Research Report

Errors Are Aversive

Defensive Motivation and the Error-Related Negativity

Greg Hajcak and Dan Foti

Stony Brook University

ABSTRACT—*The error-related negativity (ERN) is a negative deflection in the event-related potential that is maximal approximately 50 ms after the commission of an error. The ERN is generated in the anterior cingulate cortex, a region of the medial prefrontal cortex implicated in both cognitive and emotional processing. Affective and motivational variables influence the magnitude of the ERN, which suggests that the ERN may relate to emotional or motivational aspects of error detection. In the present study, we evaluated the possibility that errors prime defensive motivational responses, and that the ERN may predict the magnitude of defensive reactivity after errors. We found that (a) the defensive startle response was larger following errors than following correct responses, and (b) the magnitude of the ERN predicted the degree of startle potentiation following errors. Thus, response errors prime defensive motivation—and the ERN predicts individual differences in the aversive response to errors.*

Research on error processing has grown over the past 15 years, in part because of the discovery of the error-related negativity (ERN; Falkenstein, Hohnsbein, Hoormann, & Blanke, 1991; Gehring, Coles, Meyer, & Donchin, 1995). The ERN is a sharp negative deflection in the event-related potential (ERP) that peaks approximately 50 ms after an unintended response (Falkenstein, Hoormann, Christ, & Hohnsbein, 2000). The ERN is maximal over fronto-central recording sites, and evidence from source localization (Holroyd, Dien, & Coles, 1998), intracerebral recording (Brazdil, 2005), and magnetoencephalography (Miltner et al., 2003) indicates that the ERN is generated in the anterior cingulate cortex (ACC). The ACC is richly interconnected with both limbic and prefrontal areas of the brain, and is ideally situated to respond to, and integrate, cognitive and affective information (Bush, Luu, & Posner, 2000).

Address correspondence to Greg Hajcak, Department of Psychology, Stony Brook University, Stony Brook, NY 11794-2500, e-mail: greg.hajcak@stonybrook.edu.

Although the ERN is known to occur immediately following the commission of an error, there remains considerable debate as to its functional significance and the specific process it indexes. Computational models suggest that the ERN reflects either the coactivation of erroneous and error-correcting responses (conflict theory; Yeung, Cohen, & Botvinick, 2004) or the evaluation, based on expectations developed during learning history, that events are worse than anticipated (reinforcement-learning theory; Holroyd & Coles, 2002). Both the conflict and reinforcement-learning theories, then, highlight specific cognitive functions subserved by the ACC in their conceptualization of the ERN.

Given the involvement of the ACC in pain, depressed mood, and other distressed states, Luu, Collins, and Tucker (2000) proposed that the ERN may signify affective processing in response to errors. In fact, there is ample evidence that the ERN can be modulated by motivational and affective variables (cf. Hajcak, Moser, Yeung, & Simons, 2005). In between-groups designs, individuals with obsessive-compulsive disorder (Gehring, Himle, & Nisenson, 2000; Hajcak, Franklin, Foa, & Simons, in press; Hajcak & Simons, 2002), depression (Chiu & Deldin, 2007), and more generalized anxiety (Hajcak, McDonald, & Simons, 2003a, 2004) have exhibited increased ERNs compared with control participants—a finding consistent with the notion that affective psychopathology is associated with hyperresponsivity to negative information. Other studies have suggested a possible relation between the ERN and individual differences in punishment sensitivity (Boksem, Tops, Wester, Meijman, & Lorist, 2006; Dikman & Allen, 2000). It is important to note that most of these results cannot be accounted for simply by differences in error rates and reaction time, and they suggest that the ERN varies as a function of motivational states and traits.

In a broader context, errors are maladaptive responses that may place an organism in danger and threaten its safety. It stands to reason that error detection should prompt defensive motivation—and rapidly activate those reflexes and dispositions toward action that protect the organism (Gray, 1994). To date, however, there has been no demonstration of a direct link

between errors and defensive motivation. Although errors prompt heart rate deceleration and skin conductance responses (Hajcak, McDonald, & Simons, 2003b, 2004), this pattern of autonomic reactivity is consistent with either an orienting or a defensive response to errors (Cook & Turpin, 1997).

The human startle response, in contrast, is a well-validated and specific measure of defensive activation (Lang, 1993; Lang, Davis, & Öhman, 2000). The startle response is a reflex in which the body contracts into a defensive posture and is typically measured in humans by recording blink magnitude from the obicularis oculi in response to a sudden and loud acoustic probe. The magnitude of the startle response can be modulated by aversive stimuli and settings: Startle is potentiated when participants view threatening stimuli (Bradley, Codispoti, & Lang, 2006; Bradley, Moulder, & Lang, 2005; Lang et al., 2000) and when they anticipate an electric shock (Curtin, Patrick, Lang, Cacioppo, & Birbaume, 2001; Grillon, Ameli, Merikangas, Woods, & Davis, 1993). These findings are consistent with the notion that aversive stimuli prime defensive reflexes.

In the present study, we sought to determine whether defensive responding would similarly be increased following errors, compared with correct responses. To evaluate this possibility, we measured the human startle response 300 ms following errors and correct responses. We hypothesized that startle responses would be larger after errors than after correct responses. To evaluate the role of startle predictability, we included both predictable and unpredictable startle probes on correct trials. Additionally, we simultaneously measured error-related brain activity to determine whether individuals with larger ERNs would demonstrate greater startle potentiation following errors. If the ERN reflects a motivationally relevant correlate of error processing, larger ERNs should predict greater startle potentiation following errors.

METHOD

Thirty-one undergraduate students (26 male, 5 female) participated in this study. No participants discontinued their participation in the experiment once it had begun, and all received course credit for their participation.

The task was an arrow version of the flankers task (cf. Hajcak et al., 2005). On each trial, five horizontally aligned arrowheads were presented, and participants had to respond to the direction of the central arrowhead by pressing the left or right mouse button. On compatible trials, all five arrowheads pointed in the same direction (either left or right), and on incompatible trials, the central arrowhead pointed in the direction opposite the direction of the flanking arrowheads. Compatible and incompatible trials were equally frequent, and all stimuli were presented for 200 ms with an intertrial interval that varied randomly from 500 to 1,000 ms. Participants performed eight blocks of 30 trials. At the end of each block, participants received performance feedback designed to encourage fast and accurate responding. If performance was 75% correct or lower, the message "Please try to be more accurate" was displayed; performance above 90% correct was followed by "Please try to respond faster"; if performance was between these levels, the message "You're doing a great job" was displayed. All participants performed one practice block of 30 trials.

The startle response was measured using standard procedures for assessing defensive reactivity (Bradley et al., 2005, 2006; Curtin et al., 2001; Grillon et al., 1993; Lang et al., 2000): Startle was elicited with a 105-dB burst of white noise that had a 50-ms duration and instantaneous rise time. The startle probe was delivered on 10% of all trials in the practice block; in the actual experiment, startle probes were presented on 50% of error trials, on 50% of correct trials that followed errors, and on a random 4% of other correct trials. Thus, some startle probes were predictable (after both errors and correct trials that followed errors), whereas others were not (randomly selected correct trials). Accordingly, there were three trial types: error trials, predictable correct trials (i.e., correct trials on which the startle probe was predictable), and unpredictable correct trials (i.e., correct trials on which the startle probe was not predictable). In all cases, the startle probe was presented 300 ms after the response.

Continuous electroencephalographic (EEG) and electromyographic (EMG) activity was recorded using an ActiveTwo head cap and the ActiveTwo BioSemi system (BioSemi, Amsterdam, The Netherlands). Recordings were taken from 64 scalp electrodes based on the ten-twenty system, as well as from two electrodes placed on the left and right mastoids. The electrooculogram (EOG) generated from blinks and eye movements was recorded from four facial electrodes: two approximately 1 cm above and below the participant's right eye, one approximately 1 cm to the left of the left eye, and one approximately 1 cm to the right of the right eye. The startle response was measured with two electrodes placed approximately 12 mm apart under the participant's left eye on the obicularis muscle. As per BioSemi's design, the ground electrode during acquisition was formed by the Common Mode Sense active electrode and the Driven Right Leg passive electrode. All bioelectric signals were digitized on a laboratory microcomputer using ActiView software (BioSemi). Sampling was at 1024 Hz.

Off-line analysis was performed using Brain Vision Analyzer software (Brain Products, Gilching, Germany). For the startle data, EMG activity was band-pass filtered (28–512 Hz; 24 dB/octave roll-off), rectified, then low-pass filtered at 30 Hz (24 dB/octave) and baseline-corrected. Individual trials were examined and rejected if the startle reflex began less than 20 ms following probe onset. The number of excluded startle trials was 1.42 (SD = 1.43) for predictable correct trials, 1.42 (SD = 1.57) for unpredictable correct trials, and 2.26 (SD = 2.42) for error trials.

EEG data were rereferenced to the numeric mean of the mastoids and band-pass filtered with cutoffs of 0.1 and 30 Hz. The EEG was segmented for each trial, beginning 200 ms before

the response and continuing for 800 ms. The EEG was corrected for blinks and eye movements using the method developed by Gratton, Coles, and Donchin (1983). Specific intervals for individual channels were rejected in each trial using a semiautomated procedure, with physiological artifacts identified by the following criteria: a voltage step of more than 50.0 μV between sample points, a voltage difference of more than 300.0 μV within a trial, and a maximum voltage difference of less than 0.50 μV within a 100-ms interval. The average number of trials excluded from ERP analyses was 2.03 ($SD = 5.22$), or 0.87% of possible trials.

Startle response magnitudes and latencies were quantified in terms of the peak in the 20- to 120-ms window after the presentation of the startle probe. The ERN was defined as the average activity in a 0- to 100-ms window following response onset on error trials. The ERN and startle responses were statistically evaluated using SPSS (Version 14.0) General Linear Model software, with Greenhouse-Geisser correction applied to p values associated with multiple-*df*, repeated measures comparisons.

RESULTS

On average, participants made 23.39 ($SD = 9.68$) errors. Figure 1 presents the average startle magnitude as a function of trial type. As hypothesized, the magnitude of the startle response differed as a function of trial type, $F(2, 60) = 5.64, p < .05$. Post hoc one-tailed t tests confirmed that startle magnitudes were larger following errors than following correct trials, whether the startle was predictable, $t(30) = 2.51, p < .01$, or unpredictable, $t(30) = 2.61, p < .01$; startle magnitude did not vary as a function of predictability on correct trials, $t(30) = 0.99, p > .30$. Across participants, the magnitude of the error-potentiated startle did not correlate with the number of errors ($r = .06, p > .75$). Although the latency of the startle response was shorter when the probe was predictable (error trials: $M = 96$ ms, $SD = 24$ ms; predictable correct trials: $M = 92$ ms, $SD = 20$ ms) than when it was unpredictable (unpredictable correct trials: $M = 100$ ms, $SD = 20$ ms), this difference was not statistically significant, $F(2, 60) = 1.46, p > .20$.

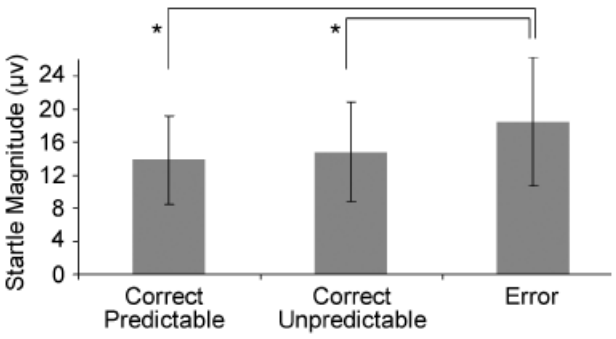

Fig. 1. Mean startle magnitude as a function of trial type. Error bars represent ± 2.04 standard errors, the 95% confidence interval. Asterisks indicate significant differences between trial types, *$p < .01$.

The illustration in Figure 2 shows the scalp topography of the difference in ERPs between error and correct trials in the 0- to 100-ms window following responses. At FCz, where the difference was numerically maximal, a one-sample t test confirmed that the ERN was significantly less than zero, $t(30) = 3.83$, $p < .001$. Grand-average ERP waveforms for correct and error trials at FCz are presented on the right in Figure 1. The ERN is evident as the negative deflection that peaks approximately 50 ms after an incorrect response.

To determine whether the magnitude of the ERN predicted the increase in startle following errors, we first defined startle potentiation for each participant as the startle magnitude after errors minus the average of startle responses on predictable correct and unpredictable correct trials. The amplitude of the ERN predicted the degree to which errors potentiated the startle response (see Fig. 2; $r = -.38, p < .05$).[1] Because the magnitude of the ERN is negative and startle potentiation following errors is positive, this negative correlation indicates that individuals with larger error-related brain activity were characterized by larger increases in startle responses after errors.

DISCUSSION

To our knowledge, this is the first study to use startle methodology to directly evaluate affective processing following errors. Our findings are consistent with the notion that errors prompt defensive motivation: Startle magnitude was larger after the commission of errors than after correct responses. Our data indicate that, like threatening stimuli and the anticipation of shock, which have previously been shown to increase the startle response (Bradley et al., 2006; Grillon et al., 1993), errors also potentiate reflexes that are inherently defensive. These data fit well within previous work demonstrating peripheral autonomic response to errors (Hajcak et al., 2003b, 2004), but go further to suggest that errors prompt both orienting and defensive responses.

In the present study, larger startle potentiation following errors was predicted by an increased ERN magnitude. Notably, the magnitude of the error-potentiated startle was not related to the size of the error positivity—the positive deflection in the ERP that follows the ERN. This suggests that early error processing indexed by the ERN may uniquely serve to prime defensive

[1]The error positivity (Pe) was scored at Pz, where it was maximal, as the average activity in a 200- to 400-ms postresponse window on error trials on which a startle probe was not presented; unlike the ERN, however, the Pe did not correlate with the magnitude of the error-potentiated startle ($r = -.04, p > .80$). Across participants, the magnitudes of the ERN and Pe were both uncorrelated with the number of errors ($r = .21, p > .30$, and $r = .33, p > .05$, respectively). To determine whether the inclusion of startle probes in the present study might have made errors unusually aversive and increased the ERN, we ran 36 participants in a paradigm that was identical except that no startle probes were presented. The mean ERN in those participants (−3.32 μV, $SD = 6.11$ μV) did not differ from the mean ERN in the present study (−3.54 μV, $SD = 6.52$ μV), $t(65) = 0.21, p > .60$. This finding suggests that the presence of startle probes did not influence the magnitude of the ERN.

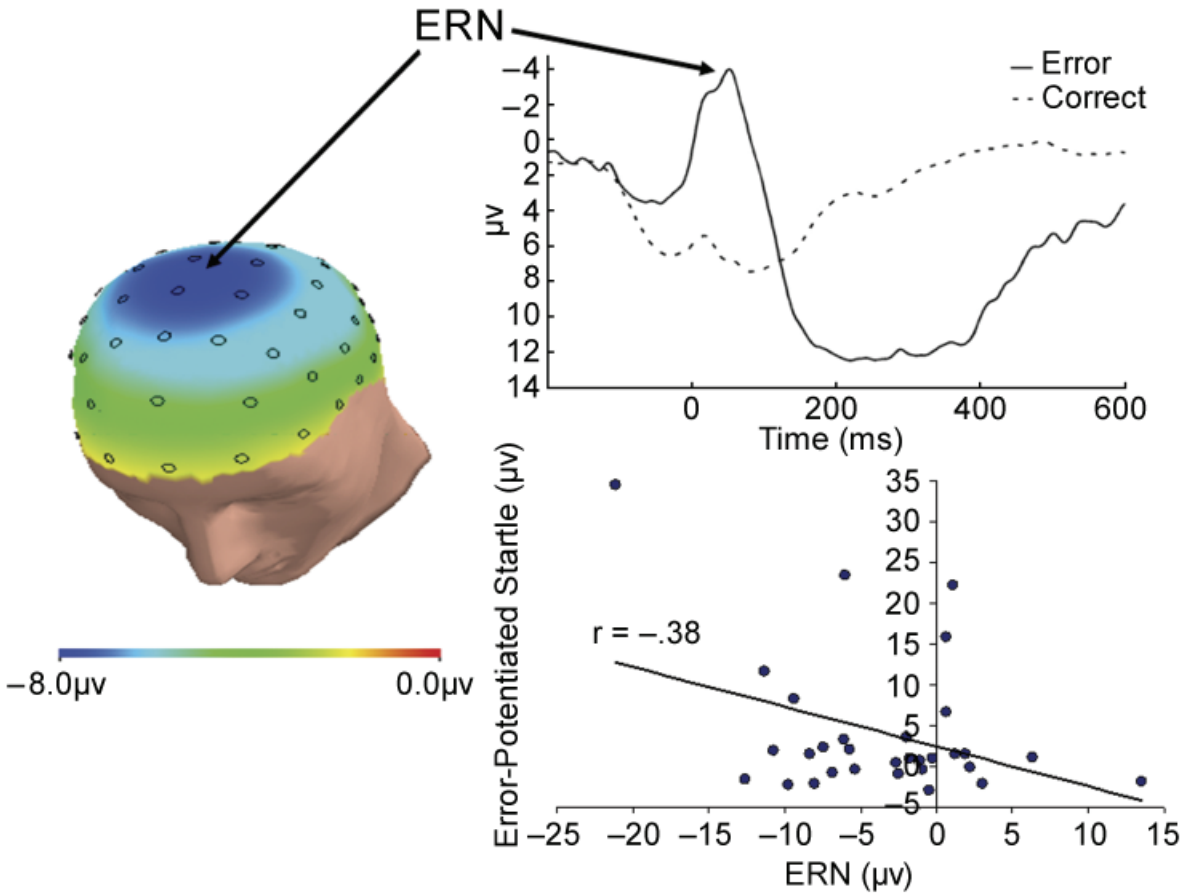

Fig. 2. Event-related potential (ERP) results. The illustration at the left shows the scalp topography of error-related brain activity from 0 to 100 ms postresponse. The graph at the top right presents response-locked ERPs for error and correct trials at FCz, where the error-related negativity (ERN) was maximal. Response onset occurred at 0 ms, and negative is plotted up. The scatter plot at the bottom right depicts magnitude of the error-potentiated startle as a function of ERN magnitude.

motivational systems. Current computational models suggest that the ERN signals the need to increase cognitive control following errors (Holroyd & Coles, 2002; Yeung et al., 2004). The present data, however, suggest that the ERN may also serve a more basic motivational function: to activate defensive systems in the service of protecting the organism. This view is consistent with existing data relating the ERN to sensitivity to negative, but not positive, stimuli and feedback (Boksem et al., 2006; Frank, Woroch, & Curran, 2005).

Several studies have provided evidence supporting a relation between the ERN and individual differences in emotionality: Individuals who experience more anxiety and sensitivity to punishment have larger ERNs (Boksem et al., 2006; Hajcak et al., 2003a, 2004; Luu et al., 2000). Increased startle responses and an increased ERN have both been related to heightened risk for anxiety disorders (Grillon, 2002). Moreover, the current data suggest that individuals with increased ERNs process errors as more threatening and aversive. That is, anxious individuals who are characterized by increased ERNs may exhibit a greater defensive response to errors compared with nonanxious individuals. Within Gray's (1994) theory, these data are consistent with the notion that the behavioral inhibition system (BIS) responds to signals of nonreward and punishment, and that functioning of the BIS relates to the experience of anxiety and fear. We are currently pursuing a line of research relating the BIS, ERN, and error-potentiated startle to further evaluate this possibility.

In the present study, startle probes were presented just 300 ms after both correct and erroneous responses. Thus, the facilitated startle following errors appears to reflect relatively rapid defensive mobilization. An important unanswered issue concerns the mechanism whereby errors rapidly increase defensive responding. In light of the association between the ERN and error-potentiated startle, one possibility is that early ACC activity directly activates circuits in the brain (e.g., the amygdala) that are critical to defensive mobilization and startle potentiation. Future studies might further examine the time course of error-potentiated startle—to determine how early the potentiated startle begins, and for how long it persists, following errors. Additional research might also examine the underlying mechanisms whereby errors increase the startle response.

Although errors may directly prime startle, it is also possible that emotional valence mediates the relation between error

detection and increased startle, and other peripheral psychophysiological measures, such as facial EMG, could be used to test this possibility. It is worth acknowledging that in the present study, startle probes were presented on a much larger percentage of error than correct trials. This was necessary in light of the fact that errors are relatively rare in speeded response tasks. Although startle probes were therefore more predictable on error than on correct trials, we manipulated predictability on correct trials. Although the magnitude of the startle response on correct trials was equivalent when the probe was predictable and when it was unpredictable, it is possible that the presence of startle probes made errors particularly aversive. This possibility could be evaluated in future studies by asking participants how they felt about making mistakes both in a condition with startle probes and in a condition without startle probes.

Overall, the current study indicates that errors are motivationally salient events that prime defensive systems: The startle response is potentiated following errors, compared with correct responses. Insofar as the magnitude of the ERN predicted the degree of startle potentiation following errors, these data further highlight the role of the ERN in affective and motivational responses to errors. Following previous work on the ERN and startle, future studies should evaluate how error-potentiated startle varies as a function of individual difference variables.

Acknowledgments—We give special thanks to E. David Klonsky for helpful comments on an earlier draft of the manuscript.

REFERENCES

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

Bradley, M.M., Codispoti, M., & Lang, P.J. (2006). A multi-process account of startle modulation during affective perception. *Psychophysiology*, *43*, 486–497.

Bradley, M.M., Moulder, B., & Lang, P.J. (2005). When good things go bad: The reflex physiology of defense. *Psychological Science*, *16*, 468–473.

Brazdil, M. (2005). Intracerebral error-related negativity in a simple go/nogo task. *Journal of Psychophysiology*, *19*, 244–255.

Bush, G., Luu, P., & Posner, M.I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences*, *4*, 215–222.

Chiu, P.H., & Deldin, P.J. (2007). Neural evidence for enhanced error detection in major depressive disorder. *American Journal of Psychiatry*, *164*, 608–616.

Cook, E.W., III, & Turpin, G. (1997). Differentiating orienting, startle, and defense responses: The role of affect and its implications for psychopathology. In P.J. Lang, R.F. Simons, & M.T. Balaban (Eds.), *Attention and orienting: Sensory and motivational processes* (pp. 137–164). Hillsdale, NJ: Erlbaum.

Curtin, J.J., Patrick, C.J., Lang, A.R., Cacioppo, J.T., & Birbaume, N. (2001). Alcohol affects emotion through cognition. *Psychological Science*, *12*, 527–531.

Dikman, Z.V., & Allen, J.J. (2000). Error monitoring during reward and avoidance learning in high- and low-socialized individuals. *Psychophysiology*, *37*, 43–54.

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.

Frank, M.J., Woroch, B.S., & Curran, T. (2005). Error-related negativity predicts reinforcement learning and conflict biases. *Neuron*, *47*, 495–501.

Gehring, W.J., Coles, M.G., Meyer, D.E., & Donchin, E. (1995). A brain potential manifestation of error-related processing. *Electroencephalography and Clinical Neurophysiology*, *S44*, 261–272.

Gehring, W.J., Himle, J., & Nisenson, L.G. (2000). Action-monitoring dysfunction in obsessive-compulsive disorder. *Psychological Science*, *11*, 1–6.

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

Gray, J.A. (1994). Three fundamental emotion systems. In P. Ekman & R.J. Davidson (Eds.), *The nature of emotion* (pp. 243–247). New York: Oxford University Press.

Grillon, C. (2002). Startle reactivity and anxiety disorders: Aversive conditioning, context, and neurobiology. *Biological Psychiatry*, *52*, 958–975.

Grillon, C., Ameli, R., Merikangas, K., Woods, S.W., & Davis, M. (1993). Measuring the time course of anticipatory anxiety using the fear-potentiated startle reflex. *Psychophysiology*, *30*, 340–346.

Hajcak, G., Franklin, M.E., Foa, E.B., & Simons, R.F. (in press). Increased error-related brain activity in pediatric OCD before and after treatment. *American Journal of Psychiatry*.

Hajcak, G., McDonald, N., & Simons, R.F. (2003a). Anxiety and error-related brain activity. *Biological Psychology*, *64*, 77–90.

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

Hajcak, G., McDonald, N., & Simons, R.F. (2004). Error-related psychophysiology and negative affect. *Brain and Cognition*, *56*, 189–197.

Hajcak, G., Moser, J.S., Yeung, N., & Simons, R.F. (2005). On the ERN and the significance of errors. *Psychophysiology*, *42*, 151–160.

Hajcak, G., & Simons, R.F. (2002). Error-related brain activity in obsessive-compulsive undergraduates. *Psychiatry Research*, *110*, 63–72.

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

Holroyd, C.B., Dien, J., & Coles, M.G. (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.

Lang, P.J. (1993). The motivational organization of emotion: Affect reflex connections. In S. van Goozen, N.E. van der Poll, & J.A. Sergeant (Eds.), *The emotions: Essays on emotion theory* (pp. 61–96). Hillsdale, NJ: Erlbaum.

Lang, P.J., Davis, M., & Öhman, A. (2000). Fear and anxiety: Animal models and human cognitive psychophysiology. *Journal of Affective Disorders*, *61*, 137–159.

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.

Miltner, W.H., Lemke, U., Weiss, T., Holroyd, C., Scheffers, M.K., & Coles, M.G. (2003). Implementation of error-processing in the human anterior cingulate cortex: A source analysis of the magnetic equivalent of the error-related negativity. *Biological Psychology*, *64*, 157–166.

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

(Received 5/15/07; Revision accepted 8/26/07)