Motivational Context Modulates Prediction Error Response in Schizophrenia

Jenna M. Reinen[*,1,2], Jared X. Van Snellenberg[3,4], Guillermo Horga[3,4], Anissa Abi-Dargham[3,4], Nathaniel D. Daw[5,7], and Daphna Shohamy[1,6,7]

[1]Department of Psychology, Columbia University, New York, NY; [2]Department of Psychology, Yale University, New Haven, CT; [3]Department of Psychiatry, Columbia University Medical Center, New York, NY; [4]Division of Translational Imaging, New York State Psychiatric Institute, New York, NY; [5]Princeton Neuroscience Institute and Department of Psychology, Princeton University, Princeton, NJ; [6]Zuckerman Mind, Brain, Behavior Institute and Kavli Center for Brain Science, Columbia University, New York, NY.

[7]These authors contributed equally to this work.

*To whom correspondence should be addressed; Department of Psychology, Yale University, 1 Prospect Street, New Haven, CT 06511, US; tel: 203-436-9449, fax: 203-432-7172, e-mail: jenna.reinen@yale.edu

Background: Recent findings demonstrate that patients with schizophrenia are worse at learning to predict rewards than losses, suggesting that motivational context modulates learning in this disease. However, these findings derive from studies in patients treated with antipsychotic medications, D2 receptor antagonists that may interfere with the neural systems that underlie motivation and learning. Thus, it remains unknown how motivational context affects learning in schizophrenia, separate from the effects of medication. *Methods:* To examine the impact of motivational context on learning in schizophrenia, we tested 16 unmedicated patients with schizophrenia and 23 matched controls on a probabilistic learning task while they underwent functional magnetic resonance imaging (fMRI) under 2 conditions: one in which they pursued rewards, and one in which they avoided losses. Computational models were used to derive trial-by-trial prediction error responses to feedback. *Results:* Patients performed worse than controls on the learning task overall, but there were no behavioral effects of condition. FMRI revealed an attenuated prediction error response in patients in the medial prefrontal cortex, striatum, and medial temporal lobe when learning to predict rewards, but not when learning to avoid losses. *Conclusions:* Patients with schizophrenia showed differences in learning-related brain activity when learning to predict rewards, but not when learning to avoid losses. Together with prior work, these results suggest that motivational deficits related to learning in schizophrenia are characteristic of the disease and not solely a result of antipsychotic treatment.

Key words: prediction error/antipsychotic medication/ reinforcement learning/model-based fMRI/ neuroimaging

Introduction

Dopamine is thought to play a role in incentive salience, reinforcement learning,[1–3] and in the pathophysiology of schizophrenia.[4–6] Consequently, it has been postulated that deficits in reward signaling and/or value representation may be related to the affective and motivational (negative) symptoms of schizophrenia,[7–12] and to abnormal attributions of salience that are characteristic of positive symptoms.[13,14]

In reinforcement learning, a prediction error (PE) is the difference between the reward received and what was expected.[1] This learning signal is a critical component in the computation and maintenance of value that supports the ability to anticipate, seek, and select cues that maximize gain and minimize loss.[15,16] PE magnitude correlates with dopaminergic signaling,[1,2] and is typically observed in regions known to be both targets of midbrain dopaminergic projections and involved in value conceptualization.[17,18] In humans, functional magnetic resonance imaging (fMRI) studies have shown that PE signals correlate with activity in the ventral tegmental area, striatum, and regions of medial/orbitofrontal prefrontal and parietal cortices.[16,17,19–21] Patients with schizophrenia demonstrate blunting of PEs in these regions,[22–26] suggesting that motivational deficits are related to abnormal dopaminergic value signals.

In schizophrenia, performance deficits and attenuated BOLD responses in reinforcement learning have been reported when patients update positive, relative to negative, outcomes.[8–11,24,27,28] This may be related to alterations in dopamine transmission in schizophrenia, as evidence from dopaminergic alterations in Parkinson's disease and pharmacological interventions in healthy controls have

demonstrated similar asymmetrical effects of dopamine on learning to predict positive vs negative outcomes.[29–31] Furthermore, larger doses of antipsychotic medication are associated with attenuated PEs in the striatum and prefrontal cortex (PFC).[27]

Because many patient participants in these studies were receiving antipsychotic medication during testing, it remains possible that antipsychotics, which are potent D2 receptor (D2R) antagonists, may enhance loss learning[32,33] or attenuate positive error signals.[30,34] Further, while the few studies that examined reinforcement learning in unmedicated patients with schizophrenia demonstrated attenuation of striatal PEs when learning from rewards,[26,35–37] none employed a loss condition to examine how motivational context, which has been described as "information regarding the attractiveness/aversiveness of the past, present, or possible future reward/threat,"[38] impacts learning in this population. Generally, motivational effects related to reward and punishment expectation and learning might manifest either trial-wise, to correct vs incorrect feedback, or contextually between situations or conditions involving reward vs loss (here, unlike most work in this area, we primarily investigate the latter possibility). Because schizophrenia is specifically associated with motivational deficits,[12] understanding the mechanisms underlying reinforcement learning across motivational contexts in the absence of pharmacological D2R blockade is especially important.

To address this, we used a reinforcement learning task with 2 separate conditions in which unmedicated patients with schizophrenia and demographically-matched controls learned to (1) pursue reward and (2) avoid loss. We sought to determine whether group differences existed in learning signals when predicting rewards vs losses at the behavioral and neural levels. Given evidence supporting deficits in approach motivation,[12] response to positive stimuli,[8] and reward context performance[28] in schizophrenia, we expected to observe blunted PE response in patients vs controls especially when learning to predict rewards and specifically in the striatum and medial prefrontal cortex (mPFC), regions thought to be critically involved in reinforcement learning and associated with value-related deficits in schizophrenia.[22,26,35,39,40]

Methods

Participants

Twenty-three controls (mean age = 33.7 y, SD = 8.6; 13 females) and 16 patients (ages 19–55, mean age = 34.3, SD = 10.5, 7 females, medication-naïve = 7, medication-free = 9) participated in the study. All spoke English. Patients were stable outpatients at the New York State Psychiatric Institute. All study procedures were approved by the Institutional Review Board and participants provided written informed consent prior to participation. Patients had a diagnosis of DSM-IV schizophrenia,

schizoaffective disorder, or schizophreniform disorder confirmed with the Diagnostic Instrument for Genetic Studies.[41] Controls were matched on age, sex, ethnicity, and parental socioeconomic status (table 1). Exclusion criteria were positive urine toxicology, pregnancy, comorbid Axis I disorder (with the exception of history of substance use with no use in the past 6 mo), neurological disorders, and current psychotropic medication use. Controls were excluded for any psychiatric disorder and family history (first-degree) of schizophrenia.

Procedures and Task

In the scanner, participants first performed a working memory battery, reported separately.[42] Participants performed a short practice session, in which they were able to ask questions, before completing 2 non-intermixed, separate, counterbalanced phases of 60 trials (120 total trials) of a probabilistic reinforcement learning task.[28] They were instructed that their goal was either to make money (gain condition), or avoid losing money endowed prior to task (loss condition, $40 endowment). On each trial (figure 1) jittered intervals separated (1) choice (3 s) between 1 of 2 geometric stimuli; (2) written verbal feedback ("Correct"/"Incorrect," 1 s); and (3) monetary outcome related to performance (high-resolution image of the currency received, 1 s). Inter-stimulus and inter-trial-intervals lasted 3–7 seconds, taken from an exponential distribution. One of the available shapes was the "optimal" stimulus that yielded "Correct" feedback 70% of the time, while the other yielded "Correct" feedback only 30% of the time. Stimuli were counterbalanced for condition and optimal shape. Participants were instructed that they would receive the actual monetary outcome shown during the task in addition to compensation for participation.

Monetary outcome was linked to feedback probabilistically. In the gain condition, when a participant was correct, they received either $1 or $0.50 (50% probability). When incorrect, they received $0.50 or no earnings. When

Table 1. Group Demographics

Demographics	Patients (SD)	Controls (SD)
N	16	23
Age	34.3 (10.5)	33.7 (8.5)
Parental socioeconomic status	42.9 (13)	39.9 (14)
Gender (M/F)	9M/7F	10M/13F
Illness duration	14.2(11.8)	N/A
PANSS (General)*	28.4 (6.9)	16.9 (1.3)
PANSS (Negative)*	16.2 (5.8)	8.7 (1.4)
PANSS (Positive)*	13.7 (4.5)	7 (0)

Note: PANSS, positive and negative syndrome scale. Significant demographic group differences are denoted with *. Parental SES (socioeconomic status) data missing for 2 controls and 1 patient. PANSS missing for 4 controls and 1 patient.
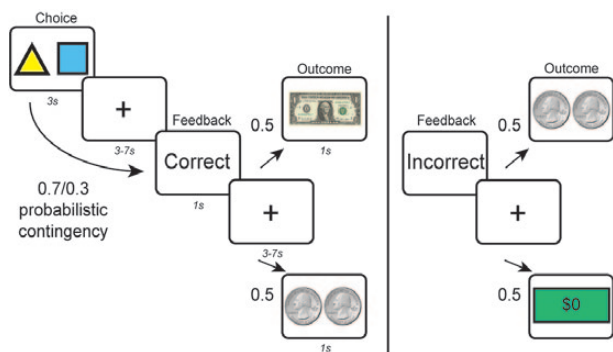
Fig. 1. Probabilistic learning task. Participants learned to predict gains and losses, which were probabilistically associated with cues in an instrumental task. A sample trial from the gain condition is shown for correct and incorrect choices: participants made a choice (triangle/square), received verbal feedback on their choice ("correct"/"incorrect"), and then received a monetary outcome. Trial structure and probabilities were identical across conditions.

a participant was correct in the loss condition, they either lost no money or $0.50, and when incorrect, they lost $1 or $0.50. This design was implemented to equate the magnitude of both positive and negative PEs across conditions, while allowing for differences in motivational context. If participants failed to respond within 3 seconds they received the worst possible outcome for that condition.

This task involves 2 feedback phases, first correct/incorrect, then monetary, which in principle, may affect choices or induce PEs. In practice, the correct/incorrect feedback is more informative and immediate than the later, noisy, monetary payoff. Accordingly, analyses (not reported here) weighing both feedback events indicated that behavioral and neural effects were predominantly driven by the initial feedback. Thus, analyses in this paper concern learning and PEs driven by verbal feedback.

Reinforcement Learning Model

To assess learning and to model BOLD PE signals, we fit a reinforcement learning model[16,43,44] to each subject's choices to estimate 2 free parameters for each subject and condition (gain/loss). On each trial (t), the model can learn from the PE (δ) at feedback (fb) about the choice (c), which is linked to the monetary reward outcome. These are used to learn the expected values (Q) associated with the choices and feedback:

$$Q_{c(t)} = Q_{c(t)} + \alpha * \delta_t$$

$$\delta_t = fb(t) - Q_{c(t)}$$

The model contains a free learning rate parameter (α) estimated separately for each subject and condition. The learning rate reflects how strongly PE affects value updating across all trials in each condition, as measured by their choices. Critically, the ability to update the cue value relies

on information in the form of PE, the difference between value expected and outcome received (δ). We further assumed that participants made their choices according to the learned values $Q_{c(t)}$ at each step using a softmax distribution with free inverse temperature parameter β, which effectively serves as a regression weight that controls the extent to which Q values determine participant choices.[44] To estimate the free parameters (and specifically the effects of group and condition on them), and to generate subject-specific trial-by-trial PE timeseries given each participant's data for subsequent fMRI analysis, we used Bayesian MCMC statistical inference.[45–47] We also examined 2 follow-up elaborations of the baseline model, in which either the learning rate or the scaling of the feedback term $fb(t)$ was allowed to vary not just as a function of condition (gain/loss), but also of the type of feedback (positive/negative) on each trial, allowing us to test for group differences on these more specific effects. Method specifications are outlined in the supplementary material.

FMRI Methods

Scanning took place at the Neurological Institute at Columbia University Medical Center on a 1.5T Philips Intera Scanner using an 8-channel SENSE head coil. Participants lay supine on the scanner bed while viewing images projected on a screen, and used a hand-held fiber optic trackball to respond to the task. T1-weighted structural images were acquired with an SPGR sequence (256 mm field of view [FOV], 200 slices, 1 mm isotropic voxels). Whole-brain functional EPIs were obtained at a 2 seconds repetition time (TR), 28 ms echo time (TE), 77° flip angle, 192 mm FOV, 40 slices, and 3 mm isotropic voxels. We collected 173 volumes in each of 6 runs (3 runs gain/3 loss, 5 min 46 s/run).

Preprocessing procedures were carried out as described elsewhere[42] and in the supplementary material. After preprocessing, first-level statistical analyses using a standard general linear model were implemented in NeuroElf (http://neuroelf.net/, last accessed April 15, 2016). The model included 4 regressors of interest including choice, comprised of the time interval from cue onset to button press, feedback, and reward outcome, which were modeled as stick functions and reported as 1/0 contrasts. Additionally, we computed a trial-by-trial and subject-by-subject parametric regressor for feedback PE, defined as the median over MCMC samples of the PE computed at each timestep (Since the samples are from the posterior distribution over the model's free parameters, for each subject, this procedure accounts for the effects on PE of between-subject variation in parameters and uncertainty in their estimation). Assessment of motion, data quality, and results concerning choice and reward outcome are discussed in the supplementary material. Finally, we carried out whole-brain analyses using family-wise error (FWE) thresholds of $P < .05$ using AlphaSim[48] with a smoothness estimated
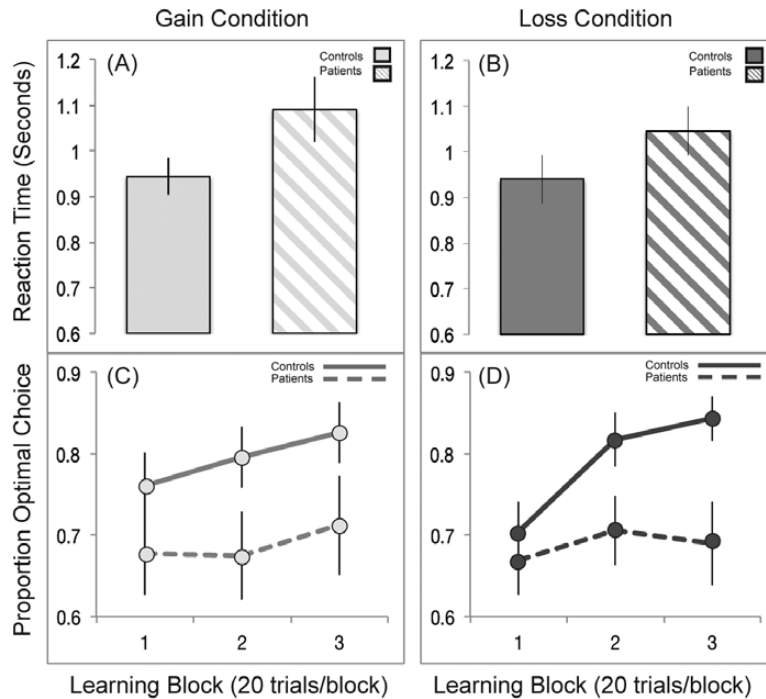
Fig. 2. Behavioral results. For gain and loss conditions, patients demonstrated slower reaction time (A.Gain/B.Loss) and lower optimal choice performance scores (C.Gain/D.Loss) relative to controls.

from the error term for each contrast to examine PE across conditions and groups.

Exploratory Analyses

Correlations between learning behavior, disease symptoms, and memory capacity,[49] and their relationship with PE response in striatum and mPFC, are reported in the supplementary material. We also explored within-condition group contrasts to examine the direction of the interaction effects (figure 4).

Results

Behavioral Results

Choice reaction time (RT)[50] for each group and condition (figures 2A and 2B) were analyzed in a 2-way ANOVA with group (patients/controls) and condition (gain/loss) as factors. We found a main effect of group ($F_{1,74} = 5.02$, $P = .03$) but not condition ($F_{1,74} = 0.19$, $P = .66$), and no interaction ($F_{1,74} = 0.13$, $P = .72$), indicating that patients were slower than controls across both conditions. The same ANOVA model was used for optimal choice (figures 2C and 2D), demonstrating a main effect of group, indicating that patients performed significantly worse overall ($F_{1,74} = 8.44$, $P < .01$), but there was no effect of condition ($F_{1,74} = 0.01$, $P = .92$) and no interaction ($F_{1,74} = 0.01$, $P = .92$).

Reinforcement Learning Model

We fit the free parameters of a reinforcement learning model to the trial-by-trial choices, using a MCMC

Table 2. Behavioral Data (Credible Intervals on Model Parameters [CI] With Bayesian *P*-values for the Model Parameters and Their Interaction With Group and Condition)

Behavioral Results (Model)	Lower CI	Upper CI	*P*
α*Group	−1.67	0.68	.21
α*Condition	−2.71	1.78	.28
α*Group*Condition	−1.29	1.27	.48
β* Group	−0.68	−0.21	<.001
β* Group*Condition	−0.18	0.37	.75
β* Condition	−0.51	0.63	.46

procedure to estimate each participant's parameters, by condition (gain/loss), and how these varied at the population level as a function of disease group (table 2, supplementary table 1, model 1). We found no effect of group, condition, or their interaction on the learning rate α ($P = .21$, $P = .28$, $P = .48$, respectively). However, we did find an effect of group on the softmax temperature parameter β ($P < .001$), suggesting patients were noisier in their choices (median β difference: −.49). There was no effect of condition on this parameter, nor did the groupwise difference interact with condition ($P = .46$, $P = .75$). For information about 2 follow-up models, including trial-specific effects, please see the supplementary material.

Functional Imaging Results

Whole-Brain Corrected Condition-By-Group Interaction. We next examined correlations of BOLD activity with a trial-by-trial PE signal generated for each subject from the fit model.

A whole brain corrected (FWE $P < .05$) condition-by-group interaction analysis (PE on gain>loss, controls>patients; figure 3, supplementary table 3) revealed several significant clusters, including 1 in the left medial temporal lobe (MTL; peak −42, −21, −9, t-max = 4.93, k = 829) extending into hippocampus and parahippocampal gyrus (−42, −33, −6, t-max = 3.31, k = 62). We also found a significant effect in the right temporal lobe (peak 27, 12, −36, t-max = 4.20, k = 571) and the striatum (local maxima in putamen: 18, 3, 0, t-max = 4.06, k = 42; and ventral striatum/caudate: 3, 9, −3, t-max = 3.54, k = 71), as well as in precuneus (12, −51, 39, t-max = 4.39, k = 472) extending to posterior cingulate (9, −36, 21, t-max = 4.20, k = 23), and in precentral gyrus (30, −15, 30, t-max = 3.50, k = 508), extending to middle frontal gyrus (30, −6, 39, t-max = 3.37, k = 24) and postcentral gyrus (42, −27, 36, t-max = 3.22, k = 58).

Exploratory Analyses

We next unpacked this analysis to explore group differences separately within each condition (gain vs loss; supplementary figure 3). In the gain condition, a direct group contrast of feedback PE revealed significantly greater PE-related activation in the controls than the patients in several clusters. The first was found in medial frontal gyrus (−15, 3, 66, t-max = 5.37, k = 11 752), precuneus (9, −51, 39, t-max = 5.33, k = 270), MTL (−21, −21, −9, t-max = 4.20, k = 99), posterior cingulate, (12,

−60, 12, t-max = 4.06, k = 91), and culmen (−6, 69, −9, t-max = 4.20, k = 123). The second extended from lateral (−33, 63, 0, t-max = 4.57, k = 669) to medial PFC (−6, 63, −12, t-max = 4.05, k = 125; figure 4). In the loss condition, 1 cluster across cerebellum (−42, −57, −27, t-max = 4.35, k = 1412), cuneus (0, −78, 39, t-max = 4.14, k = 148) and occipital lobe (−30, −90, −12, t-max = 3.84, k = 77) survived correction. In contrast to the gain condition, although no significant group difference was seen for the loss condition in mPFC (figure 4), the 2-way interaction reported above was not significant in this region.

We did not find any significant relationships between negative symptoms and learning or memory performance, or in PE response in striatum or mPFC. Details and statistics pertaining to these analyses are reported in the supplementary material.

Discussion

The findings of the present study demonstrate for the first time that unmedicated patients with schizophrenia show blunted PE responses relative to controls in brain regions including the striatum, mPFC, and MTL in the context of predicting gains, but not losses. However, no corresponding behavioral effects of gain vs loss condition or condition-by-group differences were detected, a negative result we return to below. Nevertheless, although blunted
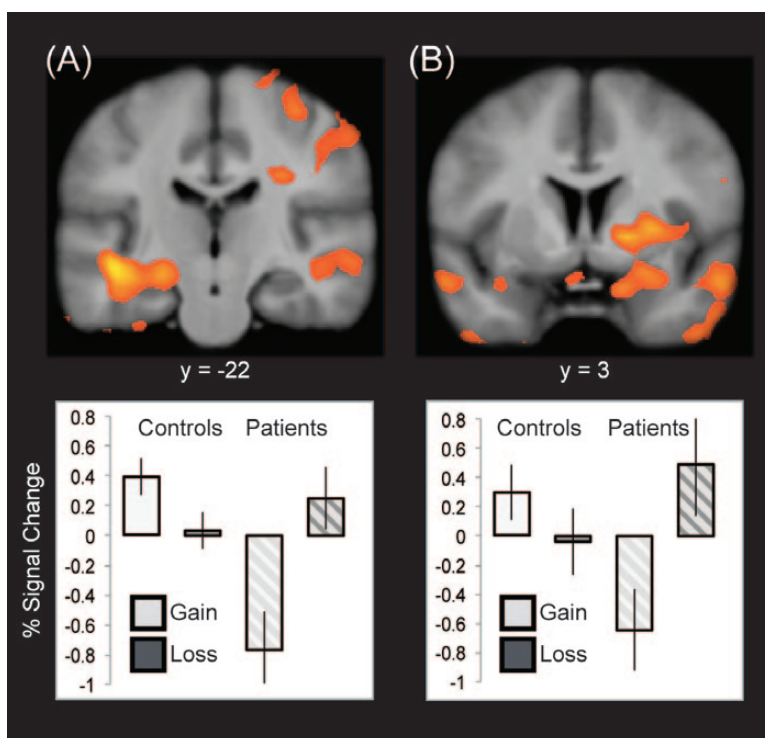
Fig. 3. Condition-group interaction. Interaction results (gain>loss, controls>patients) for prediction error (PE)-related BOLD responses to feedback are shown at a corrected threshold (whole-brain family-wise error [FWE] $P < .05$). Regression weight plots from clusters in (A) left medial temporal lobe and (B) right putamen are shown to demonstrate the direction of the interaction, and not for statistical inference.
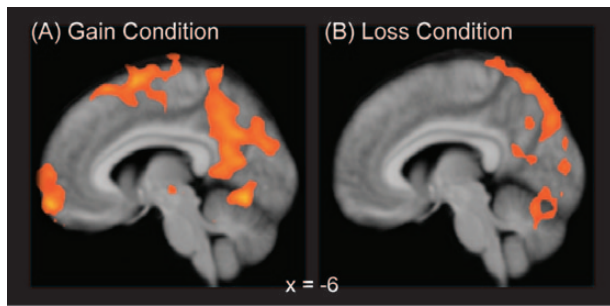
Fig. 4. Corrected condition-specific group differences. A whole-brain corrected (family-wise error [FWE] $P < .05$) contrast for controls>patients on prediction error (PE)-related BOLD responses to feedback showed a difference in medial prefrontal cortex in the gain condition. No such difference was significant in the loss condition.

reward response has been demonstrated in unmedicated patients,[26,35–37] these findings suggest that deficits in PE signaling in patients exist when motivated by rewards, indicating that motivational context modulates the engagement of neural systems supporting reinforcement learning in schizophrenia.

These data are noteworthy as prior research has implicated antipsychotic treatment as a potential mechanism contributing to deficits in reward PE.[30,34] However, because all patients in this study were unmedicated, these findings instead suggest that D2R blockade via pharmacological means is not the primary driving factor underlying the neural mechanism for blunted appetitive and intact avoidant learning signals. Rather, it is likely that the deficit specific to learning to predict positive outcomes may be characteristic of the disease itself.

The literature concerning the relationship between medication and reward learning deficits in schizophrenia has been mixed, and has been confounded with medication type and demographic factors. Antipsychotic medications, and specifically D2R antagonism, can attenuate responses to positive PEs and enhance learning from negative PEs.[29,30,39,51] Some typical antipsychotics are used at doses that achieve high D2 occupancy,[52] and are associated with greater deficits in PE signaling than atypical medications, further implicating this mechanism in reward PE blunting.[39,53–55] Differences in predicting rewards vs losses have been reported largely in medicated patients,[8–11,28] but while some studies have shown dose-related learning trends,[27] others demonstrated no relationship.[22] Imaging studies including unmedicated participants, however, reported an overall attenuation of PE response in the striatum,[26,35,36] and 1 mixed finding related to negative PE in mPFC and striatum.[37] Thus, while the present data cannot rule out the possibility that medication can exacerbate these valence effects, the cumulative evidence suggests that reward-specific PE abnormalities in schizophrenia may exist independent of medication status. Future studies

incorporating controlled comparison across medication status will be necessary to fully address this issue.

Several mechanisms have been proposed to underlie reward-specific learning deficits. For example, learning from positive and negative outcomes is thought to be related to the direct and indirect pathways in the basal ganglia that are modulated by D1 and D2 receptors, respectively,[10,29] and provide a framework by which D2R blockade by antipsychotics facilitate learning[30] from negative outcomes in schizophrenia. In patients, we observed lower β values, consistent with a noisier and qualitatively different pattern of choice behavior, which appeared to substantially reflect a blunted behavioral response to positive feedback, supporting prior findings.[8,11] The identification of deficits on the level of trial and condition in schizophrenia implies there may exist abnormalities on multiple temporal levels, potentially attributed to disrupted trial-level feedback responses due to aberrant dopamine signaling,[13] and sustained value computation abnormalties mediated by tonic signals.[56,57] Consequently, understanding these processes is an important direction in schizophrenia research.

An important caveat to the finding of neural effects of condition is that the differences reported did not have any detectable behavioral counterpart in terms of differences in patients' choice or RT performance between conditions. While one should be cautious interpreting neural effects in the absence of a corresponding behavioral effect, these neural results seem plausible in the context of the strong previous literature on these issues, suggesting the failure more likely concerns our behavioral measures. One possible reason for the lack of behavioral effects is that these patients were generally impaired in their learning performance, which may have masked more selective differences across the gain vs loss conditions. Another possibility is that this task was not sensitive enough to identify condition-specific behavioral differences; indeed, no main effect of condition was found within the control group alone (though this need not be a prerequisite for finding condition effects in patients). Prior reports of condition effects have employed paradigms that required frequent contingency updating.[28] Instead, the present task used a static 70/30 contingency to ensure participants could learn well enough to acquire a cue-specific value expectation before it can be violated, creating a PE for use with fMRI. Future studies should consider using tasks with dynamic contingencies, which may be more sensitive to reinforcement learning deficits.

The PE BOLD response we found in controls was consistent across conditions with results of meta-analyses examining PE and value.[17,58] Patients, though, showed this pattern in the loss but not gain condition, even demonstrating a negative relationship with PE in striatum and MTL. This points to a characteristic difference in how patients process stimuli warranting an approach or avoidance response. In fact, affective and motivational

states can modulate PE responses in rewarding[59] and aversive contexts.[60] For example, the role of the striatum has been established in rewarding[1] and aversive learning,[61–64] but key functional differences may exist, ranging from anatomical segregation within its subdivisions[65,66] to interactions with particular learning systems, such as amygdala[67] based on learning type and context. Further, tasks involving altered motor responses based upon framing have demonstrated differences in striatal recruitment.[68,69] Thus, understanding how neural learning signals differ depending upon response context may elucidate mechanisms underlying blunted reward-specific responses in schizophrenia.

Exploratory analyses revealed effects of condition and group in regions outside of the striatum and mPFC, such as in cingulate, MTL, and cuneus (figure 3), for which several interpretations exist. Salience and value responses have been demonstrated in cuneus[70] and posterior cingulate.[17] Further, response in posterior cingulate may be linked specifically to the representation of positive value[17] and attenuation of response to incentive value in this region has been implicated in psychosis.[71] Additionally, the MTL is subject to modulation by dopamine and rewards,[72] both of which affect mnemonic and cognitive processes related to motivation and decision making in healthy participants[73] and patients with schizophrenia.[74] Given known hippocampal abnormalities in schizophrenia[75] and importance of the posterior cingulate in reward-specific incentive salience, these findings suggest that reward-modulated cognitive processes in these regions should be investigated in future studies.

One limitation of this study stems from the difficulty in recruiting medication-free patients, as only a limited number of patient participants were tested. Additionally, while our modeling analyses fit both groups better than chance, the controls showed a better fit than patients. This may simply reflect the underlying behavioral finding—ie, that patients' behavior is less driven by (and therefore less predictable based on) the task feedback, and ultimately noisier. However, it is possible an alternative model exists which might better account for patient behavior. Importantly, because the current task does not involve a transfer phase to distinguish choices driven by learned action preferences ("policies") vs reward predictions, the actor–critic model (which has been suggested to be a better structural account of learning in schizophrenia in some circumstances[8]) behaves similar to the Q-learning model used here. Indeed, behavioral fits using an actor-critic model (results not shown) did not detectably differ in either group. Future studies should employ transfer phases or other manipulations to address this question in unmedicated patients across motivational contexts, and symptom profiles. Also, given that about half of our patient sample had prior exposure to antipsychotic medications, it is important to consider the possibility that—despite these and other[36] findings—prior antipsychotic

use may exert long-term effects on the brain, although this concern is mitigated by the absence of observed differences in PE between medication-naïve and medication-free patients (see the supplementary material). Finally, it is important to acknowledge that medication use is often correlated with disease severity. In prior studies, patients on high medication doses may have had more severe symptoms, confounding results.

To summarize, this study suggests that neural learning signals are processed differently in patients with schizophrenia dependent upon the motivational context of the task. These data indicate that patients may have a selective neural learning signal deficit when motivated to pursue rewards compared to avoid losses, which may be a disease characteristic, and not solely attributable to antipsychotic use. Along with a larger body of evidence, these findings confirm abnormalities of reward-motivated processing in schizophrenia and warrant continued investigation of this deficit and its relationship to disease symptoms.

Supplementary Material

Supplementary material is available at http://schizophreniabulletin.oxfordjournals.org.

References

1. Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. *Science*. 1997;275:1593–1599.
2. Bayer HM, Glimcher PW. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron*. 2005;47:129–141.
3. Hart AS, Rutledge RB, Glimcher PW, Phillips PE. Phasic dopamine release in the rat nucleus accumbens symmetrically encodes a reward prediction error term. *J Neurosci*. 2014;34:698–704.
4. Howes OD, Kapur S. The dopamine hypothesis of schizophrenia: version III--the final common pathway. *Schizophr Bull*. 2009;35:549–562.

5. Abi-Dargham A, Gil R, Krystal J, et al. Increased striatal dopamine transmission in schizophrenia: confirmation in a second cohort. *Am J Psychiatry*. 1998;155:761–767.

6. Nordstrom AL, Farde L, Wiesel FA, et al. Central D2-dopamine receptor occupancy in relation to antipsychotic drug effects: a double-blind PET study of schizophrenic patients. *Biol Psychiatry*. 1993;33:227–235.

7. Gold JM, Waltz JA, Prentice KJ, Morris SE, Heerey EA. Reward processing in schizophrenia: a deficit in the representation of value. *Schizophr Bull*. 2008;34:835–847.

8. Gold JM, Waltz JA, Matveeva TM, et al. Negative symptoms and the failure to represent the expected reward value of actions: behavioral and computational modeling evidence. *Arch Gen Psychiatry*. 2012;69:129–138.

9. Strauss GP, Frank MJ, Waltz JA, Kasanova Z, Herbener ES, Gold JM. Deficits in positive reinforcement learning and uncertainty-driven exploration are associated with distinct aspects of negative symptoms in schizophrenia. *Biol Psychiatry*. 2011;69:424–431.

10. Waltz JA, Frank MJ, Robinson BM, Gold JM. Selective reinforcement learning deficits in schizophrenia support predictions from computational models of striatal-cortical dysfunction. *Biol Psychiatry*. 2007;62:756–764.

11. Waltz JA, Frank MJ, Wiecki TV, Gold JM. Altered probabilistic learning and response biases in schizophrenia: behavioral evidence and neurocomputational modeling. *Neuropsychology*. 2011;25:86–97.

12. Kring AM, Barch DM. The motivation and pleasure dimension of negative symptoms: neural substrates and behavioral outputs. *Eur Neuropsychopharmacol*. 2014;24:725–736.

13. Kapur S. Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. *Am J Psychiatry*. 2003;160:13–23.

14. Horga G, Schatz KC, Abi-Dargham A, Peterson BS. Deficits in predictive coding underlie hallucinations in schizophrenia. *J Neurosci*. 2014;34:8072–8082.

15. Daw ND, Shohamy D. The cognitive neuroscience of motivation and learning. *Social Cognition*. 2008;26:593–620.

16. Schönberg T, Daw ND, Joel D, O'Doherty JP. Reinforcement learning signals in the human striatum distinguish learners from nonlearners during reward-based decision making. *J Neurosci*. 2007;27:12860–12867.

17. Bartra O, McGuire JT, Kable JW. The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage*. 2013;76:412–427.

18. Clithero JA, Rangel A. Informatic parcellation of the network involved in the computation of subjective value. *Soc Cogn Affect Neurosci*. 2014;9:1289–1302.

19. Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ. Cortical substrates for exploratory decisions in humans. *Nature*. 2006;441:876–879.

20. Hare TA, O'Doherty J, Camerer CF, Schultz W, Rangel A. Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *J Neurosci*. 2008;28:5623–5630.

21. Rutledge RB, Dean M, Caplin A, Glimcher PW. Testing the reward prediction error hypothesis with an axiomatic model. *J Neurosci*. 2010;30:13525–13536.

22. Murray GK, Corlett PR, Clark L, et al. Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. *Mol Psychiatry*. 2008;13:239, 267–276.

23. Waltz JA, Schweitzer JB, Gold JM, et al. Patients with schizophrenia have a reduced neural response to both unpredictable and predictable primary reinforcers. *Neuropsychopharmacology*. 2009;34:1567–1577.

24. Waltz JA, Schweitzer JB, Ross TJ, et al. Abnormal responses to monetary outcomes in cortex, but not in the basal ganglia, in schizophrenia. *Neuropsychopharmacology*. 2010;35:2427–2439.

25. Gradin VB, Kumar P, Waiter G, et al. Expected value and prediction error abnormalities in depression and schizophrenia. *Brain*. 2011;134:1751–1764.

26. Schlagenhauf F, Huys QJ, Deserno L, et al. Striatal dysfunction during reversal learning in unmedicated schizophrenia patients. *Neuroimage*. 2014;89:171–180.

27. Insel C, Reinen J, Weber J, et al. Antipsychotic dose modulates behavioral and neural responses to feedback during reinforcement learning in schizophrenia. *Cogn Affect Behav Neurosci*. 2014;14:189–201.

28. Reinen J, Smith EE, Insel C, et al. Patients with schizophrenia are impaired when learning in the context of pursuing rewards. *Schizophr Res*. 2014;152:309–310.

29. Frank MJ, Seeberger LC, O'Reilly RC. By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science*. 2004;306:1940–1943.

30. Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*. 2006;442:1042–1045.

31. Rutledge RB, Lazzaro SC, Lau B, Myers CE, Gluck MA, Glimcher PW. Dopaminergic drugs modulate learning rates and perseveration in Parkinson's patients in a dynamic foraging task. *J Neurosci*. 2009;29:15104–15114.

32. Centonze D, Usiello A, Costa C, et al. Chronic haloperidol promotes corticostriatal long-term potentiation by targeting dopamine D2L receptors. *J Neurosci*. 2004;24:8214–8222.

33. Wiecki TV, Riedinger K, von Ameln-Mayerhofer A, Schmidt WJ, Frank MJ. A neurocomputational account of catalepsy sensitization induced by D2 receptor blockade in rats: context dependency, extinction, and renewal. *Psychopharmacology (Berl)*. 2009;204:265–277.

34. Jocham G, Klein TA, Ullsperger M. Differential modulation of reinforcement learning by D2 dopamine and NMDA glutamate receptor antagonism. *J Neurosci*. 2014;34:13151–13162.

35. Juckel G, Schlagenhauf F, Koslowski M, et al. Dysfunction of ventral striatal reward prediction in schizophrenia. *Neuroimage*. 2006;29:409–416.

36. Nielsen MO, Rostrup E, Wulff S, et al. Alterations of the brain reward system in antipsychotic naive schizophrenia patients. *Biol Psychiatry*. 2012;71:898–905.

37. Schlagenhauf F, Sterzer P, Schmack K, et al. Reward feedback alterations in unmedicated schizophrenia patients: relevance for delusions. *Biol Psychiatry*. 2009;65:1032–1039.

38. Watanabe M, Sakagami M. Integration of cognitive and motivational context information in the primate prefrontal cortex. *Cereb Cortex*. 2007;17:i101–i109.

39. Juckel G, Schlagenhauf F, Koslowski M, et al. Dysfunction of ventral striatal reward prediction in schizophrenic patients treated with typical, not atypical, neuroleptics. *Psychopharmacology (Berl)*. 2006;187:222–228.

40. Waltz JA, Gold JM. Probabilistic reversal learning impairments in schizophrenia: further evidence of orbitofrontal dysfunction. *Schizophr Res*. 2007;93:296–303.

41. Nurnberger JI Jr, Blehar MC, Kaufmann CA, et al. Diagnostic interview for genetic studies. Rationale, unique

features, and training. NIMH Genetics Initiative. *Arch Gen Psychiatry*. 1994;51:849–859; discussion 863-844.

42. Van Snellenberg JX, Slifstein M, Read C, et al. Dynamic shifts in brain network activation during supracapacity working memory task performance. *Human Brain Mapp*. 2015;36:1245–1264.

43. Sutton RS, Barto AG. *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT; 1998.

44. Daw ND. Trial-by-trial data analysis using computational models. In: Delgado MR, Phelps EA, Robbins TW, eds. *Decision Making, Affect, and Learning*. New York: Oxford University Press; 2011: 3–38.

45. Otto AR, Raio CM, Chiang A, Phelps EA, Daw ND. Working-memory capacity protects model-based learning from stress. *Proc Natl Acad Sci U S A*. 2013;110:20941–20946.

46. Sharp ME, Foerde K, Daw ND, Shohamy D. Dopamine selectively remediates 'model-based' reward learning: a computational approach. *Brain*. 2016;139:355–364.

47. Gelman A, Carlin JB, Stern HS, Rubin DB. *Bayesian Data Analysis*. 2nd ed. London, UK: Chapman and Hall; 2003.

48. Ward BD. Simultaneous inference for fMRI data. Milwaukee, WI: Biophysics Research Institute, Medical College of Wisconsin; 2000.

49. Van Snellenberg JX, Conway AR, Spicer J, Read C, Smith EE. Capacity estimates in working memory: reliability and interrelationships among tasks. *Cogn Affect Behav Neurosci*. 2014;14:106–116.

50. Mir P, Trender-Gerhard I, Edwards MJ, Schneider SA, Bhatia KP, Jahanshahi M. Motivation and movement: the effect of monetary incentive on performance speed. *Exp Brain Res*. 2011;209:551–559.

51. Abler B, Erk S, Walter H. Human reward system activation is modulated by a single dose of olanzapine in healthy subjects in an event-related, double-blind, placebo-controlled fMRI study. *Psychopharmacology (Berl)*. 2007;191:823–833.

52. Farde L, Nordstrom AL, Wiesel FA, Pauli S, Halldin C, Sedvall G. Positron emission tomography analysis of central D1 and D2 dopamine receptor occupancy in patients treated with classical neuroleptics and clozapine. Relation to extrapyramidal side effects. *Arch Gen Psychiatry*. 1992;49:538–544.

53. Kirsch P, Ronshausen S, Mier D, Gallhofer B. The influence of antipsychotic treatment on brain reward system reactivity in schizophrenia patients. *Pharmacopsychiatry*. 2007;40:196–198.

54. Schlagenhauf F, Juckel G, Koslowski M, et al. Reward system activation in schizophrenic patients switched from typical neuroleptics to olanzapine. *Psychopharmacology (Berl)*. 2008;196:673–684.

55. Danna CL, Elmer GI. Disruption of conditioned reward association by typical and atypical antipsychotics. *Pharmacol Biochem Behav*. 2010;96:40–47.

56. Howe MW, Tierney PL, Sandberg SG, Phillips PE, Graybiel AM. Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. *Nature*. 2013;500:575–579.

57. Niv Y. Cost, benefit, tonic, phasic: what do response rates tell us about dopamine and motivation? *Ann N Y Acad Sci*. 2007;1104:357–376.

58. Garrison J, Erdeniz B, Done J. Prediction error in reinforcement learning: a meta-analysis of neuroimaging studies. *Neurosci Biobehav Rev*. 2013;37:1297–1310.

59. Adcock RA, Thangavel A, Whitfield-Gabrieli S, Knutson B, Gabrieli JD. Reward-motivated learning: mesolimbic activation precedes memory formation. *Neuron*. 2006;50:507–517.

60. Robinson OJ, Overstreet C, Charney DR, Vytal K, Grillon C. Stress increases aversive prediction error signal in the ventral striatum. *Proc Natl Acad Sci U S A*. 2013;110:4129–4133.

61. Carter RM, Macinnes JJ, Huettel SA, Adcock RA. Activation in the VTA and nucleus accumbens increases in anticipation of both gains and losses. *Front Behav Neurosci*. 2009;3:21.

62. Delgado MR, Li J, Schiller D, Phelps EA. The role of the striatum in aversive learning and aversive prediction errors. *Philos Trans R Soc Lond B Biol Sci*. 2008;363:3787–3800.

63. Delgado MR, Jou RL, Phelps EA. Neural systems underlying aversive conditioning in humans with primary and secondary reinforcers. *Front Neurosci*. 2011;5:71.

64. Niznikiewicz MA, Delgado MR. Two sides of the same coin: learning via positive and negative reinforcers in the human striatum. *Dev Cogn Neurosci*. 2011;1:494–505.

65. Seymour B, Daw N, Dayan P, Singer T, Dolan R. Differential encoding of losses and gains in the human striatum. *J Neurosci*. 2007;27:4826–4831.

66. Robinson OJ, Frank MJ, Sahakian BJ, Cools R. Dissociable responses to punishment in distinct striatal regions during reversal learning. *Neuroimage*. 2010;51:1459–1467.

67. Delgado MR, Jou RL, Ledoux JE, Phelps EA. Avoiding negative outcomes: tracking the mechanisms of avoidance learning in humans during fear conditioning. *Front Behav Neurosci*. 2009;3:33.

68. Leotti LA, Delgado MR. The value of exercising control over monetary gains and losses. *Psychol Sci*. 2014;25:596–604.

69. Wrase J, Kahnt T, Schlagenhauf F, et al. Different neural systems adjust motor behavior in response to reward and punishment. *Neuroimage*. 2007;36:1253–1262.

70. Litt A, Plassmann H, Shiv B, Rangel A. Dissociating valuation and saliency signals during decision-making. *Cereb Cortex*. 2011;21:95–102.

71. Bernacer J, Corlett PR, Ramachandra P, et al. Methamphetamine-induced disruption of frontostriatal reward learning signals: relation to psychotic symptoms. *Am J Psychiatry*. 2013;170:1326–1334.

72. Shohamy D, Adcock RA. Dopamine and adaptive memory. *Trend Cogn Sci*. 2010;14:464–472.

73. Wimmer GE, Shohamy D. Preference by association: how memory mechanisms in the hippocampus bias decisions. *Science*. 2012;338:270–273.

74. Shohamy D, Mihalakos P, Chin R, Thomas B, Wagner AD, Tamminga C. Learning and generalization in schizophrenia: effects of disease and antipsychotic drug treatment. *Biol Psychiatry*. 2010;67:926–932.

75. Lodge DJ, Grace AA. Hippocampal dysregulation of dopamine system function and the pathophysiology of schizophrenia. *Trend Pharmacol Sci*. 2011;32:507–513.