Psychophysiology, 38 (2001), 752–760. Cambridge University Press. Printed in the USA.

Error-related brain potentials are differentially related to awareness of response errors: Evidence from an antisaccade task

SANDER NIEUWENHUIS,[a] K. RICHARD RIDDERINKHOF,[a] JOS BLOM,[a]
GUIDO P.H. BAND,[b] AND ALBERT KOK[a]

[a]Department of Psychonomics, University of Amsterdam, The Netherlands
[b]Experimental and Theoretical Psychology Unit, Leiden University, The Netherlands

Abstract

The error negativity (Ne/ERN) and error positivity (Pe) are two components of the event-related brain potential (ERP) that are associated with action monitoring and error detection. To investigate the relation between error processing and conscious self-monitoring of behavior, the present experiment examined whether an Ne and Pe are observed after response errors of which participants are *unaware*. Ne and Pe measures, behavioral accuracy, and trial-to-trial subjective accuracy judgments were obtained from participants performing an antisaccade task, which elicits many unperceived, incorrect reflex-like saccades. Consistent with previous research, subjectively unperceived saccade errors were almost always immediately corrected, and were associated with faster correction times and smaller saccade sizes than perceived errors. Importantly, irrespective of whether the participant was aware of the error or not, erroneous saccades were followed by a sizable Ne. In contrast, the Pe was much more pronounced for perceived than for unperceived errors. Unperceived errors were characterized by the absence of posterror slowing. These and other results are consistent with the view that the Ne and Pe reflect the activity of two separate error monitoring processes, of which only the later process, reflected by the Pe, is associated with conscious error recognition and remedial action.

Descriptors: Error-related negativity, Error processing, ERN, Ne, Error positivity, Awareness, Antisaccades

To adjust performance appropriately to varying task demands, it is important to monitor ongoing action and associated feedback for possible errors. As has already been noted by Rabbitt (e.g., 1967), error processing, at least in the context of reaction time tasks, must involve an error detection mechanism and a set of remedial action mechanisms. The latter mechanisms are responsible for immediate error correction and for adjustments to response settings to prevent errors from recurring in the future. Recent neuroimaging studies and neuropsychological studies suggest that at least some of these error processing mechanisms may be implemented in a brain circuit involving the anterior cingulate cortex and lateral prefrontal cortex (Carter et al., 1998; Gehring & Knight, 2000; Kiehl, Liddle, & Hopfinger, 2000). Progress on identifying the *functional* characteristics of the error monitor system has been made primarily through the study of two psychophysiological indices thought to be specifically associated with error processing: (1) the error negativity (Ne; Falkenstein, Hohnsbein, & Hoormann, 1991), or error-related negativity (ERN; Gehring, Goss, Coles, Meyer, & Donchin, 1993). This is a sharp negative deflection in the event-related brain potential with a frontocentral distribution and peak approximately 80 ms after an incorrect response. The onset of the Ne can be as early as the first EMG activity leading to the incorrect response (Gehring et al., 1993); (2) the error positivity (Pe; Falkenstein et al., 1991; Falkenstein, Hohnsbein, & Hoormann, 1995). This is a slow positive wave with centroparietal distribution, which often but not necessarily follows the Ne on incorrect trials. Although most investigators in the field have focused on validating the Ne and not the Pe, most authors appear to agree that the two indices represent different aspects of error processing (see Falkenstein, Hoormann, Christ, & Hohnsbein, 2000).

The objective of the present research was to further our understanding of the psychological processes manifested by the Ne and Pe by examining the relationship between these psychophysiological indices and the presence or absence of awareness that a response error has occurred. It is likely that the Ne and Pe reflect the activity of two separate error monitoring processes. The early onset latency of the Ne with respect to the response (e.g., Gehring et al., 1993) is suggestive of an internal error monitoring system, acting rapidly on information from central (as opposed to peripheral) sources. Sensory or proprioceptive information signalling an erroneous response could not be available until after the response

The research of Dr. Ridderinkhof has been made possible by a fellowship of the Royal Netherlands Academy of Arts and Sciences. Helpful comments on earlier versions of this article were provided by Eric-Jan Wagenmakers and Michael Coles. We thank Durk Talsma for his technical assistance.

Address requests for reprints to: Sander Nieuwenhuis, Department of Psychonomics, University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam, The Netherlands. E-mail: stn20@dds.nl.

has been initiated. In contrast, the timing of the Pe allows for the possibility of a second error monitoring process, which is informed by peripheral (e.g., proprioceptive or reafferent) information indicating that an error has occurred. Note that a distinction between an "internal" and "external" error monitoring loop has often been made in models of motor control (e.g., Cooke & Diggles, 1984). Partly because of their timing with respect to the response, it is well-conceivable that the two error monitoring processes are differentially related to subjective awareness of response errors. For instance, Falkenstein et al. (1991) have proposed that the early error-related ERP component, the Ne, "reflects a (perhaps unconscious) mismatch between response selection and . . . response execution, whereas the slow wave [i.e., the Pe] reflects the conscious evaluation of the error" (p. 454). As we will argue below, there is already some indirect evidence in support of this proposal. In our experiment, the hypothesis that the processes leading to awareness of a response error are associated with the processes manifested by the Pe, but not with those underlying the Ne, was further tested.

Thus far there have been no studies of error-related brain activity that have explicitly distinguished between subjectively perceived and unperceived erroneous responses. The likely reason for this is that in almost all studies of the Ne and Pe, overt response errors have been defined by an inappropriate button press, and participants are typically aware of this type of error. Here we report an experiment in which we obtained error-related ERP components, behavioral accuracy data, and trial-to-trial subjective accuracy judgments from participants performing an antisaccade task. Performance in this type of task shows many inappropriate, reflex-like saccades which are *not* perceived by the participant, even though they are immediately corrected (Kramer, Hahn, Irwin, & Theeuwes, 2000; Mokler & Fischer, 1999; Theeuwes, Kramer, Hahn, & Irwin, 1998). In the remainder of the introduction, we provide a selective review of empirical data and theory regarding the Ne and Pe (for an extensive review of both components, see Falkenstein et al., 2000). Then, we will discuss the antisaccade task and how it can be employed to investigate awareness of response errors.

Error Negativity

The Ne can be observed in both auditory tasks and visual tasks (e.g., Falkenstein et al., 1991), and has been reported after both hand and foot errors (Holroyd, Dien, & Coles, 1998), and after failures to inhibit eye movements in a saccade variant of the Go—No-go paradigm (Van't Ent & Apkarian, 1999). Although there is some evidence that the amplitude of the Ne is related to more conservative response behavior on trials following an error (Gehring et al., 1993), the process reflected by the Ne does not seem to be involved in error correction (e.g., Scheffers, Coles, Bernstein, Gehring, & Donchin, 1996). Indeed, an Ne-like component has also been observed after failures to reach a response deadline (Luu, Flaisch, & Tucker, 2000), which allows no immediate error correction.

The available evidence of the Ne is generally consistent with the proposal (e.g., Bernstein, Scheffers, & Coles, 1995; Falkenstein et al., 1991; Scheffers et al., 1996) that the negativity following errors reflects the manifestation of an error detection system that checks actual behavior against an internal goal standard. The output of this error detection system may reflect both the cognitive and emotional significance of a deviation from the anticipated result (see, e.g., Gehring et al., 1993; Kiehl et al., 2000). In choice RT tasks, a representation of the actual response may be formed on the basis of some "efference copy" of the motor commands sent to the response effectors (see Gehring et al., 1993). A representation of the task-appropriate (goal-directed) response may become available through further processing of the imperative stimulus, beyond that which is used to guide the actual response. These two representations provide the input to the error detection system, whose output signal is some unspecified increasing function of the degree of mismatch between the two representations.

Two lines of evidence have been put forward in support of the proposal that the Ne reflects a mismatch measure as described above. First, there are several studies indicating a positive relation between Ne amplitude and the degree of error. Here, the degree of error refers to the difference between the correct and the incorrect response in terms of temporal (Luu, Flaisch, et al., 2000) or spatial parameters (Bernstein et al., 1995; Scheffers et al., 1996; Van't Ent & Apkarian, 1999; but see Gehring et al., 1993). The second type of evidence comes from studies showing that the Ne amplitude is smaller under circumstances that may be expected to affect the quality of perceptual processing. These circumstances include experimental settings in which participants have to respond under time pressure (Falkenstein, Hohnsbein, Hoormann, & Blanke, 1990; see also Falkenstein et al., 2000), have been deprived of sleep (Scheffers, Humphrey, Stanny, Kramer, & Coles, 1999), or are presented with degraded stimuli (Scheffers & Coles, 2000). The argument is that under these circumstances participants are likely to have a compromised representation of the response required by the stimulus. This, in turn, reduces the mismatch with the representation of the actual response in case of an error.

One previous study has paid attention to the relation between the Ne and participants' confidence in the accuracy of their response (Scheffers & Coles, 2000). In this study, participants performed an Eriksen flankers task in which they were required to respond to the identity of one of two briefly presented target letters (H or S). The target letter was surrounded by either four compatible flankers (e.g., SSSSS) or four incompatible flankers (i.e., letters associated with the response of the other hand, e.g., HHSHH). Importantly, the contrast between stimuli and background was reduced in order to induce uncertainty in participants about the presented target. At the end of each trial, participants rated the accuracy of their response on a five-point scale ranging from *sure incorrect* to *sure correct*. The results showed that Ne amplitude varied with the subjectively perceived accuracy of the response, irrespective of its objective accuracy. For example, on a small proportion of correct trials the response was judged to be sure incorrect by the participant, suggesting that the stimulus was misperceived and the given response was not consistent with the incorrect stimulus information. These trials were associated with a large Ne. Likewise, after incorrect responses judged to be sure correct, only a small Ne was observed.

Although these results highlight the consequences of an inaccurate representation of the correct response for the computation of the Ne, it is perhaps not surprising that Ne amplitude was so strictly tied to participants' awareness of response accuracy. Note that both the subjective accuracy ratings and Ne amplitude behaved precisely as what would be predicted from the degree of mismatch between the stimulus information extracted from the actual stimulus display, and the emitted response. The observation rather emphasizes the notion that the Ne system does not react to errors if the information necessary to identify the error is lacking. A related argument applies to the observation by Dehaene, Posner, and Tucker (1994) that an Ne occurs after the incorrect execution of an appropriate motor schema (i.e., a slip), but not after the

correct execution of an inappropriate schema due to failures to retrieve the correct schema (i.e., a mistake). Here the missing information, necessary to identify the error, concerns the correct schema. If we want to examine the relation between error-related brain activity and the awareness of *slips*—the type of error that has been the focus of psychophysiological research on error processing—we need a paradigm in which the Ne system can easily derive a representation of the correct response, but participants are not (always) aware of errors in the execution of this response. Before describing our choice of paradigm, we briefly review the empirical evidence and theoretical notions regarding the Pe.

Error Positivity

The morphology, polarity, and scalp topography of the Pe are similar to those of the P3, a positive slow wave in the stimulus-evoked ERP, which is maximal at centroparietal recording sites and peaks at 300 ms or more after task-relevant stimuli. Because of this similarity, Falkenstein et al. (1991; see also Falkenstein et al., 2000; Leuthold & Sommer, 1999) interpreted the Pe as representing a second P3, elicited by the evaluation of the incorrect response. Thus, according to this interpretation, error trials are characterized by two successive P3 components, the first of which is sensitive to stimulus evaluation and the second of which is sensitive to the evaluation of the incorrect response (which is usually a highly task-relevant event).

In contrast, there has been some debate (e.g., Falkenstein et al., 1995) as to whether the Pe might reflect a delayed (component of the) stimulus-evoked P3 complex contributing to the response-locked ERP. However, there are several arguments favoring the hypothesis of an additional P3-like positivity over this latter possibility. First, although the P3 latency may be delayed to some extent on error trials (Donchin, Gratton, Dupree, & Coles, 1988), various reports of Pe latency would imply implausibly large latency shifts of the P3 on these trials (Falkenstein et al., 1991; Kaiser, Barker, Haenschel, Baldeweg, & Gruzelier, 1997; Leuthold & Sommer, 1999). Second, sometimes two positive peaks can be discerned in the stimulus-locked ERP associated with incorrect responses: one in the latency range of the classic, stimulus-evoked P3, and one after the incorrect response (Falkenstein et al., 1991; Leuthold & Sommer, 1999). A particularly convincing demonstration of two successive P3-like positive components on error trials can be found in Luu, Collins, and Tucker (2000). Using grand-average voltage maps of the scalp, these authors show that the second, error-related positivity has a slightly more anterior distribution than the first positivity, which at the same latency is also present on correct trials, and which is presumably the stimulus-evoked P3. Note that, if the Pe reflects a second P3, it will usually overlap to some extent with the stimulus-evoked P3, certainly if the average reaction time of errors is fast, as is the case in many Ne/Pe studies. Third, the Pe appears to be more pronounced in response-locked averages than in stimulus-locked averages (e.g., Falkenstein et al., 1991; Leuthold & Sommer, 1999), suggesting that it is related to the response rather than the stimulus. And fourth, it appears that the Pe and P3 are differentially affected by several empirical conditions (see Falkenstein et al., 1991, 2000; Leuthold & Sommer, 1999). However, it is not clear to what extent this can be attributed to variable contributions of the stimulus-locked components to the response-locked ERP, due to differences in RT distribution between the empirical conditions. To summarize, the available evidence suggests that an additional, response-related positivity, the Pe, is present on incorrect response trials. Several characteristics of the Pe suggest that it is a P3-like wave, which is elicited by the error event.

Not much work has been done to establish the functional significance of the Pe. The Pe has been observed on corrected and uncorrected error trials, as well as on false alarm trials (see Falkenstein et al., 2000). Hence, it has been argued that the Pe cannot be a correlate of an error correction process (Falkenstein et al., 2000). As noted above, an alternative interpretation of the Pe is that it reflects the conscious processing of the error event (Falkenstein et al., 1991, 2000). As such, it may be directly related to the controlled adjustment of response strategies following the recognition of an error, or it may reflect the conscious recognition of the error itself. To our knowledge, the relation between the Pe and indices of response strategy adjustments (e.g., posterror slowing; Rabbitt, 1966) remains to be empirically investigated. However, some evidence suggests a relation between the Pe and conscious error recognition. First, Vidal, Hasbroucq, Grapperon, and Bonnet (2000) have reported that in a two-choice manual response task overt motor errors are followed by an Ne and a Pe. In contrast, trials that were characterized by subthreshold muscle activation of the incorrect hand followed by the appropriate overt response showed a clear Ne but no Pe. It can be argued that a Pe was missing in these trials because subthreshold errors may be less likely to reach awareness, especially if they are followed by a correct button press. Second, Kaiser et al. (1997) examined to what extent the Ne and Pe were sensitive to the effects of hypnosis. Their study focused on errors that were committed by participants performing a simple S-R compatibility task. In a group of participants characterized as highly susceptible to hypnosis, the Pe was essentially abolished under hypnosis as compared to a prehypnosis baseline that showed a clear Pe. Importantly, the Ne was fully unaffected in the hypnosis condition. To the extent that hypnosis can be regarded as a state of altered or lacking consciousness, this result appears nicely compatible with the hypothesis that the Pe is related to the conscious recognition of errors whereas the Ne is not.

The Antisaccade Paradigm

The antisaccade task (e.g., Hallet, 1978; Nieuwenhuis, Ridderinkhof, De Jong, Kok, & Van der Molen, 2000; Roberts, Hager, & Heron, 1994) provides a suitable laboratory paradigm for the study of overt motor responses that do not reach awareness. In this task, participants are instructed to generate a saccade to the opposite side of a peripheral abrupt onset cue (i.e., an *antisaccade*). Because abrupt onset stimuli are well known to capture the eyes in an automatic fashion (e.g., Theeuwes et al., 1998), correct antisaccade performance requires controlled inhibition of the reflexive saccade towards the abrupt onset cue. A direction error is defined by a saccade in the direction of the cue, even if the eyes are immediately redirected to the opposite location. Despite the task instructions, participants typically make many direction errors in the antisaccade task.

Recently, several authors have reported that participants performing an antisaccade task are often not aware of their direction errors when asked immediately after a trial (Deubel, Mokler, Fischer, & Schneider, 1999; Mokler & Fischer, 1999; for related evidence see Kramer et al., 2000; Theeuwes et al., 1998). This finding is especially surprising because the unperceived direction errors, just as their perceived counterparts, are typically followed by a corrective saccade of approximately twice the size of the initial error. In Mokler and Fischer, for instance, participants were instructed to look to the opposite side of an abrupt onset cue that was presented 4 deg either to the left or to the right from a central

fixation point. At the end of each trial, they were to indicate, by means of a button press, whether they believed they had made a direction error. The average percentage of direction errors was 19%, and almost all of them were immediately followed by a large corrective saccade, necessary to redirect the eyes towards the target location. Importantly, of all the direction errors, only 50% were perceived by the participants as such.

If most saccade direction errors are followed by a corrective saccade, why then are they so often not consciously recognized by participants as errors? At this point, it seems reasonable to recall that participants in the studies mentioned above were asked to give their awareness judgments *at the end* of each trial. Thus, in the case of an unperceived error, we cannot be completely certain whether the participant was in fact briefly aware of the error while it was happening. Awareness of an event is used here in a pragmatic sense, that is, in terms of whether information about the event is available for report briefly after it has taken place (Scheffers & Coles, 2000, refer to this as "perceived accuracy"). The question, then, is why participants fail to report so many of their errors. From a neurophysiological perspective, it has been argued (see Mokler & Fischer, 1999) that saccade direction errors have a significant chance of remaining unrecognized, because they are generated mostly subcortically (Schiller, 1998). The increased size of the required corrective antisaccade (compared to a direct antisaccade) probably remains unnoticed because voluntary saccades are specified in terms of the desired end position of the eye in the orbit (Mays & Sparks, 1980; Schiller, 1998), which is the same in either case. Another interesting explanation, at the psychological level, is offered in the Discussion.

As in Mokler and Fischer (1999), participants in the present experiment were to judge their antisaccade performance at the end of each trial. The stimulus events and their relative timing are illustrated in Figure 1. Participants were instructed to suppress a reflexive saccade, and to generate a direct antisaccade away from the abrupt onset cue as fast as they could. One second after cue onset, a cross was presented at the target location. At this point, participants were to push a response button if and only if they thought they had moved their eyes in the direction of the cue. To increase the proportion of direction errors, the task design included

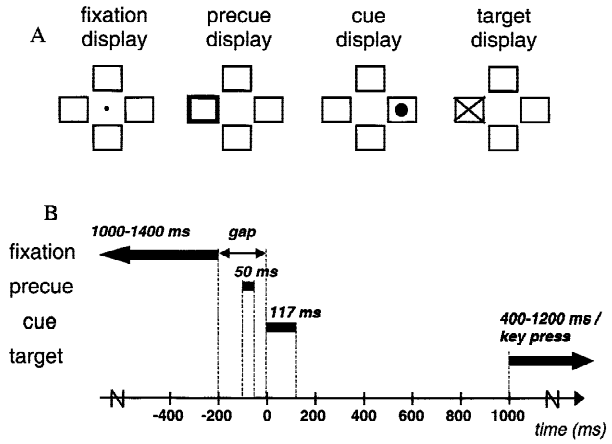

Figure 1. A: Example of stimulus displays in the present experiment. See text for actual size. B: Relative timing and presentation duration of stimulus events in the present experiment. Time $t = 0$ ms corresponds to the moment of cue onset. Note that the probability of a precue being presented was 50%.

the following features. First, saccades were required in both the horizontal and the vertical dimension. Increasing the number of possible cue locations from two to four decreases antisaccade performance when locations are positioned along two orthogonal axes (Delaney & Roberts, 1999). Second, a gap of 200 ms was introduced between fixation point offset and cue onset. This feature is well known to increase error rates in the antisaccade task (e.g., Munoz, Broughton, Goldring, & Armstrong, 1998). A third feature included to increase the error rates was the use of a precue (in approximately 50% of the trials, to prevent habituation; e.g., Fischer & Weber, 1996), validly indicating the target location on each presentation.[1]

Thus, the antisaccade paradigm permitted the study of error-related brain activity in an experimental setting in which (a) participants were always aware of the *required* response, as indicated by the rapid correction of both perceived and unperceived direction errors; (b) at some level in the participants' information processing system there was information about the *actual* incorrect response (This seems trivial because the system produces that response itself.); and (c) participants were often not aware of response errors. If the monitoring processes manifested by the Ne and Pe are related to the processes associated with participants' subjective identification of errors, then the associated components should be substantially affected after saccade direction errors that were not recognized by participants as such. More specifically, in line with previous empirical (Kaiser et al., 1997; Vidal et al., 2000) and theoretical (Falkenstein et al., 1991; 2000) contributions to the literature, we expected that awareness of response errors would significantly correlate with Pe measures but not with Ne measures.

Method

Participants

Fifteen undergraduate students (12 women) from the University of Amsterdam participated in this experiment. The participants, ranging in age from 18 to 23 years, received course credit for their participation. Three participants were replaced because their proportion of perceived or unperceived errors in the relevant (horizontal; see Data Analysis section) dimension was smaller than 2.5%.

Stimuli

Stimuli were presented on a black computer screen (see Figure 1A). The fixation display consisted of a white fixation dot, subtending 0.6 deg, surrounded by four yellow square outlines, each subtending 3.4 deg, that were symmetrically positioned above, below, to the left and right of fixation. The visual angle between fixation and the center of each square was 10 deg. The precue consisted of the brief thickening of one of the outlines. The cue consisted of a white circle, subtending 1.2 deg, and was presented in the center of a square. The "target" consisted of a white cross filling one of the squares.

Design and Procedure

The experiment involved one session that lasted approximately 1.5 hr. After having received 50 trials of practice with the task,

[1]The exact mechanism by which valid precues increase error rates is as yet unknown. Fischer and Weber (1996) propose that the orienting mechanism has an automatic tendency to orient away from stimuli when trained in an antisaccade setting. This automatic tendency is detrimental to performance when a valid precue is presented.

participants entered the experimental phase, which consisted of eight blocks of 100 trials. The order of cue locations was randomly determined for each new block. Cue location could be either above, below, to the left or to the right of fixation (10, 10, 40, 40 trials per block, respectively; see below for a justification). The probability of a precue being presented was 50% on each trial.

The temporal order of stimulus events is illustrated in Figure 1B. Participants received instructions to direct their eyes at fixation at the start of each trial. After a random duration of 1,000–1,400 ms the fixation point disappeared and, after a gap period of 200 ms, the cue was presented for 117 ms. In case a precue was presented, the outline opposite to the cue was brightened for 50 ms, starting 100 ms before cue onset. Participants were instructed to move their eyes to the target location as soon as possible after cue onset. After an interval of 1,000 ms, starting at cue onset, the cross was displayed until the space bar was pressed (indicating a perceived error), with a minimum of 400 ms (because of the feedback function of the cross—"this is where your eyes should be now") and a maximum of 1,250 ms (in case of no key press), after which the next trial began. Participants were instructed to press the space bar if, and only if, they thought they had moved their eyes in the direction of the cue. To prevent hand movements during the cue-target interval, it was emphasized that possible key presses had to be made while the cross was visible. Finally, a rest break of 10 min was allowed halfway through the experiment.

Psychophysiological Recording

Recordings of the electroencephalogram (EEG) were made from Fz, Cz, and Pz, using an ECI electrocap, and were referenced to the activity recorded at an electrode on the left earlobe. The electrooculogram (EOG) was recorded from tin electrodes placed above and below the left eye (vertical EOG) and from electrodes lateral to each eye (horizontal EOG). A ground electrode was positioned on the forehead. All electrode impedances were below 10 kΩ. The EEG and EOG signals were amplified using a Nihon-Kohden system with a time constant set to 5 s and a low-pass filter of 35 Hz. The signals were digitized with a sample rate of 500 Hz. Single trial epochs with a duration of 2,048 ms (including a 250-ms prestimulus baseline) were extracted off-line.

Eye movements were recorded with an infrared-based iView eye tracker (SMI) with 50-Hz temporal resolution and a <0.1 deg spatial resolution. The head was stabilized by means of a chin rest, which was located 48 cm from the monitor. Before the start of each experimental block, participants fixated three series of five calibration targets that were presented on the screen, one at a time and in the shape of a plus symbol. The iView system was calibrated using standard techniques involving the computation of the linear regression of target location on the average eye position signal.

Data Analysis

The single trial EEG signals were corrected for vertical EOG artifacts, using the algorithm described by Woestenburg, Verbaten and Slangen (1983). The method for dealing with horizontal EOG artifacts is described in a subsequent section. Then, for each participant and each condition, the EEG epochs were averaged in order to obtain event-related potential waveforms that were time-locked with respect to the eye movement onset on each trial. Measures of the Ne amplitude and Pe amplitude were derived from the average individual difference waveforms (perceived errors minus correct, and unperceived errors minus correct). The Ne amplitude was defined as the difference between the most negative peak of the difference signal from 0 to 150 ms after saccade onset and the most positive peak in the 150 ms preceding saccade onset. The Pe amplitude was defined as the average amplitude of the difference signal in a window from 200 to 400 ms after saccade onset, relative to the average amplitude in a window from 100 to 60 ms preceding saccade onset. The latter window was chosen to ensure that it preceded the first saccade-related activity, as recorded at the horizontal EOG electrodes.

From the eye movement data, several dependent measures were determined off-line. A saccade onset was detected if the velocity signal exceeded 25 deg/s. Saccadic reaction times (SRTs) were defined as the time between the onset of the cue and the detected saccade onset. Direction errors were limited to those trials in which the first saccade was in the direction of the cue. Corrective saccades were defined as those saccades that followed a direction error and were opposite in sign from the erratic saccade. Finally, saccadic correction time was defined as the time between the onset of the direction error and the onset of the corrective saccade (using the same velocity criterion for the corrective saccade). If a key press by the participant was recorded after a direction error, the trial was classified as a perceived error. The other direction errors were classified as unperceived. A correct eye movement followed by a key press was classified as false alarm.

The following trials were not included in the analyses: (a) trials that required a saccade in the vertical dimension (one-fifth of all trials). Note that the vertical saccade dimension was only included in the design in order to increase the overall error rate. These trials were discarded from the analyses because the EOG activity associated with vertical eye movements produces too much artifact in the EEG signals. (b) Trials on which the initial saccade was erroneously in the vertical dimension. (c) Trials on which no eye movement was detected. (d) Trials with SRTs shorter than 80 ms (anticipations) or with SRTs longer than 600 ms. Trials with long SRTs could not be used because an epoch of 400 ms after the response was needed for the computation of the Ne and the Pe. (e) Trials with recording artifacts. For reasons b–e, roughly 10% of the experimental trials (those requiring a horizontal saccade) were discarded.

Analysis of variance (ANOVA) with repeated measures was employed to assess the experimental effects on SRT, saccade size, saccadic correction time (after direction errors), Ne amplitude, and Pe amplitude. The variables in the ANOVA design were electrode site (for the psychophysiological measures), and trial type (correct, perceived error, unperceived error).

Results

Performance Data

Table 1 shows mean values of the eye movement indices for correct antisaccades, perceived errors, and unperceived to errors. No systematic differences were found between leftward and rightward saccades, so these were pooled together. The main statistical analyses yielded the following results. First, the percentage of unperceived errors was reliably larger than the percentage of perceived errors, $F(1,14) = 5.2$, $p < .05$. The majority of the perceived errors ($M = 85\%$, $SE = 4\%$) and of the unperceived errors ($M = 79\%$, $SE = 3\%$) were made on trials with a precue. Because for many participants there were not sufficient errors on trials without a precue to compute reliable averages of behavioral and electrophysiological data, the factor Precue (present/absent) was not included in any of the statistical analyses. Second, correct antisaccades and the two types of errors reliably differed with

Table 1. *Summary of Performance Data: Percentage of Trials, Saccade Size, Saccadic Reaction Time (SRT), and Saccadic Correction Time as a Function of Trial Type*

Trial type	% Trials	Size (deg visual angle)	SRT (ms)	Correction time (ms)
Correct				
M	79.9	9.8	323	N/A
SD	(9.5)	(0.8)	(52)	N/A
Perceived errors				
M	7.4	8.7	194	314
SD	(3.3)	(0.7)	(54)	(86)
Unperceived errors				
M	11.2	7.8	200	200
SD	(7.1)	(0.7)	(46)	(64)

Note: 1.5% of the trials were categorized as false alarms.

respect to the size of the saccade, $F(2,28) = 86.8$, $p < .001$. Pairwise comparisons indicated that the correct antisaccades were reliably larger than the two types of error saccades, both p's $< .001$, and that perceived error saccades were larger than unperceived error saccades, $p < .001$. Third, the main effect of trial type on SRT was reliable, $F(2,28) = 162.3$, $p < .001$. Perceived and unperceived errors were associated with essentially equal SRTs, and were initiated much faster than correct antisaccades. Finally, virtually all errors (approximately 97%) were followed by a corrective saccade. However, unperceived errors were corrected reliably faster than perceived errors, $F(1,14) = 29.3$, $p < .001$. Importantly, the mean percentage of false alarms (1.5%) was very small, indicating that participants very rarely reported an error when in fact they had made a correct antisaccade.

Figure 2 presents the scatterplots of the correction times versus the error size for perceived and unperceived errors separately. Note that the saccade size and correction time values associated with the two types of errors show a large overlap in terms of their distributions. The finding that perceived errors are, on average, corrected more slowly is due to a relatively more pronounced slow tail of the distribution. Likewise, the finding that unperceived errors are, on average, smaller in size is due to a relatively high number of saccades of very small size.

Psychophysiological Data

The analysis needed to evaluate the relationship between awareness of a response error and the error-related ERP components consisted of several steps. First, before deriving individual response-locked waveforms in each of the conditions, we selected a subset of the correct trials with SRTs matching the SRTs of the incorrect trials. Otherwise, any differences between response-locked waveforms of correct and error trials could be attributed to different contributions of the stimulus-related ERP activity overlapping the response-locked ERP activity (see, e.g., Gehring et al., 1993; Scheffers et al., 1996). Thus, to equalize the mean RT of the correct and incorrect trials, correct trials were selected if their SRT was equal to or faster than the SRT associated with the 10th fastest percentile of the SRT distribution. Averaged across participants, the resulting subset of correct trials was comparable with the two types of error trials with respect to both SRT ($M = 214$ ms, $SD = 50$ ms) and number of trials (10.6% of the overall set of included trials). In a second step, we computed the response-locked waveforms in each of the conditions of interest for leftward and rightward saccades separately. Figure 3 depicts the grand-average horizontal EOG waveforms associated with leftward and rightward saccades. Note that the EOG patterns reflect the differences between the three trial types in correction time, as reported above. Subsequently, we averaged each of the response-locked wave-

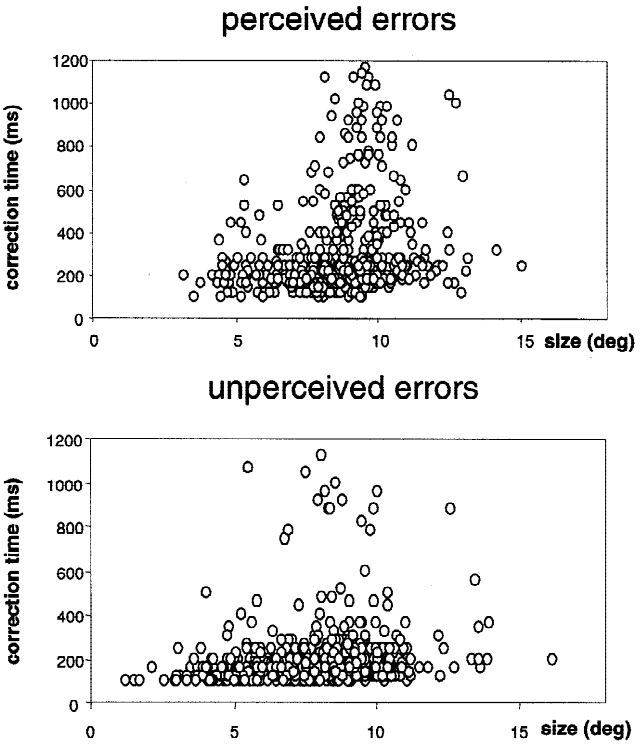

Figure 2. Scatterplots including all observations of correction time (i.e., the time between the onsets of the direction error and the corrective saccade) versus saccade size for perceived direction errors (upper panel) and unperceived direction errors (lower panel). Note that the cue was presented at 10 deg from fixation.

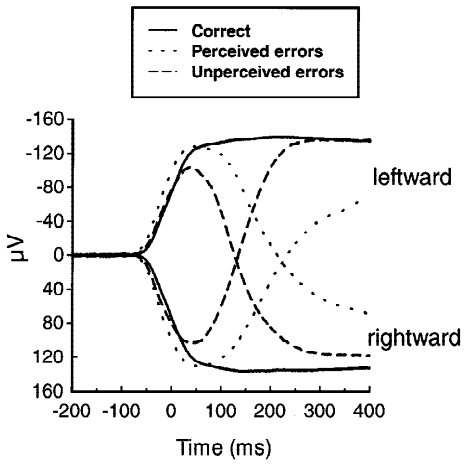

Figure 3. Response-locked horizontal EOG waveforms for leftward and rightward correct antisaccades, perceived direction errors, and unperceived direction errors.

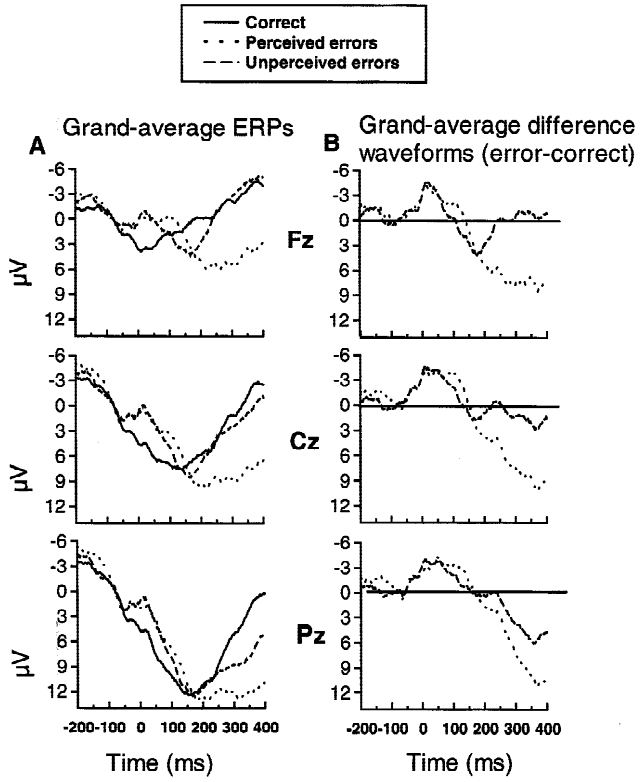

Figure 4. A: Original, grand-average response-locked ERP waveforms for each trial type. B: Grand-average response-locked ERP difference waveforms for (perceived errors minus correct) and (unperceived errors minus correct).

forms across saccade direction. Importantly, as can be inferred from Figure 3, the resulting EOG waveforms were essentially flat. Hence, their contribution to the newly obtained response-locked waveforms (averaged across saccade direction) must be negligible. Indeed, within the plotted time interval [−200 ms, 400 ms] all three EOG waveforms remained within the range [$-25\ \mu$V, $25\ \mu$V].

Figure 4A presents grand-average response-locked ERPs associated with the three midline electrodes for SRT-matched correct trials, perceived errors, and unperceived errors. Figure 4B presents the grand-average difference waveforms (perceived errors-correct) and (unperceived errors-correct). Importantly, as is evident from Figures 4A and 4B, a clear negative deflection is elicited on *both* perceived and unperceived error trials. These deflections appear to start at approximately the time of onset of the "incorrect" EOG activity (cf. Gehring et al., 1993) and peak about 80 ms later. The morphology and latency of these negative waves correspond to those of the Ne reported in previous studies (e.g., Scheffers & Coles, 2000; Van't Ent & Apkarian, 1999). An ANOVA on the Ne amplitudes of the individual difference waveforms[2] revealed a highly significant Ne following errors, $F(1,14) = 72.6$, $p < .001$, which did not differ across electrode site, $F < 1$. Importantly, there was no reliable difference between the Ne amplitudes for perceived and unperceived errors, $F(1,14) = 1.3$, $p = .28$. Also, the interaction between trial type and electrode site was not significant, $F < 1$.

Although perceived and unperceived errors did not differ with respect to the Ne, Figures 4A and 4B show that the subsequent positivity was notably different for the two types of errors. The perceived errors were associated with a prominent slow positive wave, which, as can be seen in Figure 4A, overlapped with the trailing slope of a preceding slow positive wave (probably the stimulus-evoked P3), and which we believe to be the Pe. In contrast, on unperceived error trials this additional positivity was much smaller, and most clearly so on Fz and Cz. An ANOVA on the Pe amplitudes of the individual difference waveforms revealed a significant Pe when averaged across perceived and unperceived errors, $F(1,14) = 19.2$, $p < .001$. The main effect of electrode site was not significant, $F(2,28) = 1.4$, $p = .26$. Importantly, the main effect of trial type was highly reliable, $F(1,14) = 14.7$, $p < .003$, indicating that the Pe was larger on perceived error trials. Although the effect of trial type appeared to differ substantially between the three electrode sites, the interaction was only marginally significant, $F(2,28) = 4.1$, $p = .06$.

Posterror Slowing

To determine whether the Ne or Pe on incorrect trials was related to some form of remedial action associated with the prevention of future errors, we performed an additional analysis focusing on the degree of posterror slowing (e.g., Rabbitt, 1966; 1967) of correct SRT after perceived and unperceived errors. This analysis involved a comparison between SRT on correct trials immediately following perceived errors and correct trials immediately following unperceived errors. If, as has been hypothesized (Rabbitt, 1966), posterror slowing is an intentionally controlled process, one would expect no slowing of SRT after unperceived errors. If indeed this were the case, and given that Ne amplitude was similar for perceived and unperceived errors, this would argue against the involvement of the Ne system in this aspect of remedial action (cf. Gehring et al., 1993). The analysis showed that, indeed, perceived errors were associated with substantial posterror slowing (SRT = 349 ms; recall that overall correct SRT was 323 ms), whereas there was no slowing on correct trials following unperceived errors (SRT = 321 ms). An ANOVA revealed a significant effect of trial type (perceived vs. unperceived errors) on posterror SRT, $F(1,14) = 11.9$, $p < .005$.

Discussion

The present study examined the relation between error processing, as reflected by the Ne and Pe, and conscious self-monitoring of behavior. To this end, we investigated error-related brain activity using a variant of the antisaccade paradigm. As expected, participants made many reflex-like eye-movement errors, many of which remained unperceived. The eye-movement data showed the same patterns of results as those reported by Mokler and Fischer (1999). On average, perceived and unperceived direction errors showed

[2]The reported statistical analyses of Ne and Pe amplitudes were performed on the difference waves rather than, as is usually recommendable, on the original waveforms, because the shape of the waveforms for correct trials (see Figure 4A) did not allow the detection of a negative peak in the Ne latency range. Although under some conditions difference waveforms may distort the spatial and temporal structure of the original waveforms, an additional analysis on the original waveforms (perceived vs. unperceived for the Ne, all three trial types for the Pe) yielded the same pattern of results as the analysis on the difference waveforms. In addition, pairwise comparisons indicated that, though strongly reduced compared to perceived error trials, the Pe on unperceived error trials was nevertheless marginally significant compared to the positivity on correct trials, $F(1,14) = 4.7$, $p = .047$.

similar onset latencies, but unperceived direction errors were smaller in size and were corrected faster than perceived direction errors despite a large overlap in distributions. A possible explanation for these results will be given below. Here, it suffices to note that apparently the two classes of errors differed in some aspects other than their accessibility for subjective report. This, in combination with the low false alarm rate, suggests that participants identified errors according to some meaningful, physical criterion, although it is not yet known what this criterion is.[3]

Our psychophysiological results provide the first demonstration of an Ne and Pe after saccade direction errors or *errors of choice* in the oculomotor domain. In a recent study, Van't Ent and Apkarian (1999) already reported an Ne after No-go saccade errors or *errors of action*. These findings further emphasize the notion of a generic error-processing system that acts irrespective of response modality (Holroyd et al., 1998), and type of error (Scheffers et al., 1996). Most importantly, we observed a substantial Ne after both perceived and unperceived errors. Indeed, if anything, the amplitude of the Ne on unperceived error trials was underestimated rather than overestimated, because error size (which was relatively smaller for unperceived error saccades) has previously been found to vary positively with Ne amplitude (Bernstein et al., 1995; Scheffers et al., 1996; Van't Ent & Apkarian, 1999; but see Gehring et al., 1993). In contrast, the Pe, though clearly present on perceived error trials, was remarkably reduced on unperceived error trials. Interestingly, unperceived error trials were also characterized by the absence of posterror slowing, which is an important manifestation of remedial action following errors (Rabbitt, 1966, 1967).

With regard to the Ne, the present results indicate that the processes that lead to the Ne and those that lead to awareness of the error do not necessarily have access to the same information (cf. Scheffers & Coles, 2000). More specifically, it appears that a representation of the actual response, needed for the computation of the Ne, can be derived from the subcortical structures that are involved in saccade programming (e.g., Schiller, 1998), but whose activity often escapes awareness. Note that previous Ne experiments, which have used manual responses, must have involved erroneous motor commands generated at the level of the motor cortex. In addition, because the Ne was fully intact on unperceived error trials, the absence of posterror slowing after these errors argues against the involvement of the Ne monitoring system in this manifestation of remedial action. A dissociation between Ne activity and some aspects of remedial action was also reported by Gehring and Knight (2000), who studied the effects of lateral prefrontal damage on error monitoring. It is not clear how Gehring et al.'s (1993) report of a significant, positive relation between Ne amplitude and correct RT on the next trial can be reconciled with these results. A plausible possibility is that the amount of remedial action and Ne amplitude are not directly related, but were jointly influenced by a third, hidden variable in the study by Gehring and colleagues.

Regarding the Pe, the present findings are nicely in line with the idea that the error positivity manifests the conscious recognition of an error, and that this is reflected in a P3-like wave elicited by the error event (see Falkenstein et al., 1991, 2000; Leuthold & Sommer, 1999). That is, on those trials on which participants were not aware of a direction error, the Pe was strongly reduced. The residual error-related positivity on these trials can be argued to emanate from trials on which participants were only partially aware of an error but decided not to report it. Although the finding that posterror slowing and Pe amplitude covaried in a meaningful way does not necessarily imply a causal relation between the two, it is certainly *consistent* with another hypothesis raised by Falkenstein et al. (2000), namely that the Pe reflects the adjustment of response settings after errors. One way to test this hypothesis more firmly would be to compute the within-subjects correlation between Pe amplitude and the degree of posterror slowing or the probability of an error on the subsequent trial.

Given the finding that perceived and unperceived errors differed significantly with respect to saccade size and correction time (see also Mokler & Fischer, 1999), one may ask whether the presence or absence of awareness was perhaps directly related to one of these variables. More specifically, unperceived erroneous saccades may have remained unrecognized because of their short size, or, alternatively, because they were corrected so rapidly. However, both of these explanations seem inadequate when the large overlap between the distributions of these variables of both types of errors is regarded (see Figure 2). This amount of overlap suggests that the presence or absence of awareness cannot be a simple, deterministic function of size or correction time. Recently, some authors (Deubel et al., 1999; Mokler & Fischer, 1999) have proposed that awareness depends on whether covert attention accompanies the eyes to the incorrect location or moves to the correct location at once. If attention accompanies the eyes to the correct location, then the error is recognized as such. In contrast, if attention moves directly to the correct location while the eyes go in the opposite direction, then the subjects' phenomenal experience is that of a correct saccade being made, whereas in reality the eyes arrive at the correct location with a relatively long delay. According to this view, the correction times of perceived errors are increased because it takes more time to redirect both the eyes and covert attention to the correct location than it takes to redirect the eyes alone, as in the case of unperceived errors. It is not clear how this view explains the differences in saccade size between the two types of errors. Interestingly, some initial evidence for this "attention hypothesis" has recently been reported by Deubel et al.

We conclude that our results support the existence of two error monitoring processes. At a very early stage, which is reflected by the Ne, incorrect motor commands generated at both the cortical level and subcortical level are detected via a central processing pathway. This error detection process operates independently of conscious error perception, and is not directly involved in remedial action. A later error monitoring process, reflected by the Pe, is strongly associated with awareness of the occurrence of the actual (erroneous) response. We hypothesize that the input to the system operating at this stage may be formed by peripheral (e.g., proprioceptive or reafferent) feedback about the response. The present results also suggest that the system underlying the Pe may be related to or may implement itself a remedial action mechanism. Of course, if the view that the Pe and stimulus-evoked P3 are different manifestations of the same biological and/or functional system turns out to be true, then the Pe will eventually have to be interpreted in terms of a broader theory of P3-related brain activity. Importantly, the present view can explain why an intact Ne but no Pe was observed after subthreshold motor errors by Vidal et al. (2000). Depending on one's theory of hypnosis, it also presents a

[3]However, we do not claim that the resulting classes of errors were necessarily homogenous in the sense that one class consisted only of errors of which participants were *fully* aware, and the other class consisted only of trials of which participants were *fully* unaware. Because of its binary character, the required classification was rather too crude to capture possible subtleties in participants' confidence about the correctness of their response.

possible explanation for the results reported by Kaiser et al. (1997), showing that under conditions of hypnosis highly susceptible participants showed a clearly reduced Pe but no change in Ne. Finally, although the present conclusions seem at least valid for the current paradigm and oculomotor response modality, an interesting question for future research is whether the present results generalize to other response modalities and experimental settings.

REFERENCES

Bernstein, P. S., Scheffers, M. K., & Coles, M. G. H. (1995). Where did I go wrong? A psychophysiological analysis of error detection. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 1312–1322.

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science*, *280*, 747–749.

Cooke, J. D., & Diggles, V. A. (1984). Rapid error correction during human arm movements: Evidence for central monitoring. *Journal of Motor Behavior*, *16*, 348–363.

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

Delaney, T., & Roberts, R. J. (1999). Experimental analysis of a competitive timing model of antisaccade performance in humans. Paper presented at the meeting of the 10th European Conference on Eye Movements, Utrecht, the Netherlands.

Deubel, H., Mokler, A., Fischer, B., & Schneider, W. A. X. (1999). Reflexive saccades are not preceded by shifts of visual attention: Evidence from an antisaccade task. Paper presented at the meeting of the 10th European Conference on Eye Movements, Utrecht, the Netherlands.

Donchin, E., Gratton, G., Dupree, D., & Coles, M. (1988). After a rash action: Latency and amplitude of the P300 following fast guesses. In: G. C. Galbraith, M. L. Kietzman, & E. Donchin (Eds.), *Neurophysiology and psychophysiology: Experimental and clinical applications* (pp. 173–188). Hillsdale, NJ: Erlbaum.

Falkenstein, M., Hohnsbein, J., & Hoormann, J. (1995). Event-related potential correlates of errors in reaction tasks. In: G. Karmos, M. Molnar, V. Scépe, I. Czigler, & J. E. Desmedt (Eds.), *Perspectives of event-related potentials research* (pp. 287–296). Amsterdam: Elsevier.

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

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

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

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

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

Gehring, W. J., & Knight, R. T. (2000). Prefrontal-cingulate interactions in action monitoring. *Nature Neuroscience*, *3*, 516–520.

Hallet, P. E. (1978). Primary and secondary saccades to goals defined by instructions. *Vision Research*, *18*, 1279–1296.

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

Kaiser, J., Barker, R., Haenschel, C., Baldeweg, T., & Gruzelier, J. H. (1997). Hypnosis and event-related potential correlates of error processing in a stroop-type paradigm: A test of the frontal hypothesis. *International Journal of Psychophysiology*, *27*, 215–222.

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

Kramer, A. F., Hahn, S., Irwin, D. E., & Theeuwes, J. (2000). Age differences in the control of looking behavior: Do you know where your eyes have been? *Psychological Science*, *11*, 210–217.

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

Luu, P., Collins, P., & Tucker, D. M. (2000). Mood, personality, and self-monitoring: Negative affect and emotionality in relation to frontal lobe mechanisms of error monitoring. *Journal of Experimental Psychology: General*, *129*, 43–60.

Luu, P., Flaisch, T., & Tucker, D. M. (2000). Medial frontal cortex in action monitoring. *Journal of Neuroscience*, *20*, 464–469.

Mays, L. E., & Sparks, D. L. (1980). Saccades are spatially, not retinocentrically, coded. *Science*, *208*, 1163–1165.

Mokler, A., & Fischer, B. (1999). The recognition and correction of involuntary prosaccades in an antisaccade task. *Experimental Brain Research*, *125*, 511–516.

Munoz, D. P., Broughton, J. R., Goldring, J. E., & Armstrong, I. T. (1998). Age-related performance of human subjects on saccadic eye movement tasks. *Experimental Brain Research*, *121*, 391–400.

Nieuwenhuis, S., Ridderinkhof, K. R., De Jong, R., Kok, A., & Van der Molen, M. W. (2000). Inhibitory inefficiency and failures of intention activation: Age-related decline in the control of saccadic eye movements. *Psychology and Aging*, *15*, 635–647.

Rabbitt, P. M. A. (1966). Errors and error correction in choice reaction tasks. *Journal of Experimental Psychology*, *71*, 264–272.

Rabbitt, P. M. A. (1967). Three kinds of error-signalling responses in a serial choice task. *Quarterly Journal of Experimental Psychology*, *20*, 179–188.

Roberts, R. J., Hager, L. D., & Heron, C. (1994). Prefrontal cognitive processes: Working memory and inhibition in the antisaccade task. *Journal of Experimental Psychology: General*, *123*, 374–393.

Scheffers, M. K., & Coles, M. G. H. (2000). Performance monitoring in a confusing world: Event-related brain activity, judgements of response accuracy, and types of errors. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 141–151.

Scheffers, M. K., Coles, M. G. H., Bernstein, P., Gehring, W. J., & Donchin, E. (1996). Event-related brain potentials and error-related processing: An analysis of incorrect responses to go and no-go stimuli. *Psychophysiology*, *33*, 42–53.

Scheffers, M. K., Humphrey, D. G., Stanny, R. R., Kramer, A. F., & Coles, M. G. H. (1999). Error-related processing during a period of extended wakefulness. *Psychophysiology*, *36*, 149–157.

Schiller, P. H. (1998). The neural control of visually guided eye movements. In: J. E. Richards (Ed.), *Cognitive neuroscience of attention. A developmental perspective* (pp. 3–50). Mahwah, NJ: Erlbaum.

Theeuwes, J., Kramer, A. F., Hahn, S., & Irwin, D. E. (1998). Our eyes do not always go where we want them to go: Capture of the eyes by new objects. *Psychological Science*, *9*, 379–385.

Van't Ent, D., & Apkarian, P. (1999). Motoric response inhibition in finger movement and saccadic eye movement: A comparative study. *Clinical Neurophysiology*, *110*, 1058–1072.

Vidal, F., Hasbroucq, T., Grapperon, J., & Bonnet, M. (2000). Is the 'error negativity' specific of errors? *Biological Psychology*, *51*, 109–128.

Woestenburg, J. C., Verbaten, M. N., & Slangen J. L. (1983). The removal of the eye-movement artifact from the EEG by regression analysis in the frequency domain. *Biological Psychology*, *16*, 127–147.

(Received August 21, 2000; Accepted February 1, 2001)