



Published in final edited form as:

Neuropsychology. 2011 January ; 25(1): 86–97. doi:10.1037/a0020882.

Altered probabilistic learning and response biases in schizophrenia: Behavioral evidence and neurocomputational modeling

J. A. Waltz¹, M. J. Frank^{2,3,4,5}, T. V. Wiecki², and J. M. Gold¹

¹Dept of Psychiatry, University of Maryland School of Medicine

²Dept of Cognitive & Linguistic Sciences, Brown University

³Dept of Psychology, Brown University

⁴Dept of Psychiatry and Human Behavior, Brown University

⁵Brown Institute for Brain Science, Brown University

Abstract

Objective—Patients with schizophrenia (SZ) show reinforcement learning impairments related to both the gradual/procedural acquisition of reward contingencies, and the ability to use trial-to-trial feedback to make rapid behavioral adjustments.

Method—We used neurocomputational modeling to develop plausible mechanistic hypotheses explaining reinforcement learning impairments in SZ. We tested the model with a novel Go/NoGo learning task in which subjects had to learn to respond or withhold responses when presented with different stimuli associated with different probabilities of gains or losses in points. We analyzed data from 34 patients and 23 matched controls, characterizing positive-and negative-feedback-driven learning in both a Training Phase and a Test Phase.

Results—Consistent with simulations from a computational model of aberrant dopamine input to the basal ganglia patients, SZs showed an overall increased rate of responding in the Training Phase, together with reduced response-time acceleration to frequently-rewarded stimuli across training blocks, and a reduced relative preference for frequently-rewarded training stimuli in the Test Phase. Patients did not differ from controls on measures of procedural negative-feedback-driven learning, although SZs exhibited deficits in trial-to-trial adjustments to negative feedback, with these measures correlating with negative symptom severity.

Conclusions—These findings support the hypothesis that SZ patients have a deficit in procedural “Go” learning, linked to abnormalities in DA transmission at D1-type receptors, despite a “Go bias” (increased response rate), potentially related to excessive tonic dopamine. Deficits in trial-to-trial reinforcement learning were limited to a subset of SZ patients with severe negative symptoms, putatively stemming from prefrontal cortical dysfunction.

Address for Correspondence: James A. Waltz, PhD University of Maryland School of Medicine Maryland Psychiatric Research Center P.O. Box 21247 Baltimore, MD 21228 jwaltz@mprc.umaryland.edu 410-402-6044 (Phone) 410-402-7198 (Fax) <http://www.mprc.umaryland.edu/>.

Publisher's Disclaimer: The following manuscript is the final accepted manuscript. It has not been subjected to the final copyediting, fact-checking, and proofreading required for formal publication. It is not the definitive, publisher-authenticated version. The American Psychological Association and its Council of Editors disclaim any responsibility or liabilities for errors or omissions of this manuscript version, any version derived from this manuscript by NIH, or other third parties. The published version is available at www.apa.org/pubs/journals/neu

Keywords

schizophrenia; dopamine; basal ganglia; reinforcement learning; procedural learning

Deficits in reinforcement-driven learning have been frequently observed in patients with schizophrenia (SZ; Malenka et al., 1982; Rushe et al., 1999). This has been especially true for learning tasks in which explicit hypotheses are tested and evaluated on a trial-to-trial basis as on the Wisconsin Card Sort Test (Prentice et al., 2008) and Conditional Associative Learning paradigms (Gold et al., 2000; Kemali et al., 1987). Impairments on these tasks are typically interpreted to suggest prefrontal cortical (PFC) dysfunction. The results of studies examining less explicit (e.g., procedural) forms of reinforcement learning in SZ patients, by contrast, have been mixed (see Gold et al., 2008, for a review). Patients with schizophrenia have shown intact performance on a variety of paradigms thought to rely primarily on implicit learning mechanisms, including serial reaction time tasks (Foerde et al., 2008), probabilistic classification learning tasks (Keri et al., 2000; Weickert et al., 2002), and artificial grammar learning tasks (Danion et al., 2001; Horan et al., 2008), although examples of impaired performance exist for some of these tasks, as well (Foerde et al., 2008; Horan et al., 2008; Schwartz et al., 2003). Thus, it appears likely that differences in clinical features among studied cohorts, as well as the cognitive demands of specific tasks, may have an impact on observed results, making it difficult to draw more general inferences from this body of work.

Our own previous work (Waltz et al., 2007) suggests possible performance dissociations between 1) tasks relying primarily on *positive*-feedback-driven procedural learning mechanisms and those relying primarily on *negative*-feedback-driven procedural learning mechanisms, and 2) tasks of negative-feedback-driven learning, primarily dependent on *procedural* mechanisms vs. those primarily reliant on *explicit/declarative* mechanisms (e.g., a shift to a new deterministic rule when the previous one is no longer appropriate). In short, our previous results argue against a *general* sparing of procedural learning capacities, but nevertheless suggest that some mechanisms are relatively preserved and may be able to compensate, to some extent, for those that are disrupted. Feedback-driven learning of procedures and habits depends on intact function of the basal ganglia (BG; Frank and Claus, 2006; Graybiel, 2008; Knowlton et al., 1996; Tricomi et al., 2009). Given the evidence that schizophrenia involves BG DA dysfunction, we have argued that neurocomputational models of BG function which have been developed refined and tested to account for learning as a function of BG DA manipulations, may fruitfully contribute to a more differentiated account of the relative preservation and disruption of procedural learning mechanisms in schizophrenia.

Predictions from Computational Modeling

Simulation models (Frank, 2005; Frank and Claus, 2006; Wiecki et al., 2009) have been the sources of multiple specific hypotheses about the functional consequences of particular aspects of dopamine modulation of striato-cortical circuits. In brief, phasic (transient) dopamine signals that occur during positive and negative prediction errors (Schultz et al., 1997) are required to drive changes in synaptic plasticity via D1 and D2 receptors in “Go” and “NoGo” neuronal populations, respectively (Frank, 2005; Frank et al., 2004). The “Go” pathway is thought to be critical for learning actions that are associated with rewarding, positive outcomes, whereas the “NoGo” pathway is critical for learning to avoid actions that are associated with negative outcomes. Furthermore, tonic DA levels also modulate relative activity states in these cells, with higher levels favoring activity in the Go pathway over the

NoGo pathway during response selection, affecting the speed with which responses are executed (Wiecki et al., 2009).

Based on evidence that schizophrenia involves high tonic DA levels in the BG (Abi-Dargham et al., 2000; Laruelle and Abi-Dargham, 1999), we predicted that patients should show an overall “Go bias” in the context of reinforcement learning tasks. Such a “Go bias” would be evidenced by an overall tendency to make rather than withhold motor responses even when it is disadvantageous to do so. We also predicted that patients would exhibit a *Go learning deficit*, based on the hypothesis that excessive DA tone would be associated with reduced fidelity of phasic increases, together with evidence for reduced D1-receptor transmission in SZ (Abi-Dargham et al., 2002; Abi-Dargham and Moore, 2003; Weinberger, 1987). That is, we hypothesized that a reduced ability to interpret phasic DA bursts, against the background of high DA tone, would result in a comprised ability to *learn* from positive reinforcement, and diminished tendency to *selectively* make appropriate Go responses to positive stimuli, despite an overall increased tendency to make Go responses (Go bias).

Behavioral Tests of Model Predictions

Preliminary evidence of a Go learning deficit in SZ comes from the results of a recent study by our group (Waltz et al., 2007), in which we showed that patients with schizophrenia exhibit impairment when procedural (probabilistic) learning is driven by positive feedback, but normal performance when procedural learning is driven by negative feedback. The task used in that study, however, required subjects to choose a stimulus on every trial. Thus, we were unable to test the hypothesis that SZ patients have a “Go bias” – an overall bias to respond.

In order to address both of the model predictions above, we administered a novel probabilistic “Go/NoGo” task (Frank and O’Reilly, 2006) to patients with schizophrenia and controls. This task required subjects to learn about the reinforcement properties of stimulus choices by button-pressing (“Go” responding). For some stimuli, Go responses were rewarded most of the time with points, whereas, for other stimuli, responses were punished most of the time with point-deductions. Non-responses were neither rewarded nor punished. By integrating reinforcement associated with button-presses to the different stimuli, subjects could learn which stimuli to respond to receive a reward, and which ones to avoid to responding in order to avoid losses.

Gradual “Go” learning could be assessed both in a Training Phase (by measuring changes in Go response times across blocks, predicted to speed up for the most reinforced stimuli; Moustafa et al, 2008) and in a Test/Transfer Phase administered following training (by measuring the tendency to *selectively* boost responding to the most positively-reinforced stimuli). Gradual “NoGo” learning could be assessed both in the Training Phase (by measuring changes in false-alarm rates across blocks) and in the Test/Transfer Phase, following training (by measuring the tendency to selectively withhold responses to punished stimuli). Further, this paradigm enabled us to quantify the general tendencies (“biases”) of subjects to respond (Go) and to withhold responses (NoGo) to familiar and novel stimuli in the Test/Transfer Phase.

Importantly, we were also able to assess rapid reinforcement learning using this paradigm, by quantifying learning at the beginning of the Training Phase, and by characterizing trial-by-trial adjustments in behavior. Based on previous findings from our group (Waltz et al., 2007; Waltz and Gold, 2007), we predicted that patients with schizophrenia would show deficits in *rapid* early learning of reinforcement contingencies (i.e. by hypothesis testing and working memory, presumably dependent on PFC function), even when guided by negative

feedback, despite a relatively intact ability to use negative feedback to *gradually* acquire stimulus-response contingencies.

Method

Patients

Thirty-seven outpatients with a diagnosis of schizophrenia, based on the Structured Clinical Interview for DSM-IV (SCID-I; First et al., 1997), were recruited from the Maryland Psychiatric Research Center (MPRC; Table 1). Data from three patients who did not appear to understand the task (and thus rarely withheld responses) were removed from the analysis data set. All patients were clinically stable, as determined by their treating clinician. All patients were tested while receiving stable medication regimens (no changes in type or dose within 4 weeks of study). Almost half of patients (16/34) were taking one of the second-generation antipsychotics as their only antipsychotic medication (7 on clozapine, 5 on risperidone, 3 on olanzapine, and 1 on aripiprazole). Seven patients were on first-generation antipsychotic monotherapy (4 on haloperidol, 3 on fluphenazine). Eleven patients were taking two antipsychotics (almost all clozapine with risperidone).

Control subjects

Twenty-five healthy control subjects consented to participate in the study. Data were discarded from two controls who did not appear to understand the task, leaving 23 control subjects in the analysis data set. They were recruited through a combination of newspaper advertisements and random phone number dialing and were extensively screened for Axis I and II disorders using the SCID-I (First et al., 1997) and the Structured Interview for DSM-III-R Personality Disorders (SIDP-R; Pfohl et al., 1989). Subjects were also screened for family history of psychosis and medical conditions that might impact cognitive performance, including drug use. All control subjects were free of any significant personal psychiatric and medical history, had no history of severe mental illness in first-degree relatives, and did not meet criteria for current substance abuse or dependence.

General Procedures

After explanation of study procedures, all subjects provided written informed consent for a protocol approved by the University of Maryland School of Medicine Internal Review Board. Before signing consent documents, patients had to demonstrate adequate understanding of study demands, risks, and means of withdrawing from participation in response to structured probe questions. All subjects were compensated for study participation.

In addition, we also administered a brief battery of standard neuropsychological tests for purposes of sample description and correlational analyses. Tests included measures of word reading (the Wechsler Test of Adult Reading, or WTAR; Wechsler, 2001), word list learning (Hopkins Verbal Learning Test-Revised; Brandt and Benedict, 2001), and working memory (Letter-number Span and Spatial Span; Gold et al., 1997; Wechsler, 1997).

Patients were also characterized using the Brief Psychiatric Ratings Scale (BPRS; Overall and Gorman, 1962), the Scales for the Assessment of Negative Symptoms (SANS; Andreasen, 1984), and the Calgary Depression Scale (CDS; Addington et al., 1992). The symptom and functioning ratings were conducted by masters and doctoral level clinicians. Intraclass correlation coefficients (ICCs) for these instruments ranged from 0.76 to 0.90.

Experimental Task

We used a computerized probabilistic reinforcement Go/NoGo paradigm, in which stimuli were presented one at a time and the participant had to either press a key (Go) or withhold their response (NoGo). During the Training Phase, six different patterns were presented in random order, associated with reinforcement probabilities of 90%, 80%, 70%, 30%, 20%, and 10% for button presses (Figure 1A). Stimuli were presented for 2 s, and responses were accepted for the duration of presentation. Subjects were told that some stimulus patterns would lead to point gains if selected (always 1 point), while others would cause them to lose a point, and that their goal should be to maximize point totals. After each bar press response, visual feedback was provided for 1 s (“You won a point!” written in blue or “You lost a point” written in red). No feedback was provided if subjects chose not to respond to a particular stimulus. The interval between trials was 1 s. Training trials were divided into 3 blocks of 60 trials each, with each stimulus being presented 30 times (10 presentations/block). Over time, participants learned that three of the stimuli should be associated with a button press (because their corresponding probabilities of reinforcement are greater than 50%), but that responses made to the other three will likely make them lose points.

A post-training Test/Transfer session (Figure 1B) followed the three training blocks. Subjects were told that “during this set of trials [they] will NOT receive feedback (“correct” or “incorrect”) to [their] responses” and that they would “not know [their] point totals during this phase” and therefore “try to use what [they] learned before to get the most points possible.” Subjects were also told that “besides the patterns [they] saw before, [they] may see new combinations of patterns in the test.” In these new combinations of patterns, the left and right halves of the combined pattern each represented one of the training patterns. For example, half of the composite pattern may have consisted of a familiar pattern that was 80% correct, while the other half consisted of one that was 80% incorrect, so that the combined pattern should have been equally associated with “Go” and “NoGo”. Such patterns had an expected value of zero, and thus were termed “neutral” stimuli. In other cases one of the patterns was more strongly associated with a certain outcome (i.e., 90% reinforced, combined with 70% unreinforced), and thus termed “Novel Positive” or “Novel Negative”. Stimuli were present on the screen until subjects made a response. In this phase, subjects saw 69 total trials: each of the six single patterns from the Training Phase were presented six times (36 total trials), and each of the eleven novel combined patterns were presented three times (33 total trials). Thus, 18 of the Test trials (involving patterns A, B, and C) were termed “Familiar Positive”, 12 were termed “Novel Positive”, nine were termed “Novel Neutral”, 12 were termed “Novel Negative”, and 18 (D, E, and F) were termed “Familiar Negative”.

Statistical Analysis

To characterize Go-responding in the Training Phase, we performed a 3-way analysis of variance (ANOVA; mixed model) for accuracy rates, with factors of group (2 levels), training block (3 levels), and valence (2 levels: Go/Positive and NoGo/Negative). Accuracy rates were computed as Go responses to frequently-reinforced items (A, B, and C) and NoGo responses (withheld responses) to frequently-punished items (D, E, and F). We also performed a two-way ANOVA for response times to positive stimuli, with factors of group and training block (3 levels). We calculated mean response times from the onset of the stimulus until the time of response. We did not analyze response times to negative stimuli, because many subjects made no, or very few, Go responses to negative stimuli by the third block, reflecting successful acquisition. To assess general response biases, we performed a *t*-test to compare mean Go-response rates between groups in the Training Phase, independent of stimulus condition.

Because we had evidence from previous studies (Prentice et al., 2008; Waltz et al., 2007), as well as the present study, that patients and controls show differences in rapid acquisition early in a session, we computed “win-stay” and “lose-shift” scores for each reinforcement condition during the Training Phase. “Win-stay” and “lose-shift” scores served as measures of rapid, trial-to-trial, learning, in that they characterized the tendency of subjects to respond immediately to feedback, rather than make choices based on the expected value of a stimulus, integrated over the course of *many* trials (this was assessed through changes in accuracy or RTs over the course of blocks). We computed “win-stay” scores by computing the proportion of positive feedback instances from valid trials (in which an appropriate Go response was reinforced) that were followed by another button press to the same stimulus when it was next encountered. We computed “lose-shift” scores by computing the proportion of negative feedback instances from valid trials (in which inappropriate Go response were punished) that were followed by the withholding of a response to the same stimulus when it was next encountered. We then generated total “win-stay” and “lose-shift” scores by averaging scores across stimulus conditions for each measure. Between-group differences in mean scores were then assessed using t-tests. Effect sizes were also computed (Cohen's D) and presented as supplementary data.

To determine whether participant groups differed in the gradual integration of probabilistic Go- or NoGo-learning signals across trials, we also used measures from the Test/Transfer Phase, which was designed to assess learning across the entire Training Phase (Frank and O'Reilly, 2006). Because subjects received no feedback in the Test/Transfer Phase, no rapid, trial-to-trial learning could occur in this phase. To assess subjects' tendencies to selectively boost Go responses to positively reinforced stimuli, and to selectively withhold responses to negative stimuli, we computed Go-response rates to positive/negative stimuli relative to Go response rates to the neutral stimuli (which serve as a baseline; see Figure 1B).

We used Spearman correlation analyses to assess relationships between Go/NoGo task performance and three types of characterizing variables: symptom ratings, standard neuropsychological measures, and antipsychotic medication doses (converted to haloperidol equivalent units; see Supplementary Table 1). We used four measures of Go/NoGo task performance in our correlation analyses, all of which showed group differences: the correct-reject and lose-shift rates from Training Block 1, the change in the average RT to positive stimuli from Block 1 to Block 3 of the Training Phase, and the [Familiar Positive – Novel Neutral] Go-response-rate contrast from the Test/Transfer Phase.

To separately assess psychotic and disorganized symptoms from the BPRS, sub-scores were grouped into reality distortion, disorganization, negative symptom, and anxiety/depression clusters based on the four-factor model of McMahon et al. (2002)

Computational Modeling

In brief, the model used here consists of two opposing pathways from striatum to basal ganglia output nuclei, through thalamus, and up to cortex. A direct Go-pathway facilitates execution of a cortical response, whereas an indirect NoGo-pathway suppresses competing responses. These two pathways originate in the striatum which consists of two medium spiny neuronal populations oppositely modulated by dopaminergic neurons in the Substantia Nigra pars compacta (SNc), together with GABAergic interneurons. Dopamine bursts drive Go learning in the direct pathway (via D1 receptors), promoting the selection of actions that lead to reward. Phasic dopamine dips drive NoGo learning in the indirect pathway (via D2 receptors), such that actions that lead to negative outcomes are more likely to be avoided. This same model has been applied to multiple datasets across species, tasks, and manipulations. A more detailed description of the model and empirical support for it can be found elsewhere (Cohen and Frank, 2009; Frank, 2006).

Task setup (stimulus-response-reward contingencies; training and Test/Transfer Phase with re-combined stimuli) was identical to the behavioral experiment. However, instead of ten stimulus repetitions in one block, we trained our networks with 30 repetitions in one block. (The reason for this change is that the network model's learning rate is set rather conservatively and thus needs more training to achieve a similar level of overall performance, particularly given that the model used here lacks PFC mechanisms that would support rapid trial-to-trial learning (but see Frank et al., 2004). We chose to focus on the BG-mediated learning mechanisms because this same model has been applied to a range of reinforcement learning and decision making tasks as a function of DA manipulation (Frank, 2005; Frank et al., 2004; Moustafa et al., 2008; Pizzagalli et al., 2008; Santesso et al., 2009; Wiecki et al., 2009), with the same parameters used here.

In accordance with the dopamine hypothesis of SZ and empirical data (Abi-Dargham et al., 1998; Laruelle and Abi-Dargham, 1999; Meyer-Lindenberg et al., 2002), we simulated SZ in our model by increasing tonic levels of DA by 40%, accompanied by a reduction of phasic burst activity by 25% following rewards (simulating the effects of presynaptic autoreceptor regulation of DA bursts). The dip in DA during negative feedback (change from tonic levels) was kept the same as the intact case. A total of 80 intact and 80 SZ networks with random initial synaptic weights were trained and tested in an identical fashion as in the behavioral experiment.

Results

Measures of Gradual/Procedural Learning from the Training Phase

As illustrated in Figures 2A and 2B, both patients and controls learned to withhold responses to (correctly reject) frequently-punished stimuli across the three blocks of 60 trials each, although patients with schizophrenia exhibited an overall Go bias, as indicated by a higher overall rate of Go responses ($67.4 \pm 12.5\%$ vs. $59.5 \pm 14.8\%$) and a lower overall rate of correct rejections ($45.7 \pm 17.8\%$ vs. $60.0 \pm 22.0\%$). These effects were confirmed by ANOVAs showing main effects of block [$F(2,54)=29.33$, $p<0.001$], group [$F(1,55)=4.76$, $p=0.033$], and valence [positive vs. negative stimuli; $F(2,54)=136.28$, $p<0.001$].

We also observed a significant group \times valence interaction [$F(1,55)=5.20$, $p=0.026$], due to the presence of a group difference in accuracy for negative stimuli [measured by correct rejection rates across the whole session, as mentioned above; $t(55) = 2.70$, $p=0.009$], but not for positive stimuli [SZ mean= $80.5 \pm 12.7\%$, NC mean= $78.9 \pm 15.5\%$; $t(55) = 0.43$]. Furthermore, we observed a significant block \times valence interaction [$F(2,54)=28.68$, $p<0.001$], as accuracy rates were modulated by block number for negative stimuli, but not for positive stimuli.

We did not, however, observe a significant group \times block interaction [$F(2,54)=1.35$], or a significant group \times block \times valence interaction [$F(2,54)=2.05$], which would point to group differences in learning rate. Although patients and controls differed in their *overall* correct rejection rates, and most dramatically in their correct rejection rates in the first block of trials [$t(55) = 3.60$, $p=0.001$; Figure 2A], the groups did not differ in their correct rejection rates in the final training block [$t(55)=1.397$; $p>0.10$; see Supplementary Table 2 for effect sizes of group differences]. Thus, these findings support the notion that an initial Go bias, together with an impairment in *rapid* learning from negative feedback, led to deficits in withholding responses to negative stimuli in early trials. By contrast, patients were able to use negative feedback to learn *gradually* to withhold responses to the same degree as controls did, by the end of training.

As shown in Figures 2C and 2D, when we analyzed the *speed* of Go-responses in SZ patients and controls, the two groups showed differential rates of response time change across Training blocks, with controls showing greater reductions in RTs to frequently-rewarded stimuli than patients, from the first to the last training block. This impression was supported by the results of an ANOVA, which revealed a main effect of block [$F(2,54)=5.97$, $p=0.005$], and a trend toward a main effect of group [$F(1,55)=2.91$, $p=0.094$] on response time, qualified by a group \times block interaction [$F(2,54)=3.22$, $p=0.048$]. Given that previous findings and simulations suggest that progressive response speeding to reinforced stimuli depends on striatal DA/D1 dependent processes (Frank et al., 2009; Moustafa et al., 2008), the current observations point to a specific impairment in Go-learning in SZ patients.

Measures of Gradual/Procedural Learning from the Test/Transfer Phase

As illustrated in Figure 3A, both patients and controls showed strong modulations of Test/Transfer Go-response rates by the objective reinforcement value of the Test/Transfer stimuli. Figure 3B shows, however, that, controlling for baseline Go response rates to neutral stimuli, patients exhibited less of an increase in Go responding to positive training stimuli, with no differences in the ability to withhold responding to negative stimuli. That is, patients showed reduced selectivity in their Go responding, but normal selectivity in their NoGo responding. This impression was confirmed by the results of an ANOVA, which revealed a group \times trial-type interaction [$F(3,53)=2.77$, $p=0.05$] and a main effect of trial type [$F(3,53)=51.46$, $p<0.001$], but no significant main effect of group [$F(1,55)=0.116$]. Post-hoc, between-group t-tests for each trial-type, confirmed that the group \times trial-type interaction stemmed from a group difference in the tendency to increase Go responding to positive training stimuli [$t(55)=2.03$; $p=0.048$] and the lack of group differences for the other three trial types (all t 's < 1 ; see Supplementary Table 2 for effect sizes of group differences].

Simulation of Patient performance in the Test/Transfer Phase

The gradual learning of reinforcement values needed to resolve subtle probabilistic differences in stimulus-action outcomes, as observed here, is thought to depend on striatal dopaminergic mechanisms (in contrast to rapid trial-by-trial effects during acquisition; Frank and Claus, 2006; Frank et al., 2007). As such, we subjected the basal ganglia computational model to the same analysis, varying only striatal dopaminergic function to simulate SZ and determine whether this can account for the observed pattern of data (see Methods).

As can be seen in Figure 3C, both groups of networks exhibit a roughly linear relationship between Go response rates as a function of trained stimulus value, as in the behavioral data. Figure 3D further illustrates that SZ networks showed a reduced tendency to increase Go responses to positive relative to neutral stimuli. A between-group comparison reveals that this was true for both familiar positive stimuli [SZ networks: 28.0%, control networks: 40.4%; $t(158) = 6.21$, $p < 0.001$] and novel positive stimuli [6.7% vs. 9.7%; $t(158) = 2.21$, $p = 0.03$]. By contrast, SZ and control networks did not differ in their tendency to reduce response rates to negative stimuli, for either familiar [-11.7% vs. -10.8%; $t(158) = 1.71$, $p > 0.05$] or novel negative stimuli [-6.0% vs. -3.7%; $t(158) = 0.89$]. These findings resulted from a combination of two factors: (i) elevated tonic DA levels, leading to an overall “Go bias” (and therefore increased responding across the board, including to neutral and negative stimuli), and (ii) a reduction of phasic D1 signaling, leading to impaired Go learning. In contrast, the DA dip during negative outcomes was kept the same as the intact model (relative to tonic levels). Because learning in the model is a function of relative differences between Go/NoGo activity levels due to changes in DA levels (Frank, 2005), the degree of

NoGo learning was preserved. Thus, these simulations highlight that patterns of behavioral results may emerge from an underlying mechanism that may be counter-intuitive: although patients responded more to negative stimuli (if one does not correct for response rates to neutral stimuli), this could have resulted from a mechanism whereby NoGo learning was relatively preserved. Similarly, although patients responded similarly to controls for the most positive stimuli (without correcting for neutral response rates), the simulations show that the combined pattern of data are more likely to arise from a mechanism whereby Go learning to positive outcomes is impaired.

Behavioral Measures of Rapid/Explicit Learning from the Training Phase

Because we suspected that Block 1 differences in NoGo responding might reflect differences in very early learning rates, we assessed rapid learning on a trial-by-trial basis, by computing “win-stay” and “lose-shift” measures. These measures are dissociable from incremental probabilistic reinforcement integration, and are thought to depend on PFC more than the BG (Frank and Claus, 2006; Frank et al., 2007). If subjects are behaving adaptively, they should “stay with” a response that gets reinforced (wins). Responses that yield negative outcomes, however, should lead to “shifts” in response tendencies. As shown in Figure 4A, patients’ delayed NoGo learning corresponded to a greatly reduced tendency to “lose-shift” on NoGo trials, both in Block 1 [$t(55)=2.73$, $p < 0.01$] and throughout the Training Phase [$t(55)=2.88$, $p < 0.01$]. Controls shifted their responses 45% of the time when a response to a NoGo stimulus led to a point-deduction on a valid trial in the Training Phase, whereas patients shifted only 30% of the time. Otherwise stated, controls required only 2.2 instances of negative feedback to shift their response tendency, while patients required significantly more (3.3) instances of negative feedback. In contrast, patients did not show a reduced tendency to “win-stay” on Go trials, either in Block 1 [$t(55)=0.97$], or throughout the Training Phase [$t(55)=0.97$]. Both controls and patients stayed more than 90% of the time their button presses were reinforced on valid Go trials (Figure 4B). The above finding suggests that patients and controls in this study show dramatic differences in very early (putatively PFC-dependent) learning, and that these differences diminish over time, as patients eventually show rates of (putatively BG-dependent) learning similar to those of controls, largely driven by negative feedback.

Correlation Analyses

We performed Spearman correlation analyses to assess relationships between experimental measures of performance and clinical variables of interest (assessments of avolition and global negative symptoms, from the SANS, and assessments of positive and negative symptoms from the BPRS). As reported above, we observed group differences in correct-reject rates and lose-shift rates in Block 1, both findings supportive of a deficit in rapid reinforcement learning in schizophrenia. Further analyses revealed that both of these measures correlated significantly with both total score on the SANS, and the sum of scores on avolition items (Table 2). Both of these correlations were in the negative direction, indicating that these learning deficits are most evident in SZ patients with severe negative symptoms. We also observed group differences in two measures of gradual (positive-feedback-driven) learning: RT acceleration across blocks of positive Training trials and Go-response rates to familiar positive stimuli at Test/Transfer (hits), corrected for baseline Go-response rates. Further analyses revealed no systematic relationship between gradual Go-response latency shortening and measures of negative symptoms, again supporting the notion that this measure is BG DA-dependent, whereas negative symptoms reflect PFC dysfunction. Similarly, in the Test/Transfer phase, negative symptoms were not predictive of deficits in selective responding to positive stimuli (again thought to be BG DA dependent). Rather, we observed an unexpected *positive* correlation between these measures (i.e., preferences for positive training stimuli over neutral items were greatest in SZs with the

most severe negative symptoms; Table 2). We suspect that this is a spurious result as it goes in the opposite direction from all of the other significant correlations. No significant correlations were observed between neuropsychological measures and any of the experimental measures of performance showing group differences.

Discussion

Using a novel Go/NoGo learning paradigm (Frank and O'Reilly, 2006), we found evidence that SZ patients show differential disruption of complementary systems for reinforcement learning. Consistent with findings from our previous studies (Waltz et al., 2007; Waltz and Gold, 2007), patients showed severe deficits in the ability to use negative feedback to rapidly shift behavior on a trial-to-trial basis, but nonetheless gradually learned to withhold responses over the course of extended training. In contrast, patients showed impaired integration of positive feedback (reduced striatal Go learning in our model) across trials, as evidenced by selectively-reduced Go-responding to positive stimuli during the Test/Transfer Phase, as well as the absence of RT speeding to positive stimuli across training. We observed the same effects on procedural learning in simulations with an established neurocomputational model of BG function, which has been used to account for similar patterns of findings as a function of BG DA manipulation in other studies (for review; Cohen and Frank, 2009).

The reduced rate of appropriate *Go* responding at Test/Transfer suggests a weaker ability of positive reinforcement to drive responding over the long term, via direct BG pathway activation. This effect was not attributable, in either human or model performance, to a lower *overall* rate of Go responding: patients and controls showed similar overall rates of Go responding at Test/Transfer (and SZ patients actually showed a significantly elevated rate of Go responding during training, consistent with a “Go bias”).

We view our current observation in SZ patients of a deficit in gradual reward-driven learning, in the presence of intact gradual punishment-driven learning, as consistent with the results of the recent pharmacological challenge study from (Frank and O'Reilly, 2006); using the same Probabilistic Go/NoGo paradigm. That study showed reduced Go learning with the single low dose administration of a D2 receptor agonist (cabergoline) in healthy participants. Those findings were interpreted as reflecting reduced phasic DA transmission due to activation of presynaptic D2 autoreceptors (see also Santesso et al., 2009 for a similar result and interpretation with another D2 agonist, together with a simulation using the same BG model described here).

An attenuated reward anticipation signal could result from a disruption of dopaminergic mechanisms of reinforcement learning (McClure et al., 2003), which are thought to be derived from errors in reward prediction (Schultz et al., 1997). Several recent neuroimaging studies have, in fact, provided evidence for an attenuated positive reward prediction error signal in the neostriatum in schizophrenia (Koch et al., in press; Waltz et al., 2009), which could lead to a reduced impact of positive feedback on learning.

The elevated overall rates of Go responding shown by SZ patients may be consistent with evidence of excess tonic dopamine levels in the BG in schizophrenia (Abi-Dargham et al., 2000), which could also degrade the fidelity of phasic DA signals, often linked to learning (Schultz, 1998; Schultz et al., 1997). The plausibility of this account is further supported by our modeling results which show that elevated tonic DA activity accompanied by reduced phasic bursting activity in an established computational model produces similar impairments in reward integration as those observed in SZ patients.

The lack of a group difference in measures of procedural *NoGo* learning point to the relative sparing of the D2-driven network thought to support this ability (the indirect basal ganglia pathway). As has been suggested previously (Frank et al., 2004; Waltz et al., 2007), chronic administration of D2-*antagonists* may actually benefit procedural NoGo learning by causing increased sensitivity of D2 receptors in the striatum, which enhance indirect (NoGo) pathway activation and plasticity (Day et al., 2008; Shen et al., 2008). Indeed, simulations of D2 antagonism in the BG model lead to progressive increases in avoidance behavior across days, as seen in rats treated with haloperidol (Wiecki et al., 2009).

The clear impairment in *trial-to-trial* learning based on negative outcomes, on the other hand, likely reflects a limited capacity to use explicit representations of feedback to rapidly update value representations, a faculty thought to rely on intact function and ventral and medial aspects of PFC (Frank and Claus, 2006; Frank et al., 2007; Rolls et al., 2003; Schoenbaum and Roesch, 2005). Correlation analyses indicated that the ability to rapidly integrate feedback in the service of learning to inhibit responses related closely to negative symptoms, such as avolition. In our sample, patients with the most severe negative symptoms showed the greatest impairment in the ability to “lose-shift” – to avoid a punished stimulus at its next presentation. The observation of systematic relationships between negative symptom measures and measures of rapid reinforcement learning fits with our previous findings (Waltz et al., 2007; Waltz and Gold, 2007) and further support ideas that these two phenomena share a neural substrate in the PFC (Galderisi et al., 2008; Kirkpatrick and Buchanan, 1990; Vaiva et al., 2002).

Caveats and Limitations

Our model incorporates two critical formulations regarding dopamine system architecture and function: 1) the functional segregation of D1 (direct) and D2 (indirect) pathways in the BG, and 2) the asymmetric excitability of D2 relative to D1 cells in response to corticostriatal stimulation. We acknowledge the existence of studies that emphasize a degree of D1/D2 colocalization (notably Surmeier et al., 1996), as well as alternative formulations that both direct and indirect pathways become activated in response to DA depletion (Miller, 2008). In support of our theory and model, however, we cite recent evidence from BAC transgenic mice for predominant (if not complete) segregation D1 and D2 pathways in the BG (Gong et al., 2003; Surmeier et al., 2007; Valjent et al., 2009). In addition, recent evidence suggests that dopamine depletion (and also D2 antagonism) enhances the asymmetric excitability of D2 relative to D1 cells in response to corticostriatal stimulation (Day et al., 2008; Mallet et al., 2006), as well as LTP in striatopalliodal cells (Centonze et al., 2004; Hakansson et al., 2006; Shen et al., 2008), and thus promote “NoGo learning”. In humans, evidence is less direct, but genetic data suggest independence of learning from positive and negative outcomes, relying on DARPP-32 and DRD2 genes respectively (Frank and Hutchison, 2009; Frank et al., 2007), and DA drugs induce similar opposite effects on Go and NoGo learning (Bodi et al., 2009; Frank and O'Reilly, 2006; Frank et al., 2004; Moustafa et al., 2008).

Given that all of our patients were being administered therapeutic doses of antipsychotic medications, it is plausible that, by reducing DA transmission at D2 receptors (Seeman, 1987), DRD2 antagonists may have also impacted feedback-driven learning performance in SZ patients in *our* study. In order to determine if systematic associations existed between any of our experimental outcome measures and medication dose, we performed additional correlational analyses. These analyses revealed no significant correlations between behavioral measures and antipsychotic dose. Thus, our present results do *not* suggest that performance deficits in patients are due to the chronic administration of D2-blocking medications, perhaps due to compensatory brain changes thought to occur in the course of

long-term antipsychotic drug administration (Burt et al., 1977; Joyce, 2001; Seeman et al., 2005).

Clearly, this conclusion is limited by the fact that drug type and dose were not randomly assigned, and the validity of haloperidol dose conversions for second-generation antipsychotics is open to question. It is further worth noting that the high number of patients in our sample taking clozapine (53%) suggests possible resistance to treatment with D2-antagonists in these patients. However, this high percentage needs to be interpreted cautiously, given evidence that clozapine is under-utilized in most community settings (Kelly et al., 2007; Stroup et al., 2009). As an academic clinical center with a focus on treatment research, clozapine is far more likely to be tried at the MPRC than in most community settings that have less experience in the use of this compound. Thus, we suspect, but cannot prove, that the current study cohort is less treatment-resistant than they appear to be, given the frequency of clozapine use.

We acknowledge the importance of testing the hypothesis that patients with schizophrenia have a Go response bias, due to excessive dopamine tone, in the context of controlled clinical trials, or studies in medication-free patients. However, we regard the parcellation of reinforcement learning deficits in medicated schizophrenia patients as critical to the therapeutic enterprise, in that reinforcement learning deficits appear to relate closely to negative symptoms not typically remediated by antipsychotic drugs. Understanding reinforcement learning deficits in medicated patients is a first step in optimizing treatment and improving functional outcomes in the vast majority of patients.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

- Abi-Dargham A, Gil R, Krystal J, Baldwin RM, Seibyl JP, Bowers M, van Dyck CH, Charney DS, Innis RB, Laruelle M. Increased striatal dopamine transmission in schizophrenia: confirmation in a second cohort. *Am J Psychiatry* 1998;155:761–767. [PubMed: 9619147]
- Abi-Dargham A, Mawlawi O, Lombardo I, Gil R, Martinez D, Huang Y, Hwang DR, Keilp J, Kochan L, Van Heertum R, et al. Prefrontal dopamine D1 receptors and working memory in schizophrenia. *J Neurosci* 2002;22:3708–3719. [PubMed: 11978847]
- Abi-Dargham A, Moore H. Prefrontal DA transmission at D1 receptors and the pathology of schizophrenia. *Neuroscientist* 2003;9:404–416. [PubMed: 14580124]
- Abi-Dargham A, Rodenhiser J, Printz D, Zea-Ponce Y, Gil R, Kegeles LS, Weiss R, Cooper TB, Mann JJ, Van Heertum RL, et al. Increased baseline occupancy of D2 receptors by dopamine in schizophrenia. *Proc Natl Acad Sci U S A* 2000;97:8104–8109. [PubMed: 10884434]
- Addington D, Addington J, Maticka-Tyndale E, Joyce J. Reliability and validity of a depression rating scale for schizophrenics. *Schizophr Res* 1992;6:201–208. [PubMed: 1571313]
- Andreasen, NC. *The Scale for the Assessment of Negative Symptoms (SANS)*. University of Iowa; Iowa City, IA: 1984.
- Bodi N, Keri S, Nagy H, Moustafa A, Myers CE, Daw N, Dibo G, Takats A, Bereczki D, Gluck MA. Reward-learning and the novelty-seeking personality: a between- and within-subjects study of the effects of dopamine agonists on young Parkinson's patients. *Brain* 2009;132:2385–2395. [PubMed: 19416950]
- Brandt, J.; Benedict, RHB. *The Hopkins Verbal Learning Test-Revised*. Psychological Assessment Resources, Inc.; Odessa: FL: 2001.
- Burt DR, Creese I, Snyder SH. Antischizophrenic drugs: chronic treatment elevates dopamine receptor binding in brain. *Science* 1977;196:326–328. [PubMed: 847477]

- Centonze D, Usiello A, Costa C, Picconi B, Erbs E, Bernardi G, Borrelli E, Calabresi P. Chronic haloperidol promotes corticostriatal long-term potentiation by targeting dopamine D2L receptors. *J Neurosci* 2004;24:8214–8222. [PubMed: 15385604]
- Cohen MX, Frank MJ. Neurocomputational models of basal ganglia function in learning, memory and choice. *Behav Brain Res* 2009;199:141–156. [PubMed: 18950662]
- Danion JM, Meulemans T, Kauffmann-Muller F, Vermaat H. Intact implicit learning in schizophrenia. *Am J Psychiatry* 2001;158:944–948. [PubMed: 11384904]
- Day M, Wokosin D, Plotkin JL, Tian X, Surmeier DJ. Differential excitability and modulation of striatal medium spiny neuron dendrites. *J Neurosci* 2008;28:11603–11614. [PubMed: 18987196]
- First, MB.; Spitzer, RL.; Gibbon, M.; Williams, JBW. Structured Clinical Interview for DSM-IV- Axis I Disorders (SCID-I). American Psychiatric Press; Washington, DC: 1997.
- Foerde K, Poldrack RA, Khan BJ, Sabb FW, Bookheimer SY, Bilder RM, Guthrie D, Granholm E, Nuechterlein KH, Marder SR, Asarnow RF. Selective corticostriatal dysfunction in schizophrenia: examination of motor and cognitive skill learning. *Neuropsychology* 2008;22:100–109. [PubMed: 18211159]
- Frank MJ. Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. *J Cogn Neurosci* 2005;17:51–72. [PubMed: 15701239]
- Frank MJ. Hold your horses: a dynamic computational role for the subthalamic nucleus in decision making. *Neural Netw* 2006;19:1120–1136. [PubMed: 16945502]
- Frank MJ, Claus ED. Anatomy of a Decision: Striato-Orbitofrontal Interactions in Reinforcement Learning, Decision Making, and Reversal. *Psychological Review* 2006;113:300–326. [PubMed: 16637763]
- Frank MJ, Doll BB, Oas-Terpstra J, Moreno F. Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. *Nat Neurosci* 2009;12:1062–1068. [PubMed: 19620978]
- Frank MJ, Hutchison K. Genetic contributions to avoidance-based decisions: striatal D2 receptor polymorphisms. *Neuroscience* 2009;164:131–140. [PubMed: 19393722]
- Frank MJ, Moustafa AA, Haughey HM, Curran T, Hutchison KE. Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proc Natl Acad Sci U S A* 2007;104:16311–16316. [PubMed: 17913879]
- Frank MJ, O'Reilly R.C. A mechanistic account of striatal dopamine function in cognition: Psychopharmacological studies with cabergoline and haloperidol. *Behav Neurosci* 2006;120:497–517. [PubMed: 16768602]
- Frank MJ, Seeberger LC, O'Reilly R.C. By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* 2004;306:1940–1943. [PubMed: 15528409]
- Galderisi S, Quarantelli M, Volpe U, Mucci A, Cassano GB, Invernizzi G, Rossi A, Vita A, Pini S, Cassano P, et al. Patterns of structural MRI abnormalities in deficit and nondeficit schizophrenia. *Schizophr Bull* 2008;34:393–401. [PubMed: 17728266]
- Gold JM, Bish JA, Iannone VN, Hobart MP, Queern CA, Buchanan RW. Effects of contextual processing on visual conditional associative learning in schizophrenia. *Biol Psychiatry* 2000;48:406–414. [PubMed: 10978724]
- Gold JM, Carpenter C, Randolph C, Goldberg TE, Weinberger DR. Auditory working memory and Wisconsin Card Sorting Test performance in schizophrenia. *Arch Gen Psychiatry* 1997;54:159–165. [PubMed: 9040284]
- 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. [PubMed: 18591195]
- Gong S, Zheng C, Doughty ML, Losos K, Didkovsky N, Schambra UB, Nowak NJ, Joyner A, Leblanc G, Hatten ME, Heintz N. A gene expression atlas of the central nervous system based on bacterial artificial chromosomes. *Nature* 2003;425:917–925. [PubMed: 14586460]
- Graybiel AM. Habits, rituals, and the evaluative brain. *Annu Rev Neurosci* 2008;31:359–387. [PubMed: 18558860]

- Hakansson K, Galdi S, Hendrick J, Snyder G, Greengard P, Fisone G. Regulation of phosphorylation of the GluR1 AMPA receptor by dopamine D2 receptors. *J Neurochem* 2006;96:482–488. [PubMed: 16336634]
- Horan WP, Green MF, Knowlton BJ, Wynn JK, Mintz J, Nuechterlein KH. Impaired implicit learning in schizophrenia. *Neuropsychology* 2008;22:606–617. [PubMed: 18763880]
- Joyce JN. D2 but not D3 receptors are elevated after 9 or 11 months chronic haloperidol treatment: influence of withdrawal period. *Synapse* 2001;40:137–144. [PubMed: 11252025]
- Kelly DL, Kreyenbuhl J, Buchanan RW, Malhotra AK. Why Not Clozapine? *Clinical Schizophrenia & Related Psychoses* 2007;1:92–95.
- Kemali D, Maj M, Galderisi S, Monteleone P, Mucci A. Conditional associative learning in drug-free schizophrenic patients. *Neuropsychobiology* 1987;17:30–34. [PubMed: 3627390]
- Keri S, Kelemen O, Szekeres G, Bagoczky N, Erdelyi R, Antal A, Benedek G, Janka Z. Schizophrenics know more than they can tell: probabilistic classification learning in schizophrenia. *Psychol Med* 2000;30:149–155. [PubMed: 10722185]
- Kirkpatrick B, Buchanan RW. The neural basis of the deficit syndrome of schizophrenia. *J Nerv Ment Dis* 1990;178:545–555. [PubMed: 2203878]
- Knowlton BJ, Mangels JA, Squire LR. A neostriatal habit learning system in humans. *Science* 1996;273:1399–1402. [PubMed: 8703077]
- Koch K, Schachtzabel C, Wagner G, Schikora J, Schultz C, Reichenbach JR, Sauer H, Schlosser RG. Altered activation in association with reward-related trial-and-error learning in patients with schizophrenia. *Neuroimage* 50:223–232. in press. [PubMed: 20006717]
- Laruelle M, Abi-Dargham A. Dopamine as the wind of the psychotic fire: new evidence from brain imaging studies. *J Psychopharmacol* 1999;13:358–371. [PubMed: 10667612]
- Malenka RC, Angel RW, Hampton B, Berger PA. Impaired central error-correcting behavior in schizophrenia. *Arch Gen Psychiatry* 1982;39:101–107. [PubMed: 6119966]
- Mallet N, Ballion B, Le Moine C, Gonon F. Cortical inputs and GABA interneurons imbalance projection neurons in the striatum of parkinsonian rats. *J Neurosci* 2006;26:3875–3884. [PubMed: 16597742]
- McClure SM, Daw ND, Montague PR. A computational substrate for incentive salience. *Trends Neurosci* 2003;26:423–428. [PubMed: 12900173]
- McMahon RP, Kelly DL, Kreyenbuhl J, Kirkpatrick B, Love RC, Conley RR. Novel factor-based symptom scores in treatment resistant schizophrenia: implications for clinical trials. *Neuropsychopharmacology* 2002;26:537–545. [PubMed: 11927178]
- Meyer-Lindenberg A, Miletich RS, Kohn PD, Esposito G, Carson RE, Quarantelli M, Weinberger DR, Berman KF. Reduced prefrontal activity predicts exaggerated striatal dopaminergic function in schizophrenia. *Nat Neurosci* 2002;5:267–271. [PubMed: 11865311]
- Miller, R. *A Neurodynamic Theory of Schizophrenia (and related disorders)*. Lulu.com; New Zealand: 2008.
- Moustafa AA, Cohen MX, Sherman SJ, Frank MJ. A role for dopamine in temporal decision making and reward maximization in parkinsonism. *J Neurosci* 2008;28:12294–12304. [PubMed: 19020023]
- Overall JE, Gorman DR. The Brief Psychiatric Rating Scale. *Psychological Reports* 1962;10:799–812.
- Pfohl, B.; Blum, N.; Zimmerman, M.; Stangl, D. *Structured Interview for DSM-III-R Personality Disorders (SIDP-R)*. University of Iowa, Department of Psychiatry; Iowa City, IA: 1989.
- Pizzagalli DA, Evins AE, Schetter EC, Frank MJ, Pajtas PE, Santesso DL, Culhane M. Single dose of a dopamine agonist impairs reinforcement learning in humans: behavioral evidence from a laboratory-based measure of reward responsiveness. *Psychopharmacology (Berl)* 2008;196:221–232. [PubMed: 17909750]
- Prentice KJ, Gold JM, Buchanan RW. The Wisconsin Card Sorting impairment in schizophrenia is evident in the first four trials. *Schizophr Res* 2008;106:81–87. [PubMed: 17933496]
- Rolls ET, O'Doherty J, Kringelbach ML, Francis S, Bowtell R, McGlone F. Representations of pleasant and painful touch in the human orbitofrontal and cingulate cortices. *Cereb Cortex* 2003;13:308–317. [PubMed: 12571120]

- Rushe TM, Woodruff PW, Murray RM, Morris RG. Episodic memory and learning in patients with chronic schizophrenia. *Schizophr Res* 1999;35:85–96. [PubMed: 9988844]
- Santesso DL, Evins AE, Frank MJ, Schetter EC, Bogdan R, Pizzagalli DA. Single dose of a dopamine agonist impairs reinforcement learning in humans: evidence from event-related potentials and computational modeling of striatal-cortical function. *Hum Brain Mapp* 2009;30:1963–1976. [PubMed: 18726908]
- Schoenbaum G, Roesch M. Orbitofrontal cortex, associative learning, and expectancies. *Neuron* 2005;47:633–636. [PubMed: 16129393]
- Schultz W. Predictive reward signal of dopamine neurons. *J Neurophysiol* 1998;80:1–27. [PubMed: 9658025]
- Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. *Science* 1997;275:1593–1599. [PubMed: 9054347]
- Schwartz BL, Howard DV, Howard JH Jr, Hovaguimian A, Deutsch SI. Implicit learning of visuospatial sequences in schizophrenia. *Neuropsychology* 2003;17:517–533. [PubMed: 12959517]
- Seeman P. Dopamine receptors and the dopamine hypothesis of schizophrenia. *Synapse* 1987;1:133–152. [PubMed: 2905529]
- Seeman P, Weinshenker D, Quirion R, Srivastava LK, Bhardwaj SK, Grandy DK, Premont RT, Sotnikova TD, Boksa P, El-Ghundi M, et al. Dopamine supersensitivity correlates with D2High states, implying many paths to psychosis. *Proc Natl Acad Sci U S A* 2005;102:3513–3518. [PubMed: 15716360]
- Shen W, Flajolet M, Greengard P, Surmeier DJ. Dichotomous dopaminergic control of striatal synaptic plasticity. *Science* 2008;321:848–851. [PubMed: 18687967]
- Stroup TS, Lieberman JA, McEvoy JP, Davis SM, Swartz MS, Keefe RS, Miller AL, Rosenheck RA, Hsiao JK. Results of phase 3 of the CATIE schizophrenia trial. *Schizophr Res* 2009;107:1–12. [PubMed: 19027269]
- Surmeier DJ, Ding J, Day M, Wang Z, Shen W. D1 and D2 dopamine-receptor modulation of striatal glutamatergic signaling in striatal medium spiny neurons. *Trends Neurosci* 2007;30:228–235. [PubMed: 17408758]
- Surmeier DJ, Song WJ, Yan Z. Coordinated expression of dopamine receptors in neostriatal medium spiny neurons. *J Neurosci* 1996;16:6579–6591. [PubMed: 8815934]
- Tricomi E, Balleine BW, O'Doherty JP. A specific role for posterior dorsolateral striatum in human habit learning. *Eur J Neurosci*. 2009
- Vaiva G, Cottencin O, Llorca PM, Devos P, Dupont S, Mazas O, Rasclé C, Thomas P, Steinling M, Goudemand M. Regional cerebral blood flow in deficit/nondeficit types of schizophrenia according to SDS criteria. *Prog Neuropsychopharmacol Biol Psychiatry* 2002;26:481–485. [PubMed: 11999898]
- Valjent E, Bertran-Gonzalez J, Herve D, Fisone G, Girault JA. Looking BAC at striatal signaling: cell-specific analysis in new transgenic mice. *Trends Neurosci* 2009;32:538–547. [PubMed: 19765834]
- 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. [PubMed: 17300757]
- Waltz JA, Gold JM. Probabilistic reversal learning impairments in schizophrenia: further evidence of orbitofrontal dysfunction. *Schizophr Res* 2007;93:296–303. [PubMed: 17482797]
- Waltz JA, Schweitzer JB, Gold JM, Kurup PK, Ross TJ, Salmeron BJ, Rose EJ, McClure SM, Stein EA. Patients with Schizophrenia have a Reduced Neural Response to Both Unpredictable and Predictable Primary Reinforcers. *Neuropsychopharmacology* 2009;34:1567–1577. [PubMed: 19052540]
- Wechsler, D. Wechsler Memory Scale. 3rd ed.. The Psychological Corporation; San Antonio, TX: 1997.
- Wechsler, D. Wechsler Test of Adult Reading (WTAR). The Psychological Corporation; San Antonio, TX: 2001.

- Weickert TW, Terrazas A, Bigelow LB, Malley JD, Hyde T, Egan MF, Weinberger DR, Goldberg TE. Habit and skill learning in schizophrenia: evidence of normal striatal processing with abnormal cortical input. *Learn Mem* 2002;9:430–442. [PubMed: 12464703]
- Weinberger DR. Implications of normal brain development for the pathogenesis of schizophrenia. *Arch Gen Psychiatry* 1987;44:660–669. [PubMed: 3606332]
- 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. [PubMed: 19169674]

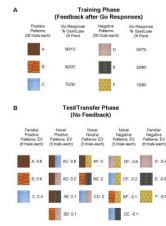


Figure 1.

Probabilistic Go/NoGo task. (A) Training Phase: Six patterns associated with reinforcement probabilities of 90%, 80%, 70%, 30%, 20% and 10% were presented 30 times each, with feedback. (B) In a Test/Transfer Phase, training (familiar) stimuli were presented 6 times each without feedback, along with 11 novel stimuli, consisting of combinations of the training patterns. By virtue of their reinforcement probabilities, training stimuli had expected values (EVs) of between +0.8 and -0.8 points [the expected value for the 90%-win stimulus, for example, was $(0.9 * 1) + (0.1 * -1) = 0.8$, and the expected value for the 70%-win stimulus was $(0.7 * 1) + (0.3 * -1) = 0.4$]. Based on the reinforcement probability of their constituent patterns during the Training Phase, novel/combination stimuli had expected values of between -0.6 and +0.6 points [the expected value for the combination of the 90%-win stimulus and the 70%-win stimulus, for example, was the mean of 0.8 and 0.4, or 0.6]. Novel patterns were presented 3 times each.

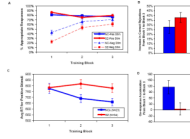


Figure 2.

Learning measures from the Training Phase. In all panels, blue lines/bars = controls; red lines/bars = patients. A. Rates of hits (solid lines) and correct rejections (dashed lines) during the Training Phase, by subject group and block. B. Changes in correct rejection rates by group, from training block 1 to training block 3. C. Average RTs to positive stimuli during the Training Phase, by subject group and block. D. Changes in average RTs to positive stimuli by group, from training block 1 to training block 3.

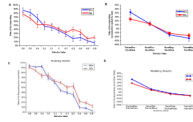


Figure 3.

Performance of subjects in the post-training Test/Transfer Phase. In all panels, blue lines = controls; red lines = patients/degraded networks. A. Rates of Go responding plotted against the estimated expected value of Test/Transfer stimuli. The expected values of combination stimuli (see Figure 2) were computed by averaging the expected values of the two component stimuli. B. Measures of Go and NoGo response selectivity at Test/Transfer. Adjusted for response rates to neutral novel stimuli, patients and controls show a similar tendency to withhold responses to familiar *negative* stimuli at Test/Transfer, but patients show a reduced tendency to “Go” to familiar *positive* stimuli. C and D show the performance of artificial neural networks on Test/Transfer items. Like SZ patients, degraded networks show a reduced preference for positive training stimuli over neutral stimuli, when compared with control networks. Degraded and control networks do not differ in their tendencies to avoid negative training stimuli, relative to neutral stimuli.

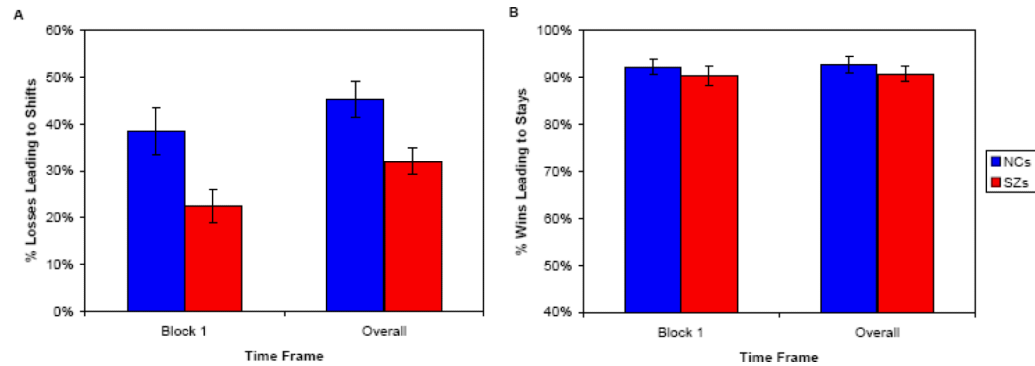


Figure 4.

Measures of rapid reinforcement learning in subjects. A. Rates of shifting in response to negative feedback (point-losses following choices) in the first block (10 trials with each stimulus) and in the entire Training Phase. Patients show a reduced tendency to “lose-shift”, relative to controls, leading to a lower rate of correct rejections early in the task. B. Rates of staying in response to positive feedback (point-gains following choices) in the first block (10 trials with each stimulus) and in the entire Training Phase. Patients and controls do not differ in their tendencies to “win-stay”.

Table 1

Characterizing information for patients and controls.

Measure	Control Mean (SD)	Patient Mean (SD)	p-value
Age	45.7 (11.5)	44.7 (9.2)	0.722
Age at Illness Onset	-	23.2 (8.0)	
Education (years)	15.3 (2.4)	12.8 (2.4)	<0.001
Paternal Education (years)	14.9 (3.4)	13.9 (4.2)	0.379
Gender (M:F)	12:11	21:13	0.587
Race			0.186
African American	4	11	
Caucasian	19	21	
Other	0	2	
Standard Neuropsychology			
Wechsler Test of Adult Reading	111.6 (9.6)	100.0 (14.1)	0.002
Hopkins Verbal Learning Score	29.5 (4.2)	21.6 (5.8)	<0.001
Spatial Span Scaled Score	11.7 (2.2)	7.9 (3.1)	<0.001
Letter-number Sequencing	16.1 (3.1)	12.1 (3.2)	<0.001
Clinical Ratings			
BPRS	-	36.3 (10.7)	
SANS	-	32.8 (15.6)	
Calgary Depression Scale	-	3.5 (3.7)	

Abbreviations: SANS=Scale for the Assessment of Negative Symptoms; BPRS=Brief Psychiatric Rating Scale.

Table 2

Results of Correlation Analyses between Experimental Measures and Symptom Assessments.

	Spearman's Rho	p-value
SANS Total		
Block 1 Correct-reject Rate	-0.373	0.030
Block 1 Lose-shift Rate	-0.461	0.006
B1-B3 Go Response-time Acceleration	-0.196	ns
[Familiar Positive - Novel Neutral] Contrast at Test	0.407	0.017
SANS Avolition		
Block 1 Correct-reject Rate	-0.379	0.027
Block 1 Lose-shift Rate	-0.431	0.011
B1-B3 Go Response-time Acceleration	-0.109	ns
[Familiar Positive - Novel Neutral] Contrast at Test	0.365	0.034
SANS Anhedonia		
Block 1 Correct-reject Rate	-0.258	ns
Block 1 Lose-shift Rate	-0.338	0.051
B1-B3 Go Response-time Acceleration	-0.070	ns
[Familiar Positive - Novel Neutral] Contrast at Test	0.302	0.083
BPRS Negative Symptom Cluster		
Block 1 Correct-reject Rate	-0.381	0.026
Block 1 Lose-shift Rate	-0.403	0.018
B1-B3 Go Response-time Acceleration	0.227	ns
[Familiar Positive - Novel Neutral] Contrast at Test	0.328	ns
BPRS Reality Distortion Cluster		
Block 1 Correct-reject Rate	-0.183	ns
Block 1 Lose-shift Rate	-0.251	ns
B1-B3 Go Response-time Acceleration	-0.149	ns
[Familiar Positive - Novel Neutral] Contrast at Test	0.178	ns

Abbreviations: SANS=Scale for the Assessment of Negative Symptoms; BPRS=Brief Psychiatric Rating Scale; Neg Cluster=Negative Symptom Cluster (Sum of 3 items: Emotional Withdrawal, Motor Retardation, Blunted Affect); RD Cluster=Reality Distortion Cluster (Sum of 4 items: Grandiosity, Suspiciousness, Hallucinations, Unusual Thought Content).