They used two different incentives to encourage participants to adopt either a high (report an error only if you are very sure) or a low criterion for reporting their errors. Steinhauser and Yeung’s main results were consistent with ours: when the high criterion was imposed, the Ne was significantly larger for Aware Errors than Unaware Errors. Inconsistent with this (as well as with our above predictions) there was no difference between the Ne elicited by Aware Errors under the low and high criterion conditions. It is hard to reconcile these two results, and further experiments will be needed to reveal whether the type of task (threshold discrimination of brightness vs. discrimination of highly visible objects) is critical in determining the effect of criterion on the Ne, or whether the direct manipulation of error signaling criterion (“Type 2 decision”) is indeed equivalent to manipulating the subjective level of confidence and awareness of the error. Putatively, once a Type 2 criterion is set and an error signal is elicited, if this criterion is crossed, confidence in this Type 2 decision (a type of 3rd level decision) may still be as variable, regardless of the initial criterion.

THE FUNCTIONAL SIGNIFICANCE OF THE Ne

Does the sensitivity of the Ne to subjective error awareness contribute to the debate about its functional significance? Three major models have been suggested to explain the Ne: (1) Error Detection (Falkenstein et al., 1991; Coles et al., 2001), (2) Reinforcement Learning (Holroyd and Coles, 2002) and (3) Conflict Detection (Botvinick et al., 2004; Yeung et al., 2004). According to the Error Detection account, a dedicated module for detecting errors exists in the brain. This module detects the mismatch between the actual response committed and the intended response, and the Ne is seen to reflect the size of this mismatch. The Reinforcement Learning theory also suggests a generic error processing system, comprised of a comparator and of a remedial action system. Upon detecting an error, the comparator, located at a subcortical level (the basal ganglia) transmits a learning signal via the mesencephalic dopaminergic system to the remedial action system situated in the ACC, indicating that the outcome of the current action was worse than expected (negative reward). The Ne is generated by the arrival of the error signal at the remedial action system. According to this account, the ACC does not monitor errors per se, but uses the learning signal to adapt the response selection process ⁵. According to the Conflict Detection account, the Ne reflects the amount of response conflict present after an error has been committed. The post-error response conflict is a consequence of continued processing of the stimulus that leads to post-error activation of the correct response and therefore, conflict with the incorrect response just produced⁶. The presence of response conflict indicates situations in which errors are likely to occur and hence require attention. Our finding of a larger Ne for certain errors seems to be compatible with both the Error-Detection and the Reinforcement Learning accounts. When an error is certain, there should be a large mismatch between the actual response and the required response compared to an uncertain error (hence larger error-detection signal), and indeed the Ne amplitude is larger for certain than uncertain errors. Concomitantly, evident errors are probably considered worse outcomes (i.e., negative reward) than unsure errors, requiring readjustments (hence learning). Are the results also consistent with the Conflict Detection account? Scheffers and Coles (2000) argued that Ne variation with subjective confidence argues against this model. They reasoned that the largest conflict should occur when subjects were uncertain of their primary response, whereas in fact, in their hands and ours, uncertainty yields smaller Ne. However, the relationship between confidence and conflict is not straightforward in our view. The response (associated with the Ne/CRN) and the confidence judgment are dissociated in time, and both stimulus and

⁵A recent study suggests that whenever an outcome of an action violates expectancies, not necessarily related to reward, the ACC appears to be engaged in evaluating the necessity of behavioral adaptation (Wessel et al., 2012). Using both EEG and fMRI, Wessel and colleagues established that the Ne and the novelty N2 which indexes processing of generally infrequent or surprising stimuli share a common neuroanatomical substrate in the ACC. The authors thus suggest generalizing the Reinforcement Learning theory of the Ne beyond reward expectancies. However, as the definition of reward is highly flexible, one can argue that novelty or surprise are rewarding (either positively or negatively) and remain with the original theory.

⁶Note that response conflict is different from perceptual conflict. There are cases where the perceptual conflict is large but the response conflict is small (e.g., Danielmeier et al., 2009).
response processing continue to evolve between the primary response and the accuracy judgment. In one scenario, a large conflict at the time of the response may lead to uncertainty regarding the response at the time the accuracy judgment is given. In another scenario, the same conflict may have been fully resolved by the time of the accuracy judgment, leading to high certainty. Thus, confidence in accuracy judgments does not correspond in a straightforward manner to high or low conflict at the time of the response. Considering only the confident responses however, it is conceivable that a higher conflict occurred in cases of Aware Errors (if both the correct and incorrect responses were activated) than in the case of Unaware Errors or Correct responses, in which one response, either the correct or the incorrect one, dominated. Therefore, in our view, the findings of our recent experiment (Shalgi and Deouell, 2012) and those of Scheffers and Coles (2000) do not unequivocally distinguish between the Error Detection, Reinforcement Learning, and the Conflict Detection hypotheses of the Ne. They do however, challenge the notion of error detection without awareness.

UNCONSCIOUS ERROR PROCESSING?

The latency of the Ne, immediately or even at the same time as the commission of an error, suggests that it must be based on quickly available feed-forward information and not require other inputs such as proprioceptive or sensory feedback (Allain et al., 2004; de Bruijn et al., 2004; Mathalon et al., 2004). The early latency is what led to the initial proposal that the Ne was an index of automatic, unaware error processing. This is because conscious awareness is thought to be a slower process (Rabbitt, 2002), which comes about only after the accumulating evidence about erroneous behavior, based on numerous sources of information, exceeds some threshold and activates sufficiently large neural networks (Dehaene et al., 2006). In this respect, the variation of the Ne with the level of awareness of the error is surprising. We propose (in line with Wessel, 2012) that the initial error signal indexed by the Ne does not reflect error awareness per se, but rather that it is a prerequisite of this process, and thus, correlated with it. Possibly, if some Ne is generated, a second, compound internal error signal based on other sources of information (including later events, such as proprioceptive feedback, autonomic responses and sensory input) may exceed the threshold for error awareness.

Thus, as opposed to prevailing views, the Ne seems to be correlated with, and possibly a prerequisite to, awareness of an error. Does this mean that without an Ne one cannot be aware of an error? A very recent study by Charles et al. (2013) suggests some types of awareness can occur without eliciting an Ne. Their results showed above-chance error awareness, detected using a forced choice technique, to stimuli that were effectively masked. For these correctly detected (in fact guessed) errors, no Ne was recorded. This finding was interpreted in light of an extension of the dual-route model of decision-making proposed by Del Cul et al. (2009; Figure 3). According to the dual-route model, two parallel routes, with different noise levels and thresholds, accumulate sensory evidence toward a categorical decision on the same input stimulus. One is a fast, non-conscious sensorimotor route (bottom row, Figure 3), and one is a slower conscious decision route (high-level route, top row). A motor response is emitted by the route that first reaches its decision threshold. Usually, we try to guide our actions by our conscious intention, but sometimes our actions start earlier through the fast route, and then we might slip and make an error. According to Charles et al. (2013), the Ne is generated as a result of a discrepancy between the responses.
computed by the two routes (the difference between intended and executed action).

By this model, the Ne is generated only when a conscious intention exists, i.e., when the conscious route has crossed its threshold. If no conscious intention is generated (e.g., when the stimulus is below threshold), no comparison can be initiated, yet a “hunch,” generated based on rudimentary (unconscious) evaluation of evidence, might be used later for the above chance guessing. Thus, the model can explain both the correlation between error awareness and the Ne when the stimulus is consciously perceived, and attenuated above-chance error awareness without an Ne when the stimulus is not consciously perceived and forced-choice error signaling is used.

**METHODOLOGICAL IMPLICATIONS**

Our findings underscore the importance of careful evaluation when it comes to the treacherous issue of conscious awareness vs. subliminal processing Simply asking for subjective report is notoriously unreliable. In signal detection theory terms, it confounds discriminability (of an error in our case) and the criterion for reporting. The wagering process we used does not solve the problem altogether, but it provides a way for determining the distance from the criterion: reluctance to bet high suggests a decision close to the criterion, and so the determination of awareness in these trials is murky. In contrast, willingness to bet high suggests the decision was made well above the criterion, making the exact position of the criterion less critical.

The addition of additional measures, like wagering or the forced choice method, is often cumbersome and disrupts the flow of the task even more than simple accuracy judgments. Therefore, an important next step in the study of the Ne as an index of subjective error processing is to identify those factors that influence overall error reporting (e.g., task difficulty, stimulus ambiguity), and the individual differences (e.g., traits) that lead to variable error reporting.

**CONCLUSION**

Error awareness is a fascinating topic. It underlies numerous philosophical and cognitive-psychological questions such as the possibility of acting without awareness (the “zombie” mode; Koch and Crick, 2001) and the meaning of awareness in executive function. In Shalgi and Deouell (2012), we presented evidence that the error-related component previously taken as support for the existence of subliminal error processing (the Ne) is actually modulated by subjective confidence that an error has occurred, and in fact correlates with conscious awareness of the error. This leaves us with no clear-cut electrophysiological evidence for the unconscious processing of errors.

---

**REFERENCES**

Agam, Y., Hamalainen, M. S., Lee, A. K., Dyckman, K. A., Friedman, J. S., Isom, M., et al. (2011). Multimodal neuroimaging dissociates hemodynamic and electrophysiological correlates of error processing. *Proc. Natl. Acad. Sci. U.S.A.* 108, 17556–17561. doi: 10.1073/pnas.1103475108

Allain, S., Hasbroucq, T., Burle, B., Grapperon, J., and Vidal, F. (2004). Response monitoring without sensory feedback. *Clin. Neurophysiol.* 115, 2014–2020. doi: 10.1016/j.clinph.2004.04.013

Baars, B. J. (2005). Global workspace theory of consciousness: toward a cognitive neuroscience of human experience. *Prog. Brain Res.* 130, 45–53. doi: 10.1016/S0079-6123(05)50004-9

Bargh, J. A., and Morsella, E. (2008). The unconscious mind. *Perspect. Psychol. Sci.* 3, 73–79. doi: 10.1111/j.1745-6916.2008.00064.x

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. doi: 10.1016/j.tibs.2004.10.003

Brafeldt, 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. doi: 10.1007/s00221-002-1201-y

Brazdil, M., de Bruiijn, E. R., Buiten, B. H., von Borries, A. K., van Lankveld, J. J., Buiteijar, J. K., et al. (2009). Early and later components of error monitoring in violent offenders with psychosis. *Biol. Psychiatry* 65, 137–143. doi: 10.1016/j.biopsych.2008.08.011

Charles, L., Van Opstal, F., Marti, S., and Dehaene, S. (2013). Distinct brain mechanisms for conscious versus subliminal error detection. *Neuroimage* 73C, 80–94. doi: 10.1016/j.neuroimage.2013.01.054

Clarke, F. R. (1959). Two types of ROC curves and definitions of parameters. *Acoust. Soc. Am.* J. 31, 629. doi: 10.1121/1.1907764

Clifford, C. W., Arbabzadeh, E., and Harris, J. A. (2008). Getting technical about awareness. *Trends Cogn. Sci.* 12, 54–58. doi: 10.1016/j.tibs.2007.11.009

Coles, M. G., Scheffers, M. K., and Holroyd, C. B. (2001). Why is there an ERP/Ne on correct trials. Response representations, stimulus-related components, and the theory of error-processing. *Biol. Psychol.* 56, 173–189. doi: 10.1016/S0301-0511(01)00076-X

Costello, P., Jiang, Y., Baartman, B., McGlennan, K., and He, S. (2009). Semantic and subword priming during binocular suppression. *Conscious Cogn.* 18, 375–382. doi: 10.1016/j.concog.2009.02.003

Danielmeier, C., Wessel, J. R., Steinhauser, M., and Ullsperger, M. (2009). Modulation of the error-related negativity by response conflict. *Psychophysiology* 46, 1288–1298. doi: 10.1111/j.1469-8986.2009.00860.x

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. doi: 10.1523/JNEUROSCI.3286-05.2005

de Bruijn, E. R., 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.

Dehaene, S., Changeux, J. P., Naccache, L., Sakur, J., and Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends Cogn. Sci.* 10, 204–211. doi: 10.1016/j.tibs.2006.03.007

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. doi: 10.1111/j.1467-9280.1994.tb00630.x

Del Cul, A., Dehaene, S., Reyes, P., Bravo, E., and Slachkovsky, A. (2009). Causal role of prefrontal cortex in the threshold
for access to consciousness. Brain 132(Pt 9), 2531–2540. doi: 10.1093/brain/awp111
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 countermatching task. J. Psychophysiol. 19, 275–280. doi: 10.1027/0269-8803.19.4.275
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
Endrass, T., Reuter, B., and Kathmann, N. (2007). ERP correlates of conscious and unconscious recognition: aware and unaware errors in an antisaccade task. Eur. J. Neurosci. 26, 1714–1720. doi: 10.1111/j.1460-9568.2007.05785.x
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. doi: 10.1016/0013-4694(91)90062-9
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. doi: 10.1016/S0301-0511(99)00381-9
Fleming, S. M., and Dolan, R. J. (2010). Effects of loss aversion on post-decision wagering: implications for measures of awareness. Conscious. Cogn. 19, 352–363. doi: 10.1016/j.concog.2009.11.002
Ford, J. M. (1999). Schizophrenia: the broken P300 and beyond. Psychophysiology 36, 667–682. doi: 10.1111/1469-8986.360667
Galvin, S. J., Podd, J. V., Draga, V., and Whitmore, J. (2003). Type 2 tasks in the theory of signal detectability: discrimination between correct and incorrect decisions. Psychon. Bull. Rev. 10, 843–876. doi: 10.3758/BF03196546
Gehring, W. J., Gross, 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. doi: 10.1111/1467-9280.1993.tb00586.x
Green, D. M., and Swets, J. A. (1966). Signal Detection Theory and Psychophysics. New York, NY: Wiley.
Hassin, R. R. (2013). Yes it can: on the functional abilities of the human unconscious. Perspect. Psychol. Sci. 8, 195–207. doi: 10.1177/1745691612460864
Herrmann, M. J., Rommeler, J., Ehlis, A. C., Heidrich, A., and Fallgatter, A. J. (2004). Source localization (LORETA) of the error-related negativity (ERN/Ne) and positivity (Pe). Brain Res. Cogn. Brain Res. 20, 294–299. doi: 10.1016/j.cogbrainres.2004.02.013
Hester, R., Foxe, J. J., Molholm, S., Shapner, M., and Garavan, H. (2005). Neural mechanisms involved in error processing: a comparison of errors made with and without awareness. Neuroimage 27, 602–608. doi: 10.1016/j.neuroimage.2005.04.035
Hester, R., Nestor, I., and Garavan, H. (2009). Impaired error awareness and anterior cingulate cortex hypoxia in chronic cannabis users. Neuropsychopharmacology 34, 2450–2458. doi: 10.1038/npp.2009.67
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. doi: 10.1111/j.1469-8986.2011.02092.x
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. doi: 10.1037/0033-295X.109.4.679
Hughes, G., and Yeung, N. (2011). Dissociable correlates of response conflict and error awareness in error-related brain activity. Neuropsychologia 49, 405–415. doi: 10.1016/j.neuropsychologia.2010.10.036
Klein, T. A., Endrass, T., Kathmann, N., Neumann, J., von Cranon, D. Y., and Ullsperger, M. (2007). Neural correlates of error awareness. Neuroimage 34, 1774–1781. doi: 10.1016/j.neuroimage.2006.11.014
Koch, C. (2004). The Quest for Consciousness: A Neurobiological Approach. Englewood, CO: Roberts and Company Publishers.
Koch, C., and Crick, F. (2001). The zombie within. Nature 411, 893. doi: 10.1038/35082161
Kouider, S., and Dehaene, S. (2007). Levels of processing during non-conscious perception: a critical review of visual masking.
Philos. Trans. R. Soc. Lond. B Biol. Sci. 362, 857–875. doi: 10.1098/rstb.2007.2093
Kunimoto, C., Miller, J., and Pashler, H. (2001). Confidence and accuracy of near-threshold discrimination responses. Conscious. Cogn. 10, 294–340. doi: 10.1016/S0931-5597(00)00494-7
Maier, M., Steinhauer, 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. doi: 10.1162/jocn.2008.20159
Mandler, G. (2002). Consciousness Recovered: Psychological Functions and Origins of Conscious Thought. Amsterdam: Philadelphia: John Benjamins Publishing Company.
Maniscalco, B., and Lau, H. (2012). A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. Conscious. Cogn. 21, 422–430. doi: 10.1016/j.concog.2011.09.021
Mathalon, D. H., Colrain, I., Gray, M., and Ford, J. M. (2004). “It’s not my fault: ERPs to induced errors,” 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), 27–35.
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. doi: 10.1177/0956797611408736
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. doi: 10.1111/1469-8986.3850752
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. doi: 10.1016/j.neuropsychologia.2009.01.011
O’Connell, R. G., Dockree, P. M., Bellgrove, M. A., Kelly, S. P., Hester, R., Garavan, H., et al. (2009). The neural correlates of deficient error awareness in attention-deficit hyperactivity disorder (ADHD). Neuropsychologia 47, 1149–1159. doi: 10.1016/j.neuropsychologia.2009.01.011
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. doi: 10.1016/S1053-8119(03)00351-3
Rabbitt, P. M. (1966). Error and error correction in choice-response tasks. J. Exp. Psychol. 71, 264–272. doi: 10.1037/h0022853
Rabbitt, P. M. (2002). Consciousness is slower than you think. Q. J. Exp. Psychol. A 55, 1081–1092. doi: 10.1080/0272498024400080
Rabbitt, P. M., and Vyas, S. (1981). Processing a display even after you make a response to it. How perceptual errors can be corrected.
correlates of subjective confidence level about a decision. Clin. Neurophysiol. 123, 1158–1167. doi: 10.1016/j.clinph.2011.10.037

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. doi: 10.1111/j.1460-9589.2009.06690.x

Shalgi, S., and Deouell, L. Y. (2010). “Is there a hemispatial bias in detecting errors and in error awareness?” Paper Presented at The Israel Society for Neuroscience 19th Annual Meeting Einat (Eilat).

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

Simons, R. F. (2009). The way of our errors: theme and variations. Psychophysiology 47, 1–14. doi: 10.1111/j.1469-8986.2009.0929.x

Sikdar, A. Y., Levy, N., Goldstein, A., Mandel, R., Maril, A., and Hassin, R. R. (2012). Reading and doing arithmetic nonconsciously. Proc. Natl. Acad. Sci. U.S.A. 109, 19614–19619. doi: 10.1073/pnas.1211645109

Steinhauser, M., and Yeung, N. (2010). Decision processes in human performance monitoring. J. Neurosci. 30, 15643–15653. doi: 10.1523/JNEUROSCI.1899-10.2010

Ullsperger, M., Haray, 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. doi: 10.1007/s00429-010-0261-1

van Boxtel, G. J. M., van der Molen, M. W., and Jennings, J. R. (2005). Differential involvement of the anterior cingulate cortex in performance monitoring during a stop-signal task. J. Psychophysiol. 19, 1–10. doi: 10.1027/0269-8803.19.1.1

Van’t Ent, D., and Apkarian, P. (1999). Motoric response inhibition in finger movement and saccadic eye movement: a comparative study. Clin. Neurophysiol. 110, 1058–1072. doi: 10.1016/S1388-2457(98)00306-4

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

van Gaal, S., and Lamme, V. A. F. (2012). Unconscious high-level information processing. Neuron 18, 287–301. doi: 10.1177/1073858411404079

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. doi: 10.1162/08989290260045837

Vidal, F., Hashbroucq, T., Grapperon, J., and Bonnet, M. (2000). Is the ‘error negativity’ specific to errors. Biol. Psychol. 51, 109–128. doi: 10.1016/S0301-0511(99)00032-0

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. doi: 10.1523/JNEUROSCI.6352-11.2012

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. doi: 10.1162/jocn.2011.21635

Woodman, G. F. (2010). Masked targets trigger event-related potentials indexing shifts of attention but not error detection. Psychophysiology 47, 410–414. doi: 10.1111/j.1469-8986.2009.00948.x

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. doi: 10.1037/0033-295X.111.4.931

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: 06 April 2013; paper pending publication: 27 May 2013; accepted: 02 August 2013; published online: 29 August 2013.

Citation: Shalgi S and Deouell LY (2013) Is there any electrophysiological evidence for subliminal error processing? Front. Neurosci. 7:150. doi: 10.3389/fnins.2013.00150

This article was submitted to the journal Frontiers in Neuroscience.

Copyright © 2013 Shalgi and Deouell. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.