Archival Report

Biological Psychiatry: CNNI

Probabilistic Reinforcement Learning in Patients With Schizophrenia: Relationships to Anhedonia and Avolition

Erin C. Dowd, Michael J. Frank, Anne Collins, James M. Gold, and Deanna M. Barch

ABSTRACT

BACKGROUND: Anhedonia (a reduced experience of pleasure) and avolition (a reduction in goal-directed activity) are common features of patients with schizophrenia that have substantial effects on functional outcome but are poorly understood and treated. We examined whether alterations in reinforcement learning may contribute to these symptoms in patients with schizophrenia by impairing the translation of reward information into goal-directed action.

METHODS: Thirty-eight stable outpatients with schizophrenia or schizoaffective disorder and 37 healthy control subjects underwent functional magnetic resonance imaging scans during a probabilistic stimulus selection reinforcement learning task with dissociated choice- and feedback-related activation, followed by a behavioral transfer task allowing separate assessment of learning from positive versus negative outcomes. A *Q*-learning algorithm was used to examine functional activation relating to prediction error at the time of feedback and to expected value at the time of choice.

RESULTS: Behavioral results suggested a reduction in learning from positive feedback in patients; however, this reduction was unrelated to anhedonia/avolition severity. On analysis of the functional magnetic resonance imaging scans, prediction error–related activation at the time of feedback was highly similar between patients and control subjects. During early learning, patients activated regions in the cognitive control network to a lesser extent than control subjects. Correlation analyses revealed reduced responses to positive feedback in dorsolateral prefrontal cortex and caudate among those patients higher in anhedonia/avolition.

CONCLUSIONS: These results suggest that anhedonia/avolition are as strongly related to cortical learning or higher-level processes involved in goal-directed behavior, such as effort computation and planning, as to striatally mediated learning mechanisms.

Keywords: Anhedonia, Motivation, Prediction error, Reinforcement learning, Schizophrenia, Striatum

http://dx.doi.org/10.1016/j.bpsc.2016.05.005

Negative symptoms are major contributors to disability and poor quality of life among individuals with schizophrenia but are poorly understood and treated (1,2). Anhedonia (a reduced ability to experience pleasure) and avolition (a reduced motivation to initiate or persist in goal-directed activity) together comprise a dissociable factor of negative symptomatology (3) that has garnered increasing attention for a possible association with abnormalities in reward processing. In previous work, we described several processes required for the translation of reward information into goal-directed behavior; any disruption of these processes could lead to anhedonia or avolition (4). The work described here examines one of these processes, reinforcement learning (RL), and its relationship to anhedonia and avolition in patients with schizophrenia.

Numerous behavioral studies have suggested that RL is intact in patients with schizophrenia when learning is fairly implicit [although Siegert *et al.* (5) found evidence of impaired serial reaction time task learning] but more impaired when learning tasks require explicit representations of stimulus–reward contingencies (4,6). This pattern has given rise to the theory that the striatally mediated gradual RL system may be intact in patients with schizophrenia while more rapid, on-line, cortically mediated learning systems are impaired (6,7). Support for this theory is drawn from probabilistic reversal learning studies that show intact acquisition of probabilistic reward contingencies (which are thought to be striatally mediated) coupled with impaired reversal learning (which is thought to be cortically mediated) (8,9). Similarly, several studies using the weather prediction task have shown a relatively intact learning rate but impaired asymptotic performance, which provides mixed evidence for striatal learning impairments (7,10–12). However, a study with a larger sample size found lower learning rates in patients with schizophrenia than control subjects, suggesting possible impairments in striatally mediated learning (13). The behavioral literature therefore provides a mixed picture on whether striatally mediated learning is intact in patients with schizophrenia.

Another approach to studying RL is to ask whether the pattern of functional activation in regions receiving dopaminergic projections is consistent with a prediction error (PE)

signal. PEs are thought to be coded by dopaminergic projections to the basal ganglia, which signal the difference between predicted and received rewards and drive learning by iteratively updating reward predictions (14). In the schizophrenia literature, this approach has revealed some evidence for altered striatal PE activity among patients with schizophrenia using both Pavlovian and instrumental reward-learning tasks and for both monetary and liquid rewards (15–18), with some suggestion that positive PEs may be more affected than negative PEs (19,20) and some suggestion that the effects may be more apparent in nonmedicated (21) compared with medicated patients.

The findings reviewed above suggest the hypothesis that impairments in learning from positive outcomes related to reductions in striatal signaling of positive PEs or impaired cortical learning systems may contribute to motivational deficits in patients with schizophrenia. We tested these hypotheses by examining brain activity during a probabilistic RL paradigm, allowing examination of activation during both choice execution and feedback and the separate assessment of learning from positive versus negative outcomes. We used a model of the role of dopamine in RL proposed by Frank *et al.* (22–24), which emphasizes the separate contributions of D_1 and D_2 receptors in the striatum to "Go" and "NoGo" learning, respectively. Two previous studies have used this framework to examine Go learning (i.e., learning from rewarding outcomes) and NoGo learning (i.e., learning from nonrewarding outcomes) in medicated patients with schizophrenia, and found evidence of impaired Go learning but intact NoGo learning (25,26)—although one other study found impairments in both Go and NoGo learning (27). These findings are consistent with the hypothesis that the effectiveness of phasic dopamine signals in response to positive feedback is reduced in patients with schizophrenia, thereby impairing Go learning. These studies also examined the relationship between negative symptoms and RL impairments and showed correlations between negative symptom severity and measures of rapid explicit learning, suggesting a role for deficits in cortical learning systems in negative symptomatology. In addition, in a modeling study by Gold *et al.* (28), the behavior of patients with high negative symptoms was best captured by a computational model of striatal learning only, while a model with both striatal and cortical components best captured the behavior of patients lower in negative symptoms.

METHODS AND MATERIALS

Participants

Study participants included 49 stable outpatients with schizophrenia or schizoaffective disorder as defined by the DSM-IV and 41 healthy control subjects with no personal or family history of psychosis. Both medicated and nonmedicated patients were recruited from the community, and medication status and dose was required to have been stable for at least 2 weeks. Participants were group matched on sex, age, race, parental education, handedness (29), and smoking status. Inclusion criteria included 1) patients 18 to 50 years of age and 2) the ability to give informed consent. Exclusions can be found in the Supplement and include patients who had been diagnosed with major depressive disorder or dysthymia in the past year as defined by the DSM-IV. Ten individuals with schizophrenia and four healthy control subjects were excluded for excessive movement (described below), and an additional patient was excluded for having >50% nonresponse trials, yielding a final sample size of 37 control subjects and 38 patients (29 patients with schizophrenia and nine patients with schizoaffective disorder). All procedures were approved by the Washington University Human Research Protection Office.

Diagnosis and Clinical Assessment

Participant diagnoses were based on a Structured Clinical Interview for DSM-IV-TR (30) conducted by a master's-level clinician. See Supplement for details on clinical assessments and measures, which generated both clinician-rated and self-reported measures of anhedonia/amotivation.

Task

The experimental paradigm was a modified version of the probabilistic stimulus selection task (Figure 1) (22), consisting of an acquisition phase, during which functional magnetic resonance imaging (fMRI) scanning took place, and a test phase that was completed outside of the scanner. During acquisition, participants were presented on each trial with one of three pairs of stimuli (i.e., "AB," "CD," or "EF") in pseudorandomized order and were instructed to choose the stimulus that they believed was "correct" based on feedback received over time. Stimuli were displayed for 2000 ms, during which the participant was required to choose one of the stimuli via button press. After a jittered interstimulus interval ranging from 2000 to 6000 ms, feedback consisting of the words "Correct! +$" in green text, "Incorrect $0" in red text, or "Too Slow!" were presented on screen for 2000 ms. Subjects were told that they would win money for each correct choice, up to $20 (in actuality, all subjects were paid an additional $20 upon completion). For stimulus pair AB, the choice of A was rewarded 80% of the time, while B was rewarded 20% of the time; for pair CD, C was rewarded 70% of the time and D was rewarded 30% of the time; and for pair EF, E was rewarded 60% of the time and F was rewarded 40% of the time. Feedback was followed by an intertrial interval jittered

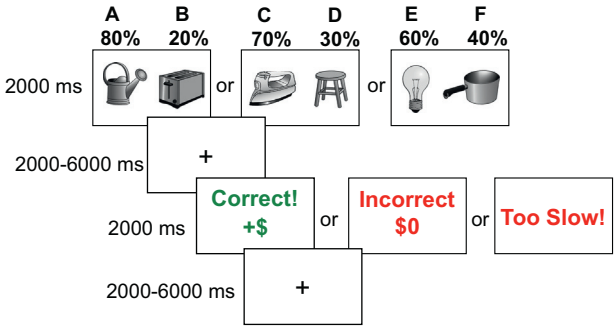

Figure 1. Experimental paradigm. The trial types and timing of the acquisition phase of the probabilistic stimulus selection task are shown. Both interstimulus and intertrial intervals were jittered to allow reconstruction of the blood oxygen level–dependent response at the time of both choice and feedback.