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

updated iteratively by a PE, which represents the difference between the expected value of that outcome ($Q_i(t)$) and the observed outcome (r), which was coded 1 or 0 for positive and negative feedback, respectively. The learning rate parameter (α) that fit best to each participant's choice behavior reflects the degree to which previous reinforcement outcomes affect subsequent Q values.

$$Q_i(t+1) = Q_i(t) + \alpha(r(t) - Q_i(t))$$

The probability of selecting one stimulus over another was then computed using the softmax function:

$$P_A(t) = \frac{e^{\frac{Q_{A(t)}}{\beta}}}{e^{\frac{Q_{A(t)}}{\beta}} + e^{\frac{Q_{B(t)}}{\beta}}}$$

Where β, the softmax temperature, reflects the stochasticity of the softmax function. Thus, higher β values indicate exploration (i.e., choosing the lower Q-value option) and low β values indicate exploitation (i.e., choosing the higher value option).

Finally, to determine participants that adequately fit the computational model we conducted analyses to determine if the predictive probability of each individual's choice behavior under the Q-learning model was greater than chance, similar to Schlagenhauf *et al.* (20).

fMRI Data Analysis

Images were acquired for both samples using a 3T MR scanner (see the Supplement for preprocessing). To examine the neural correlates of trialwise PEs (derived from the Q-learning model with subject specific learning rates), we implemented a general linear model with a parametric design. We modeled each trial without separating cue and outcome. In this design, the BOLD activation during each trial was modulated by subject-specific trialwise PE magnitudes and convolved with a canonical hemodynamic response function to provide an amplitude-modulated regressor in the general linear model. Between-group whole-brain comparisons were then conducted to determine differences in PE-related BOLD activation patterns using independent-samples *t* tests. Whole-brain analyses were corrected for multiple comparisons (WUSTL: $p = .006$, cluster size = 945 mm^3; Maryland: voxelwise threshold $p = .005$, cluster size = 675 mm^3), as determined by Monte Carlo simulations to achieve a significance level of $p = .05$, correcting for multiple comparisons over the whole brain. Regions demonstrating significant effects in the fMRI analyses were then correlated with task and symptom variables. Finally, to test specific hypotheses about hypoactivation of VS RPE signaling in SZ we conducted a volume of interest (VOI) analysis using an 8-mm sphere (center: ±10, 8, –4).

Power Analysis

Each sample yielded approximately 80% power to detect a medium effect of diagnostic group on VS RPE signaling (Cohen's $d = 0.7$). Further, when combining both samples we had 98% power to detect such effect sizes, and 83% power to detect a Cohen's $d = 0.5$. Thus, we had adequate power to detect effect sizes previously reported in the literature (18,20,26); however, we did not have adequate power to detect small effect sizes.

RESULTS

Sample Characteristics

Table 1 shows participant demographic information and task behavior. As previously presented in both samples, the SZ group was slower learning the initial reward contingencies, displayed decreased accuracy, and achieved fewer reversals compared with CN participants (28,30).

Computational Modeling of Behavior

In the WUSTL sample, the computational model fit SZ choice behavior poorer than CN as measured by negative log likelihood (Table 2). However, the two groups demonstrated similar learning rates and exploration parameter values after removing outliers (±3 SDs). When assessing whether the predictive probability of each participant's data was significantly greater under the Q-learning model than under a chance model, we found that 10 CN and 35 SZ did not fit the Q-learning model significantly better than they did the chance model. Poor-fit participants demonstrated worse task performance and decreased intellectual functioning than did good-fit participants, but no differences in symptoms were observed (see the Supplement).

In the Maryland sample, the computational model was shown to have a poorer fit to patient choice behavior, relative to CN participants (Table 2). The exploration parameter differed between groups. However, both groups demonstrated similar learning rates. Eleven SZ participants did not fit the Q-learning model significantly greater than they did the chance model in the Maryland sample. Poor-fit participants demonstrated worse overall task performance, but no differences in symptoms or intellectual functioning were observed between good- and poor-fitting SZ participants (see the Supplement).

Neural Correlates of PE

In both samples, SZ and CN groups displayed robust VS RPE signaling in a whole-brain analysis (Figure 1, Tables 3 and 4). Further, no significant between-group differences in RPE-related BOLD activations were found in the VS or other regions (Tables 3 and 4). These results remained consistent when performing analyses using a VS VOI (for WUSTL, right VS: Cohen's $d = –0.43$; left VS: Cohen's $d = –0.03$; for Maryland, left VS: Cohen's $d = –0.2$; right VS: Cohen's $d = –0.08$; for combined, left VS: Cohen's $d = –0.3$; right VS: Cohen's $d = –0.03$) with positive values indicated greater activation for CN participants.

Supplementary Bayesian analyses revealed that the null hypothesis, no difference in VS RPE signaling between groups, was 2–5 times more likely compared with the alternative [right VS: Cohen's $d = 0.5$; left VS: Cohen's $d = 0.3$; taken from Radua *et al.* (17)] for each sample. Further, when combining our samples we found that our data was 4–5 times more likely under the null hypothesis, providing moderate support for the null (see Supplemental Table S1).

Table 1. Participant Demographic, Clinical, Self-Report Measures, and Task Behavior

	WUSTL Sample		Maryland Sample	
Variables	CN	SZ	CN	SZ
Demographics				
Age, years	36.6 ± 9.2	37.0 ± 8.6	39.6 ± 10.5	39.6 ± 10.0
Male	52.8	66.7	71.4	82.8
Non-Caucasian	66.7	61.4	33.3	37.9
Personal education, years	14.2 ± 2.1	13.0 ± 2.2	15.1 ± 2.1	13.4 ± 1.7
Parental education, years	12.8 ± 1.5	12.9 ± 3.2	14.2 ± 3.3	14.2 ± 3.4
Medication Status				
Atypical antipsychotics	NA	75.4	NA	100.0
Typical antipsychotics	NA	5.3	NA	0.0
Typical and atypical	NA	7.0	NA	0.0
Not medicated	NA	12.3	NA	0.0
Clinical Rating (Item Average)				
General psychiatric (BPRS)	NA	NA	NA	1.9 ± 0.3
Positive (BPRS)	NA	NA	NA	2.5 ± 1.3
Positive (SAPS)	NA	0.7 ± 0.7	NA	NA
Disorganization (SAPS)	NA	0.4 ± 0.4	NA	NA
Negative (SANS)	NA	1.2 ± 0.6	NA	1.5 ± 0.9
Intellectual Function				
WTAR scaled score	98.7 ± 16.1	95.1 ± 16.6	109.7 ± 11.7	101.3 ± 17.1
Task Behavior				
Reversals	13 ± 9.1	8.4 ± 7.8	7.3 ± 1.3	4.5 ± 2.9
Number of errors	133.1 ± 48.6	149.5 ± 44.5	87.3 ± 18.1	122.7 ± 37.0
Initial acquisition trials	23.5 ± 15.2	31.7 ± 15.0	14.0 ± 3.0	21.7 ± 8.7

Values are mean ± SD or %.

BPRS, Brief Psychiatric Rating Scale; CN, control; NA, not applicable; SAPS, Scale for the Assessment of Positive Symptoms; SZ, schizophrenia; WTAR, Wechsler Test of Adult Reading; WUSTL, Washington University in St. Louis.

To increase our confidence that the VS activity we observed in previous analysis was indexing RPE signaling (as opposed to a value signal), we conducted analyses in which both trialwise RPE magnitude and trialwise Q-value magnitude of the stimulus that the participant chose were simultaneously entered as predictors of BOLD activity (see the Supplement). We found robust VS activation as a function of RPE magnitude when controlling for Q-value magnitude, providing support for our interpretation that this striatal activation is at least, partially, indexing trialwise RPE magnitude. Results were similar when excluding poor-fit participants.

To assess the contribution of model fit to the previous results, we separated SZ and CN groups into subgroups of individuals whose choice behavior fit the model better than chance and individuals whose behavior did not (for WUSTL, SZ good fit: $n = 23$; SZ poor fit: $n = 35$; CN good fit: $n = 30$; CN poor fit: $n = 10$; for Maryland, SZ good fit: $n = 18$; SZ poor fit: $n = 11$; CN good fit: $n = 23$; CN poor fit: $n = 0$). At WUSTL, robust VS RPE signaling was found for good-fit CN and SZ as well as poor-fit SZ, but not poor-fit CN. At Maryland, robust VS RPE signaling was found for good-fit CN and SZ, but not for poor-fit SZ (see the Supplement). However, when examining group differences in VS RPE signaling no significant differences were found between poor-fitting and good-fitting subjects in either sample.

Correlates Between Clinical Variables and VS RPE Signaling

In the WUSTL sample, no significant correlations were found between RPE signaling and negative symptoms (SANS total, left VS: $r = -.132$, $p = .56$; right VS: $r = -.047$, $p = .83$),

Table 2. Parameter Estimates and Fit Indices of Computational Model

	WUSTL Sample			Maryland Sample		
	SZ	CN	*p* Value	SZ	CN	*p* Value
Negative Log Likelihood	263 ± 70	229 ± 82	.04	220 ± 31	187 ± 25	.001
Learning Rate (α)	0.58 ± 0.3	0.60 ± 0.4	.79	0.56 ± 0.2	0.63 ± 0.1	.17
Exploration (β)	1.13 ± 1.2	1.65 ± 1.5	.1	1.77 ± 0.6	1.14 ± 0.5	.001
Pseudo-R^2	0.17	0.30	NA	0.18	0.29	NA

Values are mean ± SD.

CN, control; SZ, schizophrenia; NA, not applicable; WUSTL, Washington University in St. Louis.

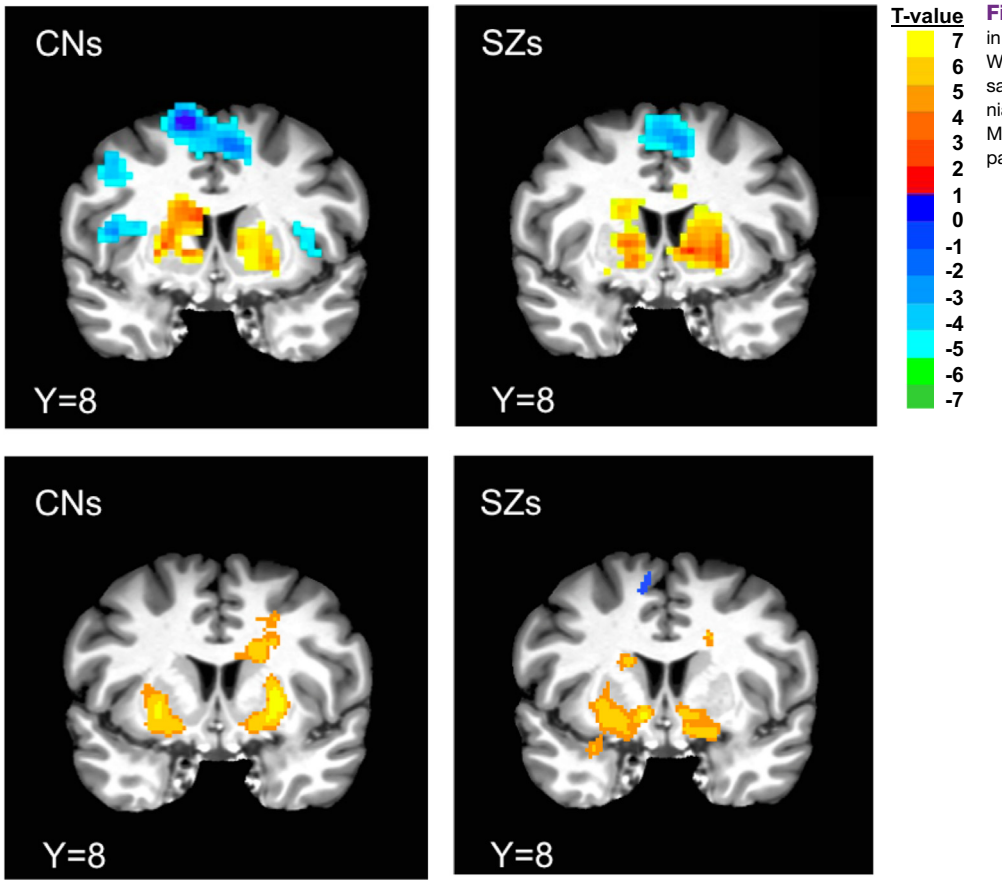

Figure 1. Prediction error signaling in the ventral striatum. Top row: Washington University in St. Louis sample control (CN) and schizophrenia (SZ) participants. Bottom row: Maryland sample CN and SZ participants.

positive symptoms (Scale for the Assessment of Positive Symptoms total, left VS: $r = .14$, $p = .54$; right VS: $r = -.12$, $p = .61$), task variables (reversals achieved and accuracy), antipsychotic dose, or model fit. In the Maryland sample, significant correlations were observed between VS RPE signaling and negative symptoms (left VS and SANS total: $r = -.377$, $p = .044$; right VS and the sum of SANS avolition and anhedonia scores: $r = -.405$, $p = .029$; left VS and the sum of SANS avolition and anhedonia scores: $r = -.447$, $p = .015$; Figure 2). A significant positive correlation was also observed between left VS RPE activation and positive symptoms (Brief Psychiatric Rating Scale psychosis factor: $r = .405$, $p = .029$). However, these correlations did fail to meet more stringent multiple comparison correction ($p = .05/14 = .003$). No significant correlations were found among VS RPE activations and task variables (reversals achieved and accuracy), antipsychotic dose, or model fit.

DISCUSSION

The goal of this report was to leverage computational approaches to examine the contribution of VS RPE signaling to reward-learning deficits in SZ. Thus, we conducted neuroimaging analyses, with parametric regressors generated from trial-by-trial RPE values derived from computational models of behavior. As previously reported, SZ patients showed poorer task performance in both samples (28,30). Contrary to previous reports (18–20,22), however, as a group, both SZ and CN displayed robust VS RPE activations. Further, as a group, we found no significant differences in VS RPE activation between SZ patients and CN participants. These results were consistent across whole-brain and VS VOI analyses, as well as when controlling for computational model fit. VS RPE signals did not show strong relationships to antipsychotic dosage or task performance. However, correlations between symptom severity and VS RPE magnitude were observed in the Maryland sample.

What may explain the discrepancies between our group level results and previous reports of abnormal VS RPE signals in SZ patients (18–20,22)? First, several previous reports (18,19,21,22) of VS RPE signaling involved tasks without contingency reversals [although see Schlagenhauf *et al.* (20)]. In PRL tasks, the most salient PEs tends to be negative given rarity of negative feedback. Recent evidence suggests abnormal positive but intact negative PE signaling in SZ patients (22). Thus, differences may have been observed if we had utilized a task more sensitive to positive PE signaling. Further, our claims generalize to RPEs only and not to other forms of PEs known to be disrupted in SZ (i.e., causal PE signaling in dorsolateral prefrontal cortex) (36). Second, our samples included chronic medicated patients, whereas several studies reporting blunted VS RPE signaling have examined antipsychotic naive patients (19,20) [however, see Morris *et al.* (18) and Koch *et al.* (22)]. Acute psychosis is associated with increased striatal dopamine neurotransmission, increased dopamine synthesis capacity, and chaotic dopaminergic firing

Table 3. Prediction Error–Related Blood Oxygen Level–Dependent Activations for Washington University in St. Louis Sample

Region	Vol (mm)	BA	x	y	z	CN	SZ	CN-SZ
Frontal								
Left middle frontal gyrus	2997	6	–38	–3	46	–2.6532[a]	–2.7165[a]	–0.4664
Left superior frontal gyrus	3429	6	–12	2	67	–3.152[a]	–3.0218[a]	–0.5873
Left precentral gyrus	1728	9	–44	21	34	–3.5627[a]	–2.3151[a]	–0.6984
Right paracentral lobule	5103	31	5	–31	44	1.6304	3.6285[a]	–1.0928
Right middle frontal gyrus	2619	9	39	28	29	–4.5824[a]	–1.7199[a]	–2.0543[a]
Right middle frontal gyrus	4779	6	42	6	45	–3.4866[a]	–2.0265[a]	–1.1371
Right superior frontal gyrus	13500	8	2	15	49	–4.7239[a]	–3.1633[a]	–1.7488[a]
Temporal								
Left inferior temporal gyrus	4887	37	–50	–65	1	1.5956	4.543[a]	–1.099
Right middle temporal gyrus	4374	37	51	–59	2	1.9313[a]	3.8982[a]	–1.1048
Parietal								
Left precuneus	2943	31	–10	–61	21	0.817	4.0002[a]	–1.7432[a]
Left inferior parietal lobule	5049	40	–59	–35	33	2.5657[a]	4.7808[a]	–0.9932
Right precuneus	2511	31	9	–52	34	1.2488	3.4923[a]	–0.939
Right inferior parietal lobule	4347	40	63	–34	29	2.5262[a]	4.6371[a]	–0.9624
Cingulate								
Left anterior cingulate	1998	24	–3	38	0	1.5471	3.603[a]	–0.5447
Left anterior cingulate	3078	24	0	18	22	1.9084[a]	2.1456[a]	–0.5008
Right cingulate gyrus	2025	8	23	20	24	3.3659[a]	2.6404[a]	0.592
Insula								
Left insula	4077	13	–34	16	7	–3.1586[a]	–0.9892	–1.1964
Right insula	2079	13	39	15	10	–3.9067[a]	–2.0955[a]	–1.7585[a]
Striatum								
Left putamen	18360		–20	–1	6	3.8084[a]	4.9123[a]	–0.7236
Right caudate	3078		22	–24	22	1.9049[a]	3.6445[a]	–0.4097
Right putamen	11421		16	3	6	5.4539[a]	5.8313[a]	0.0761
Right claustrum	6993		33	–7	12	2.0469[a]	4.4293[a]	–1.1575
Thalamus								
Left thalamus	5508		–25	–25	7	1.812[a]	4.1495[a]	–1.0366
Right thalamus	432		3	–18	21	2.5351[a]	3.795[a]	–0.559
Occipital								
Left superior occipital gyrus	1836	19	–44	–78	26	2.4556[a]	3.0894[a]	0.0523
Left middle occipital gyrus	5670	19	–28	–89	15	2.3829[a]	3.6727[a]	–0.4318
Left fusiform gyrus	2889	18	–25	–89	–16	1.7909[a]	3.8984[a]	–0.9688
Right middle occipital gyrus	3240	18	26	–90	12	1.7231[a]	3.1759[a]	–1.1256

BA, Brodmann area; CN, control; SZ, schizophrenia; Vol, volume.
[a]Significant at $p < .05$.

patterns (11,12). Atypical antipsychotics have been hypothesized to normalize these altered aspects of neurotransmission through dopamine receptor antagonism (37). Indeed, reports have shown that neural correlates of reward processing, particularly reward anticipation, are improved after administration of atypical antipsychotics (38–40). Thus, atypical antipsychotics may normalize striatal dopaminergic signaling, contributing to the robust VS RPE activation we see, at a group level, in the current report. Another way of viewing this discrepancy is that RPE signaling disruption may play and important role in the emergence of psychosis but less of a role in its maintenance.

Although both groups displayed robust VS RPE signaling, SZ patients demonstrated poorer performance on the PRL task, suggesting an alternative mechanism contributing to behavioral deficits. PRL tasks are known to involve salience network regions such as insula and task positive regions such as the ventrolateral prefrontal cortex and dorsal anterior cingulate (41). In addition, successful performance on PRL tasks likely involves the explicit learning of rules and engagement of dorsal prefrontal and parietal regions. Thus, PRL may not be optimal for detecting subtle differences in implicit/procedural and, therefore, more basal ganglia–driven learning. As such, it is possible that group behavior deficits in PRL performance stem, in part, from the disruption of prefrontal and parietal circuitry implemented in maintaining task representations (42–44). In fact, considerable evidence suggests that reward-learning deficits in SZ emerge when tasks become complex, requiring increased resources to maintain value and rule representations (22,29,45–48). In a conventional imaging

Table 4. Prediction Error–Related Blood Oxygen Level–Dependent Activations for Maryland Sample

Region	Vol (mm^3)	BA	x	y	z	CN	SZ	CN-SZ
Frontal								
Left medial frontal gyrus	857	10	–7	55	10	3.8230[a]	2.7020[a]	1.1490
Left middle frontal gyrus	888	10	–31	44	–4	3.1790[a]	4.2100[a]	–0.3410
Left middle frontal gyrus	1299	6	–29	–6	52	–3.1070[a]	–2.9860[a]	–0.6270
Left frontopolar cortex	4739	6	–20	28	54	3.8210[a]	3.4140[a]	0.6320
Right dorsomedial PFC	7226	6	4	13	53	–3.2940[a]	–3.1300[a]	–0.8960
Left precentral gyrus	864	4	–57	–9	23	2.8880[a]	2.6700[a]	0.7660
Parietal								
Left postcentral gyrus	1279	40	–57	–30	20	3.2470[a]	2.5250[a]	1.1880
Right superior parietal lobule	2309	7	21	–51	63	4.8740[a]	2.1580[a]	2.1070[a]
Left angular gyrus	3679	39	–46	–65	25	2.6770[a]	3.1640[a]	–0.0450
Right angular gyrus	847	39	49	–71	23	3.2050[a]	2.3940[a]	0.9860
Cingulate								
Left posterior cingulate	28434	31	–2	–43	27	3.1340[a]	3.9350[a]	0.0020
Insula								
Left anterior insula	2680	13	–34	22	6	–3.4020[a]	–3.0610[a]	–1.1150
Right anterior insula	3284	13	35	20	8	–3.8430[a]	–3.0940[a]	–1.1840
Striatum								
Left caudate/putamen	27550		–25	–4	4	4.2920[a]	4.3120[a]	–0.2930
Right caudate/putamen	38846		29	–6	8	4.1930[a]	4.2600[a]	0.3260
Occipital								
Right middle occipital gyrus	786	37	43	–61	–3	3.1920[a]	2.8070[a]	0.6140
Cerebellum								
Left cerebellum	770		–19	–45	–53	3.6490[a]	4.6900[a]	0.3430

BA, Brodmann area; CN, control; PFC, prefrontal cortex; SZ, schizophrenia; Vol, volume.
[a]Significant at $p < .05$.

analysis of the current data (28) we found that group differences in behavior on a PRL task were related to hypoactivation of a frontoparietal network of brain regions, whereas striatal regions were not. Waltz *et al.* (30) and Walter *et al.* (23) reported similar findings demonstrating hypoactivation of an executive control network of brain regions for SZ patients during performance of reward-learning tasks. Furthermore, Waltz *et al.* (49) found that neural abnormalities associated with outcome processing in SZ were largely cortical. In short, evidence suggests that much of the impairment exhibited by

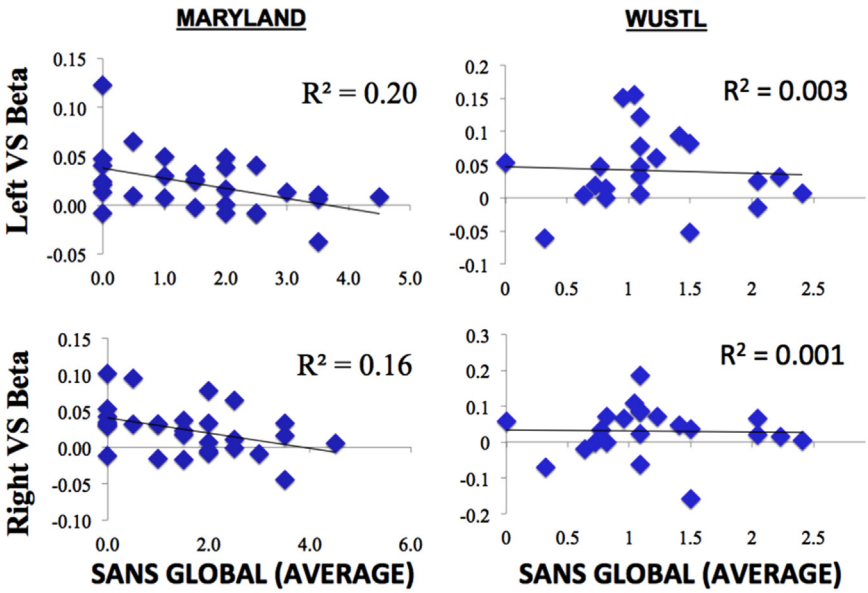

Figure 2. Relationship between ventral striatal (VS) reward prediction error signaling and negative symptoms. Item scores from the Scale for the Assessment of Negative Symptoms (SANS) were averaged for this analysis. WUSTL, Washington University in St. Louis.

SZ patients on PRL tasks may stem from a problem of outcome-driven learning, and not from a problem of feedback processing per se, and as a consequence, aberrant prefrontal circuitry appears to make a robust contribution to reward-learning deficits reported in chronic medicated patients.

This distinction between striatal and prefrontal mechanisms has recently been defined under two distinct computational learning rules, model based and model free, which collectively guide decision making (50,51). In model-free learning, the values of actions are learned by increasing the value of previously rewarded actions. This form of learning has been most closely tied to VS RPE signaling (52,53). Alternatively, model-based learning relies on prospective information such as the future consequences of actions to guide decision making. This form of learning has been associated with the dorsolateral prefrontal cortex (54) and ventromedial and orbital frontal cortices (54–56), in addition to the VS (52,56), and shows strong relationship to higher-order cognitive abilities (57). Our group recently implemented a task assessing model-based versus model-free learning and used computational approaches to quantify the relative contributions of both of these value-learning mechanisms (52) in medicated SZ patients. At a group level, we found that patients were similar to CN participants in their use of model-free learning (suggesting intact VS RPE signaling) but displayed reductions in model-based learning suggesting a deficit in the maintenance of rule and task representations (58). Interestingly, model-free learning estimates were negatively correlated with self-reported anhedonia, suggesting interaction between striatal and prefrontal mechanisms in high negative symptom patients. Further exploration with such computational frameworks will be instrumental in further specifying mechanisms of reward dysfunction in SZ.

Although we did not find group differences in VS RPE signaling in either sample, we did find a correlation between increased negative symptom severity and VS RPE signal hypoactivation, as well as increased positive symptoms and exaggerated VS RPE signaling in the Maryland sample. The inconsistency between samples may be due to numerically greater negative and positive symptom severity of patients in the Maryland sample. Importantly, this relationship between negative symptom severity and VS RPE magnitude converges with several other reports finding relationships between VS hypoactivation during reward-processing tasks and negative symptoms (18,25,30,39,59,60). This finding also points to the possibility that both striatal and prefrontal mechanisms may be involved in PRL deficits in high negative symptom patients. Such findings are consistent with the aforementioned report by our group assessing model-based and model-free learning. Further understanding how striatal and prefrontal mechanisms integrate to produce reward-learning impairment remains an important question for future research.

Applications to Other Datasets

Although the current report utilized a particular reinforcement-learning modeling framework (Q-learning) to examine task data (PRL) in a particular population (SZ), the approach implemented in the current investigation is broadly applicable. A myriad of reward-learning tasks exist in the basic and clinical science literatures for which data analysis could be enriched through fitting choice behavior to computational models. Such models have a number of advantages, as they allow for the estimation of unobservable aspects of reward learning such as PE magnitude and the expected value of options on a trial-by-trial basis. Quantifying these aspects of reinforcement learning, particularly those closely tied to biology, allows for new windows into the etiology and underlying mechanisms of psychiatric disorders. The current approach may also prove useful in understanding how reward dysfunction may vary across diagnostic boundaries, or in differentiating mechanisms associated with certain phases of illness.

Limitations

Our neuroimaging design (rapid event related), as well as the lack of jittered timing between choice and feedback, makes parsing component reward signals difficult in this study. Thus, it is possible that the VS signals we indexed contained component reward signals or motor signals unrelated to RPE. However, analyses including both value signals and RPEs into the same model did provide some specificity for RPE-related VS activation (see the Supplement). Second, PRL tasks may not be the most suitable paradigm in eliciting purely model-free VS RPE signaling as adequate task performance likely involves substantial model-based contributions. Third, although Q-learning algorithms have been used in many previous reports to index VS RPE in SZ (26,61,62), such algorithms may be limited in their account of choice behavior for PRL tasks as they fail to account for implicit task structure. Thus, more sophisticated models may have revealed group differences. Finally, although not specific to the current report, concerns have been raised regarding the sensitivity of BOLD activation to computationally derived regressors (63). For example, Wilson and Niv (63) demonstrated that assigning arbitrary learning rate values to individual subject data and deriving trialwise RPEs does not drastically change results when regressing trialwise VS RPE onto the BOLD signal. This insensitivity can make interpreting group differences challenging. Although our design has a large number of trials, a factor that may mitigate these insensitivities, our results may still be affected by such inherent limitations of model-based imaging approaches.

Summary

We used a reinforcement-learning framework to examine the contribution of VS RPE to reward dysfunction in SZ. Contrary to previous reports in unmedicated patients (20,26), we found, in two large samples, that although patients demonstrated worse task performance compared with CN participants, they did not demonstrate abnormal VS RPE signaling. One alternative mechanism for reward-learning dysfunction is disruption of prefrontal and parietal regions critically implemented in the maintenance of task representations. Future studies utilizing tools such as model-based reinforcement-learning algorithms are needed to assess the contribution of these mechanisms to reward learning in SZ.

ACKNOWLEDGMENTS AND DISCLOSURES

This study was supported by National Institute of Medical Health Grant Nos. R01 MH066031 (to DMB), K12 RR023250 (to JAW), and R01 MH080066

(to JAW); a project grant from HHSN271200599091C/ADB Contract # N01DA-5-9909, and by the National Institute on Drug Abuse Intramural Research Program (to JAW).

We thank our staff and the participants who gave their time for this study. We would also like to acknowledge Michael Frank and Lorenz Deserno for helpful guidance on some of the model fitting analyses presented in the current manuscript.

Dr. Waltz has consulted for Hoffman LaRoche. Dr. Barch has consulted for Pfizer, Amgen, Roche and Takeda, and has a contract to analyze imaging data for Pfizer. All other authors report no biomedical financial interests or potential conflicts of interest.

ARTICLE INFORMATION

From the Department of Psychological and Brain Sciences (AJC, AW, DMB) and Department of Psychiatry & Radiology (DMB), Washington University in St. Louis, St. Louis, Missouri; and the Department of Psychiatry and Maryland Psychiatric Research Center, University of Maryland School of Medicine, Baltimore, Maryland (JAW, ZX).

Address correspondence to Deanna M. Barch, Ph.D., Washington University in St. Louis, One Brookings Drive, Box 1125, St. Louis, MO 63130; E-mail: dbarch@wustl.edu.

Received Jan 31, 2016; revised July 7, 2016; accepted July 26, 2016.

Supplementary material cited in this article is available online at http://dx.doi.org/10.1016/j.bpsc.2016.07.007.

REFERENCES

1. Deserno L, Boehme R, Heinz A, Schlagenhauf F (2013): Reinforcement learning and dopamine in schizophrenia: Dimensions of symptoms or specific features of a disease group? Front Psychiatry 4: 172.
2. Gold JM, Waltz JA, Prentice KJ, Morris SE, Heerey EA (2008): Reward processing in schizophrenia: A deficit in the representation of value. Schizophr Bull 34:835–847.
3. Waltz JA, Gold JM (2015): Motivational deficits in schizophrenia and the representation of expected value. Curr Top Behavi Neurosci 27: 375–410.
4. Sutton RS, Barto AG (1998): Reinforcement Learning: An Introduction. Cambridge, MA: MIT Press.
5. Schultz W, Dayan P, Montague PR (1997): A neural substrate of prediction and reward. Science 275:1593–1599.
6. O'Doherty JP (2004): Reward representations and reward-related learning in the human brain: insights from neuroimaging. Curr Opin Neurobiol 14:769–776.
7. Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD (2006): Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. Nature 442:1042–1045.
8. Huys QJ, Moutoussis M, Williams J (2011): Are computational models of any use to psychiatry? Neural Netw 24:544–551.
9. Montague PR, Dolan RJ, Friston KJ, Dayan P (2012): Computational psychiatry. Trends Cogn Sci 16:72–80.
10. Wang X-J, Krystal JH (2014): Computational psychiatry. Neuron 84: 638–654.
11. Fusar-Poli P, Meyer-Lindenberg A (2012): Striatal presynaptic dopamine in schizophrenia, part II: Meta-analysis of [18F/11C]-DOPA PET studies. Schizophr Bulls 39:33–42.
12. Howes OD, Kambeitz J, Kim E, Stahl D, Slifstein M, Abi-Dargham A, *et al*. (2012): The nature of dopamine dysfunction in schizophrenia and what this means for treatment: Meta-analysis of imaging studies. Arch Gen Psychiatry 69:776–786.
13. Heinz A (2002): Dopaminergic dysfunction in alcoholism and schizophrenia–psychopathological and behavioral correlates. Eur Psychiatry 17:9–16.
14. Heinz A, Schlagenhauf F (2010): Dopaminergic dysfunction in schizophrenia: Salience attribution revisited. Schizophr Bull 36: 472–485.
15. Knutson B, Bjork JM, Fong GW, Hommer D, Mattay VS, Weinberger DR (2004): Amphetamine modulates human incentive processing. Neuron 43:261–269.
16. Kapur S (2003): Psychosis as a state of aberrant salience: A framework linking biology, phenomenology, and pharmacology in schizophrenia. Am J Psychiatry 160:13–23.
17. Radua J, Schmidt A, Borgwardt S, Heinz A, Schlagenhauf F, McGuire P, *et al*. (2015): Ventral striatal activation during reward processing in psychosis: A neurofunctional meta-analysis. JAMA Psychiatry 72: 1243–1251.
18. Morris R, Vercammen A, Lenroot R, Moore L, Langton J, Short B, *et al*. (2012): Disambiguating ventral striatum fMRI-related bold signal during reward prediction in schizophrenia. Mol Psychiatry 17:280–289.
19. Murray G, Corlett P, Clark L, Pessiglione M, Blackwell A, Honey G, *et al*. (2007): Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. Mol Psychiatry 13:267–276.
20. Schlagenhauf F, Huys QJ, Deserno L, Rapp MA, Beck A, Heinze H-J, *et al*. (2014): Striatal dysfunction during reversal learning in unmedicated schizophrenia patients. Neuroimage 89:171–180.
21. Gradin VB, Kumar P, Waiter G, Ahearn T, Stickle C, Milders M, *et al*. (2011): Expected value and prediction error abnormalities in depression and schizophrenia. Brain 134:1751–1764.
22. Koch K, Schachtzabel C, Wagner G, Schikora J, Schultz C, Reichenbach JR, *et al*. (2010): Altered activation in association with reward-related trial-and-error learning in patients with schizophrenia. Neuroimage 50:223–232.
23. Walter H, Kammerer H, Frasch K, Spitzer M, Abler B (2009): Altered reward functions in patients on atypical antipsychotic medication in line with the revised dopamine hypothesis of schizophrenia. Psychopharmacology (Berl) 206:121–132.
24. Waltz JA, Schweitzer JB, Gold JM, Kurup PK, Ross TJ, Salmeron BJ, *et al*. (2009): Patients with schizophrenia have a reduced neural response to both unpredictable and predictable primary reinforcers. Neuropsychopharmacology 34:1567–1577.
25. Wolf DH, Satterthwaite TD, Kantrowitz JJ, Katchmar N, Vandekar L, Elliott MA, *et al*. (2014): Amotivation in schizophrenia: Integrated assessment with behavioral, clinical, and imaging measures. Schizophr Bull 40:1328–1337.
26. Murray G, Corlett P, Clark L, Pessiglione M, Blackwell A, Honey G, *et al*. (2008): Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. Mol Psychiatry 13:267–276.
27. Button KS, Ioannidis JP, Mokrysz C, Nosek BA, Flint J, Robinson ES, *et al*. (2013): Power failure: Why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci 14:365–376.
28. Culbreth AJ, Gold JM, Cools R, Barch DM (2015): Impaired activation in cognitive control regions predicts reversal learning in schizophrenia. Schizophr Bull 42:484–493.
29. Waltz JA, Gold JM (2007): Probabilistic reversal learning impairments in schizophrenia: Further evidence of orbitofrontal dysfunction. Schizophr Res 93:296–303.
30. Waltz JA, Kasanova Z, Ross TJ, Salmeron BJ, McMahon RP, Gold JM, *et al*. (2013): The roles of reward, default, and executive control networks in set-shifting impairments in schizophrenia. PLoS One 8:e57257.
31. First MB, Spitzer RL, Gibbon M, Williams JB (2001): Structured Clinical Interview for DSM-IV-TR Axis I Disorders—Patient Edition (SCID-I/P. 2/2001 Revision). New York: Biometrics Research Department, New York State Psychiatric Institute.
32. Andreasen NC (1983): Scale for the Assessment of Negative Symptoms. Iowa City, IA: University of Iowa Press.
33. Andreasen NC (1984): Scale for the Assessment of Positive Symptons (SAPS). Iowa City, IA: University of Iowa Press.
34. Overall JE, Gorham DR (1962): The brief psychiatric rating scale. Psychol Rep 10:799–812.
35. Wechsler D (2001): Wechsler Test of Adult Reading (WTAR). San Antonio, TX: Psychological Corporation.
36. Corlett PR, Murray GK, Honey GD, Aitken MR, Shanks DR, Robbins TW, *et al*. (2007): Disrupted prediction-error signal in psychosis: Evidence for an associative account of delusions. Brain 130: 2387–2400.

37. Howes O, Egerton A, Allan V, McGuire P, Stokes P, Kapur S (2009): Mechanisms underlying psychosis and antipsychotic treatment response in schizophrenia: Insights from PET and SPECT imaging. Curr Pharm Des 15:2550–2559.
38. Nielsen MO, Rostrup E, Wulff S, Bak N, Broberg BV, Lublin H, *et al*. (2012): Improvement of brain reward abnormalities by antipsychotic monotherapy in schizophrenia. Arch Gen Psychiatry 69:1195–1204.
39. Juckel G, Schlagenhauf F, Koslowski M, Filonov D, Wüstenberg T, Villringer A, *et al*. (2006): Dysfunction of ventral striatal reward prediction in schizophrenic patients treated with typical, not atypical, neuroleptics. Psychopharmacology (Berl) 187:222–228.
40. Schlagenhauf F, Juckel G, Koslowski M, Kahnt T, Knutson B, Dembler T, *et al*. (2008): Reward system activation in schizophrenic patients switched from typical neuroleptics to olanzapine. Psychopharmacology (Berl) 196:673–684.
41. Greening SG, Finger EC, Mitchell DG (2011): Parsing decision making processes in prefrontal cortex: Response inhibition, overcoming learned avoidance, and reversal learning. Neuroimage 54:1432–1441.
42. Heinrichs RW, Zakzanis KK (1998): Neurocognitive deficit in schizophrenia: A quantitative review of the evidence. Neuropsychology 12:426–445.
43. Lee J, Park S (2005): Working memory impairments in schizophrenia: A meta-analysis. J Abnorm Psychol 114:599–611.
44. Semkovska M, Bédard M-A, Godbout L, Limoge F, Stip E (2004): Assessment of executive dysfunction during activities of daily living in schizophrenia. Schizophr Res 69:289–300.
45. Gold JM, Waltz JA, Matveeva TM, Kasanova Z, Strauss GP, Herbener ES, *et al*. (2012): Negative symptoms and the failure to represent the expected reward value of actions: Behavioral and computational modeling evidence. Arch Gen Psychiatry 69:129–138.
46. Morris SE, Heerey EA, Gold JM, Holroyd CB (2008): Learning-related changes in brain activity following errors and performance feedback in schizophrenia. Schizophr Res 99:274–285.
47. Yılmaz A, Simsek F, Gonul AS (2012): Reduced reward-related probability learning in schizophrenia patients. Neuropsychiatr Dis Treat 8:27–34.
48. Collins AGE, Brown JK, Gold JM, Waltz JA, Frank MJ (2014): Working memory contributions to reinforcement learning impairments in schizophrenia. J Neurosci 34:13747–13756.
49. Waltz JA, Schweitzer JB, Ross TJ, Kurup PK, Salmeron BJ, Rose EJ, *et al*. (2010): Abnormal responses to monetary outcomes in cortex, but not in the basal ganglia, in schizophrenia. Neuropsychopharmacology 35:2427–2439.
50. Balleine BW, Daw ND, O'Doherty JP (2008): Multiple forms of value learning and the function of dopamine. In: Glimcher PW, editor. Neuroeconomics: Decision Making and the Brain. New York: NYU Press, 367–385.
51. Daw ND, Niv Y, Dayan P (2005): Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat Neurosci 8:1704–1711.
52. Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ (2011): Model-based influences on humans' choices and striatal prediction errors. Neuron 69:1204–1215.
53. Gläscher J, Daw N, Dayan P, O'Doherty JP (2010): States versus rewards: Dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. Neuron 66:585–595.
54. Smittenaar P, FitzGerald TH, Romei V, Wright ND, Dolan RJ (2013): Disruption of dorsolateral prefrontal cortex decreases model-based in favor of model-free control in humans. Neuron 80:914–919.
55. Lee SW, Shimojo S, O'Doherty JP (2014): Neural computations underlying arbitration between model-based and model-free learning. Neuron 81:687–699.
56. McDannald MA, Lucantonio F, Burke KA, Niv Y, Schoenbaum G (2011): Ventral striatum and orbitofrontal cortex are both required for model-based, but not model-free, reinforcement learning. J Neurosci 31:2700–2705.
57. Otto AR, Skatova A, Madlon-Kay S, Daw ND (2014): Cognitive control predicts use of model-based reinforcement learning. J Cogn Neurosci 27:319–333.
58. Culbreth A, Westbrook A, Daw N, Botvinick M, Barch D (2016): Reduced model-based decision-making in schizophrenia [published online ahead of print May 12]. J Abnorm Psychol.
59. Juckel G, Schlagenhauf F, Koslowski M, Wüstenberg T, Villringer A, Knutson B, *et al*. (2006): Dysfunction of ventral striatal reward prediction in schizophrenia. Neuroimage 29:409–416.
60. Kirschner M, Hager OM, Bischof M, Hartmann MN, Kluge A, Seifritz E, *et al*. (2015): Ventral striatal hypoactivation is associated with apathy but not diminished expression in patients with schizophrenia. J Psychiatry Neurosci 41:152–161.
61. Dowd EC, Frank MJ, Collins A, Gold JM, Barch DM (2016): Probabilistic reinforcement learning in schizophrenia: Relationships to anhedonia and avolition [published online ahead of print Aug 5]. Biol Psychiatry Cogn Neurosci Neuroimaging.
62. Reinen JM, Van Snellenberg JX, Horga G, Abi-Dargham A, Daw ND, Shohamy D (2016): Motivational context modulates prediction error response in schizophrenia [published online ahead of print April 22]. Schizophr Bull.
63. Wilson RC, Niv Y (2015): Is model fitting necessary for model-based fMRI? PLoS Comput Biol 11:e1004237.