

# Selective Reinforcement Learning Deficits in Schizophrenia Support Predictions from Computational Models of Striatal-Cortical Dysfunction

James A. Waltz, Michael J. Frank, Benjamin M. Robinson, and James M. Gold

**Background:** Rewards and punishments may make distinct contributions to learning via separate striatal-cortical pathways. We investigated whether fronto-striatal dysfunction in schizophrenia (SZ) is characterized by selective impairment in either reward- (Go) or punishment-driven (NoGo) learning.

**Methods:** We administered two versions of a probabilistic selection task to 40 schizophrenia patients and 31 control subjects, using difficult to verbalize stimuli (experiment 1) and nameable objects (experiment 2). In an acquisition phase, participants learned to choose between three different stimulus pairs (AB, CD, EF) presented in random order, based on probabilistic feedback (80%, 70%, 60%). We used analyses of variance (ANOVAs) to assess the effects of group and reinforcement probability on two measures of contingency learning. To characterize the preference of subjects for choosing the most rewarded stimulus and avoiding the most punished stimulus, we subsequently tested participants with novel pairs of stimuli involving either A or B, providing no feedback.

**Results:** Control subjects demonstrated superior performance during the first 40 acquisition trials in each of the 80% and 70% conditions versus the 60% condition; patients showed similarly impaired (<60%) performance in all three conditions. In novel test pairs, patients showed decreased preference for the most rewarded stimulus (A;  $t = 2.674$ ;  $p = .01$ ). Patients were unimpaired at avoiding the most negative stimulus (B;  $t = .737$ ).

**Conclusions:** The results of these experiments provide additional evidence for the presence of deficits in reinforcement learning in SZ, suggesting that reward-driven learning may be more profoundly impaired than punishment-driven learning.

**Key Words:** Basal ganglia, dopamine, orbitofrontal, prefrontal, reinforcement, schizophrenia

Cognitive deficits are widely recognized as central features of schizophrenia (SZ) (Barch 2005; Wilk *et al.* 2005). Of the impairments documented in the literature, deficits involving the use of feedback to guide decision making and learning are highly reliable and sometimes clinically dramatic. Patients' poor performance on many of these tasks like the Wisconsin Card Sort Test (WCST) (Goldberg *et al.* 1987) and conditional associative learning paradigms (Gold *et al.* 2000) is often interpreted as evidence of dysfunction in either dorsolateral regions of prefrontal cortex (PFC) (Weinberger *et al.* 1986) or lateral and medial areas of ventral prefrontal cortex (Boettiger and D'Esposito 2005), also called orbitofrontal cortex (OFC).

In contrast, several (but not all) studies of procedural, or habit, learning (Keri *et al.* 2000, 2005; Weickert *et al.* 2002) have documented surprisingly normal learning among SZ patients. These tasks also employ feedback to guide learning but tend to involve gradual learning of difficult-to-discern probabilistic response-outcome relationships. Both functional imaging and studies of patient populations such as Parkinson disease (PD) suggest that the basal ganglia (BG) play a critical role in this gradual learning of stimulus-response mappings (Knowlton *et al.* 1996; Seger and Cincotta 2005).

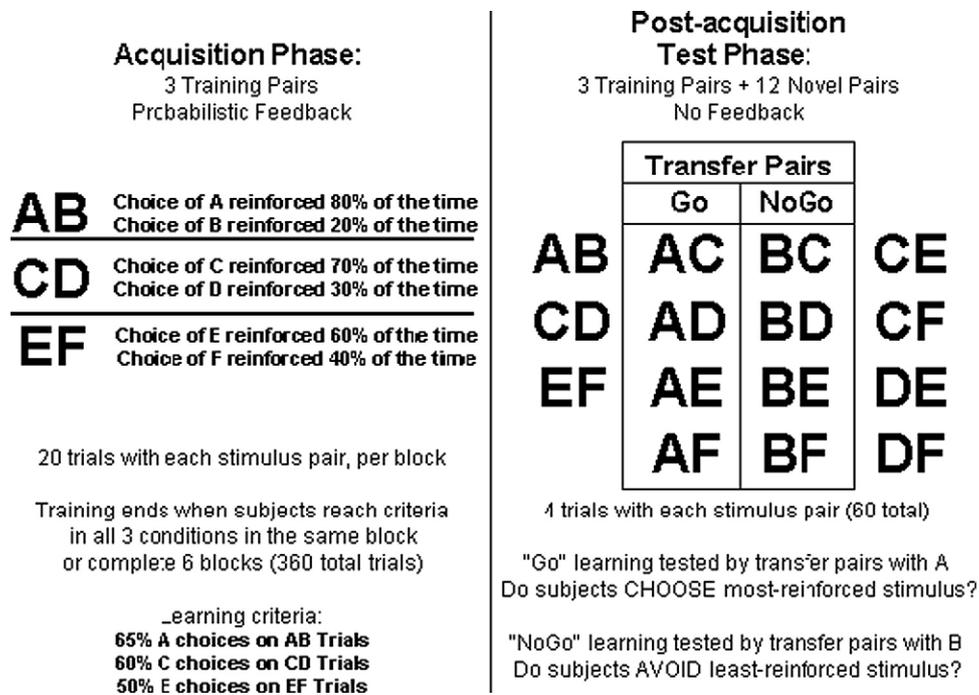
From the Maryland Psychiatric Research Center (JAW, BMR, JMG), University of Maryland School of Medicine, Baltimore, Maryland; and Department of Psychology (MJF), University of Arizona, Tucson, Arizona.

Address reprint requests to James A. Waltz, Ph.D., University of Maryland School of Medicine, Maryland Psychiatric Research Center, PO Box 21247, Baltimore, MD 21228; E-mail: jwaltz@mprc.umaryland.edu.

Received June 14, 2006; revised September 20, 2006; accepted September 25, 2006.

Explaining the differential impairment of these learning processes in schizophrenia is difficult, given the evidence that brain dopamine (DA) systems are known to play a critical role in both PFC-mediated and BG-dependent reinforcement-learning processes. One possible explanation for the relative sparing of habit learning in SZ is that some DA pathways in the BG are largely intact. To investigate this question, we adopted the experimental methods and computational framework of Frank *et al.* (2004), who examined learning performance in a group of PD patients studied both on and off L-Dopa. Frank *et al.* (2004) used a probabilistic stimulus selection (PSS) task, where subjects are initially presented with three different stimulus pairs (AB, CD, EF) and have to learn to choose the most frequently reinforced stimulus from each pair using probabilistic feedback (Figure 1). After achieving the learning criterion in this acquisition phase, subjects are then presented with the original stimuli in novel pairings in a postacquisition test phase. This design provides a means of studying the contributions of positive and negative feedback to probabilistic learning, in that it enables the assessment of whether subjects have a bias for choosing frequently reinforced stimuli or for avoiding frequently punished stimuli.

Frank *et al.* (2004) demonstrated that unmedicated PD patients showed considerable impairment in learning driven by positive feedback when compared with their performance in the medicated state. Importantly, their learning driven by negative feedback was superior to that in the medicated state. These results were interpreted in the context of computational models of reward-based learning (Frank 2005; Frank *et al.* 2001) that formalize ideas about the role of dopaminergic signaling in the BG. These signals are thought to communicate information about reward contingencies in the environment that guide action selection and learning. A degree of functional segregation characterizes pathways in the BG, such that activity in the direct pathway sends a "Go" signal to facilitate the execution of a



**Figure 1.** The probabilistic stimulus selection (PSS) task. The task consists of two phases. During an acquisition phase, subjects are presented with three training pairs and instructed to identify which stimulus from each pair is more frequently reinforced. In AB trials, for example, a choice of stimulus A leads to positive feedback in 80% of trials, whereas a B choice is reinforced on the remaining 20%. Learning the most frequently rewarded stimulus in each pair can be accomplished either by learning that one of the stimuli leads to positive feedback or that the other leads to negative feedback (or both). Subjects are told to choose that stimulus as often as possible. Once subjects reach criterion on all three training pairs or complete 360 total trials, they proceed to a postacquisition test phase, during which they are presented with four trials each of the three training pairs, along with 12 new pairs created from all unused combinations of the training stimuli. The eight new stimulus pairs involving A and B are called the transfer pairs and are used to gauge Go and NoGo learning. If positive feedback was more effective, they should reliably choose stimulus A in all novel test pairs in which it is present; if they learned more from negative feedback, they should avoid stimulus B. PSS, probabilistic stimulus selection.

response considered in cortex, whereas activity in the indirect pathway sends a "NoGo" signal to suppress inappropriate responses (Figure 2) (Centonze *et al.* 2001; Nishi *et al.* 1997). Furthermore, dopaminergic innervation of these pathways is thought to be relatively distinct, such that the direct pathway is excited via D1 receptors by bursting activity in dopamine neurons, while the indirect pathway is tonically inhibited via D2 receptors. Phasic DA bursts are thought to support Go learning to reinforce rewarding choices by enhancing neural activity and plasticity in the direct (D1) pathway following reinforcement and enhancing inhibition of the indirect (D2) pathway. Transient cessations of DA cell firing, following negative feedback, are thought to have the opposite effect: they release inhibition of the indirect pathway and cause reductions of activity in the direct pathway, thereby supporting NoGo learning to avoid unrewarding choices (Frank 2005; O'Reilly and Frank 2006). These authors concluded that in unmedicated PD patients, DA depletion attenuates the impact of DA bursts. In medicated PD patients, the impact of DA "dips" is attenuated due to overall increased levels of synaptic DA.

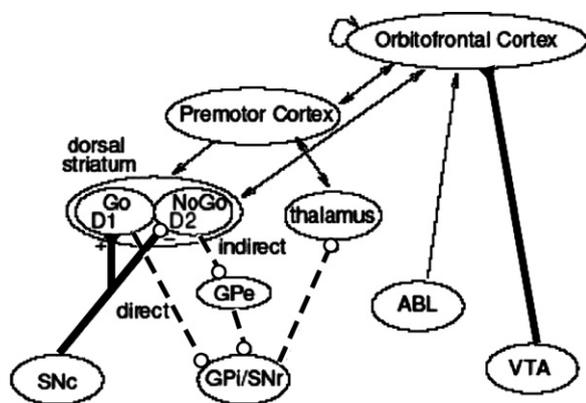
While learning at the level of the basal ganglia is thought to occur on a gradual time scale, Go and NoGo signals emanating from the BG are hypothesized to impact the rapid learning of changing reinforcement contingencies in the frontal cortex via parallel striatal-cortical circuits by updating working memory (WM) representations required for representing differences in relative magnitude of reinforcement online (Frank *et al.* 2001; O'Reilly and Frank 2006). This idea extends earlier computa-

tional work emphasizing the role of phasic DA in driving the updating of PFC WM representations (Braver and Cohen 2000; Cohen *et al.* 1996). The idea that OFC figures critically in the online representation of reward and punishment magnitudes and thus subserves a kind of working memory is supported by recent evidence (Rolls *et al.* 2003; Schoenbaum and Roesch 2005). Simulations by Frank and Claus (2006) have shown that models capable of instantaneously updating WM representations of reward value in OFC and using them to bias behavior via efferent projections to the BG and motor cortical areas show rapid acquisition of probabilistic contingencies, whereas models with OFC damage exhibit much slower learning because they can only acquire probabilistic contingencies via changes in synaptic weights in the BG.

**Relevance of Dopamine System Function Models to SZ**

This framework has the potential to offer a differentiated account of feedback-driven learning deficits in SZ. Whereas PD involves mainly BG hypofunction brought on by dopamine depletion, SZ may be characterized by DA dysfunction in both PFC and the BG. While the severity and consequences of PFC hypofunction in schizophrenia appear to be profound (Weinberger 1987; Weinberger and Berman 1988), BG dysfunction in schizophrenia may be more mild, based on findings of relatively intact procedural learning (Keri *et al.* 2005; Kern *et al.* 1997; Weickert *et al.* 2002).

We tested three specific hypotheses by applying the paradigm used by Frank *et al.* (2004) in their study of PD patients. We



**Figure 2.** The cortico-striato-thalamo-cortical loops, including the direct and indirect pathways of the basal ganglia. The cells of the striatum are divided into two subclasses based on differences in biochemistry and efferent projections. The Go cells project directly to the GPi/SNr, and their activity disinhibits the thalamus, thereby facilitating the execution of a cortical response. The NoGo cells are part of the indirect pathway to the GPi/SNr and have an opposing effect, suppressing actions from getting executed. Dopamine from the SNc projects to the dorsal striatum, differentially modulating activity in the direct and indirect pathways by activating different receptors. The Go cells express the D1 receptor, and the NoGo cells express the D2 receptor. The orbitofrontal cortex is thought to maintain reinforcement-related information in working memory and provide top-down biasing on the more primitive BG system, in addition to direct influencing of response selection processes in premotor cortex. The OFC receives information about relative magnitude of reinforcement values from the ABL, which it can also maintain in working memory. Dopamine from the VTA projects to ventral striatum (not shown) and orbitofrontal cortex. Gpi, internal segment of globus pallidus; Gpe, external segment of globus pallidus; SNc, substantia nigra pars compacta; SNr, substantia nigra pars reticulata; OFC, orbitofrontal cortex; BG, basal ganglia; ABL, basolateral amygdala; VTA, ventral tegmental area.

expected the relative severity of prefrontal cortical versus basal ganglia dysfunction in schizophrenia to have two specific effects on the performance of probabilistic reinforcement learning tasks. First, due to PFC hypofunction, we expected that patients would show marked deficits in the initial learning of the most favorable stimulus in each pair, which is critically dependent on the rapid updating of reward value representations. Second, based on prior studies showing relatively intact procedural learning in SZ, we expected that patients would show delayed but eventual acquisition of the stimulus pairs. Finally, we speculatively hypothesized that SZ patients would show the pattern of reduced Go learning seen in unmedicated PD patients. This might occur if the fidelity of burst-driven phasic signaling is reduced in SZ but the sensitivity of D2 receptors (which are sensitive to small decreases in dopamine) (Frank and O'Reilly 2006) is enhanced (Curran *et al.* 2004; Seeman *et al.* 2005).

## Methods and Materials

### Patients

Forty 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). 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). Most patients (28/40) were on antipsychotic monotherapy, while 12 patients were

taking two antipsychotics (almost all clozapine with risperidone) (Table 1).

### Control Subjects

Thirty-one healthy control subjects, recruited through newspaper advertisements and random phone number dialing, participated in the study. They were extensively screened for Axis I and Axis 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. 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.

Data collection occurred through a battery of standard and experimental neuropsychological tests. Tests included measures of word reading, word list learning, and working memory. 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).

### Cognitive Task

We used the PSS task from Frank *et al.* (2004) described above (Figure 1). Blocks consisted of 60 trials (20 per condition), in which three different stimulus pairs (AB, CD, EF) were presented in pseudorandom order, and the acquisition phase was terminated when participants achieved criterion on all three stimulus pairs in the same block or when subjects had completed 360 trials, whichever occurred first. The discontinuation criterion was 65% correct in the AB (80:20) condition, 60% in the CD (70:30) condition, and 50% in the EF (60:40) condition. This liberal criterion was intended to prevent overlearning of contingencies prior to the postacquisition test phase.

In the postacquisition test phase, no feedback was provided. To examine whether subjects had preferentially learned through the use of positive or negative feedback, we analyzed performance on the 32 test trials involving novel combinations of stimulus pairs involving either an A or a B. All subjects who failed to demonstrate initial learning of the AB contingency by choosing A at least three times in the four test trials were excluded from the analysis of transfer performance.

We performed two experiments using this paradigm. All 40 patients completed both experiments, while 28 of 31 control subjects who completed experiment 1 also completed experiment 2. In experiment 1, we used the Hiragana characters from Frank *et al.* (2004). In experiment 2, we used clip art images of common objects (flashlight, clock, etc). The second experiment was initiated after approximately half of the participants had finished the first experiment, because we had found extremely poor learning in the patient group, with few meeting the criterion for the transfer analysis. Using images of common objects was

**Table 1.** Characterizing Information for Patients and Controls Subjects

Measure	Control Mean (SD)	Patient Mean (SD)	<i>p</i> -Value
Age	44.63 (10.38)	44.43 (9.24)	.933
Age at Illness Onset	—	23.28 (7.854)	
Education (years)	14.77 (2.43)	12.80 (2.28)	.001
Paternal Education (years)	13.11 (4.65)	13.88 (4.29)	.492
Gender (M:F)	14:17	25:15	.145
Race			.192
African American	9	16	
Caucasian	22	22	
Other	0	2	
Antipsychotic Medication Regimen			
Haloperidol or Fluphenazine Only		8	
Clozapine Only		7	
Other Second-Generation Only		13	
Clozapine + Another Antipsychotic		12	
Clinical Ratings			
BPRS	—	35.92 (8.61)	
SANS	—	32.59 (15.78)	
Calgary Depression Scale	—	3.28 (3.54)	
Standard Neuropsychology			
Wechsler Test of Adult Reading	109.77 (13.00)	96.35 (17.53)	<.001
Verbal List Learning	28.87 (4.49)	21.58 (5.41)	<.001
Spatial Span Forward	8.61 (1.98)	6.85 (1.86)	<.001
Spatial Span Backward	8.19 (1.25)	6.08 (1.97)	<.001
Spatial Span Scaled Score	11.52 (2.20)	7.80 (2.91)	<.001
Letter-Number Sequencing	15.81 (3.24)	11.90 (3.28)	<.001

BPRS, Brief Psychiatric Ratings Scale; SANS, Scales for the Assessment of Negative Symptoms.

intended to address the possibility that poor performance was the result of difficulty encoding the Hiragana characters. Thus, all subjects received the Hiragana version first, followed by the clip art version, with the two testing occasions separated by up to 9 months (mean = 3.18 months). In experiment 2, all subjects performed at least two training blocks to facilitate examination of early acquisition, with the same discontinuation criteria applied to subsequent blocks. In this experiment, 8 of 40 patients and 4 of 28 controls failed to demonstrate initial learning of the AB contingency.

### Statistical Analyses

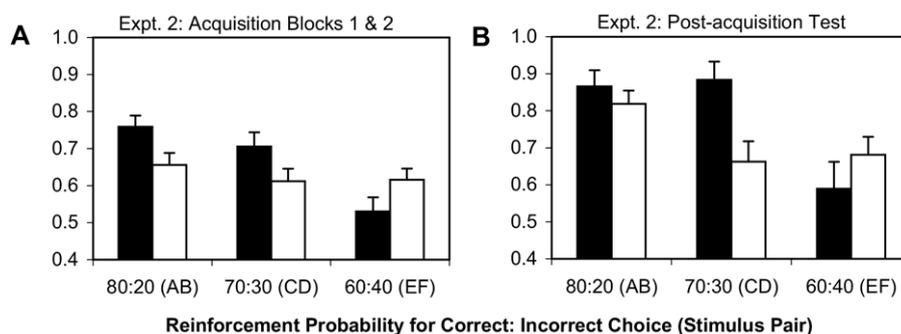
In comparing the acquisition of contingencies between patients and control subjects, we performed two-way analyses of variance (ANOVAs), as well as appropriate post hoc tests, on subjects' proportions of correct responses in the first two blocks of the acquisition phase of each experiment with factors of group and reinforcement probability. For experiment 1, proportion-correct scores from the first acquisition block (first 20 trials of each type) were carried forward to the second block for eight control subjects and two patients who reached the acquisition criterion after only one block. In the results, we term the first two blocks (the first 40 trials of each type) of the acquisition phase "early acquisition."

The learning of probabilistic contingencies was also assessed at the postacquisition test using ANOVAs with factors of group and reinforcement probability. These scores reflect learning of contingencies after up to six blocks of training, rather than just the first two. Whereas acquisition was assessed using the three training pairs repeated during the test phase, group differences in transfer performance in the test phase were assessed using *t* tests for measures of Go and NoGo learning generated from subjects'

cumulative test scores on the four novel pairs involving A (Go) and the four novel pairs involving B (NoGo).

We interpret performance on the postacquisition test items to reflect the gradual, habit-like acquisition of contingencies, largely dependent on the BG (Frank *et al.* 2004, 2006). A second kind of reinforcement learning involves the (PFC-dependent) ability to represent and integrate feedback online to rapidly learn contingencies. To directly assess the contribution of online feedback integration to the rapid acquisition of probabilistic contingencies, we computed "win-stay" and "lose-shift" scores for each reinforcement condition in block 1 (i.e., early in the acquisition phase). We computed win-stay scores by computing the proportion of repeated stimulus selections in a given condition that followed reinforced choices. We computed lose-shift scores by computing the proportion of switched stimulus selections in a given condition that followed nonreinforced choices. We then generated total win-stay and lose-shift scores by averaging scores across conditions for each, and between-group differences in mean scores were assessed using *t* tests. We also computed effect sizes (Cohen's *d* scores) to characterize between-group differences in means by dividing each mean difference by the pooled standard deviation.

We used Pearson correlation analyses to assess relationships between probabilistic selection performance and three types of characterizing variables: symptom ratings, antipsychotic medication doses (converted to haloperidol equivalent units) (Supplementary Table 1), and standard neuropsychological measures and probabilistic selection performance. To do so, we created a summary measure of probabilistic selection performance by averaging the proportion of correct responses from all three conditions in the first two acquisition blocks. To separately assess psychotic and disorganized symptoms from the BPRS,



**Figure 3.** Acquisition of probabilistic contingencies by patients and controls in experiment 2. (A) In blocks 1 and 2. (B) Performance on training pairs at postacquisition test. The proportion of correct responses was defined as the proportion of trials on which the most frequently reinforced stimulus was chosen. In both panels, black bars = control subjects, white bars = patients.

subscores were grouped into reality distortion, disorganization, negative symptom, and anxiety/depression clusters based on the four-factor model of McMahon *et al.* (2002).

## Results

### Acquisition of Contingencies

In our first experiment, patients demonstrated dramatic impairment in the acquisition of probabilistic contingencies, whereas healthy subjects demonstrated clear learning of the two most frequently rewarded stimuli. Two-way ANOVAs for data from both early acquisition and postacquisition tests revealed main effects of group (see Supplementary text and Supplementary Figure 1 for details), indicating that patients performed worse than control subjects in experiment 1, regardless of reinforcement condition. Because fully 50% of patients failed to reach criterion, however, we did not analyze the transfer results from experiment 1 due to concern that the patients who did meet criterion were unrepresentative of the total group. Thus, experiment 1 provided robust evidence of marked reward-processing impairments in patients, but we were unable to address whether this impairment resulted from a more selective deficit in the processing of positive or negative outcomes.

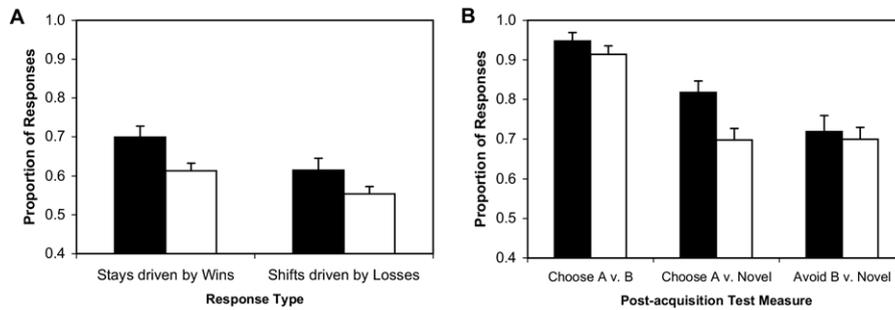
Our entire sample of subjects performed better during the acquisition phase in experiment 2 than in experiment 1, reflecting greater ease of encoding verbalizable stimuli (see Supplementary text). Proportions of correct responses given by subjects in the first two blocks of the acquisition phase and the postacquisition test phase in experiment 2 are shown in Figure 3. The ANOVA for acquisition measures from experiment 2 revealed no main effect of group [ $F(1,66) = 1.43$ ], but a main effect of reward contingency [ $F(2,132) = 10.27$ ;  $p < .001$ ], and a significant group  $\times$  reward contingency interaction [ $F(2,132) = 4.54$ ;  $p = .012$ ]. Post hoc tests revealed that control subjects performed better than patients in the 80% reward probability condition [ $t(132) = 2.32$ ;  $p = .022$ ], while there was a trend in the direction of control subjects performing better than patients in the 70% reward probability condition [ $t(132) = 1.80$ ;  $p < .10$ ]. In the 60% condition, there was a trend for the patients to perform better than control subjects [ $t(132) = 1.78$ ;  $p < .10$ ]. Control subjects in experiment 2 demonstrated robust early acquisition performance in the 80% and 70% conditions clearly superior to that in the 60% condition ( $t > 2.75$  for both the 80% vs. 60% and 70% vs. 60% comparisons). In contrast, patients showed no difference in performance among the three conditions, with  $<70\%$  accuracy in all conditions.

The ANOVA of training-pair performance in the postacquisition test phase of experiment 2 (Figure 3B) also failed to show a main effect of group [ $F(1,66) = 1.92$ ;  $p > .010$ ], although it revealed a main effect of reward contingency [ $F(2,132) = 10.00$ ;  $p < .001$ ] and a significant group  $\times$  reward contingency interaction [ $F(2,132) = 3.60$ ;  $p = .030$ ]. This interaction resulted from the fact that both groups performed similarly on the easiest [80%;  $t(132) = .85$ ] and most difficult [60%;  $t(132) = 1.07$ ] pairs, while control subjects continued to outperform patients on the 70% pairs [ $t(132) = 3.06$ ,  $p = .003$ ]. Within-group analyses revealed that control subjects performed significantly better on AB and CD pairs than on the EF pairs [ $t(132) > 3$  for both], whereas patients only showed significantly better performance on AB pairs relative to EF pairs [ $t(132) = 2.84$ ,  $p < .005$ ;  $t(132) = .24$  for the 70:30 vs. 60:40 comparison]. Thus, patients were only able to discriminate the easiest from the hardest pairs.

As a test of the influence of general neuropsychological functioning on experimental task performance, we used subjects' scores on the Wechsler Test of Adult Reading (WTAR; the standard neuropsychological measure showing the strongest association with probabilistic learning scores) as a covariate in analyses of covariance (ANCOVAs) with factors of group and reinforcement probability. We found that, although there was evidence of an association between WTAR scores and experimental task performance, the use of WTAR scores as a covariate in an ANCOVA did not substantially alter the effects of group and reward contingency on task performance (see Supplementary text). These results argue against the possibility that group differences in experimental task performance primarily reflect differences in global neuropsychological functioning.

### Impact of Feedback on Rapid Early Learning

Feedback had a greatly reduced impact on the subsequent choices of patients, relative to those of control subjects, in block 1 of the acquisition phase. For data from experiment 1,  $t$  tests revealed that patients were much less likely to repeat reinforced stimulus selections (win-stay) than control subjects [ $t(69) = 4.06$ ,  $p < .001$ , Cohen's  $d = .969$ ; Supplementary Figure 2]. Patients were also much less likely than control subjects to choose the alternative stimulus after being told they were incorrect [lose-shift;  $t(69) = 3.74$ ,  $p < .001$ ,  $d = .880$ ]. For data from experiment 2,  $t$  tests revealed trends for both of these effects [ $t(66) = 1.90$  for win-stay comparison,  $d = .482$ ;  $t(64) = 1.79$  for lose-shift,  $d = .436$ ; Figure 4A]. Thus, we observed large effect sizes for experiment 1 and medium effect sizes for experiment 2, indicating that patients still had difficulty using feedback to rapidly modify choice behavior.



**Figure 4.** Performance of subjects on two measures of feedback-driven learning from experiment 2. In both plots, black bars = control subjects, white bars = patients. **(A)** Impact of trial-by-trial task feedback on subsequent choices in a given condition in first acquisition block (20 trials in each stimulus condition). Win-stay scores reflect the proportion of repeated stimulus selections in a given condition following reinforced choices. Lose-shift scores reflect the proportion of switched stimulus selections in a given condition following nonreinforced choices. Total win-stay and lose-shift scores were generated by averaging scores across conditions for each. **(B)** Performance of 24 control subjects and 32 patients qualified for transfer analysis in the postacquisition test phase. This analysis only included subjects who demonstrated acquisition of the 80:20 contingency by choosing A on at least 75% of AB test trials, and thus, the groups showed similar performance on the AB (80:20) test pair. Go learning was assessed using novel pairs involving the 80% reinforced stimulus (choose A vs. novel), as choosing A depends on having learned from positive feedback. NoGo learning was assessed using novel pairs involving the 20% reinforced stimulus (avoid B vs. novel), as avoiding B depends on having learned from negative feedback.

## Performance on Transfer Measures

**Assessment of Procedural Go and NoGo Learning.** In experiment 2, we included 32 patients (80%) and 24 control subjects (86%) in the transfer analysis who met the 75% correct criterion on the AB test trials. A *t* test revealed that control subjects more consistently chose A (the most frequently rewarded stimulus) when presented in novel pairs than did patients [ $82 \pm 3\%$  vs.  $70 \pm 3\%$ ;  $t(54) = 2.852$ ;  $p = .01$ ; Figure 4B]. This result is consistent with our operationalization of impaired Go learning. By contrast, patients ( $70 \pm 3\%$ ) did not show a decreased avoidance of the least frequently rewarded stimulus when presented in novel pairs, when compared with control subjects [ $72 \pm 4\%$ ;  $t(54) = .397$ ], consistent with our operationalization of intact NoGo learning.

Note that the lose-shift results described above appear to contradict this evidence of intact NoGo learning. However, if one views the measures as assessments of two different kinds of feedback-driven learning, it is entirely plausible that a between-group difference might be evident in reward-driven or punishment-driven learning in one case but not the other. We interpret this result to indicate that SZ patients can gradually integrate negative outcomes to generalize and avoid poor choices over many trials (BG-dependent), whereas they are impaired at the online/cortical-dependent use of a single instance of negative feedback to modify behavior in the very next trial.

**Correlations Between Characterizing Variables and Experimental Measures.** Pearson correlation analyses between performance measures from the PSS paradigm and clinical and standard neuropsychological ratings revealed a moderate relationship between total proportion correct during acquisition phase and total score on the SANS ( $r = -.372$ ,  $p = .020$ ). Correlations between our combined measure of probabilistic selection performance and total scores on the Calgary Depression Scale ( $-.063$ ) and BPRS ( $-.161$ ) did not achieve significance. Only the negative symptom subscore of the BPRS correlated with the total proportion correct during early acquisition at the trend level ( $r = -.299$ ,  $p = .061$ ). None of the reality distortion ( $r = .067$ ), disorganization ( $r = .007$ ), or depression ( $r = -.021$ ) subscores of the BPRS showed any evidence of a systematic relationship with our combined measure of probabilistic selection performance. No correlations between PSS performance and standard neuropsychological measures were

significant, with Pearson coefficients ranging from .197 for our spatial short-term memory span measure to .242 for the Wechsler Test of Adult Reading ( $p > .10$ ). This result further suggests that patients' poor performance on the probabilistic learning task is not simply a product of impaired neuropsychological performance in general.

## Discussion

### PSS Task Performance

We examined the performance of patients and control subjects on two probabilistic learning and transfer tasks. In the first version, using Hiragana characters, patients exhibited profound impairment in the acquisition of probabilistic contingencies. This seemed to reflect impairments in the use of feedback to modify behavior on a trial-by-trial basis, consistent with models of PFC/OFC dysfunction. In the second experiment (using clip art stimuli), patients showed impairment in the early acquisition stages of the task, but demonstrated eventual learning of the easiest (80:20) discrimination. However, even the patients who learned the 80:20 discrimination showed a less robust preference than control subjects for the 80% stimulus when it was presented in new pairings. Patients exhibited normal performance in avoiding the least frequently rewarded stimulus when it was presented in novel pairings, successfully generalizing from repeated exposure to negative outcomes.

Thus, patients did not exhibit a simple failure in generalization but rather more selective difficulty in learning from positive outcomes. This dissociation cannot be easily explained by the presence of generally lower levels of neuropsychological performance in patients relative to control subjects, as no standard neuropsychological measure correlated significantly with probabilistic contingency acquisition and none of the main effects or interactions from the ANOVAs for acquisition data were modified substantially by the inclusion of WTAR scores as a covariate in ANCOVAs.

Within the context of the computational model described above, the deficit exhibited by patients may result primarily from dysfunction of the direct (Go) BG pathway linking the dorsal striatum and the globus pallidus interna, which is thought to be driven largely by activity at D1 receptors, whereas the intact NoGo learning exhibited by patients can be interpreted as

evidence of preserved function of the indirect BG pathway, which is driven largely by activity at D2 receptors (Aubert *et al.* 2000; Hernandez-Lopez *et al.* 2000). Thus, it is possible that SZ patients have a compromised ability to use dopamine bursts to drive behavior in habit-learning type tasks but a surprisingly intact ability to use momentary cessations of dopamine cell firing (dips) that may signal the absence of expected reinforcement. A possible consequence of disrupted reward-driven learning, as McClure *et al.* (2003) have noted, can be the inappropriate attribution of incentive salience to a stimulus. Consistent with this formulation, Juckel *et al.* (2006b) found evidence of reduced activity in the BG during reward processing and anticipation in unmedicated SZ patients.

The patient deficits in early learning and failure to generalize from positive feedback after training may represent two different types of DA dysfunction. In the model, cortical DA hypofunction should be expected to interfere with rapid learning of relative reward value of different responses, precisely what we observed in the first two blocks of both experiments. This behavioral finding is consistent with a broad body of evidence suggestive of PFC dysfunction in schizophrenia (Heinz *et al.* 2003; Weinberger *et al.* 2001), with D1 hypofunction widely considered to be a critical contributor to the deficit (Goldman-Rakic 1994; Weinberger 1987). The finding of a significant correlation between the early acquisition of probabilistic contingencies and ratings of negative symptoms, also thought to reflect prefrontal dysfunction (Kirkpatrick and Buchanan 1990), further supports the claim that intact PFC is critical for the rapid learning of changing reinforcement contingencies.

The impairment in Go learning observed using the transfer measure may reflect a second, albeit related, abnormality of DA function: excessive DA release in the neostriatum (Abi-Dargham *et al.* 1998; McGowan *et al.* 2004). Several reports (Bertolino *et al.* 2000; Meyer-Lindenberg *et al.* 2002), in fact, point to a systematic relationship between PFC hypofunction and striatal hyperactivity. Based on evidence that tonic DA levels regulate the level of phasic DA release via inhibitory presynaptic D2 autoreceptors (O'Donnell and Grace 1998), it has been proposed that elevated levels of tonic DA in SZ reduce the effectiveness of phasic DA signals (Bilder *et al.* 2004; Grace 2000). Such a mechanism might provide an explanation for why schizophrenia has been associated with both tonic hyperactivity in the BG in blood flow studies (Abi-Dargham *et al.* 1998; McGowan *et al.* 2004) and reduced stimulus-evoked activity in the BG during task performance in magnetic resonance imaging (MRI) studies (Juckel *et al.* 2006b; Kumari *et al.* 2002).

Given the large body of work showing impaired use of error information in SZ, what might explain the lack of a NoGo learning deficit in patients in this experiment? Many feedback-driven learning tasks, such as the WCST, require the online representation of feedback in WM and its use in resolving response conflicts, appearing to rely on regions of prefrontal cortex. The selection task used here involves learning probabilistic discriminations over many trials by integrating the overall frequency of reinforcement. Rapid updating after unexpected feedback would serve to actually impair performance. One possibility is that a dissociation exists between the ability to rapidly use single instances of negative feedback versus the more gradual learning of the most and least advantageous choices, which may reflect the properties of different learning systems as suggested by several theoretical accounts (Rolls 2004; Schoenbaum and Roesch 2005).

Importantly, our patients did not exhibit a complete inability to learn the three training pairs, especially in experiment 2 where most patients reached criterion in all three reinforcement conditions. Rather, patients showed evidence of delayed acquisition of probabilistic reward contingencies. As noted above, several findings (Keri *et al.* 2000; Weickert *et al.* 2002) indicate that SZ patients have a relatively intact ability to learn in this fashion. Importantly, these studies did not necessarily demonstrate fully normal performance compared with control subjects but only roughly normal improvement from initial, impaired performance.

How might cessations of dopamine activity still be effective in providing a learning signal in schizophrenia? One possible answer is that D2 receptors are supersensitive in patients with chronic schizophrenia (Curran *et al.* 2004; Seeman *et al.* 2005). It is also possible that the D2-blocking effects of antipsychotic medications (when chronically administered) actually increase activity and plasticity in the indirect BG pathway (Centonze *et al.* 2004), facilitating NoGo learning (Amtage and Schmidt 2003). We do not argue that D2 blockade might be beneficial for reward processing and procedural learning in general. Indeed, the results of several recent studies (Beninger *et al.* 2003; Juckel *et al.* 2006a) indicate that D2 blockade, especially by conventional antipsychotics, has an overall detrimental effect. We suggest only that blocking D2 transmission may benefit NoGo learning, thought to depend on dips in DA levels.

One important issue to address in future studies would be the effects that different antipsychotic medications have on reward-driven and punishment-driven learning. Antipsychotic medications vary widely in their affinities for different receptor types, with second-generation antipsychotics generally having weaker, lower affinities for D2 receptors than first-generation antipsychotics but greater affinities for D1 and serotonin receptors (Kapur and Seeman 2001). Unfortunately, we were not able to study medication effects in this study in a systematic way, as medication effects are fully confounded with the patient clinical characteristics that lead clinicians to choose specific drugs. For clinical historical reasons, subgroups of patients on similar medication regimens are not well matched in terms of demographics or symptom profiles.

Furthermore, less than 20% of patients at MPRC are taking one of the first-generation antipsychotics. Patients taking conventional agents have largely chosen to do so because they are doing well clinically and are unwilling to risk the instability that would follow a treatment change. Only eight of our patients were on typical neuroleptics alone, and these patients, in fact, tended to be the youngest (mean age = 39.71 ± 3.39) and most treatment-responsive patients in our sample (mean BPRS = 28.00 ± 4.59). They also experienced the least severe negative symptoms (mean SANS = 22.63 ± 5.46). Therefore, it was impossible in our study to determine whether typical neuroleptics, relative to second-generation antipsychotics, had a more severe impact on BG function, as Beninger *et al.* (2003) and Juckel *et al.* (2006a) have shown. The unique effects that individual antipsychotic medications have on aspects of reinforcement learning need to be studied in the context of controlled clinical trials.

While the study of unmedicated patients would also certainly be of interest to inform our understanding of the extent to which the deficits documented here can be attributed to illness as opposed to treatment effects, we argue that the current results observed in medicated patients are clinically relevant. Almost all patients with schizophrenia are treated with antipsychotics that block D2 receptors, and the clinical challenge facing the field is to develop novel treatment approaches to the reward-processing

deficits observed in patients receiving medications that block DA receptors.

Nonetheless, our results indicate that the primary BG-dependent learning impairment in SZ is a deficit in Go learning (learning in response to positive feedback), which may not be remedied by D2 blockade. Furthermore, the results of Frank *et al.* (2004) indicate that treatment with a dopamine precursor reverses a Go learning deficit. Given that dopamine agonists have been associated with mild improvement of various cognitive impairments in SZ (Barch and Carter 2005; Goldberg *et al.* 1991), it is possible that performance on probabilistic reinforcement learning tasks would benefit from treatment with dopamine agonists, with the risk of exacerbating positive symptoms (Levy *et al.* 1993; van Kammen *et al.* 1982). Thus, the key to adequately treating both the positive symptoms and cognitive deficits of schizophrenia may lie in the relative agonistic and antagonistic properties of drugs acting at different dopamine receptor types in the BG-PFC action selection system.

*This research was made possible by Grant # P30 MH068580-01 and Grant # 1 R24 MH72647-01A1 from the National Institutes of Mental Health. We have no conflicts of interest, financial or otherwise, to report.*

*Mary Ramsey, Pablo Diego, Sharon August, and Kimberly Warren assisted with the collection of experimental and characterizing data.*

*These data were presented, in part, at the 61st Annual Convention of the Society of Biological Psychiatry in Toronto, Ontario, Canada, on May 19, 2006.*

*Supplementary material cited in this article is available online.*

- Abi-Dargham A, Gil R, Krystal J, Baldwin RM, Seibyl JP, Bowers M, *et al.* (1998): Increased striatal dopamine transmission in schizophrenia: Confirmation in a second cohort. *Am J Psychiatry* 155:761–767.
- Addington D, Addington J, Maticka-Tyndale E, Joyce J (1992): Reliability and validity of a depression rating scale for schizophrenics. *Schizophr Res* 6:201–208.
- Amtage J, Schmidt WJ (2003): Context-dependent catalepsy intensification is due to classical conditioning and sensitization. *Behav Pharmacol* 14: 563–567.
- Andreasen NC (1984): *The Scale for the Assessment of Negative Symptoms (SANS)*. Iowa City, IA: University of Iowa.
- Aubert I, Ghorayeb I, Normand E, Bloch B (2000): Phenotypical characterization of the neurons expressing the D1 and D2 dopamine receptors in the monkey striatum. *J Comp Neurol* 418:22–32.
- Barch DM (2005): The cognitive neuroscience of schizophrenia. *Annu Rev Clin Psychol* 1:321–353.
- Barch DM, Carter CS (2005): Amphetamine improves cognitive function in medicated individuals with schizophrenia and in healthy volunteers. *Schizophr Res* 77:43–58.
- Beninger RJ, Wasserman J, Zanibbi K, Charbonneau D, Mangels J, Beninger BV (2003): Typical and atypical antipsychotic medications differentially affect two nondeclarative memory tasks in schizophrenic patients: A double dissociation. *Schizophr Res* 61:281–292.
- Bertolino A, Breier A, Callicott JH, Adler C, Mattay VS, Shapiro M, *et al.* (2000): The relationship between dorsolateral prefrontal neuronal N-acetylaspartate and evoked release of striatal dopamine in schizophrenia. *Neuropsychopharmacology* 22:125–132.
- Bilder RM, Volavka J, Lachman HM, Grace AA (2004): The catechol-O-methyltransferase polymorphism: Relations to the tonic-phasic dopamine hypothesis and neuropsychiatric phenotypes. *Neuropsychopharmacology* 29:1943–1961.
- Boettiger CA, D'Esposito M (2005): Frontal networks for learning and executing arbitrary stimulus-response associations. *J Neurosci* 25:2723–2732.
- Braver TS, Cohen JD (2000): On the control of control: The role of dopamine in regulating prefrontal function and working memory. In: Monsell S, Driver J, editors. *Control of Cognitive Processes: Attention and Performance, vol. XVIII*. Cambridge, MA: MIT Press, 713–737.
- Centonze D, Picconi B, Gubellini P, Bernardi G, Calabresi P (2001): Dopaminergic control of synaptic plasticity in the dorsal striatum. *Eur J Neurosci* 13:1071–1077.
- Centonze D, Usiello A, Costa C, Picconi B, Erbs E, Bernardi G, *et al.* (2004): Chronic haloperidol promotes corticostriatal long-term potentiation by targeting dopamine D2L receptors. *J Neurosci* 24:8214–8222.
- Cohen JD, Braver TS, O'Reilly RC (1996): A computational approach to prefrontal cortex, cognitive control and schizophrenia: Recent developments and current challenges. *Philos Trans R Soc Lond B Biol Sci* 351:1515–1527.
- Curran C, Byrappa N, McBride A (2004): Stimulant psychosis: Systematic review. *Br J Psychiatry* 185:196–204.
- First MB, Spitzer RL, Gibbon M, Williams JBW (1997): *Structured Clinical Interview for DSM-IV-Axis I Disorders (SCID-I)*. Washington, DC: American Psychiatric Press.
- Frank MJ (2005): Dynamic dopamine modulation in the basal ganglia: A neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. *J Cogn Neurosci* 17:51–72.
- Frank MJ, Claus ED (2006): Anatomy of a decision: Striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. *Psychol Rev* 113:300–326.
- Frank MJ, Loughry B, O'Reilly RC (2001): Interactions between frontal cortex and basal ganglia in working memory: A computational model. *Cogn Affect Behav Neurosci* 1:137–160.
- Frank MJ, O'Reilly RC (2006): A mechanistic account of striatal dopamine function in cognition: Psychopharmacological studies with cabergoline and haloperidol. *Behav Neurosci* 120:497–517.
- Frank MJ, O'Reilly RC, Curran T (2006): When memory fails, intuition reigns: Midazolam enhances implicit inference in humans. *Psychol Sci* 17:700–707.
- Frank MJ, Seeberger LC, O'Reilly RC (2004): By carrot or by stick: Cognitive reinforcement learning in parkinsonism. *Science* 306:1940–1943.
- Gold JM, Bish JA, Iannone VN, Hobart MP, Queern CA, Buchanan RW (2000): Effects of contextual processing on visual conditional associative learning in schizophrenia. *Biol Psychiatry* 48:406–414.
- Goldberg TE, Bigelow LB, Weinberger DR, Daniel DG, Kleinman JE (1991): Cognitive and behavioral effects of the coadministration of dextroamphetamine and haloperidol in schizophrenia. *Am J Psychiatry* 148:78–84.
- Goldberg TE, Weinberger DR, Berman KF, Pliskin NH, Podd MH (1987): Further evidence for dementia of the prefrontal type in schizophrenia? A controlled study of teaching the Wisconsin Card Sorting Test. *Arch Gen Psychiatry* 44:1008–1014.
- Goldman-Rakic PS (1994): Working memory dysfunction in schizophrenia. *J Neuropsychiatry Clin Neurosci* 6:348–357.
- Grace AA (2000): The tonic/phasic model of dopamine system regulation and its implications for understanding alcohol and psychostimulant craving. *Addiction* 95(suppl 2):S119–S128.
- Heinz A, Romero B, Gallinat J, Juckel G, Weinberger DR (2003): Molecular brain imaging and the neurobiology and genetics of schizophrenia. *Pharmacopsychiatry* 36(suppl 3):S152–S157.
- Hernandez-Lopez S, Tkatch T, Perez-Garci E, Galarraga E, Bargas J, Hamm H, *et al.* (2000): D2 dopamine receptors in striatal medium spiny neurons reduce L-type Ca<sup>2+</sup> currents and excitability via a novel PLC[ $\beta$ 1-IP3-calcineurin-signaling cascade. *J Neurosci* 20:8987–8995.
- Juckel G, Schlagenhauf F, Koslowski M, Filonov D, Wustenberg T, Villringer A, *et al.* (2006a): Dysfunction of ventral striatal reward prediction in schizophrenic patients treated with typical, not atypical, neuroleptics. *Psychopharmacology (Berl)* 187:222–228.
- Juckel G, Schlagenhauf F, Koslowski M, Wustenberg T, Villringer A, Knutson B, *et al.* (2006b): Dysfunction of ventral striatal reward prediction in schizophrenia. *Neuroimage* 29:409–416.
- Kapur S, Seeman P (2001): Does fast dissociation from the dopamine D(2) receptor explain the action of atypical antipsychotics? A new hypothesis. *Am J Psychiatry* 158:360–369.
- Keri S, Kelemen O, Szekeres G, Bagoczky N, Erdelyi R, Antal A, *et al.* (2000): Schizophrenics know more than they can tell: Probabilistic classification learning in schizophrenia. *Psychol Med* 30:149–155.
- Keri S, Nagy O, Kelemen O, Myers CE, Gluck MA (2005): Dissociation between medial temporal lobe and basal ganglia memory systems in schizophrenia. *Schizophr Res* 77:321–328.

- Kern RS, Green MF, Wallace CJ (1997): Declarative and procedural learning in schizophrenia: A test of the integrity of divergent memory systems. *Cogn Neuropsychiatry* 2:39–50.
- Kirkpatrick B, Buchanan RW (1990): The neural basis of the deficit syndrome of schizophrenia. *J Nerv Ment Dis* 178:545–555.
- Knowlton BJ, Mangels JA, Squire LR (1996): A neostriatal habit learning system in humans. *Science* 273:1399–1402.
- Kumari V, Gray JA, Honey GD, Soni W, Bullmore ET, Williams SC, *et al.* (2002): Procedural learning in schizophrenia: A functional magnetic resonance imaging investigation. *Schizophr Res* 57:97–107.
- Levy DL, Smith M, Robinson D, Jody D, Lerner G, Alvir J, *et al.* (1993): Methylphenidate increases thought disorder in recent onset schizophrenics, but not in normal controls. *Biol Psychiatry* 34:507–514.
- McClure SM, Daw ND, Montague PR (2003): A computational substrate for incentive salience. *Trends Neurosci* 26:423–428.
- McGowan S, Lawrence AD, Sales T, Quedsted D, Grasby P (2004): Presynaptic dopaminergic dysfunction in schizophrenia: A positron emission tomographic [18F]fluorodopa study. *Arch Gen Psychiatry* 61:134–142.
- McMahon RP, Kelly DL, Kreyenbuhl J, Kirkpatrick B, Love RC, Conley RR (2002): Novel factor-based symptom scores in treatment resistant schizophrenia: Implications for clinical trials. *Neuropsychopharmacology* 26:537–545.
- Meyer-Lindenberg A, Miletich RS, Kohn PD, Esposito G, Carson RE, Quarantelli M, *et al.* (2002): Reduced prefrontal activity predicts exaggerated striatal dopaminergic function in schizophrenia. *Nat Neurosci* 5:267–271.
- Nishi A, Snyder GL, Greengard P (1997): Bidirectional regulation of DARPP-32 phosphorylation by dopamine. *J Neurosci* 17:8147–8155.
- O'Donnell P, Grace AA (1998): Dysfunctions in multiple interrelated systems as the neurobiological bases of schizophrenic symptom clusters. *Schizophr Bull* 24:267–283.
- O'Reilly RC, Frank MJ (2006): Making working memory work: A computational model of learning in the frontal cortex and basal ganglia. *Neural Comput* 18:283–328.
- Overall JE, Gorman DR (1962): The Brief Psychiatric Rating Scale. *Psychol Rep* 10:799–812.
- Pfohl B, Blum N, Zimmerman M, Stangl D (1989): *Structured Interview for DSM-III-R Personality Disorders (SIDP-R)*. Iowa City, IA: University of Iowa, Department of Psychiatry.
- Rolls ET (2004): Convergence of sensory systems in the orbitofrontal cortex in primates and brain design for emotion. *Anat Rec A Discov Mol Cell Evol Biol* 281:1212–1225.
- Rolls ET, O'Doherty J, Kringelbach ML, Francis S, Bowtell R, McGlone F (2003): Representations of pleasant and painful touch in the human orbitofrontal and cingulate cortices. *Cereb Cortex* 13:308–317.
- Schoenbaum G, Roesch M (2005): Orbitofrontal cortex, associative learning, and expectancies. *Neuron* 47:633–636.
- Seeman P, Weinschenker D, Quirion R, Srivastava LK, Bhardwaj SK, Grandy DK, *et al.* (2005): Dopamine supersensitivity correlates with D2High states, implying many paths to psychosis. *Proc Natl Acad Sci U S A* 102:3513–3518.
- Seger CA, Cincotta CM (2005): The roles of the caudate nucleus in human classification learning. *J Neurosci* 25:2941–2951.
- van Kammen DP, Bunney WE Jr, Docherty JP, Marder SR, Ebert MH, Rosenblatt JE, *et al.* (1982): d-Amphetamine-induced heterogeneous changes in psychotic behavior in schizophrenia. *Am J Psychiatry* 139:991–997.
- Weickert TW, Terrazas A, Bigelow LB, Malley JD, Hyde T, Egan MF, *et al.* (2002): Habit and skill learning in schizophrenia: Evidence of normal striatal processing with abnormal cortical input. *Learn Mem* 9:430–442.
- Weinberger DR (1987): Implications of normal brain development for the pathogenesis of schizophrenia. *Arch Gen Psychiatry* 44:660–669.
- Weinberger DR, Berman KF (1988): Speculation on the meaning of cerebral metabolic hypofrontality in schizophrenia. *Schizophr Bull* 14:157–168.
- Weinberger DR, Berman KF, Zec RF (1986): Physiologic dysfunction of dorsolateral prefrontal cortex in schizophrenia. I. Regional cerebral blood flow evidence. *Arch Gen Psychiatry* 43:114–124.
- Weinberger DR, Egan MF, Bertolino A, Callicott JH, Mattay VS, Lipska BK, *et al.* (2001): Prefrontal neurons and the genetics of schizophrenia. *Biol Psychiatry* 50:825–844.
- Wilk CM, Gold JM, McMahon RP, Humber K, Iannone VN, Buchanan RW (2005): No, it is not possible to be schizophrenic yet neuropsychologically normal. *Neuropsychology* 19:778–786.