International Journal of Psychophysiology xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

International Journal of Psychophysiology

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

The application of reward learning in the real world: Changes in the reward positivity amplitude reflect learning in a medical education context☆

Chad C. Williams[a,*], Kent G. Hecker[b,c], Michael K. Paget[c], Sylvain P. Coderre[c], Kelly W. Burak[c], Bruce Wright[d], Olave E. Krigolson[a]

[a] *Centre for Biomedical Research, University of Victoria, Canada*
[b] *Faculty of Veterinary Medicine, University of Calgary, Canada*
[c] *Cumming School of Medicine, University of Calgary, Canada*
[d] *Division of Medical Science, University of Victoria, Canada*

ABSTRACT

Evidence ranging from behavioural adaptations to neurocognitive theories has made significant advances into our understanding of feedback-based learning. For instance, over the past twenty years research using electroencephalography has demonstrated that the amplitude of a component of the human event-related brain potential – the reward positivity – appears to change with learning in a manner predicted by reinforcement learning theory (Holroyd and Coles, 2002; Sutton and Barto, 1998). However, while the reward positivity (also known as the feedback related negativity) is well studied, whether the component reflects an underlying learning process or whether it is simply sensitive to feedback evaluation is still unclear. Here, we sought to provide support that the reward positivity is reflective of an underlying learning process and further we hoped to demonstrate this in a real-world medical education context. In the present study, students with no medical training viewed a series of patient cards that contained ten physiological readings relevant for diagnosing liver and biliary disease types, selected the most appropriate diagnostic classification, and received feedback as to whether their decisions were correct or incorrect. Our behavioural results revealed that our participants were able to learn to diagnose liver and biliary disease types. Importantly, we found that the amplitude of the reward positivity diminished in a concomitant manner with the aforementioned behavioural improvements. In sum, our data support theoretical predictions (e.g., Holroyd and Coles, 2002), suggest that the reward positivity is an index of a neural learning system, and further validate that this same system is involved in learning across a wide range of contexts.

1. Introduction

Converging evidence has made significant advances into understanding how humans learn from feedback. Whereas pioneer research has described how behaviours change in response to rewards and punishments (Skinner, 1958), more recent studies have theorized the neural mechanisms that underlie reward learning systems within the brain (Holroyd and McClure, 2015). In particular, neuroimaging studies have discovered that there are at least two neurocognitive mechanisms to learning from feedback. First, it has become evident that there is an early, unconscious system that is sensitive to violations of expectancy (Holroyd and Coles, 2002; Holroyd and Krigolson, 2007; Krigolson et al., 2014; Sutton and Barto, 1998). Second, there also appears to be a later conscious system responsible for updating mental representations of the environment in order to adapt behaviours and predictions (Sato et al., 2005; Yeung and Sanfey, 2004). The former of these processes has been theorized to be driven by the midbrain dopamine system which delivers signals that reflect reward prediction errors – the degree to which the predictions of outcomes do not match the actual outcomes – to the anterior cingulate cortex (ACC; Holroyd and McClure, 2015; Schultz et al., 1997). More precisely, within this specific framework (i.e., Holroyd and Coles, 2002), prediction errors are computed within the basal ganglia, and projected to the ACC via the midbrain dopamine system. Computational theories describe the ACC to be a 'controller' of cognitive resources in that it integrates these dopamine signals and directs how to best use resources across the brain in order to learn from the environment (Holroyd and McClure, 2015).

Over the past twenty years there has been a large body of work

---

☆ Sources of support: Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant (RGPIN 2016-0943), and the Neuroeducation Network (007).

* Corresponding author at: School of Exercise Science, Physical & Health Education, University of Victoria, P.O. Box 17000, STN CSC, Victoria, British Columbia V8W 2Y2, Canada.
E-mail address: ccwillia@uvic.ca (C.C. Williams).

http://dx.doi.org/10.1016/j.ijpsycho.2017.10.010
Received 31 July 2017; Received in revised form 22 September 2017; Accepted 26 October 2017

examining the electroencephalographic (EEG) responses of these reward signals. In 1997, Miltner, Braun, and Coles first reported the feedback related negativity (FRN), a component of the human event-related brain potential (ERP) evoked by the processing of outcome feedback that is now theorized by some to reflect the arrival of dopamine signals at the ACC (Holroyd and Coles, 2002; Holroyd and McClure, 2015; Holroyd and Yeung, 2012; Schultz et al., 1997). More recently, it has been suggested that the FRN should be framed as reward positivity reflecting the sensitivity of this component to positive as opposed to negative outcomes (Foti et al., 2011; Holroyd et al., 2008; Proudfit, 2015). The reward positivity component arises in frontal-central regions of the scalp 250 to 350 ms following performance feedback (Proudfit, 2015). Specifically, it is theorized to be the ERP analog of reward prediction error dopamine signals arriving at the ACC (Holroyd and Coles, 2002; Holroyd and McClure, 2015; Holroyd and Yeung, 2012).

If the Holroyd and Coles hypothesis is true, it seems logical that the amplitude of the reward positivity should reflect underlying learning processes – yet, to date, findings are mixed. For instance, Krigolson et al. (2014) demonstrated that the amplitude of the reward positivity diminished with learning, a result also reported by the same group in 2009 (Krigolson, Pierce, Tanaka, & Holroyd) and by others (Bellebaum and Colosio, 2014; Bellebaum and Daum, 2008; Eppinger et al., 2008; Luque et al., 2012; Sailer et al., 2010). Other studies, however, have found that behavioural and neural changes linked to learning did not always coincide (Bellebaum et al., 2010; Eppinger et al., 2009; Groen et al., 2007; Hämmerer et al., 2010; Holroyd and Coles, 2002; Nieuwenhuis et al., 2002; see Walsh and Anderson, 2012 for a review). As such, it is unclear if the reward positivity reflects an underlying learning process or whether it is simply sensitive to feedback evaluation.

One potential explanation for the conflicting findings may relate to whether the information in experimental paradigms is relevant and/or learnable. For example, in some of the gambling paradigms typically used to study the reward positivity no learning can actually occur. This was explored by Bellebaum and Colosio (2014) who had participants make decisions about alphabetic characters in which feedback for some stimuli was contingent on participant responses (learning could occur), while for other stimuli it was not (learning could not occur). They found that the reward positivity amplitude decreased across the task only for the contingent stimuli. As such, it appears to be important that we study the reward positivity in tasks where learning can occur. Related to that, is the relationship between information and outcomes. Specifically, in the aforementioned studies participants had to learn about shapes (Bellebaum and Daum, 2008; Bellebaum et al., 2010; Krigolson et al., 2009; Krigolson et al., 2014; Sailer et al., 2010), simple objects (Eppinger et al., 2008; Eppinger et al., 2009; Groen et al., 2007; Holroyd and Coles, 2002; Luque et al., 2012), and alphabetic characters (Bellebaum and Colosio, 2014; Hämmerer et al., 2010; Nieuwenhuis et al., 2002). However, in none of these experiments did the stimuli naturally lead to a correct answer. In other words, the stimulus–response mappings were in a sense both arbitrary and meaningless. Put another way, what was learned by participants in these studies could never be used in, nor ever arise from, any natural environments.

In contrast, behavioural research has explored learning in real-world contexts. For example, two recent studies have demonstrated the efficacy of reinforcement learning in medical education (Anderson et al., 2016; Xu et al., 2016). Anderson et al. (2016) used a reinforcement learning paradigm to enhance the teaching of neuroanatomy to medical students. Specifically, they had participants learn to identify neuroanatomical structures via a computer based trial and error shaping process – participants saw an image with a label, determined whether the structure and label were correctly matched, and were provided with feedback about the accuracy of their response. Importantly, participants learned to identify multiple neuroanatomical structures as was indicated by increasing accuracy rates and decreasing response times (Anderson et al., 2016). Further evidence supporting this in a medical education context comes from Xu et al. (2016) who used a similar approach to teach students to correctly categorize melanoma. These paradigms are progressing in the correct direction, yet still rely on simple stimuli (e.g., an image). We propose that to truly understand how learning occurs organically it is important to extend these findings to learning more complex real-world material while at the same time investigating the neural processes involved.

Here, we seek to demonstrate that changes in the reward positivity are related to an actual learning process and moreover that the system underlying this component plays a role when learning complex real-world material. In the current study, participants were to learn to diagnose liver and biliary diseases by making judgments on patient case-studies and utilizing simple performance feedback while electroencephalographic data were recorded. We hypothesized that participants would be able to learn complex data structures in order to categorize clinical cases through the use of a reinforcement learning paradigm. Specifically, we predicted that accuracy rates would be higher and reaction times (i.e., viewing the patient card and viewing the diagnostic options) would be quicker late in each phase, when learning has occurred, as opposed to early in each phase. Further, we predicted that participants would score higher than chance on a retention test. In regards to neural data, we hypothesized that performance feedback would elicit a reward positivity – indicating the processing of said feedback. Importantly, we also predicted that the amplitude of the reward positivity would diminish with learning – a result in line with previous work and theoretical predictions (i.e., Sutton and Barto, 1998).

2. Methods

2.1. Participants

Thirty undergraduate students with no medical training (23 female, mean age 20 years old [CI: ± 1 year]) from the University of Victoria participated in the experiment. All participants had normal or corrected-to-normal vision, no neurological impairments, and volunteered for extra course credit in a psychology course. Four participants were removed as they did not progress past the first phase (see below) resulting in twenty-six participants (19 female, mean age 20 years old [CI: ± 1 year]). All participants provided informed consent approved by the Human Research Ethics Board at the University of Victoria, and the study followed ethical standards as prescribed in the 1964 Declaration of Helsinki.

2.2. Apparatus and procedure

Participants were seated in a sound dampened room in front of a 19″ LCD computer monitor and used a handheld 5-button RESPONSEPixx (VPixx, Vision Science Solutions, Quebec, Canada) controller to complete an adaptation of the Cards reinforcement learning paradigm (Bannister et al., 2016; Burak et al., 2015; Horrey et al., 2016; Kazoleas, 2016; Tang et al., 2016) written in MATLAB (Version 8.6, Mathworks, Natick, U.S.A.) using the Psychophysics Toolbox extension (Brainard, 1997).

Cards teaches participants through the application of reinforcement learning principles. In our experiment, participants were presented with physiological data (e.g., liver enzyme values) which they then used to make clinical decisions. Specifically, participants learned to classify five types of liver and biliary diseases: cholestatic intrahepatic, cholestatic extrahepatic, mild hepatocellular, moderate hepatocellular, and severe hepatocellular. This classification mimics the first step of cognitive organization structures called "schemes", a process particularly ascribed to expertise (Coderre et al., 2003). During each clinical case (i.e., trial) of the experiment, participants were shown a patient case-study card followed by a multiple-choice presentation of the diagnostic