Research Report

A NEURAL SYSTEM FOR ERROR DETECTION AND COMPENSATION

William J. Gehring,[1] Brian Goss,[1] Michael G.H. Coles,[1] David E. Meyer,[2] and Emanuel Donchin[1]

[1]University of Illinois and [2]University of Michigan

Abstract—*Humans can monitor actions and compensate for errors. Analysis of the human event-related brain potentials (ERPs) accompanying errors provides evidence for a neural process whose activity is specifically associated with monitoring and compensating for erroneous behavior. This error-related activity is enhanced when subjects strive for accurate performance but is diminished when response speed is emphasized at the expense of accuracy. The activity is also related to attempts to compensate for the erroneous behavior.*

A fundamental characteristic of human cognition is its fallibility. People rarely perform tasks perfectly, even though the costs of imperfection can be devastating (Norman, 1988; Reason, 1990). It is plausible to assume that the prevalence of errors, and their high cost, has led to the evolution of mechanisms that monitor the accuracy of actions and attempt to correct, or compensate for, errors. That such mechanisms exist is, indeed, assumed explicitly or implicitly in many theories of cognition. For example, concepts of error monitoring are included in theories of action (MacKay, 1987), learning (Adams, 1971; Rumelhart, Hinton, & Williams, 1986), speaking (Levelt, 1989), and consciousness (Kosslyn & Koenig, 1992). Monitoring mechanisms are also implied by theories of executive or supervisory cognitive control systems (Logan, 1985; Shallice, 1988; Stuss & Benson, 1986).

Given the frequency with which the concept of error monitoring is invoked, it is remarkable that there is little direct neurophysiological evidence for the existence of error-detection and -compensation systems (but see Gemba, Sasaki,

& Brooks, 1986, for an exception). Much of the work investigating the issue has inferred the existence of a monitoring and compensation apparatus from behavior that appears to be compensatory, as when subjects execute an error and then quickly execute the correct response (Rabbitt, 1966, 1968) or slow down subsequent to errors (Laming, 1968; Rabbitt, 1966). These phenomena, while consistent with the existence of an error-monitoring system, are not conclusive, however, because they could occur without the presence of an error-detection system: The apparent correction could simply be a correct response produced in parallel with, but more slowly than, the error. Furthermore, a response on a trial after an error could be slow because of a persistence of the processing problem that caused the error.

More direct evidence for an error-monitoring mechanism comes from descriptions of an event-related brain process that appears to be evoked contemporaneously with the commission of erroneous responses. We (Gehring, Coles, Meyer, & Donchin, 1990) have reported that an error-related negativity (ERN) appears selectively on error trials in choice reaction time experiments. The ERN takes the form of a sharp, negative-going deflection of up to 10 μV in amplitude and is largest at electrodes placed over the front and middle of the scalp. Its onset is shortly after the onset of electromyographic (EMG) activity detected in the limb that is about to make an error, and it peaks about 100 ms following its onset. A similar observation was made, independently, by Falkenstein, Hohnsbein, Hoormann, and Blanke (1990).

In this report, we present evidence that the ERN is a manifestation of the activity of a system associated with monitoring the accuracy of the response system and with compensating for errors. Our test is predicated on the assumption that if the ERN manifests the activity of such a system, it will be more active

when response accuracy is important to the subject. We predicted that the amplitude of the ERN will vary with the relative weight the subject's task assigns to accuracy and speed. Furthermore, if the ERN is a manifestation of an error-compensation mechanism, there ought to be a relationship between its amplitude and the dynamics of the erroneous responses. We varied, therefore, the speed and accuracy requirements placed upon the subject, and we measured several performance parameters that may reflect compensatory activity, including the force with which the subject executes a response, the probability of correcting the error, and the speed of responses following the error. We embedded these manipulations and measures in a task known from previous research to produce erroneous response activation (see Coles, Gratton, Bashore, Eriksen, & Donchin, 1985; Gratton, Coles, Sirevaag, Eriksen, & Donchin, 1988).

METHOD

Subjects

Six University of Illinois students (4 men and 2 women) between the ages of 18 and 26 served as subjects. All were right-handed and had normal or corrected-to-normal vision. They received $3.50 per hour plus bonuses for participation.

Stimuli

Stimuli were presented on a Hewlett-Packard computer display (#1310A). Subjects sat 1 m from the screen, such that each letter subtended approximately 0.5° of visual angle. One of four arrays occurred on each trial: The two compatible arrays were HHHHH and SSSSS, and the two incompatible arrays were SSHSS and HHSHH. The probability of each of these arrays was .25. A fixation

dot, visible throughout the experiment, appeared 0.1° below the central letter of the array. An asterisk appeared as a warning stimulus 1,000 ms prior to the onset of each stimulus array.

Procedure

Subjects were required to respond with the left or right hand according to the identity (H or S) of the letter at the center of the letter array. Subjects responded by squeezing zero-displacement dynamometers (as in Gratton et al., 1988). (For details of the effects of the compatibility manipulation—the flanking letters—see Coles et al., 1985.)

Subjects received financial penalties for errors and bonuses for responses faster than a particular deadline. These values were varied such that in the speed condition subjects responded quickly with little regard for errors, in the accuracy condition subjects responded slowly and accurately, and in the neutral condition subjects responded at an intermediate level of speed and accuracy. In each of three experimental sessions, subjects had to perform 520 trials under each of the three speed-accuracy conditions. The order of conditions was counterbalanced.

Responses

A criterion overt response was defined as a squeeze that exceeded 25% of maximum squeeze force (determined for each subject and each hand separately), and reaction time (used to compute bonuses) was defined as the time at which the force exceeded this criterion. Several hundred initial trials were deemed "practice" trials, and on these trials only, subjects received feedback, in the form of an auditory click, whenever their squeeze response exceeded the criterion.

For the analyses in this article, we used the latency and side of the first detectable EMG activity to classify the response as correct or incorrect on each trial. This classification allowed us to subdivide the error category according to the level of squeeze amplitude attained on each trial (see below). We also noted whether a second EMG response

occurred in the other arm, to identify errors that were corrected. EMG onset latency was defined as the first point, following the stimulus, that exceeded 4 standard deviations of the baseline, prestimulus activity.

Psychophysiological Recording

The electroencephalogram (EEG) was recorded from midline and lateral scalp electrode sites: F_z, C_z, P_z, C3' (4 cm to the left of C_z), and C4' (4 cm to the right of C_z) according to the 10/20 system, referenced to linked mastoids. Horizontal and vertical eye movements and forearm flexor EMG activity were recorded from standard locations (see Gratton et al., 1988). All electrodes were Medical Associates disposable Ag/AgCl electrodes affixed with Grass EC2 electrode cream. Impedance for EEG and electrooculogram (EOG) electrodes was below 10 KOhm; EMG electrode impedance was less than 20 KOhm.

EEG, EOG, and EMG signals were amplified and conditioned, and eye movement artifacts were corrected, as in Gratton et al. (1988). The derived Voltage × Time functions were digitized at 100 Hz for 2,100 ms, starting 100 ms prior to the presentation of each warning stimulus.

RESULTS AND DISCUSSION[1]

Performance Data

EMG data revealed that the speed-accuracy instructions were effective. EMG onset latency on correct trials and accuracy (percentage correct) were as follows: for the speed condition, 236.9 ms, 67.6%; for the neutral condition, 271.8 ms, 78.8%; for the accuracy condition, 304.4 ms, 89.4%. These values (proportion correct values were transformed with the arc sine transform; Neter, Wasserman, & Kutner, 1985) were

---

1. The Greenhouse-Geisser correction was applied when appropriate to correct for possible violations of the analysis of variance assumption of sphericity. The text lists corrected p values.

submitted to separate 3 (Speed Condition) × 2 (Compatibility) repeated measures analyses of variance (ANOVAs). For correct EMG onsets, the main effect of speed condition was statistically significant, $F(2, 10) = 57.21$, $p < .0001$, $MS_e = 239.01$, as was the planned analysis of linear trend, $F(1, 5) = 65.589$, $p < .0005$, $MS_e = 416.8$. For proportion correct, the same effects were evident (main effect of speed condition: $F[2, 10] = 69.62$, $p < .0001$, $MS_e = 0.014$; linear trend: $F[1, 5] = 126.73$, $p < .0001$, $MS_e = 0.016$).

Error-Related Negativity

The event-related potential (ERP) data confirmed the presence of error-related brain electrical activity. Response-locked EEG activity was obtained by extracting an epoch of 1,000 ms in duration that began 400 ms before the first detectable EMG on each trial. The EEG records for correct and incorrect responses were then averaged separately to yield the waveforms for the C_z electrode (located at the center of the scalp), shown in Figure 1. The ERP associated with an incorrect response was characterized by a negative-going deflection (the ERN) that began at around the time of the incorrect response and peaked about 100 ms later. The deflection was not evident in the averages based on trials in which the subjects responded correctly.

To quantify this activity, we filtered the EEG using a 59-point equal-ripple, zero-phase-shift, optimal finite impulse response low-pass filter with a passband cutoff frequency of 8 Hz and a stopband cutoff frequency of 10 Hz (cf. Farwell, Martinerie, Bashore, Rapp, & Goddard, 1993). The amplitude of the ERN was defined as the difference between the average amplitude of the waveform within a 50-ms window centered at 100 ms post-EMG onset and the average amplitude of a baseline for the 50 ms immediately preceding EMG onset. We submitted the data to a 5 (Electrode) × 3 (Speed Condition) × 2 (Response Accuracy) × 2 (Compatibility) ANOVA. The main effect of accuracy was $F(1, 5) = 120.58$, $p < .0001$, $MS_e = 5,236.69$.

To evaluate the relationship between the ERN and the importance of errors to
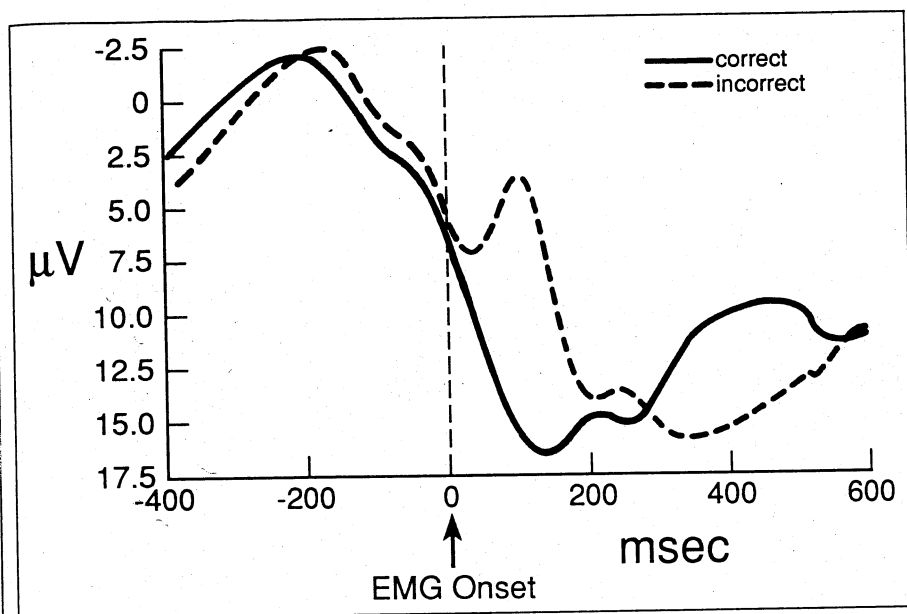
Fig. 1. Comparison of response-locked event-related potential activity, recorded at the C_z electrode, for correct and incorrect trials.

discriminant function procedure and the resulting posterior probability measures were sensitive to the distinctive wave-shape seen in Figures 1 and 2.

We found the amplitude of the ERN to be related to three measures of the subjects' compensatory behavior. Figure 3 (right panel, top) shows the mean amplitude of the error response squeeze for each posterior probability quartile. The figure suggests that the larger the ERN, the smaller the error squeeze. To confirm this pattern, we submitted these measures to a 3 (Speed Condition) × 4 (Quartile) repeated measures ANOVA. A main effect of quartile ($F[3, 15] = 11.66, p < .01, MS_e = 831.63$) and a significant analysis of linear trend ($F[1, 5] = 19.22, p < .01, MS_e = 1,505.12$) both support the observation that large ERNs were associated with small error squeezes, suggesting that the ERN might
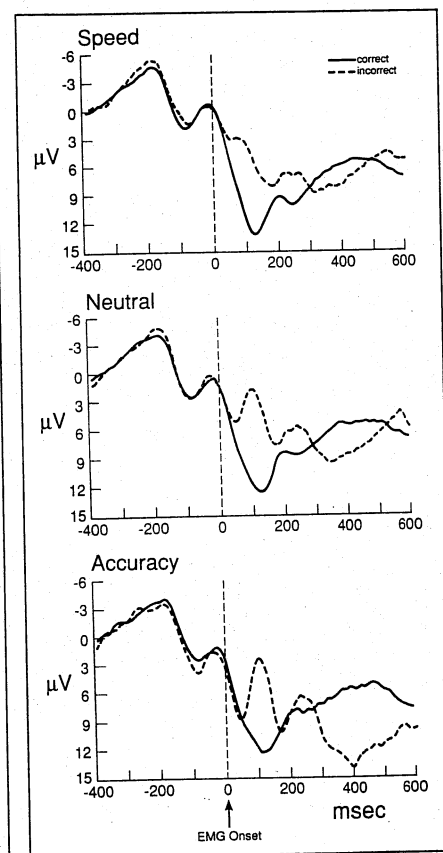
the subject, we compared the ERNs derived from the three accuracy conditions. Because any effect of speed-accuracy emphasis could be attributed to the overall difference in response speed between conditions, we selected from each condition trials whose reaction times fell within the same 50-ms reaction time window. For each subject, the reaction time bin was either 250 to 300 ms or 200 to 250 ms. As can be seen in Figure 2, the ERN was largest for the accuracy condition and smallest for the speed condition, with the neutral condition intermediate. We submitted the mean amplitude measures (at C_z) from error trials to a single-factor (speed condition) repeated measures ANOVA. A significant effect of speed condition ($F[2, 10] = 6.19, p < .05, MS_e = 371.48$) and a significant analysis of linear trend ($F[1, 5] = 8.44, p < .05, MS_e = 507.7$) confirmed that the ERN increased in amplitude from speed to neutral to accuracy conditions. These data are consistent with the view that the ERN is associated with an error-related processing mechanism, whose activity is modulated by the degree to which accuracy is important to the subject.

We now turn to additional data that suggest the ERN is related to attempts to compensate for errors. To analyze this relationship, we used a measure of ERN amplitude derived from stepwise discriminant analysis (SWDA) of the ERP data. SWDA produces a discriminant function—vector of weights—that when cross-multiplied with the actual ERP waveforms produces scores whose values optimally distinguish correct from incorrect trials. To build the functions, we used a randomly chosen subset of correct and incorrect trials in the accuracy condition. We limited the analysis to the epoch from 100 ms prior to EMG onset to 200 ms following EMG onset for the F_z (frontal) and C_z (central) channels (relative to a 100-ms prestimulus baseline). For most subjects, the major contribution to the discrimination between correct and incorrect trials was provided by activity at the C_z electrode in the postresponse portion of the epoch.

For each trial in the experiment, the SWDA computed the posterior probability that the response on a trial was in fact an error. We used these posterior probability values as estimates of the amplitude of the ERN on each trial. For the analyses below, we partitioned the data according to four quartiles of the distribution of posterior probabilities, which yielded four different levels of ERN amplitude. Figure 3 (left panel) shows the average ERPs (at the C_z electrode) for these four levels of ERN amplitude on error trials. These data confirm that the

Fig. 2. The effects of different speed-accuracy instructions on the error-related negativity recorded at C_z. For each of the three conditions, trials with the same reaction time were used (see text).
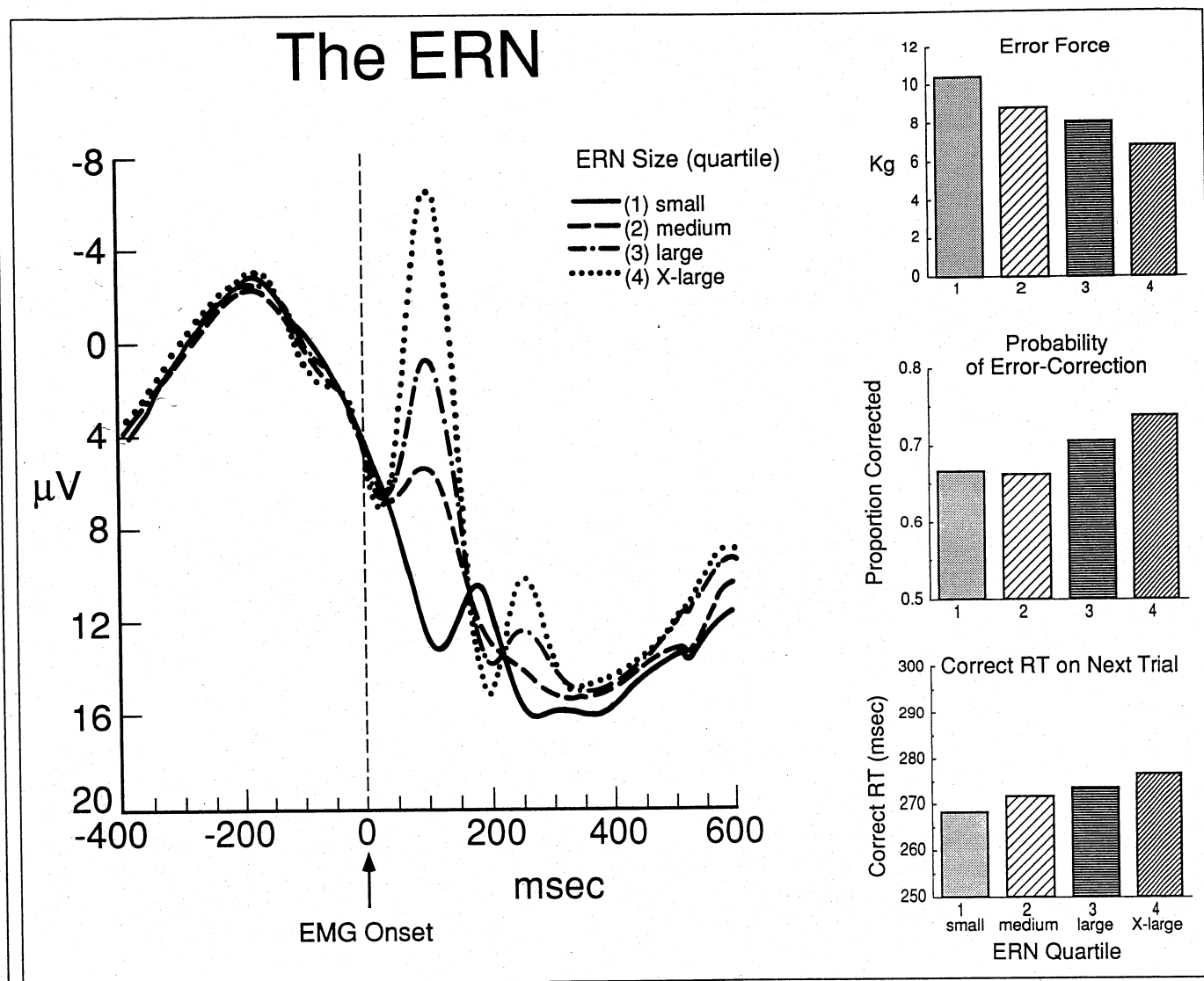
Fig. 3. Relationship between error-related negativity (ERN) amplitude and three measures of compensatory behavior. Left panel: Average event-related potentials at the C_z electrode as a function of the four levels of the posterior probability measure of ERN amplitude. Right panel, top: Error squeeze force in Kg as a function of the four ERN levels. Right panel, middle: Probability of error correction as a function of the four ERN levels. Right panel, bottom: Correct reaction time on the trial following an error as a function of the four ERN levels.

reflect an attempt to brake the erroneous response.

Figure 3 (right panel, center) further suggests that the larger the ERN, the greater the probability that the error would be followed by a correct response on the same trial. We submitted the probability that a response was corrected (converted with the arc sine transform) to a 3 (Speed Condition) × 2 (Response Accuracy) × 4 (Posterior Probability Quartile) repeated measures ANOVA. A significant main effect of quartile, $F(3, 15) = 5.94$, $p < .05$, MS_e

= 0.059, and a subsequent analysis of linear trend, $F(1, 5) = 6.54$, $p < .06$, MS_e = 0.15, confirmed that the probability that a response was corrected increased with the size of the ERN.[2]

---

2. We included correct trials in this analysis because we found the ERN to be present on correct trials in which the correct response was followed by an error. If such correct responses were themselves "corrected," then the presence of the ERN on these trials is consistent with our hypothesis that the ERN manifests the operation of a system for error

monitoring and compensation. This interpretation is, however, clouded by the possibility that the ERNs on these trials could be elicited by the second, erroneous response. The close temporal proximity of the correct response and the subsequent error on these trials makes it difficult to determine which response elicited the ERN. Indeed, our attempts to disambiguate these interpretations (e.g., by examining trials in which the interval between the correct response and the subsequent error was relatively large) were not successful, in part because of the small number of trials available for such an analysis.

Finally, the larger the ERN, the slower the subjects' response on the immediately following correct trial (Fig. 3, right panel, bottom). We examined the correct reaction times on trials subsequent to the error trials, submitting these reaction times to a 3 (Speed Condition) × 4 (Previous Trial's ERN Amplitude Quartile) repeated measures ANOVA. The main effect of amplitude quartile, $F(3, 15) = 5.58$, $p < .05$, $MS_e = 39.78$, and a planned test of linear trend, $F(1, 5) = 11.14$, $p < .025$, $MS_e = 59.15$, were both significant. These results suggest that the tendency to adopt a more conservative strategy following an error (originally described by Rabbitt, 1966, and Laming, 1968) is related to the size of the ERN on the error trial.

CONCLUSIONS

Taken together, these data are consistent with the existence of a brain system for error detection and compensation, whose behavior is manifested at the scalp in a measure of brain potential activity, the ERN. We confirmed our previous observation that the ERN occurs at around the time that subjects make erroneous responses. In addition, we found that the magnitude of the ERN was affected by the degree to which accurate performance was emphasized (see Falkenstein et al., 1990, for a similar observation) and was related to three measures of error-compensation activity.

These results place several constraints on theorizing about the process of error detection and compensation. Theorists have postulated a number of sources for the input representations used in monitoring for errors (see Adams, 1971; Angel, 1976; Gibbs, 1965; Rabbitt, 1968; Schmidt & Gordon, 1977). The fact that the onset of the ERN is contemporaneous with the error response (as also observed by Falkenstein et al., 1990) suggests that the error-detection system does not use sensory or proprioceptive information, since such information could not be available until after the response has been initiated (see Higgins & Angel, 1970, for a similar analysis of error-correction latencies). Rather, it appears that information is available when the response is initiated. This finding is consistent with models

(e.g., Angel, 1976) postulating that the brain retains a neural record of the motor commands it sends to the effectors—an efference copy—that is used to judge the accuracy of the movement.

ERN amplitude measures provide additional constraints. The relationship between the speed-accuracy manipulation and ERN activity suggests that the process manifested by the ERN is influenced by the importance of errors to the subject: The more costly the error, the greater the likelihood or strength of the activity manifested by the ERN. The association between the ERN and a variety of compensatory behaviors suggests that the error-detection system provides input to different compensatory systems. These mechanisms appear to include fast-acting systems that can inhibit and correct the error as it occurs, as well as systems that control response strategies, whose effects are evident on future trials.

It is important to note that our data are consistent with an "active" error-compensation process that is invoked when an error is detected. Theories can often account for behavior that seems compensatory without having to postulate error-detection mechanisms (e.g., Dell, 1986). In this particular experiment, subjects might generate a weak error response or a response that appears to correct the error without actually detecting the error. It is not clear, however, how such a "passive" system could produce the observed relationship between error-related neural activity and compensatory behavior.

The data from the present study do not permit us to identify the neural locus of this detection or compensation activity. Research using animals (Gemba et al., 1986) and current theories of motor function (Goldberg, 1985), however, suggests that a system involving the anterior cingulate cortex and supplementary motor areas is one possible locus. These areas might also be likely candidates to implement error-related activity given their connectivity with frontal lobe regions commonly associated with executive or supervisory information processing activity (cf. Goldberg, 1985; Goldman-Rakic, 1987; Shallice, 1988; Stuss & Benson, 1986). Indeed, the phenomenon observed in the present experiment may bear some relationship to activity ob-

served in monkeys performing go/no-go tasks: No-go trials elicit field potential activity in the prefrontal cortex that may be involved in response inhibition (Sasaki & Gemba, 1986; Sasaki, Gemba, & Tsujimoto, 1989). Gemba and Sasaki (1989) observed analogous scalp-recorded potentials in humans.

Whatever its neural substrate, the ERN provides a way to assess the behavior of the error-detection and -compensation system. This measure should enable further development and refinement of a comprehensive theory of error detection and compensation and, more generally, of the executive control of information processing.

REFERENCES

Adams, J.A. (1971). A closed-loop theory of motor learning. *Journal of Motor Behavior, 3,* 111–150.

Angel, R.W. (1976). Efference copy in the control of movement. *Neurology, 26,* 1164–1168.

Coles, M.G.H., Gratton, G., Bashore, T.R., Eriksen, C.W., & Donchin, E. (1985). A psychophysiological investigation of the continuous flow model of human information-processing. *Journal of Experimental Psychology: Human Perception and Performance, 11,* 529–553.

Dell, G.S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review, 93,* 283–321.

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.

Farwell, L.A., Martinerie, J.M., Bashore, T.B., Rapp, P.E., & Goddard, P.H. (1993). Optimal digital filters for long latency event-related brain potentials. *Psychophysiology, 30,* 306–315.

Gehring, W.J., Coles, M.G.H., Meyer, D.E., & Donchin, E. (1990). The error-related negativity: An event-related brain potential accompa-

nying errors. *Psychophysiology, 27,* S34. (Abstract)

Gemba, H., & Sasaki, K. (1989). Potential related to no-go hand movement task with color discrimination in human. *Neuroscience Letters, 101,* 263–268.

Gemba, H., Sasaki, K., & Brooks, V.B. (1986). "Error" potentials in limbic cortex (anterior cingulate area 24) of monkeys during motor learning. *Neuroscience Letters, 70,* 223–227.

Gibbs, C.B. (1965). Probability learning in step-input tracking. *British Journal of Psychology, 56,* 233–242.

Goldberg, G. (1985). Supplementary motor area structure and function: Review and hypotheses. *Behavioral and Brain Sciences, 8,* 567–616.

Goldman-Rakic, P.S. (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In F. Plum (Ed.), *Handbook of physiology: Section 1. The nervous system. Vol. 5. Higher functions of the brain. Part 1* (pp. 373–417). Bethesda, MD: American Physiological Society.

Gratton, G., Coles, M.G.H., Sirevaag, E., Eriksen, C.W., & Donchin, E. (1988). Pre- and poststimulus activation of response channels: A psychophysiological analysis. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 331–344.

Higgins, J.R., & Angel, R.W. (1970). Correction of errors without sensory feedback. *Journal of Experimental Psychology, 84,* 412–416.

Kosslyn, S.M., & Koenig, O. (1992). *Wet mind: The new cognitive neuroscience.* New York: Free Press.

Laming, D.R.J. (1968). *Information theory of choice-reaction times.* New York: Academic Press.

Levelt, W.J.M. (1989). *Speaking: From intention to articulation.* Cambridge, MA: MIT Press.

Logan, G.D. (1985). Executive control of thought and action. *Acta Psychologica, 60,* 193–210.

MacKay, D. (1987). *The organization of perception and action: A theory for language and other cognitive skills.* New York: Springer.

Neter, J., Wasserman, W., & Kutner, M.H. (1985). *Applied linear statistical models* (2nd ed.). Homewood, IL: Irwin.

Norman, D.A. (1988). *The design of everyday things.* New York: Basic Books.

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

Rabbitt, P.M.A. (1968). Three kinds of error-signalling responses in a serial choice task. *Quarterly Journal of Experimental Psychology, 20,* 179–188.

Reason, J. (1990). *Human error.* Cambridge, England: Cambridge University Press.

Rumelhart, D.E., Hinton, G., & Williams, R. (1986). Learning internal representations by error propagation. In D.E. Rumelhart & J.L. McClelland (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 1. Foundations* (pp. 318–362). Cambridge, MA: MIT Press.

Sasaki, K., & Gemba, H. (1986). Electrical activity in the prefrontal cortex specific to no-go reaction of conditioned hand movement with colour discrimination in the monkey. *Experimental Brain Research, 64,* 603–606.

Sasaki, K., Gemba, H., & Tsujimoto, T. (1989). Suppression of visually initiated hand movement by stimulation of the prefrontal cortex in the monkey. *Brain Research, 495,* 100–107.

Schmidt, R.A., & Gordon, G.B. (1977). Errors in motor responding, "rapid" corrections, and false anticipations. *Journal of Motor Behavior, 9,* 101–111.

Shallice, T. (1988). *From neuropsychology to mental structure.* Cambridge, England: Cambridge University Press.

Stuss, D.T., & Benson, D.F. (1986). *The frontal lobes.* New York: Raven Press.