NeuroImage 42 (2008) 807–816

Contents lists available at ScienceDirect

NeuroImage

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

Individual differences in reinforcement learning: Behavioral, electrophysiological, and neuroimaging correlates

Diane L. Santesso, Daniel G. Dillon, Jeffrey L. Birk, Avram J. Holmes, Elena Goetz, Ryan Bogdan, Diego A. Pizzagalli *

Department of Psychology, Harvard University, 1220 William James Hall, 33 Kirkland Street, Cambridge, MA 02138, USA

ARTICLE INFO

Article history:
Received 10 April 2008
Revised 17 May 2008
Accepted 20 May 2008
Available online 2 July 2008

Keywords:
Reinforcement learning
Anterior cingulate cortex
Basal ganglia
Reward
Feedback-related negativity
Probabilistic learning

ABSTRACT

During reinforcement learning, phasic modulations of activity in midbrain dopamine neurons are conveyed to the dorsal anterior cingulate cortex (dACC) and basal ganglia (BG) and serve to guide adaptive responding. While the animal literature supports a role for the dACC in integrating reward history over time, most human electrophysiological studies of dACC function have focused on responses to single positive and negative outcomes. The present electrophysiological study investigated the role of the dACC in probabilistic reward learning in healthy subjects using a task that required integration of reinforcement history over time. We recorded the feedback-related negativity (FRN) to reward feedback in subjects who developed a response bias toward a more frequently rewarded ("rich") stimulus ("learners") versus subjects who did not ("non-learners"). Compared to non-learners, learners showed more positive (i.e., smaller) FRNs and greater dACC activation upon receiving reward for correct identification of the rich stimulus. In addition, dACC activation and a bias to select the rich stimulus were positively correlated. The same participants also completed a monetary incentive delay (MID) task administered during functional magnetic resonance imaging. Compared to non-learners, learners displayed stronger BG responses to reward in the MID task. These findings raise the possibility that learners in the probabilistic reinforcement task were characterized by stronger dACC and BG responses to rewarding outcomes. Furthermore, these results highlight the importance of the dACC to probabilistic reward learning in humans.

Introduction

Optimal behavior relies on the ability to internally monitor responses and to evaluate external reinforcements in order to learn about the appropriateness of those responses. Mounting evidence suggests that this reinforcement learning may depend on the basal ganglia and midbrain dopamine system. Accordingly, non-human primate studies have shown that negative reinforcement elicits phasic decreases in neuronal activity of midbrain dopaminergic neurons (i.e., negative prediction error), whereas positive reinforcement elicits increases of dopaminergic activity (i.e., positive prediction error) (Montague et al., 2004; Schultz, 2007). These phasic modulations are thought to act as teaching signals for the anterior cingulate cortex (ACC) and basal ganglia (BG) to implement goal-directed behaviors and update predictions of success or failure (Holroyd and Coles, 2002). This model has received support in the human electrophysiology literature with respect to negative reinforcement (Holroyd and Coles, 2002; Holroyd and Krigolson, 2007; Hajcak et al., 2007), but fewer studies have examined positive reinforcement. In particular, the role of the human dorsal region of the ACC (dACC) in probabilistic reward learning is not well understood.

The dACC appears critical for encoding rewards and using reinforcement histories to guide behavior (Akitsuki et al., 2003; Amiez et al., 2006; Ernst et al., 2004; Rushworth et al., 2007). In non-human primates, ACC lesions impair the ability to integrate reinforcement history over time and choose advantageous responses (Kennerley et al., 2006). In humans, modulation of behavior by reinforcement history can be investigated using two-alternative probabilistic reward tasks in which correct responses to the two stimuli are differentially rewarded; the development of a response bias towards the more frequently rewarded ("rich") stimulus indicates reward sensitivity (Pizzagalli et al., 2005, 2008). Impaired learning on this task has been demonstrated in anhedonic individuals (Pizzagalli et al., 2005), mood disorder patients characterized by dysfunctional reward processing (Pizzagalli et al., in press-a, in press-b), and in healthy participants receiving a pharmacological challenge hypothesized to disrupt phasic DA signaling

* Corresponding author. Fax: +1 617 495 3728.
E-mail address: dap@wjh.harvard.edu (D.A. Pizzagalli).

1053-8119/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2008.05.032

(Pizzagalli et al., 2008). This task thus appears suitable for examining reward learning mediated by the midbrain dopamine system. Consistent with this assumption, in a computational model of striatal–cortical function (Frank, 2005), blunted response bias was accounted for by reduced DA bursts to reward (Santesso et al., unpublished), suggesting that this task is sensitive to learning mediated by the midbrain DA system. The primary goal of the present study was to examine reward learning during this task using the feedback-related negativity (FRN) as an electrophysiological index of ACC reward-related activity.

We recorded the feedback-related negativity (FRN) as an index of ACC reward-related activity. The FRN peaks 200–400 ms following feedback and has been localized to various regions of the cingulate cortex, including the dorsal ACC (dACC; Miltner et al., 1997; Gehring and Willoughby, 2002), medial prefrontal cortex (Muller et al., 2005; Nieuwenhuis et al., 2005; Van Veen et al., 2004), and the posterior cingulate cortex (PCC), particularly in response to positive versus negative feedback (Muller et al., 2005; Nieuwenhuis et al., 2005). The FRN is thought to reflect transmission of a DA signal from the BG (Holroyd and Coles, 2002). Although commonly used to study negative reinforcement, the FRN is reliably elicited by positive feedback (Hajcak et al., 2005; Holroyd and Coles, 2008; Muller et al., 2005; Oliveira et al., 2007), and appears as a relatively more positive ERP deflection (compared to that elicited by negative feedback). We predicted that (1) reward feedback delivered after correctly identifying the rich stimulus would elicit more positive FRNs and greater dACC activation in individuals who developed a response bias toward the rich stimulus ("learners") versus those who did not ("non-learners"); and (2) dACC activation would correlate positively with reward learning and the FRN.

A secondary goal of this study was to test whether "learners" and "non-learners" would differ in brain activation in the BG, which includes the globus pallidus and three striatal regions (nucleus accumbens, caudate, and putamen), in response to reward feedback. We were able to address this issue because a sub-set of the ERP participants also participated in an fMRI session that featured a monetary incentive delay (MID) task, which has been used to probe reward-related activity in the BG (Dillon et al., 2008; Knutson et al., 2003). Relevant to the present study, recent neuroimaging findings indicate that optimal performance in probabilistic reward-learning tasks is accompanied by recruitment of striatal regions. Accordingly, in a probabilistic reward-learning task, learners (but not non-learners) showed significant correlations between prediction errors and fMRI signal in dorsal and ventral striatal regions (Schonberg et al., 2007). Along similar lines, participants who learned contingencies between specific cues and the reward probabilities and used them adaptively in a gambling task showed robust striatal responses to reward feedback, particularly at early stages of learning (Delgado et al., 2005). Based on these findings, we predicted that, relative to non-learners, learners in the probabilistic reward task would show larger BG responses to reward feedback during the MID task.

Materials and methods

Participants

Two hundred and thirty-seven adults between 18 and 40 years old (105 men, mean age = 24.5 years) were recruited from Harvard University and the surrounding community for a larger study investigating the neurobiology and molecular genetics of reward processing. Participants meeting the following criteria were excluded: present medical or neurological illness (ADHD, head injury, loss of consciousness, seizures), current alcohol/substance abuse or smoking, claustrophobia, use of psychotropic medications during the last 2 weeks, and pregnancy. All eligible participants were right-handed (Chapman and Chapman, 1987).

The study included three sessions. During the first session, all participants completed the probabilistic reward task at the Affective Neuroscience Laboratory, Harvard University. Sixty-seven subjects were excluded due to failure to meet inclusion criteria (n = 31), prior task exposure (n = 4), non-compliance and/or performance below chance level (n = 31), and outlier status (n = 1). Of the remaining 170 eligible subjects, 47 were invited to complete an electroencephalogram (EEG) and fMRI session (the order of which was counterbalanced). These 47 subjects were selected to cover a wide range of individual differences in reward learning, which was measured by a response bias difference score (block 3 – block 1; see below). To this end, we first selected participants in the upper and lower 20% of the distribution of reward learning; next, remaining subjects were selected in order to achieve a continuum in reward learning, so that selected participants would be representative of the general population. Of the 47 participants, 41 agreed to perform the probabilistic reward task while EEG was recorded, whereas 38 completed the monetary incentive delay (MID) task during functional scan acquisition at the Martinos Center for Biomedical Imaging. For both the EEG and fMRI datasets, 30 participants had usable data; data from remaining participants were lost due to an insufficient number of artifact-free EEG trials, equipment failure, incomplete data, non-compliance, motion artifacts (fMRI), and technical difficulties. Of the 30 participants with EEG data, 21 had usable data from all three sessions.

Participants received $5 for the first session plus $5.80–$6.20 in earnings in the probabilistic reward task. For the EEG session, participants received $20 plus $24.60 (fixed amount) in task earnings. For the fMRI session, participants received $60 plus $20–$22 in earnings for the MID task. Participants provided written informed consent. All procedures were approved by the Committee on the Use of Human Subjects at Harvard University and the Partners-Massachusetts General Hospital Internal Review Board.

Procedures and tasks

Probabilistic reward task (EEG session)

During the EEG session, participants repeated the reward-learning task used during subject selection, which has been described in detail elsewhere (e.g., Pizzagalli et al., 2005, 2008; see also Tripp and Alsop, 1999). Briefly, the task included 300 trials, divided into 3 blocks of 100 trials. Each trial started with the presentation of a fixation point for 1400 ms. A mouthless cartoon face was then presented for 500 ms followed by the presentation of this face with either a short mouth or a long mouth for 100 ms. Participants were asked to indicate whether a short or long mouth was presented by pressing one of two keys (counterbalanced across subjects). For each block, only 40 correct responses were followed by positive feedback ("Correct!! You won 20 cents"), displayed for 1500 ms in the center of the screen followed by a blank