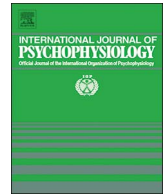




Contents lists available at ScienceDirect

International Journal of Psychophysiology

journal homepage: www.elsevier.com/locate/ijpsycho

Feedback information and the reward positivity

Jeffrey Cockburn^{a,*}, Clay B. Holroyd^b^a Division of Humanities and Social Sciences, California Institute of Technology, 1200 East California Blvd, Pasadena 91125, CA, USA^b Department of Psychology, P. O. Box 1700, STN CSC, University of Victoria, Victoria V8W 2Y2, British Columbia, Canada

ARTICLE INFO

Keywords:

Reward positivity
 Event-related brain potential
 Reinforcement learning
 Graded feedback
 Time estimation
 Feedback-related negativity

ABSTRACT

The reward positivity is a component of the event-related brain potential (ERP) sensitive to neural mechanisms of reward processing. Multiple studies have demonstrated that reward positivity amplitude indices a reward prediction error signal that is fundamental to theories of reinforcement learning. However, whether this ERP component is also sensitive to richer forms of performance information important for supervised learning is less clear. To investigate this question, we recorded the electroencephalogram from participants engaged in a time estimation task in which the type of error information conveyed by feedback stimuli was systematically varied across conditions. Consistent with our predictions, we found that reward positivity amplitude decreased in relation to increasing information content of the feedback, and that reward positivity amplitude was unrelated to trial-to-trial behavioral adjustments in task performance. By contrast, a series of exploratory analyses revealed frontal-central and posterior ERP components immediately following the reward positivity that related to these processes. Taken in the context of the wider literature, these results suggest that the reward positivity is produced by a neural mechanism that motivates task performance, whereas the later ERP components apply the feedback information according to principles of supervised learning.

1. Introduction

The ability of humans to approach good things and avoid bad things derives from the fundamental ability to distinguish between the two in the first place. Over the past two decades, a component of the human event-related brain potential (ERP) has provided evidence for the neural mechanism that underpins this process. Originally called the “feedback error-related negativity” because of the presence of a negative-going deflection in the ERP about 250 ms following error feedback (Miltner et al., 1997), it has recently been re-termed the “reward positivity” in recognition that the difference in the ERPs to positive and negative feedback results more from reward processing than from error processing (Proudfit, 2015; Holroyd and Umemoto, 2016). The logic underlying this inference stems from an analysis approach that removes confounding ERP components by subtracting one ERP from another (Luck, 2014). On this view, task-related stimuli in general – whether error-related or not – elicit a negative-going ERP deflection called the N2 (Holroyd, 2004). However, an exception occurs to reward feedback, which does not elicit the N2 – indicating that the difference in the ERPs to positive and negative feedback is specifically due to reward processing (Holroyd et al., 2008b; see also Baker and Holroyd, 2011; Foti et al., 2011b; Holroyd et al., 2011; Warren and Holroyd, 2012).

Much has been learned about the reward positivity in the 20 years

since its discovery (for reviews see Sambrook and Goslin, 2015; Walsh and Anderson, 2012). In particular, we proposed that the ERP component reflects a specific type of reward signal called a reward prediction error (Holroyd and Coles, 2002). Reward prediction error signals are modulated by feedback expectancy but have opposite signs for unexpected positive events relative to unexpected negative events (Sutton and Barto, 1998). By contrast, “surprise” or “salience” signals are larger to unexpected events relative to expected events, but have the same sign irrespective of the valence of the outcomes. Although controversial (e.g., Cohen et al., 2011; Ferdinand et al., 2012; Hajcak et al., 2005; Holroyd et al., 2012; Oliveira et al., 2007; Sallet et al., 2013; Talmi et al., 2013; Ullsperger et al., 2014), numerous studies have indicated that the reward positivity behaves as a reward prediction error signal (e.g., Heydari and Holroyd, 2016; Holroyd and Krigolson, 2007; Holroyd et al., 2003; Holroyd et al., 2009; Martin et al., 2009), as confirmed by a meta-analysis of over 55 data sets (Sambrook and Goslin, 2015). It is less clear whether the ERP component is sensitive to reward magnitude (Hajcak et al., 2006; Holroyd et al., 2006), but the meta-analysis also suggests this to be the case (Sambrook and Goslin, 2015). The reward positivity is also sensitive to the dimension of feedback valence that is most salient (Nieuwenhuis et al., 2004) and to the context in which the rewards are delivered (e.g., Holroyd et al., 2004), and is generally larger following responses (Hajcak et al., 2007)

* Corresponding author.

E-mail addresses: jcockbur@caltech.edu (J. Cockburn), holroyd@uvic.ca (C.B. Holroyd).<https://doi.org/10.1016/j.ijpsycho.2017.11.017>Received 1 July 2017; Received in revised form 25 November 2017; Accepted 29 November 2017
0167-8760/© 2017 Elsevier B.V. All rights reserved.

and in tasks with controllable outcomes (e.g., Holroyd et al., 2009; Li et al., 2011; Sambrook and Goslin, 2015; Warren and Holroyd, 2012; Yeung et al., 2005).

Reward prediction error signals provide a computationally powerful means for the adaptive modification of behavior according to principles of reinforcement learning (Sutton and Barto, 1998). A recent study provides a striking demonstration that the reward positivity functions as such a reinforcer: In the absence of overt human behavior, a brain-computer interface utilized the reward positivity to train an artificial agent to perform a task (Zander et al., 2016). For this reason, it is somewhat puzzling that variation in reward positivity amplitude has been inconsistently associated with behavioral adaptation across studies (for reviews see Holroyd and Umemoto, 2016; Walsh and Anderson, 2012). For example, in one notable study, task instructions immediately modified participants' behavior but not reward positivity amplitude, illustrating a dissociation between the two (Walsh and Anderson, 2011). Likewise, the reward positivity can be elicited even in the absence of overt behavior that immediately precedes the feedback (e.g., Yeung et al., 2005) and is sensitive to task contexts that predict reward (Umemoto et al., 2017).

One possibility for this seeming discrepancy is that the reinforcing signal modulates a higher-level decision making mechanism that is relatively unconcerned about low-level regulation of the action production system (Holroyd and Coles, 2002). In fact, a growing literature suggests that the reward positivity is associated with motivational factors related to task performance (Holroyd and Umemoto, 2016). In particular, the reward positivity is sensitive to individual differences in personality (e.g., Cherniawsky and Holroyd, 2013; Umemoto and Holroyd, 2017; Schmidt et al., in press), emotions (Foti et al., 2011a; Hewig et al., 2011), and psychiatric disorders (e.g., Baker et al., 2011; Hewig et al., 2010; Holroyd et al., 2008a; Umemoto et al., 2014) related to motivational factors associated with reward processing and task engagement (for reviews see Holroyd and Umemoto, 2016; Proudfit, 2015). These considerations dovetail with the proposal that anterior cingulate cortex, where the reward positivity is believed to be generated (Becker et al., 2014; Holroyd and Umemoto, 2016; Walsh and Anderson, 2012), is concerned with motivational control over extended behaviors (Holroyd and Yeung, 2012; Holroyd and McClure, 2015).

Although a great deal has been learned over the past decades about the response of the reward positivity to different types of feedback, surprisingly little is known about how it responds to the degree of information conveyed by the feedback. At the most granular level, feedback can either convey information or not; neutral stimuli that indicate neither whether a response was correct or incorrect elicit an N2 that is as large as (Holroyd et al., 2006) or larger than (Li et al., 2016) the negativity to error feedback, variation that appears to be due to individual differences in neuroticism (Hirsh and Inzlicht, 2008) and anxiety (Gu et al., 2010). At a finer scale, feedback information can convey not only whether a response was correct or incorrect, but also the degree of accuracy. For example, a study that compared near misses with full misses in a gambling task found that reward positivity amplitude was larger to the former than to the latter (Ulrich and Hewig, 2014), though this study measured the component base-to-peak rather than as a difference wave, complicating the interpretation. Reward positivity amplitude also scaled with feedback accuracy in a dart-throwing paradigm (Frömer et al., 2016), and with the degree of deviation from a target in the brain-computer interface experiment described above (Zander et al., 2016).

Notably, in an adaptation of the classic time-estimation task first used to demonstrate the existence of the reward positivity (Miltner et al., 1997), Mars et al. (2004) found that the reward positivity amplitude was larger to binary performance feedback compared to feedback indicating that subjects were either too fast or too slow, and to feedback indicating the degree of error. They also found that reward positivity amplitude was unrelated to post-feedback behavioral adjustments. On the other hand, another time estimation task study found

diametrically opposite results: reward positivity amplitude increased with increasing information content of the outcomes, and was related to post-feedback adjustments (Grundler et al., 2010). And still another time estimation task study that used graded feedback – that indicated by exactly how much the errors were either too fast or too slow – found larger reward positivity amplitudes for larger errors, and that the size of the component was unrelated to learning outcomes (Luft et al., 2014). These observations indicate that the interrelationship between reward positivity amplitude, the amount of information provided by feedback, and behavioral adaptation remains opaque.

To investigate this issue, we examined the reward positivity in a modified time estimation task in which on each trial subjects received, depending on condition, 1) *binary feedback* indicating that subjects were on time or not, 2) *directional feedback* indicating that they responded either too slowly, too quickly, or were on time, 3) *magnitude feedback* indicating that they were either on time, or by how much they were not on time, and 4) *full feedback* that provided information about both the direction and magnitude of the errors. We predicted that subjects' remedial behaviors would mirror the quality of the feedback, that reward positivity amplitude would be inversely related to the degree of information content conveyed by the feedback, and that reward positivity amplitude would be unrelated to behavior. Finally, in a series of exploratory analyses, we also examined the effects of the feedback information on later ERP components elicited by the feedback.

2. Methods

2.1. Participants

25 right-handed undergraduate students from the University of Victoria participated in the experiment. One participant was excluded due to language barriers and a failure to understand the instructions. All of the participants were volunteers who received extra credit in a first- or second-year 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.

2.2. Apparatus and procedure

Participants were seated in front of a computer monitor in an electromagnetically shielded booth to perform the time estimation task (written in Matlab 7.1 using the psychophysics toolbox extension, Brainard and Vision, 1997). The task was similar to that employed by Miltner et al. (1997) in which participants were asked to estimate a duration of 1 s (Fig. 1). Each trial began with a centrally positioned fixation cross (duration 500 ms) followed by an auditory cue (1500 Hz, 65 dB, duration 50 ms). Participants were asked to respond by pressing the spacebar once they believed 1 s had elapsed since the time of the tone. Feedback was displayed (duration 1000 ms) 500 ms following their response. An estimate was considered to be 'correct' if it was within a dynamic temporal window centered at 1000 ms after the tone, and was considered 'incorrect' otherwise.

The temporal window of accuracy was initialized at 1000 ms \pm 200 ms. Thus, each participant was required to respond between 800 ms and 1200 ms following the auditory cue to receive correct feedback on the first trial of each condition. The accuracy window was adjusted after each trial using a staircase procedure to ensure that participants received approximately equal proportions of correct and incorrect feedback in all conditions. The width of the window was increased or decreased by 10 ms following correct or incorrect responses respectively.

A centrally presented yellow-filled square served as correct feedback in all conditions. Incorrect feedback stimuli varied according to experimental condition (Fig. 1). In the *binary condition*, participants

Download English Version:

<https://daneshyari.com/en/article/10153485>

Download Persian Version:

<https://daneshyari.com/article/10153485>

[Daneshyari.com](https://daneshyari.com)