

Errors in reward prediction are reflected in the event-related brain potential

Clay B. Holroyd,^{CA} Sander Nieuwenhuis, Nick Yeung and Jonathan D. Cohen

Department of Psychology, Green Hall, Princeton University, Princeton, NJ 08544, USA

^{CA}Corresponding Author: cholroyd@Princeton.EDU

Received 31 March 2003; accepted 9 April 2003

DOI: 10.1097/01.wnr.0000099601.41403.a5

The error-related negativity (ERN) is a negative deflection in the event-related brain potential associated with error processing. A recent theory holds that the ERN is elicited by the impact of a reward prediction error signal carried by the mesencephalic dopamine system on anterior cingulate cortex. The theory predicts that larger ERNs should be elicited by unexpected unfavorable outcomes than by expected unfavorable outcomes. We tested the

theory in an experiment in which the frequency of occurrence of reward was varied by condition, reasoning that the system that produces the ERN would come to expect non-reward when rewards were infrequent. Consistent with the theory, we found that larger ERNs were elicited by unexpected absences of reward. *NeuroReport* 14:2481–2484 © 2003 Lippincott Williams & Wilkins.

Key words: Anterior cingulate cortex; Dopamine; Error-related negativity; Prediction error; Reinforcement learning; Reward

INTRODUCTION

People sometimes correct their mistakes. The process of error correction is mediated in part by a cognitive system that monitors and optimizes behavior. In the short term the system detects performance errors and initiates remedial actions on the fly [1], and in the long term the system identifies rewarding and punishing events and reinforces adaptive behaviors [2]. Recently, studies of a component of the event-related brain potential (ERP) have provided neurophysiological evidence for the existence of this system (for a review of ERPs see [3]). It has been proposed [4,5] that this ERP component, called the error-related negativity (ERN), comes in two varieties: the response ERN and the feedback ERN. The response ERN is elicited in speeded response time tasks ~100 ms following error commission [6,7], and the feedback ERN is elicited in reinforcement learning tasks ~250 ms following presentation of stimuli that indicate incorrect performance [5]. Both the response ERN [8] and the feedback ERN [5,9,10] appear to be generated by an error-processing system involving anterior cingulate cortex (for reviews see [4,11–13]), although there exists disagreement on this point [14]. We have recently proposed that the ERN is produced by the impact of reinforcement learning signals carried by the mesencephalic dopamine system (MDS) on anterior cingulate cortex (hereafter called the reinforcement learning theory of the ERN, or the RL-ERN theory) [4]. The theory is based on evidence indicating that the MDS carries predictive error signals that indicate when ongoing events are better than expected and worse than expected [15,16] (for review see [17]). According

to the RL-ERN theory, a monitoring system located in the basal ganglia evaluates both external information in the environment and internal efference copies [18] of self-generated behaviors. From this information, the monitoring system predicts whether ongoing events will end in success or failure. When the monitoring system revises its predictions for the better, it induces a phasic increase in the activity of the MDS, and when the monitoring system revises its predictions for the worse, it induces a phasic decrease in the activity of the MDS. These positive and negative reward prediction error signals are carried by the MDS to various brain areas, where they are used to improve performance on the task at hand in line with principles of reinforcement learning [19]. Moreover, the error signals are used by the basal ganglia to improve its predictions of future reward and non-reward [15–17], such that events that tend to precede reward come to predict reward and events that tend to precede non-reward come to predict non-reward. According to the RL-ERN theory, the ERN is generated when a phasic decrease in mesencephalic dopaminergic activity (indicating that ongoing events are worse than expected) disinhibits the apical dendrites of neurons in anterior cingulate motor cortex; conversely, a positivity in the ERP is produced when a phasic increase in mesencephalic dopaminergic activity (indicating that ongoing events are better than expected) inhibits the apical dendrites of neurons in anterior cingulate motor cortex.

In this experiment we conducted a simple test of the RL-ERN theory. Participants engaged in a pseudo-reinforcement learning task in which, on each trial, they chose

between four response options. Following each response, the participants were presented with stimuli indicating the correctness (positive feedback) or incorrectness (negative feedback) of the response. Unknown to participants, the type of feedback was delivered at random: in a reward condition, the participants were exposed mostly to positive feedback, and in a non-reward condition, the participants were exposed mostly to negative feedback. We reasoned that the monitoring system would come to expect non-reward in the non-reward condition and to expect reward in the reward condition. Thus, we predicted that the negative feedback in the non-reward condition would be associated with small ERNs (because presentation of negative feedback would confirm the expectations of the monitoring system), whereas the negative feedback in the reward condition would be associated with large ERNs (because presentation of negative feedback would violate the expectations of the monitoring system).

MATERIALS AND METHODS

Participants: Ten undergraduate students at Princeton University (four male, age 21 ± 2 years) participated in the experiment either for course credit or for payment (\$10.00/h). In addition, all participants were paid about \$10.00 in bonus money at the end of the session (see below). The experiment was approved by the institutional review panel at Princeton University and was conducted in accordance with the Declaration of Helsinki. All participants took part in the experiment on the basis of informed consent.

Task: Participants sat comfortably about 1.5 m in front of a computer screen in an electromagnetically shielded room. On each trial of the task, participants saw an imperative stimulus (0.6° high, 5.0° wide, yellow against a black background) consisting of four circles in a row (O O O O). Participants were asked to imagine that these circles were balloons, and were told that one of the balloons contained 5 cents. The imperative stimulus remained on the screen until the participant selected a balloon by pressing one of four buttons on a response pad. A black screen was then presented for 0.5 s, followed by a feedback stimulus (0.6° high, 5.0° wide, yellow against a black background, 2 s duration). Positive feedback consisted of four dollar signs in a row (\$ \$ \$ \$), whereas negative feedback consisted of four Xs in a row (X X X X). Because the ERN is insensitive to the physical nature of the eliciting stimulus [4,5], the positive and negative feedback stimuli were not counterbalanced across participants. The interstimulus interval (ISI) between the offset of the feedback stimulus and the onset of the imperative stimulus was 0.5 s. Participants were told that presentation of a positive feedback stimulus indicated that the balloon they chose on that trial contained the 5 cents, and that presentation of a negative feedback stimulus indicated that the balloon they chose on that trial was empty. They were also told that at the end of the experiment they would be rewarded all money they found, and that they should respond in a way that would maximize the total amount of money earned. Unknown to the participants, on each trial the type of feedback stimulus was selected at random. In a reward condition, subjects received positive feedback on about 75% of the trials, and in a non-reward

condition, subjects received negative feedback on about 75% of the trials. Each condition consisted of one block of 200 trials, the order of which was counterbalanced across participants. At the end of the first condition the participants were informed of the total amount of bonus money earned by that time, and at the end of the session the participants were awarded the total amount of money earned in the experiment (about \$10.00).

Data acquisition: An electrode cap with Ag/AgCl electrodes was applied to each participant. The EEG was recorded along the midline according to the 10-20 system from channels FPz, Fz, FCz, Cz, CPz, and Pz. Other electrodes were placed on the right mastoid, above and below the right eye, and on the outer canthi of both eyes. The electrode common was placed on the chin or on the cheek. All electrode recordings were referenced to an electrode placed on the left mastoid. EEG data were recorded with Sensorium Inc. (Charlotte, VT) EPA-6 128 Channel Electro-Physiology Amplifiers at a sample rate of 250 Hz. Experimental control and data acquisition were controlled by E-Prime (Psychology Software Tools, Inc., Pittsburgh, PA) and Cogniscan (Newfoundland, NJ), respectively.

Participants completed a short questionnaire upon completion of the experiment.

Data analysis: For each feedback stimulus, a 1 s epoch of data (200 ms baseline) was extracted from the continuous data file for analysis. Ocular artifact was corrected with an eye-movement correction algorithm [20]. The EEG data were re-referenced off-line to linked-mastoid electrodes and baseline corrected by subtracting, from each sample of data recorded at each channel, one-half the activity recorded at the right mastoid and the average activity of that channel during the baseline period. Single-trial EEG data were lowpass filtered below 12 Hz with the Interactive Data Language (Research Systems, Inc., Boulder, CO) digital filter algorithm.

ERPs were obtained by averaging, for each channel, the EEG data according to feedback type and reward condition. The ERN was evaluated at channel FCz, where it reaches maximum amplitude [6,7]. Thus, for each subject, we obtained four ERPs for each combination of feedback type and reward condition: positive feedback in the reward condition (frequent reward), negative feedback in the non-reward condition (frequent non-reward), positive feedback in the non-reward condition (infrequent reward), and negative feedback in the reward condition (infrequent non-reward). For each of these four ERPs, the negativity associated with the ERN was measured base-to-peak according to the following algorithm. First, the algorithm identified the sample associated with the most positive value of the ERP within a 160–240 ms window following the presentation of the feedback stimulus. The latency of this sample was taken as the time of onset of the negativity. Then, the algorithm identified the sample associated with the most negative value of the ERP within a window extending from the onset of the negativity to 325 ms following the presentation of the feedback stimulus. If the latency of this sample was 325 ms (i.e. at the edge of the window), then the ERP component was considered to be a

positivity and the amplitude of the negativity was taken as $0\text{ }\mu\text{V}$. Otherwise, the latency of the sample was taken as the time of maximum component amplitude. The amplitude of the negativity was then defined by the difference in the ERP values associated with the component maximum and the component onset.

Importantly, measurements of ERN amplitude tend to be confounded by another ERP component called the P300 (Holroyd, Larsen, and Cohen, submitted), and P300 amplitude is strongly determined by the frequency of occurrence of the eliciting stimulus (for review see [21]). Thus, we expected that base-to-peak measures of ERN amplitude associated with frequent and infrequent non-rewards would be contaminated by P300-related activity. For this reason, we defined ERN amplitude as the difference between the component amplitudes associated with frequent reward and frequent non-reward, and as the difference between the component amplitudes associated with infrequent reward and infrequent non-reward. Given that the two conditions associated with each subtraction differed only in the valence of their associated feedback (positive and negative), and not in the frequency of occurrence of the feedback, this procedure afforded us a relatively pure means for measuring the activity of the reward-related processes that give rise to the ERN.

Four grand average waveforms were obtained by averaging the four ERPs associated with each combination of feedback type and reward condition across participants. The reliability of the ERN amplitude measure was assessed with confidence intervals and paired *t*-tests.

RESULTS

Fig. 1 shows the ERPs elicited by the feedback stimuli (channel FCz). Fig. 1a illustrates the ERPs associated with each of the infrequent outcomes (infrequent reward and infrequent non-reward) and Fig. 1b illustrates the ERPs associated with each of the frequent outcomes (frequent reward and frequent non-reward). Note that the ERPs that compose each pair of waveforms are associated with different task conditions. As can be seen, the neural system that produces the ERN was differentially activated by the presentation of positive and negative feedback for both the frequent and infrequent outcomes. The ERN was larger for the infrequently presented feedback than for the frequently presented feedback ($t(9)=3.1$, $p=0.01$), and confidence intervals (0.95) confirmed that the ERN amplitudes were statistically reliable (infrequent condition: mean = $-5.3\text{ }\mu\text{V}$, interval = $-6.3\text{ }\mu\text{V}$, $-4.3\text{ }\mu\text{V}$; frequent condition: mean = $-2.8\text{ }\mu\text{V}$, interval = $-3.6\text{ }\mu\text{V}$, $-2.1\text{ }\mu\text{V}$).

DISCUSSION

The RL-ERN theory proposes that the ERN is elicited by the impact of phasic mesencephalic dopaminergic activity on anterior cingulate cortex [4]. The theory is based on recent evidence indicating that phasic activity of the MDS reflects an error in reward prediction, such that when ongoing events are worse than expected, a phasic decrease in mesencephalic dopaminergic activity occurs, and when ongoing events are better than expected, a phasic increase in mesencephalic dopaminergic activity occurs [15,16] (for

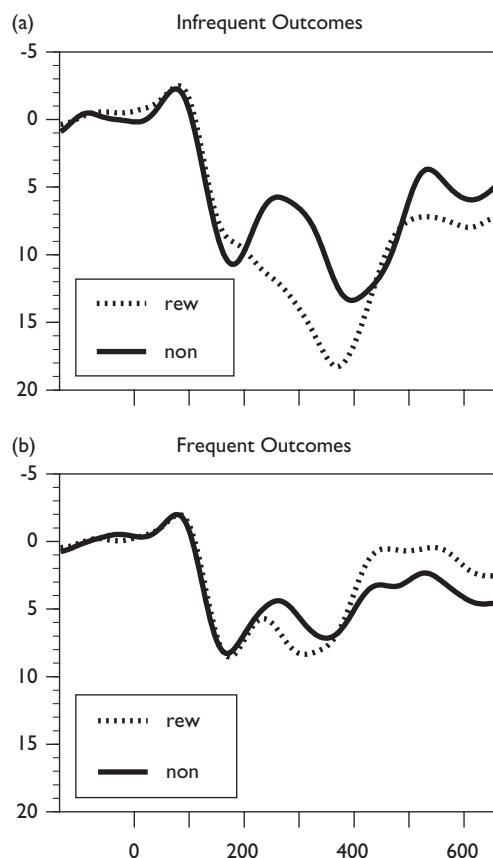


Fig. 1. ERPs recorded at channel FCz. (a) ERPs associated with infrequent rewards and infrequent non-rewards. (b) ERPs associated with frequent rewards and frequent non-rewards. rew, reward; non, non-reward. 0 ms corresponds to time of stimulus onset.

review see [17]). According to the RL-ERN theory, the ERN is generated following error commission when a negative prediction error signal carried by the MDS disinhibits anterior cingulate cortex; conversely, a positive deflection in the ERP occurs following correct responses when a positive prediction error signal carried by the MDS inhibits anterior cingulate cortex. These signals are used to train anterior cingulate cortex to improve performance on the task at hand according to principles of reinforcement learning [19]. The results of the present experiment confirm a fundamental prediction of the RL-ERN theory: larger ERNs are elicited by unexpected unfavorable outcomes than by expected unfavorable outcomes. According to the theory, the monitoring system came to expect non-reward in the non-reward condition and to expect reward in the reward condition. Thus the non-rewards in the non-reward condition elicited small ERNs because the non-rewards were consistent with the system's expectations. In contrast, the non-rewards in the reward condition elicited large ERNs because these non-rewards were inconsistent with the system's expectations. These results parallel previous findings that the amplitude of the response ERN increases with increasing accuracy [7], because the system expects correct responses under conditions in which errors are few [4]. When taken together with the results of previous studies

[4,22,23] (but see also [24]), these findings support the hypothesis that the ERN reflects a reward prediction error associated with the mesencephalic dopamine system.

However, it has also been proposed that the feedback ERN reflects the affective significance or emotional valence of the eliciting stimulus [9,14]. For example, it could be argued that infrequent errors elicit large-amplitude ERNs because infrequent errors are more salient than frequent errors and thus elicit a larger affective response. Although future research is needed to investigate this issue, the RL-ERN theory appears to be compatible with the idea that the ERN is associated with emotional processing. For example, Antonio Damasio has proposed that human decision making may be biased by a somatic marker that reflects an emotional reaction to ongoing events. The somatic marker forces attention on the negative outcome to which a given action may lead, and functions as an automated alarm signal that says: Beware of danger ahead if you choose the option that leads to this outcome [25]. Moreover, Damasio has suggested that these markers may be conveyed to cortex by subcortical catecholaminergic systems such as the MDS [25]. Given that the RL-ERN theory holds that the MDS conveys error signals associated with the predicted outcomes of ongoing events [15–17], and that these signals bias action-selection processes mediated by anterior cingulate motor cortex, the RL-ERN theory may constitute a formal instantiation of the somatic marker hypothesis.

CONCLUSION

The ERN is an ERP component associated with error processing. These results suggest that the ERN is elicited by a reward prediction error, such that unpredicted non-rewards elicit the largest ERNs. The findings are consistent with the hypothesis that the ERN is generated by the impact of reinforcement learning signals carried by the MDS on anterior cingulate cortex [4].

REFERENCES

- Rabbitt PMA. Error correction time without external error signals. *Nature* 438 (1966).
- Thorndike EL. Laws and hypotheses for behavior. In: Thorndike EL (ed.). *Animal Intelligence*. Darien, CT: Hafner Publishing Co; 1970, pp. 241–281.
- Coles MGH and Rugg MD. Event-related brain potentials: an introduction. In: Rugg MD and Coles MGH (eds). *Electrophysiology of Mind*. Oxford: Oxford University Press; 1995, pp. 1–26.
- Holroyd CB and Coles MGH. The neural basis of human error processing: reinforcement learning, dopamine, and the Error-related Negativity. *Psychol Rev* 109, 679–709 (2002).
- Miltner WHR, Braun CH and Coles MGH. Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a "generic" neural system for error detection. *J Cogn Neurosci* 9, 788–798 (1997).
- Falkenstein M, Hohnsbein J, Hoormann J and Blanke L. Effects of errors in choice reaction tasks on the ERP under focused and divided attention. In: Bruna C, Gaillard A and Kok A (eds). *Psychophysiological Brain Research*. Tilburg, The Netherlands: Tilburg University Press; 1990, pp. 192–195.
- Gehring WJ, Goss B, Coles MGH, Meyer DE and Donchin E. A neural system for error detection and compensation. *Psychol Sci* 4, 385–390 (1993).
- Dehaene S, Posner MI and Tucker DM. Localization of a neural system for error detection and compensation. *Psychol Sci* 5, 303–305 (1994).
- Gehring WJ and Willoughby AR. The medial frontal cortex and the rapid processing of monetary gains and losses. *Science* 295, 2279–2282 (2002).
- Ullsperger M and von Cramon DY. Error monitoring using external feedback: specific roles of the habenular complex, the reward system, and the cingulate motor area revealed by fMRI. *J Neurosci* 23, 4308–4314 (2003).
- Coles MGH, Scheffers MK and Holroyd C. Berger's dream? The error-related negativity and modern cognitive psychophysiology. In: Witten H, Zweiner U, Schack B and Döring A (eds). *Quantitative and Topological EEG and MEG analysis*. Jena-Erlangen: Druckhaus Mayer Verlag; 1998, pp. 96–102.
- Coles MGH, Scheffers MK and Holroyd C. Why is there an ERN/Ne on correct trials? Response representations, stimulus-related components, and the theory of error-processing. *Biol Psychol* 56, 191–206 (2001).
- Falkenstein M, Hoormann J, Christ S and Hohnsbein J. ERP components on reaction errors and their functional significance: a tutorial. *Biol Psychol* 51, 87–107 (2000).
- Luu P, Tucker DM, Derryberry D, Reed M and Poulsen C. Electrophysiological responses to errors and feedback in the process of action regulation. *Psychol Sci* 14, 47–53 (2003).
- Barto AG. Adaptive critics and the basal ganglia. In: Houk J, Davis J and Beiser D (eds). *Models of Information Processing in the Basal Ganglia*. Cambridge, MA: MIT Press; 1995, pp. 215–232.
- Montague PR, Dayan P and Sejnowski TJ. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci* 16, 1936–1947 (1996).
- Schultz W. Getting formal with dopamine and reward. *Neuron* 36, 241–263 (2002).
- Angel RW. Efference copy in the control of movement. *Neurology* 26, 1164–1168 (1976).
- Sutton RS and Barto AG. *Reinforcement Learning: An Introduction*. Boston: MIT Press, 1998.
- Gratton G, Coles MGH and Donchin E. A new method for off-line removal of ocular artifact. *Electroencephalogr Clin Neurophysiol* 55, 468–484 (1983).
- Donchin E and Coles MGH. Is the P300 component a manifestation of context updating? *Behav Brain Sci* 11, 355–372 (1988).
- Falkenstein M, Hielscher H, Dziobek I, Schwarzenau P, Hoormann J, Sundermann B et al. Action monitoring, error detection, and the basal ganglia: an ERP study. *Neuroreport* 12, 157–161 (2001).
- Nieuwenhuis S, Ridderinkhof KR, Talsma D, Coles MGH, Holroyd CB, Kok A et al. A computational account of altered error processing in older age: Dopamine and the Error-related Negativity. *Cogn Affect Behav Neurosci* 2, 19–36 (2002).
- Holroyd CB, Praamstra P, Plat E and Coles MGH. Spared error-related potentials in mild to moderate Parkinson's disease. *Neuropsychologia* 40, 2116–2124 (2002).
- Damasio AR. *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Avon Books, 1994.

Acknowledgements: We are grateful to Jack Gelfand and Dustin Reagan for technical assistance. This research was supported in part by Mental Health (NIMH) grant MH62196 and postdoctoral fellowship MH63550. The research of S.N. was supported by the Netherlands Organization for Scientific Research (NWO).