Psychophysiology, 44 (2007), 913–917. Blackwell Publishing Inc. Printed in the USA.

DOI: 10.1111/j.1469-8986.2007.00561.x

BRIEF REPORT

Reward prediction error signals associated with a modified time estimation task

CLAY B. HOLROYD AND OLAVE E. KRIGOLSON
Department of Psychology, University of Victoria, Victoria, British Columbia, Canada

Abstract

The feedback error-related negativity (fERN) is a component of the human event-related brain potential (ERP) elicited by feedback stimuli. A recent theory holds that the fERN indexes a reward prediction error signal associated with the adaptive modification of behavior. Here we present behavioral and ERP data recorded from participants engaged in a modified time estimation task. As predicted by the theory, our results indicate that fERN amplitude reflects a reward prediction error signal and that the size of this error signal is correlated across participants with changes in task performance.

Descriptors: Feedback error-related negativity, Reward prediction error, Time estimation task, Reinforcement learning

In a seminal study, Miltner, Braun, and Coles (1997) demonstrated that error feedback stimuli in a time estimation task, which indicated that participants' responses were not "on time," elicited a component of the event-related brain potential (ERP) later termed the feedback error-related negativity (fERN). Following an additive factors approach, the authors measured fERN amplitude as the maximum difference between the ERPs elicited by error and correct feedback. The two types of feedback stimuli were delivered equiprobably, so if the participants came to expect the error and correct feedback about equally, then the subtraction would have removed from the fERN any pure effect of expectancy on the ERP. Importantly, this difference wave approach is indifferent to the source of the variance between the ERPs (Luck, 2005): fERN amplitude so defined can depend on the error feedback, on the correct feedback, or on both. In fact, although the fERN is commonly understood to be elicited by error feedback, we have recently found that variance in fERN amplitude may stem from the superposition on correct trials of a positive-going ERP component over a negative-going ERP component that is present on both correct trials and error trials (the N200; Pakzad-Vaezi, Krigolson, & Holroyd, 2006; see also Holroyd, 2004).

We have previously proposed that the fERN indexes a reward prediction error signal associated with reinforcement learning (Holroyd & Coles, 2002; Nieuwenhuis, Holroyd, Mol, & Coles, 2004). According to this hypothesis, the amplitude of the fERN (measured as a difference wave) is modulated by the unexpectedness of the feedback, such that the difference between unexpected error and correct feedback is larger than the difference between expected error and correct feedback. Here we tested this prediction by modifying the original time estimation task to include "easy" and "hard" conditions. Specifically, we predicted the following: One, participants would commit fewer errors, and thus would come to expect fewer errors, in the easy condition relative to the hard condition; two, their expectations would be violated more by errors in the easy condition than in the hard condition, and more by correct responses in the hard condition than in the easy condition; three, fERN amplitude, measured from difference waves created across conditions, would be larger for unexpected feedback (error feedback in the easy condition – correct feedback in the hard condition) than for expected feedback (error feedback in the hard condition – correct feedback in the easy condition); and four, if the associated reward prediction error signals were indeed used for the purpose of reinforcement learning, then the size of the fERN would be correlated across participants with changes in their behaviors following the feedback.

Methods

Participants

Seventeen undergraduate students (8 male; 19.6 ± 2.8 years old) participated in the experiment. All of the participants were volunteers who received extra credit in a first- or second-year

This research was supported in part by National Sciences and Engineering Research Council Discovery Grant RGPIN 312409-05, Michael Smith Foundation for Health Research Fellowship ST-SGS-349 (05-1) CLIN, the University of Victoria Fellowship Program, and a Petch Research Scholarship. We are grateful to the research assistants in the Brain and Cognition Laboratory and to the students enrolled in Psychology 576C, Spring 2006, for help with data collection.

Address reprint requests to: Clay Holroyd, Department of Psychology, University of Victoria, P.O. Box 3050 STN CSC, Victoria, B.C. V8W 3P5, Canada. E-mail: holroyd@uivic.ca

psychology course for their participation and provided written, informed consent. The study was conducted in accordance with the ethical standards prescribed in the Declaration of Helsinki and was approved by the human subjects review board at the University of Victoria.

Apparatus and Procedure

Participants were seated comfortably in front of a computer monitor in an electromagnetically shielded booth and performed a time estimation task. The task was similar to that employed by Miltner et al. (1997) in which participants were required to estimate the duration of 1 s. Each trial commenced with an auditory cue (1500 Hz, 65 dB) that lasted for 50 ms. When the participants believed that 1 s had elapsed, they responded by pressing the left mouse button. Participants received feedback indicating the accuracy of their estimate 600 ms following the response. A trial was considered on time if the participants' response occurred within a window of time centered around 1 s (see below), and was considered not on time otherwise. The feedback stimuli consisted of a white plus sign and a white zero (3°, 1000 ms) presented on a high contrast black background. The mappings of the feedback stimuli with valence (correct or error) were counterbalanced across participants. Following the offset of the feedback stimulus a blank screen was presented for either 1400, 1500, or 1600 ms (equivalent probability of each).

The time window was initialized at 1000 ms $\pm$ 100 ms. Thus, each participant was required to respond between 900 and 1100 ms following the auditory cue to receive correct feedback on the first trial. Following each trial the size of the time window decreased if the response landed within the window and increased otherwise. The amount of this change depended on three experimental conditions: *control*, *easy*, and *hard*. In the control condition the window size increased by 10 ms on error trials and decreased by 10 ms on correct trials. In the easy condition the window size increased by 12 ms on error trials and decreased by 3 ms on correct trials. In the hard condition the window size increased by 3 ms on error trials and decreased by 12 ms on correct trials.

Participants began the experiment by completing 25 practice trials in the control condition. Then, they completed two blocks of 75 trials in the control condition. The control condition was followed by two blocks of 75 trials in each of the easy and hard conditions, the order of which was counterbalanced across participants. Thus, across the three experimental conditions there were 450 trials total. The purpose of the control condition was threefold: first, to replicate the standard fERN phenomenon; second, to establish a stable time window before participants engaged in the subsequent conditions (see below); and third, to ensure that participants practiced the task sufficiently before engaging in the hard condition. Participants were informed that some blocks would be more difficult than others, but were not told specifically which blocks were hard or easy. Importantly, the size of the time window on each block was initialized with the value that corresponded to the end of the previous block. Participants relaxed during self-paced rest periods between blocks.

Data Acquisition

Response time (in milliseconds) and accuracy (on time vs. not on time) were recorded on each trial using a standard USB mouse. The electroencephalogram (EEG) was recorded from 41 electrode locations using BrainVision Recorder software (Version 1.3, Brainproducts, GmbH, Munich, Germany). The electrodes were mounted in a fitted cap with a standard 10–20 layout and were referenced to the average voltage across channels. The vertical and horizontal electrooculograms were recorded from electrodes placed above and below the right eye and on the outer canthi of the left and right eyes, respectively. Electrode impedances were kept below 10 kΩ. The EEG data were sampled at 250 Hz, amplified (Quick Amp, Brainproducts, GmbH, Munich, Germany), and filtered through a passband of 0.017 Hz–67.5 Hz (90 dB octave roll off).

Data Analysis

Mean response times, accuracies, and window sizes were calculated for each participant for each condition. To gauge the impact of feedback valence on behavior, the mean absolute changes in response times following each error trial and following each correct trial were calculated for each participant for each condition.

The EEG data were filtered off-line through a (0.1 Hz–20 Hz passband) phase-shift-free Butterworth filter and re-referenced to linked mastoids. Ocular artifacts were removed using the algorithm described by Gratton, Coles, and Donchin (1983). Trials in which the change in voltage at any channel exceeded 35 μV per sampling point were also discarded. In total, less than 5% of the data were discarded. An 800-ms epoch data (from 200 ms before the feedback stimuli to 600 ms after the feedback stimuli) was extracted from the continuous EEG for each trial, channel, and participant for each of the three experimental conditions (control, easy, hard). These epochs were baseline corrected relative to the 200-ms segment preceding feedback stimulus onset. ERPs were created by averaging the EEG data by condition for each electrode channel and participant.

To minimize overlap between the fERN and other ERP components, we created "difference waves" by subtracting the correct ERPs from the incorrect ERPs. Specifically, for each participant and channel we created three fERN difference waves by (a) subtracting the correct ERP in the control condition from the error ERP in the control condition, creating a control difference wave; (b) subtracting the correct ERP in the hard condition from the error ERP in the easy condition (i.e., infrequent error – infrequent correct), creating an "unexpected" difference wave; and (c) subtracting the correct ERP in the easy condition from the error ERP in the hard condition (i.e., frequent error – frequent correct), creating an "expected" difference wave. This practice removes activity related purely to event probability, while retaining activity related to event valence and/or to the interaction of event valence with event probability (Holroyd, 2004). The amplitude of each difference wave was measured for each participant and electrode as the most negative deflection within the 600 ms following feedback stimulus onset. Note that if error feedback and correct feedback associated with a given level of expectancy did not differentially effect the ERP, then the amplitude of the difference wave would equal zero. Further, if the valence of the feedback did not interact with expectancy, then the expected and unexpected difference waves would share the same amplitude. Finally, the curvatures of the scalp distributions were estimated by finding polynomial functions (up to order 7) that best fit each difference wave along the midline (Fpz, Fz, FCz, Cz, CPz, Pz, POz, Oz) and lateral (FT9, T5, FC1, FCz, FC2, T6, FT9) electrodes sites. To confirm that fERN amplitude was not confounded by overlap with the P300, paired *t* tests were carried out on the difference wave values where these ERP components are maximal (channels FCz and Pz, respectively).

Results

Behavioral Data

In the control condition, participants were correct on about half of the trials (51%) and the mean size of the response window was 103 ms. The absolute change in response time was larger on trials that immediately followed error trials (155 ms) than on trials that immediately followed correct trials (113 ms), $t(16) = -7.20$, $p < .001$, indicating that feedback valence differentially affected subsequent behavior. Participants made more errors in the hard condition (76%) than in the easy condition (23%), $t(16) = -41.39$, $p < .001$, consistent with the mean size of the response window, which was smaller in the hard condition (52 ms) than in the easy condition (160 ms), $t(16) = -11.39$, $p < .001$. Further, a 2×2 repeated measures ANOVA on expectancy (expected, unexpected) and valence (correct, error) associated with the absolute change in response time on the following trial revealed a main effect of valence, $F(1,16) = 50.74$, $p < .001$, $\eta_p^2 = .76$, and an interaction of valence with expectancy, $F(1,16) = 47.58$, $p < .001$, $\eta_p^2 = .75$, but no main effect of expectancy, $F(1,16) = 0.20$, $p = .66$, $\eta_p^2 = .01$ (Figure 1A). To compare this behavioral measure with the fERN, we computed the difference between the absolute change in response times associated with unexpected trials (absolute response time change following easy error trials – absolute response time change following hard correct trials; right pair of bars in Figure 1A) with the absolute change in response time associated with expected trials (absolute response time change following hard error trials – absolute response time change following easy correct trials; left pair of bars in Figure 1A). The difference between the absolute change in response times following unexpected (easy) error trials and unexpected (hard) correct trials (73 ms) was larger than the difference in the absolute change in response times following expected (hard) error trials and expected (easy) correct trials (25 ms), $t(16) = -6.90$, $p < .001$.

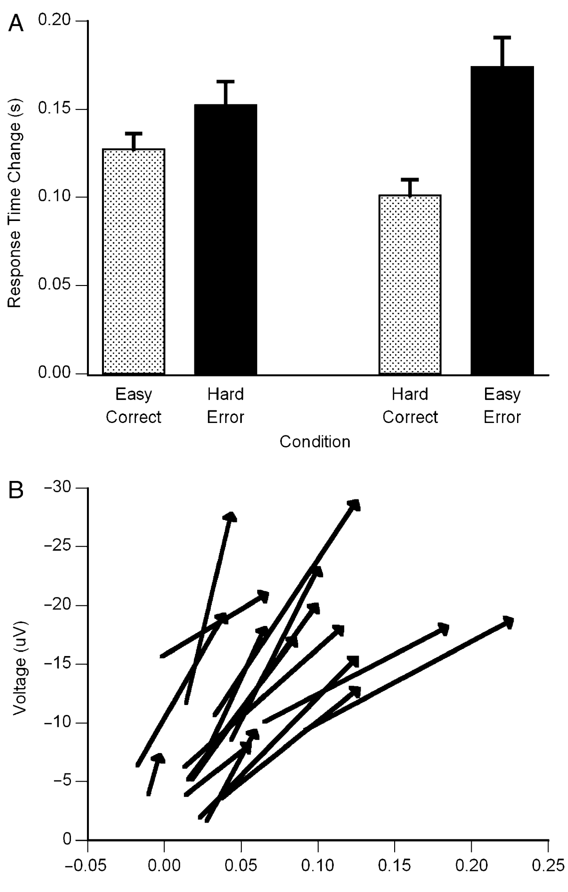

Figure 1. Performance data. A: Absolute change in response time following expected outcomes (easy correct trials and hard error trials) and unexpected outcomes (hard correct trials and easy error trials). B: Difference wave amplitudes as a function of the difference in the absolute change in response times. Each line corresponds to the data of a single participant; arrowheads point in the direction of increasing unexpectedness.

Electrophysiological Data

For the control condition, the scalp distribution of the difference wave between error trials and correct trials was maximal at frontal-central areas of the scalp, at electrode position FCz (-11.0 μV), 288 ± 5 ms following the onset of the feedback. The scalp distribution was significantly curved (Table 1) and was significantly larger at channel FCz (-11.0 μV) than at channel Pz (-8.0 μV), $t(16) = -5.14$, $p < .001$. The latency and scalp distribution of the difference wave and the morphology of the correct and error ERPs were consistent with the fERN (Miltner et al., 1997).

Figure 2A, B illustrates the scalp distributions associated with the unexpected and expected difference waves, respectively. The distributions reached maximum amplitude at channel FCz 288 ± 6 ms (unexpected) and 263 ± 10 ms (expected) following feedback onset. Both distributions were significantly curved (Table 1) and were significantly larger at channel FCz than at channel Pz (unexpected: -10.8 μV vs. -8.5 μV, $t[16] = -2.29$, $p < .05$; expected: -6.7 μV vs. -4.8 μV, $t[16] = -3.83$, $p < .005$). Figure 2C shows the correct and error ERPs recorded at channel FCz for the easy and hard conditions; the associated expected and unexpected difference waves are depicted in Figure 2D. Importantly, the unexpected difference wave (-10.8 μV) was larger than the expected difference wave (-6.7 μV), $t(16) = 3.13$, $p < .01$, indicating that the unexpected correct and error feedback differentially impacted the ERP more than did the expected correct and error feedback. Further, the distribution of the difference between the unexpected and expected difference waves was significantly curved across the scalp (Table 1) and was maximal at channel FCz (-5.94 μV), although this difference was not significantly larger than at channel Pz (-3.99 μV), $t(16) = 1.60$, $p > .05$. Together, these results indicate that the

Table 1. *Polynomial Fits to Medial and Lateral Dimensions of Difference Wave Scalp Distributions*

Type	Dimension	Best fit	*F*
Control	medial	cubic	94.5**
	lateral	quadratic	75.7**
Unexpected	medial	quadratic	33.9**
	lateral	quadratic	86.9**
Expected	medial	cubic	43.4**
	lateral	quadratic	74.2**
Unexpected–Expected	medial	quadratic	7.62*
	lateral	quadratic	18.4**

Note. For all fits, $df = 1,16$.
*$p < .05$; **$p < .001$.

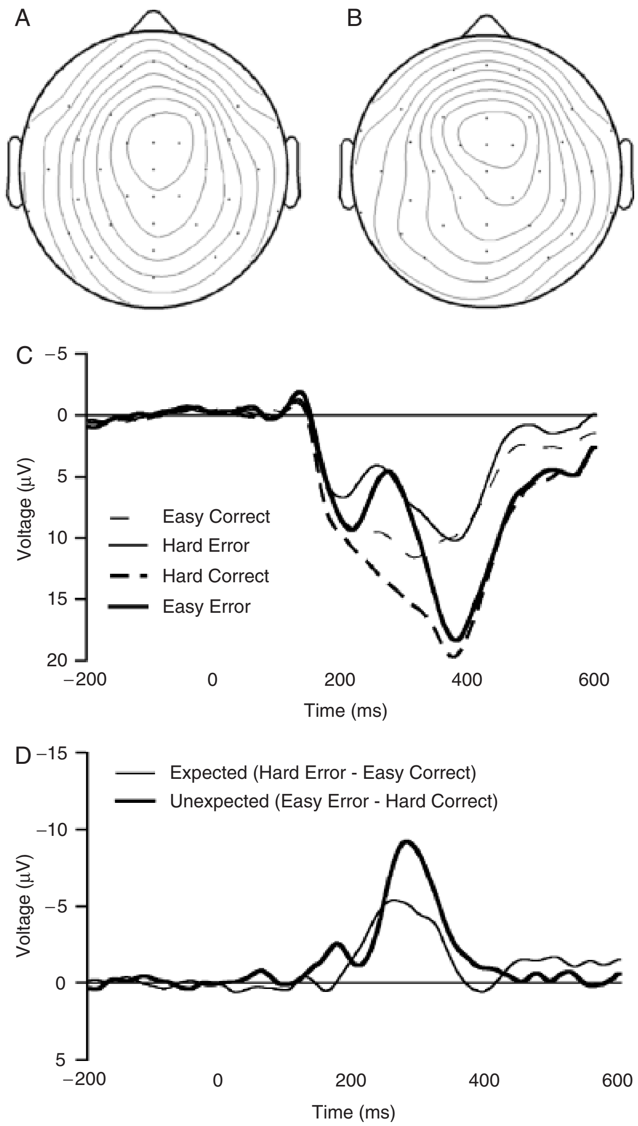

Figure 2. ERP data associated with the easy and hard conditions. A: Scalp distribution of unexpected difference wave; the change in potential between adjacent isopotential contours is -1.36 μV. B: Scalp distribution of expected difference wave; the change in potential between adjacent isopotential contours is -0.69 μV. C: ERPs recorded at channel FCz. D: Difference waves associated with channel FCz. Zero on abscissa indicates time of feedback onset. Note that negative voltages are plotted upward by convention.

relative increase in amplitude between the frequent and infrequent conditions resulted from an increase in the amplitude of the fERN, rather than from overlap with a different ERP component (such as the P300).

The amplitude of the difference wave was positively correlated across participants with the difference in the absolute change in response time, both of which increased with increasing unexpectedness of the eliciting feedback ($r = .45$, $p < .01$). Figure 1B plots the amplitude of the difference waves as a function of the difference in the absolute change in response times for the expected and unexpected events associated with each participant. Inspection of the figure revealed that, for every participant, both the amplitude of the fERN (measured as the difference between error and correct trials) and the absolute change in behavior following the trial (also measured as the difference between error and correct trials) was larger when the outcomes were unexpected (tip of the arrow) compared to when the outcomes were expected (base of the arrow). Thus, for every participant, unexpected compared to expected outcomes elicited relatively large fERNs and were associated with relatively large changes in behavior on subsequent trials.

Discussion

Consistent with the predictions of a recent theory (Holroyd & Coles, 2002), we found that fERN amplitude was differentially modulated more by unexpected than by expected correct and error feedback stimuli. Further, we found that differences in the behavioral adjustments following the feedback were larger for unexpected trials than for expected trials. Taken together, these results suggest that the fERN reflects the production of a reward prediction error signal for the adaptive modification of behavior (see also Frank, Woroch, & Curran, 2005).

These results confirm and extend previous findings on the fERN, which have been mixed (e.g., Hajcak, Moser, Holroyd, & Simons, 2006). Notably, in an experiment that was conceptually similar to that of the present study, participants selected among four ‘‘balloons’’ appearing on a computer screen to find a reward (Holroyd, Nieuwenhuis, Yeung, & Cohen, 2003). In one experimental condition, three of the four balloons contained a reward (and so the reward was likely), whereas in a second experimental condition, only one of the four balloons contained a reward (and so the reward was unlikely). As in the present study, it was found that the amplitude of the fERN was modulated more by unexpected outcomes than by expected outcomes. However, these previous results were relatively less robust compared to those of the present study, for the following reasons. First, in the previous study the fERN was measured as the difference between the base-to-peak amplitudes of the negativity occurring on error trials and on correct trials. Although this method captures variance associated with negative-going ERP components, it is insensitive to positive-going ERP components and may in fact confound fERN amplitude with the P300. In contrast, the fERN in the present study was determined as the maximum of the difference between the ERPs associated with error trials and correct trials, which is not affected by this problem. Second, the scalp distribution of the fERN was not fully characterized in the previous study, whereas here we demonstrated not only that the scalp distributions were frontal-central, but also that the difference between the distributions shared a comparable topography. Third, whereas in the previous study the relationship between fERN amplitude and behavior could not be assessed because the task lacked a useful performance measure, in the present experiment we demonstrated that the fERN amplitude was in fact correlated across participants with performance, such that relatively large fERNs were associated with relatively large changes in behavior. Although the results of several other fERN studies have been similarly suggestive (Nieuwenhuis et al., 2004), we believe that the present findings constitute the most solid evidence to date that the fERN reflects a reward prediction error signal for reinforcement learning.

REFERENCES

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

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

Hajcak, G., Moser, J., Holroyd, C. B., & Simons, R. F. (2006). The feedback-related negativity reflects the binary evaluation of good versus bad outcomes. *Biological Psychology*, *71*, 148–154.

Holroyd, C. B. (2004). A note on the Oddball N200 and the feedback ERN. In M. Ullsperger & M. Falkenstein (Eds.), *Errors, conflicts, and the brain*. *Current opinions on performance monitoring* (pp. 211–218). Leipzig: MPI of Cognitive Neuroscience.

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

Holroyd, C. B., Nieuwenhuis, S., Yeung, N., & Cohen, J. D. (2003). Errors in reward prediction are reflected in the event-related brain potential. *NeuroReport*, *14*, 2481–2484.

Luck, S. (2005). *The event-related potential technique*. Cambridge, MA: MIT Press.

Miltner, W. H. R., Braun, C. H., & Coles, M. G. H. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: Evidence for a "generic" neural system for error detection. *Journal of Cognitive Neuroscience*, *9*, 788–798.

Nieuwenhuis, S., Holroyd, C. B., Mol, N., & Coles, M. G. H. (2004). Reinforcement-related brain potentials from medial frontal cortex: Origins and functional significance. *Neuroscience & Biobehavioral Reviews*, *28*, 441–448.

Pakzad-Vaezi, K., Krigolson, O. E., & Holroyd, C. B. (2006). Underlying unity: The relationship between the N200 and the feedback ERN. *Psychophysiology*, *43*, S75.

(Received March 14, 2007; Accepted May 27, 2007)