Archival Report

Biological Psychiatry: CNNI

Intact Ventral Striatal Prediction Error Signaling in Medicated Schizophrenia Patients

Adam J. Culbreth, Andrew Westbrook, Ziye Xu, Deanna M. Barch, and James A. Waltz

ABSTRACT

BACKGROUND: Midbrain dopaminergic neurons code a computational quantity, reward prediction error (RPE), which has been causally related to learning. Recently, this insight has been leveraged to link phenomenological and biological levels of understanding in psychiatric disorders, such as schizophrenia. However, results have been mixed, possibly due to small sample sizes. Here we present results from two studies with relatively large sample sizes to assess ventral striatum (VS) RPE in schizophrenia.

METHODS: In the current study we analyzed data from two independent studies, involving a total of 87 chronic medicated schizophrenia patients and 61 control subjects. Subjects completed a probabilistic reinforcement-learning task in conjunction with functional magnetic resonance imaging scanning. We fit each participant's choice behavior to a Q-learning model and derived trialwise RPEs. We then modeled blood oxygen level–dependent (BOLD) signal data with parametric regressor functions using these values to determine whether patient and control groups differed in prediction error–related BOLD signal modulations.

RESULTS: Both groups demonstrated robust VS RPE BOLD activations. Interestingly, these BOLD activation patterns did not differ between groups in either study. This was true when we included all participants in the analysis, as well as when we excluded participants whose data was not sufficiently fit by the models.

CONCLUSIONS: These data demonstrate the utility of computational methods in isolating or testing underlying mechanisms of interest in psychiatric disorders. Importantly, similar VS RPE signal encoding across groups suggests that this mechanism does not drive task deficits in these patients. Deficits may instead stem from aberrant prefrontal or parietal circuits associated with maintenance and selection of goal-relevant information.

Keywords: Computational psychiatry, fMRI, Prediction error, Reward processing, Schizophrenia, Ventral striatum

http://dx.doi.org/10.1016/j.bpsc.2016.07.007

Schizophrenia (SZ) is associated with a diminished ability to use reward history to adaptively guide behavior. These deficits have been shown across a wide variety of tasks and have been associated with important aspects of the illness (1). Although previous studies have established a broad reward-learning deficit in SZ, most rely solely on standard tasks metrics (e.g., task accuracy) and neuroimaging approaches, which do not always clearly delineate underlying psychological and neural mechanisms. This limits etiological understanding of the specific neural circuits, neurotransmitters, and cognitive/emotional processes that give rise to these deficits. Methods in computational psychiatry hold the particular advantage of specifying such broadly defined deficits by isolating underlying mechanisms of interest.

Importantly, reward-learning impairment could reflect a number of underlying mechanisms. For example, it could emerge from abnormal representations of the expected value of actions, or from disrupted signaling of mismatches between expected and obtained outcomes, i.e., reward prediction errors (RPEs) (2,3). Reinforcement learning is a powerful framework for quantifying and linking such mechanisms to underlying biology (4). For example, a robust finding in the animal literature is that RPEs are coded by the phasic firing of midbrain dopaminergic neurons (5). More recently, human functional magnetic resonance imaging (fMRI) studies have demonstrated RPE encoding in the ventral striatum (VS) (a target region of midbrain dopaminergic neurons) extending findings observed in animals (6,7). These findings have been instrumental because they link adaptive learning to dopaminergic signaling through an intermediate computational mechanism (RPE signaling). In the current article, we aim to demonstrate how methods in the field of computational psychiatry, particularly reinforcement-learning algorithms, hold particular promise in clarifying the role of specific mechanisms potentially contributing to reward-learning impairments. Specifically, we use two relatively large samples to examine the integrity of neural indicators of RPE in SZ (8–10).

Dopamine dysregulation is associated with SZ, including increased striatal dopamine neurotransmission and synthesis capacity (11,12). Increased baseline dopaminergic activity in SZ has been proposed to introduce computational noise in the reinforcement-learning system, blunting RPE signaling and resulting in poor reward learning (13,14). This hypothesis is bolstered by evidence that pharmacological manipulations increasing dopamine tone in control (CN) participants yield blunted VS RPE signals (15). SZ has also been associated with

chaotic dopamine firing along with elevations in baseline dopaminergic activity (13). These abnormal firing patterns are thought to simulate inappropriate RPE signaling to otherwise nonsalient, neutral stimuli and may underlie the formation of delusions (13,16). The idea that abnormal VS RPE signaling might cause events to be perceived as unduly salient is an intriguing hypothesis, linking phenomenological and biological domains of understanding in SZ.

Several studies have examined RPE signaling in SZ (17). These studies have yielded mixed findings, with some reports demonstrating decreased VS RPE signaling for SZ patients compared with CN participants (18–20), and others not (21–25). There are several reasons why these reports may be mixed. First, there may be heterogeneity in the phase of illness studied: Some reports recruited first-episode unmedicated patients (19,20) and others chronic medicated patients (21–25), with some evidence suggesting that blunted VS RPE signaling may be more pronounced in unmedicated patients (20,26). This literature is also hampered by small sample sizes, with most studies recruiting fewer than 20 subjects per group (18,21–24,26). These small samples are problematic because positive findings with small samples represent estimates of effect sizes with high uncertainty (27). Another issue is heterogeneity in methods used to quantify VS RPE signaling. Some reports have examined VS RPE signaling by performing blood oxygen level–dependent (BOLD) contrasts between task conditions (18,22–25), for example, contrasting trials where reward was expected from trials where reward was unexpected. However, such approaches may lack sensitivity, as RPE magnitudes are not calculated on a trial-by-trial basis. In contrast, others have fit participant choice behavior to a reinforcement-learning algorithm to generate trialwise prediction error (PE) estimates (19–21). Finally, for those studies that implemented reinforcement-learning algorithms, few studies have performed tests to ensure that these models fit choice behavior significantly better than chance (19,21,22). This consideration is important, as parameter estimates from poor-fitting individuals are difficult to interpret and may be misinterpreted as aberrant RPE signaling. In summary, evidence is mixed for VS RPE signaling as a mechanism for reward-learning dysfunction in SZ, particularly for chronic medicated patients.

In the current study, we utilized computational approaches to examine VS RPE signaling in two independent samples of chronic medicated outpatients with SZ and CN participants, testing the assertion that aberrant VS RPE signals underlie reward-learning dysfunction. We used a probabilistic reversal learning (PRL) task that has been well validated in the basic and clinical science literatures (20,28–30). To examine trialwise PE we fit each participant's choice behavior to a Q-learning model, and entered trialwise PE values as regressors in our imaging analyses to index VS reactivity (20).

METHODS AND MATERIALS

Participants

Participants were recruited from two independent study sites: Washington University in St. Louis (WUSTL) (SZ patients = 58, CN participants = 40) and the Maryland Psychiatric Research Center at the University of Maryland School of Medicine (SZ patients = 35; CN participants = 23). Data from each of these samples using conventional fMRI analyses were presented in Culbreth *et al.* (28) and Waltz *et al.* (30), respectively. Each site received approval from their respective institutional review boards, and all subjects provided informed consent. In the Maryland sample, 6 SZ patients and 2 CN participants were excluded due to poor task performance; however, no participants were excluded due to excessive movement (see the Supplement). In the WUSTL sample, 1 SZ patient and 4 CN participants were excluded due to excessive movement during scanning (movement based on root mean square was greater than 0.2 across the run), yielding final sample sizes of 93 and 50 across sites.

Clinical Assessments

Diagnoses were determined using the Structured Clinical Interview for DSM-IV-TR (31). Negative symptoms were assessed using the Scale for the Assessment of Negative Symptoms (SANS) (32). Positive and disorganized symptoms were assessed using the Scale for the Assessment of Positive Symptoms (33) at WUSTL, and the Brief Psychiatric Rating Scale (34) at Maryland. All participants passed a drug screen. General intellectual functioning was assessed at both sites using the Wechsler Test of Adult Reading (35).

Probabilistic Reversal Learning Task

Similar PRL tasks were presented at the two sites, both in conjunction with fMRI scanning (see the Supplement). On each trial of the task, two abstract visual patterns are shown to participants, one commonly (80%) and one rarely (20%) rewarded. Participants are not told these precise percentages. Subjects are instructed to guess which pattern is most likely to yield reward. They are instructed that occasionally the reward contingencies reverse and the alternative stimulus is associated with a high probability of reward. The chosen response is highlighted upon response and participants are given feedback (correct or incorrect) on each trial. Each run consists of an initial acquisition where participants learn values for each choice. After the reward contingencies are learned—that is, participants met a performance threshold of selecting the correct response eight of 10 previous trials in the WUSTL and nine of 10 in the Maryland sample—contingencies reversed. Probabilistic negative feedback is implemented such that a correct response for each trial receives negative feedback 20% of the time. All subjects practiced the task prior to scanning. Participants won bonus money for increased task accuracy.

Behavioral Data Analysis

Independent-samples *t* tests were conducted to determine group differences in the number of errors committed and the number of reversals achieved. The initial acquisition phase of each run was also analyzed to determine the number of trials participants needed to learn the reward contingencies.

Computational Modeling of Behavior

We fit a standard Q-learning model to individual choice behavior. For each trial (t), this model estimates the value (Q) of each action (i). The action value of the chosen action is