Reward positivity elicited by predictive cues

Clay B. Holroyd[a], Olav E. Krigolson[b] and Seung Lee[c]

A recent theory holds that a component of the human event-related brain potential called the reward positivity reflects a reward prediction error signal. We investigated this idea in gambling-like task in which, on each trial, a visual stimulus predicted a subsequent rewarding or nonrewarding outcome with 80% probability. Consistent with earlier results, we found that the reward positivity was larger to unexpected than to expected outcomes. In addition, we found that the predictive cues also elicited a reward positivity, as proposed by the theory. These results indicate that the reward positivity reflects the initial assessment of whether a trial will end in success or failure and the reappraisal of that information once the outcome actually occurs. ***NeuroReport*** **22:249–252 © 2011 Wolters Kluwer Health | Lippincott Williams & Wilkins.**

NeuroReport 2011, **22**:249–252

Keywords: event-related brain potential, feedback error-related negativity, predictive cues, reinforcement learning, reward positivity, reward prediction error

[a]Department of Psychology, University of Victoria, Victoria, [b]Department of Psychology and Neuroscience, Dalhousie University and [c]School of Population and Public Health, University of British Columbia, British Columbia, Canada

Correspondence to Clay B. Holroyd, PhD, Department of Psychology, University of Victoria, Cornett Building A236, 3800 Finnerty Road (Ring Rd), Victoria, BC V8P 5C2, Canada
Tel: +1 250 853 3910; fax: +1 250 721 8929; e-mail: holroyd@uivic.ca

Received 21 January 2011 **accepted** 1 February 2011

Introduction

Research on the neural mechanisms of reinforcement learning and decision-making in humans has been informed by studies of a component of the event-related brain potential (ERP) that is alternatively called the reward positivity [1,2] and the feedback error-related negativity [3]. This ERP component is typically measured as a negative deflection in the ERP following negative performance feedback relative to positive performance feedback, peaking approximately 250 ms poststimulus over frontal-central regions of the scalp [3,4]. However, recent studies have indicated that the difference between the ERPs elicited by feedback with positive and negative valence results mainly from a positive-going deflection (the reward positivity) elicited by a reward-related neural process, rather than by a negative deflection elicited by an error-related process [1] (see Refs [5–8]). The reward positivity is elicited by a neural system that evaluates in a context-sensitive manner whether or not a task goal has been achieved [9,10]. We have proposed earlier that the reward positivity reflects a reward prediction error (RPE) signal, being largest for unexpected rewards and no-rewards relative to expected rewards and no-rewards [11]. This prediction has been confirmed across several experiments (e.g. [4,11–15]).

Much attention in the literature has been directed at reward positivities elicited by feedback stimuli in reinforcement leaning and guessing tasks. The RPE theory holds that the reward positivity is elicited by reward-related events that deviate from expectation, being more positive for good relative to bad events when these events are equally unexpected. This assertion holds for antecedent events that reliably predict trial outcomes and for the outcomes themselves, such that the amplitudes of the reward positivity to predictive events and subsequent outcomes are inversely related [11]. For example, in trial and error learning tasks with speeded response deadlines, the reward positivity is larger for predictive responses than for unpredictive responses and larger for unpredicted feedback stimuli than for predicted feedback stimuli (e.g. [6,11]). Yet surprisingly few studies have investigated whether the reward positivity is also elicited by predictive cues [16–19], and there has not been a clear demonstration that this ERP component is elicited by both cues and outcomes when the former probabilistically predict the latter.

We addressed this issue by replicating a reward positivity experiment in which on each trial the stimulus cue predicted the outcome with 80% probability. Specifically, in a study by Potts *et al.* [8], participants were required on each trial to view passively two stimuli presented in succession on a computer screen, each of which consisted of either an image of a gold bar or a lemon, and were informed that they would win some money whenever the second stimulus was a gold bar. The first stimulus in each pair predicted the second stimulus with an 80% probability such that 80% of the encounters with an initial gold bar resulted in a reward and 80% of encounters with an initial lemon resulted in no reward. Crucially, the researchers examined the ERPs to the outcomes but not to the predictive cues. Here we replicated and extended their findings by inspecting the ERPs to the predictive cues, and to the outcomes, for the presence of the reward positivity.

Methods

Participants

Eighteen undergraduate students (seven male; mean age = 19.7 ± 2.5 years) with no known neurological impairments and with normal or corrected-to-normal vision participated in the experiment. All the participants were volunteers who received extra credit in a first or second year

0959-4965 © 2011 Wolters Kluwer Health | Lippincott Williams & Wilkins DOI: 10.1097/WNR.0b013e328345441d

psychology course for their participation, and provided written, informed consent. Participants also received a task-related bonus at the end of the experiment (see below). The study was conducted in accordance with the ethics standards prescribed in the Declaration of Helsinki and was approved by the human participants 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 the gambling-like task used by Potts *et al.* [8]. On each trial participants observed two visual images of a lemon and/or a gold bar appearing in sequence (5°, 500 ms) in the center of a computer screen. The first stimulus on each trial (S1) was a gold bar on 50% of the trials and a lemon on the remaining trials; all probabilities were random without replacement (thereby enforcing exactly equal numbers of trials across conditions, for example, 30 gold bars and 30 lemons to S1 for each block). Further, S1 predicted the second stimulus (S2) with 80% probability, that is, if S1 was a gold bar then S2 was a gold bar on 80% of the trials and was a lemon on the remaining 20% of trials, and if S1 was a lemon then S2 was a lemon on 80% of the trials and a gold bar on the remaining 20% of the trials. Participants were informed at the start of the experiment that they would earn 10 cents CAN each time S2 was a gold bar stimulus, and that they would earn nothing each time S2 was a lemon stimulus. Each trial started with the display of a fixation stimulus (a ' + ' image, 1.3°, 300 ms), followed by S1, a second fixation stimulus (300 ms), S2, a third fixation stimulus (300 ms), and a feedback image that indicated the reward earned on that trial and the total reward accumulated across blocks (600 ms). The task was divided into eight blocks of 60, 2500 ms-long trials, with short rest periods given between blocks. Participants pressed the space bar of a standard USB keyboard to initiate each block but otherwise the task was entirely passive. On completion they were asked to complete a debriefing questionnaire and were provided with their accumulated bonus money (approximately $24).

Data acquisition

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 electrooculogram 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 0.017–67.5 Hz (90 dB octave roll off).

Data analysis

The EEG data were filtered offline through a 0.1–20 Hz passband phase shift free Butterworth filter and re-referenced mathematically to linked mastoid electrodes. Ocular artefacts were removed using the Gratton, Coles, and Donchin method. Trials in which the change in voltage at any channel exceeded 35 μVs per sampling point were also discarded.

ERPs elicited by S1 and S2 were created by extracting 800 ms epochs from the continuous EEG for each trial, channel and participant, baseline corrected to the 200 ms preceding the onset of each stimulus. Average ERPs were created by averaging the EEG data for each electrode channel and participant according to six conditions: S1 ERPs: Predicted reward (S1 = gold) and Predicted no-reward (S1 = lemon); S2 ERPs: Unexpected reward (S1 = lemon, S2 = gold), Unexpected no-reward (S1 = gold, S2 = lemon), Expected reward (S1 = gold, S2 = gold), and Expected no-reward (S1 = lemon, S2 = lemon). For the purpose of illustration (Fig. 1a), the Expected and Unexpected reward and no-reward ERPs were created with 1800 ms epochs of EEG baseline-corrected to the 200 ms preceding S1.

The reward positivity was evaluated at channel FCz, in which it typically reaches maximum amplitude [3,4]. Following Holroyd and Krigolson [4], for each participant an 'Unexpected' difference wave was created by subtracting the Unexpected Reward ERP from the Unexpected no-reward ERP, and an 'Expected' difference wave was created by subtracting the Expected reward ERP from the Expected no-reward ERP. Further, a 'Predicted' difference wave was created by subtracting the Predicted reward ERP from the Predicted no-reward ERP. The amplitude of each difference wave was measured for each participant as the most negative deflection within the 200–300 ms after stimulus onset. This choice of time window followed Potts *et al.* [8]; further investigation using wider time windows yielded comparable results (data not shown).

Difference wave values were tested for statistical significance using the paired *t* tests. For the purpose of display, grand-average ERPs and difference waves were created by averaging the ERP and difference wave data across participants. Scalp distributions of difference waves were created by averaging across participants the values of the differences at each electrode location, at the time of peak difference at channel FCz, and were plotted using spherical spline interpolation.

Results

Behavioral results

The passive nature of this task did not provide a performance measure. Participants' responses to a debriefing questionnaire indicated moderate motivation to carry out the task. Specifically, asked to rate on a scale of 1 (low) and 5 (high) the extent to which they felt money motivated them to pay close attention to the stimuli, the average response was 2.8 (± 1.2). Participants nevertheless accurately reported that the first stimulus in each pair was the same as the second stimulus on approximately 80% of the trials ($77.1 \pm 11.7\%$), indicating that they paid attention to the task.

Electrophysiological results

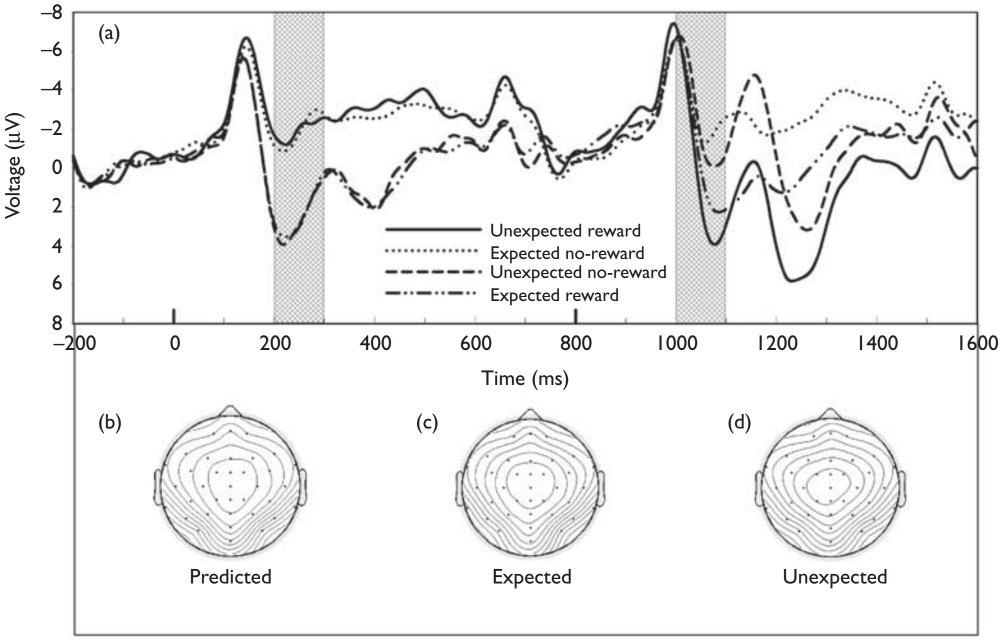
Figure 1a illustrates the ERPs elicited by the cues predicting Reward and No-reward, separately according to whether the outcome was consistent (Expected) or inconsistent (Unexpected) with the predictions; the scalp distributions associated with the Predicted, Expected, and Unexpected difference waves are indicated below (Fig. 1b–d). As observed in previous studies, the paired t test indicated that the amplitude of the Unexpected difference wave ($-7.6\,\mu\text{V}$) was significantly larger than the amplitude of the Expected difference wave ($-5.5\,\mu\text{V}$), $t(17) = 2.3$, P value of less than 0.05, Cohen's $D = 0.82$. Importantly, the Predicted difference wave showed a negative deflection in the time range of the reward positivity ($-6.4\,\mu\text{V}$) that was significantly different from $0\,\mu\text{V}$, $t(17) = 13.7$, P value of less than 0.001, Cohen's $D = 4.55$. All three difference waves were distributed frontal-centrally over the scalp (Fig. 1b–d).

Discussion

The RPE theory holds that the amplitude of the reward positivity is sensitive to the expectedness of reward-related events, such that the size of the difference in the ERPs elicited by positive and negative events is positively correlated with the degree that those events are unpredicted. Although aspects of this theory have been confirmed across a variety of studies (e.g. [4–8,11–19]), to our knowledge this experiment is the first to examine the reward positivity elicited by cues that probabilistically predict upcoming Rewards and No-rewards. Our results show that such cues impact the ERP in a manner that is consistent with the production of a reward positivity (Fig. 1). When the outcome information is subsequently delivered, these predictions are revised as appropriate, such that the reward positivity is larger to unexpected than expected outcomes [4,11–15]. These results indicate that the reward positivity is sensitive to the delicate interplay between the evaluation and reevaluation of ongoing events by a neural system that provides initial assessments of future reward and then revises those predictions as appropriate.

Note that the ERPs to the unexpected outcomes elicit a negative deflection approximately 300 ms postfeedback (corresponding to 1100–1200 ms in Fig. 1a) that is consistent with its identification as the N200 ERP component [1,2]. Measured 'base to peak', the presence of this ERP component might suggest that the reward positivity or feedback error-related negativity is elicited

Fig. 1



Event-related brain potential (ERP) data. (a) ERPs recorded at channel FCz averaged according to whether the Reward and No-reward outcomes were consistent (Expected) or inconsistent (Unexpected) with the information conveyed by the predictive cues. Negative is plotted up by convention. Bold tick marks indicate cue and feedback onset at 0 and 800 ms, respectively. Shaded areas indicate periods during which the reward positivity was evaluated. (b–d) Scalp distributions of the reward positivity to the predictive cues, expected feedback stimuli, and unexpected feedback stimuli.

by unexpected events in general, as opposed to unexpected errors or rewards in particular [20]. However, the reward positivity is elicited by unexpected rewarding events but not by unexpected nonrewarding events [1,2,5–8], whereas the N200 is elicited by unexpected events irrespective of valence [20]. As the reward positivity normally coincides temporally and spatially with the N200, it either cancels out the N200 or shifts the entire ERP during this period in the positive direction, as illustrated in Fig. 1a [1]. For this reason we isolated the reward positivity using a difference wave approach that removes the main effect of expectancy associated with the N200 (and the subsequent ERP component, the P300), thereby capturing the interaction of expectancy and valence [4]. Note that this approach departs from the difference wave method used by Potts *et al.* who computed 'worse than expected' difference waves by subtracting the Predicted reward ERPs from the Unpredicted no-reward ERPs, and 'better than expected' difference waves by subtracting the Predicted no-reward ERPs from the Unpredicted reward ERPs. These subtractions confound the interaction of valence and expectancy with the main effect of expectancy.

As this experiment required participants to view stimuli passively without responding, the results support those of earlier studies that also identified reward positivities in the absence of overt task-related behavior [8,12,21]. Further, although the effect size associated with the difference between the expected and unexpected reward positivities was small (Cohen's $D = 0.82$), this finding is consistent with the results of a series of recent studies that have shown that the amplitude of the reward positivity depends sensitively on the response generation process [21–23]. For example, we have shown that the RPE effect to outcomes (such that the reward positivity is larger to unexpected than expected events) is largest under conditions in which the response-feedback relationship is meaningful (i.e. nonrandom) [14], and appears only if the response is produced before, rather than after, the prediction is made [13]. Further, the sense of personal responsibility over choice behavior seems to be a critical factor underlying this process [24].

Conclusion

These findings indicate that the reward positivity reflects an initial appraisal of information that predicts future reward and the subsequent reappraisal of that information on reward delivery. The underlying neural system is therefore sensitive to task performance as each trial progresses, even in the absence of task-related behaviors and before forthcoming rewards.

Acknowledgements

This study was supported in part by Natural Sciences and Engineering Research Council of Canada Discovery Grants RGPIN 312409-05 and RGPIN 386139-2010. The authors are grateful to the research assistants in the Brain and Cognition Laboratory for help with data collection.

References

1 Holroyd CB, Pakzad-Vaezi KL, Krigolson OE. The feedback correct-related positivity: sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology* 2008; **45**:688–697.

2 Baker TE, Holroyd CB. Dissociated roles of the anterior cingulate cortex in reward and conflict processing as revealed by feedback error-related negativity and N200. *Biol Psychol* (in press).

3 Miltner WHR, Braun CH, 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* 1997; **9**:788–798.

4 Holroyd CB, Krigolson OE. Reward prediction error signals associated with a modified time estimation task. *Psychophysiology* 2007; **44**:913–917.

5 Cohen MX, Elger CE, Ranganath C. Reward expectation modulates feedback-related negativity and EEG spectra. *Neuroimage* 2007; **35**:968–978.

6 Eppinger B, Kray J, Mock B, Mecklinger A. Better or worse than expected? Aging, learning, and the ERN. *Neuropsychologia* 2008; **46**:521–539.

7 Foti D, Weinberg A, Dien J, Hajcak G. Event-related potential activity in the basal ganglia differentiates reward from nonrewards: temporalspatial principal components analysis and source localization of the feedback negativity. *Hum Brain Mapp* (in press).

8 Potts GF, Martin LE, Burton P, Montague PR. When things are better or worse than expected: the medial frontal cortex and the allocation of processing resources. *J Cogn Neurosci* 2006; **18**:1112–1119.

9 Holroyd CB, Larsen JT, Cohen JD. Context dependence of the event-related brain potential associated with reward and punishment. *Psychophysiology* 2004; **41**:245–253.

10 Holroyd CB, Hajcak G, Larsen JT. The good, the bad and the neutral: electrophysiological responses to feedback stimuli. *Brain Res* 2006; **1105**:93–101.

11 Holroyd CB, Coles MGH. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol Rev* 2002; **109**:679–709.

12 Donkers FCL, Van Boxtel GJM. Mediofrontal negativities to averted gains and losses in the slot-machine task: a further investigation. *J Psychophysiology* 2005; **19**:256–262.

13 Hajcak G, Moser JS, Holroyd CB, Simons RF. It's worse than you thought: the feedback negativity and violations of subjective expectancy. *Psychophysiology* 2007; **44**:905–912.

14 Holroyd CB, Krigolson OE, Baker R, Lee S, Gibson J. When is an error not a prediction error? An electrophysiological investigation. *Cogn Affect Behav Neurosci* 2009; **9**:59–70.

15 Holroyd CB, Nieuwenhuis S, Yeung N, Cohen JD. Errors in reward prediction are reflected in the event-related brain potential. *Neuroreport* 2003; **14**:2481–2484.

16 Krigolson OE, Holroyd CB. Predictive information and error processing: the role of medial-frontal cortex during motor control. *Psychophysiology* 2007; **44**:586–595.

17 Dunning JP, Hajcak G. Error-related negativities elicited by monetary loss and cues that predict loss. *Neuroreport* 2007; **18**:1875–1878.

18 Yu R, Zhou X. Brain potentials associated with outcome expectation and outcome evaluation. *Neuroreport* 2006; **17**:1649–1653.

19 Baker TE, Holroyd CB. Which way do I go? Neural activation in response to feedback and spatial processing in a virtual T-maze. *Cereb Cortex* 2009; **19**:1708–1722.

20 Oliveira FTP, McDonald JJ, Goodman D. Performance monitoring in the anterior cingulate is not all error related: expectancy deviation and the representation of action-outcome associations. *J Cogn Neurosci* 2007; **19**:1994–2004.

21 Yeung N, Holroyd CB, Cohen JD. ERP correlates of feedback and reward processing in the presence and absence of response choice. *Cereb Cortex* 2005; **15**:535–544.

22 Peterson DA, Lotz DT, Halgren E, Sejnowski TJ, Poizner H. Choice modulates the neural dynamics of prediction error processing during reward learning. *Neuroimage* 2011; **54**:1385–1394.

23 Zhou Z, Yu R, Zhou X. To do or not to do? Action enlarges the FRN and P300 effects in outcome evaluation. *Neuropsychologia* 2010; **48**:3606–3613.

24 Li P, Han C, Lei Y, Holroyd CB, Li H. Responsibility modulates neural mechanisms of outcome processing: an ERP study. *Psychophysiology* (in press).