

## Mild Reinforcement Learning Deficits in Patients With First-Episode Psychosis

Wing Chung Chang<sup>\*,1,2,4</sup>, James A. Waltz<sup>3,4</sup>, James M. Gold<sup>3</sup>, Tracey Chi Wan Chan<sup>1</sup>, and Eric Yu Hai Chen<sup>1,2</sup>

<sup>1</sup>Department of Psychiatry, The University of Hong Kong, Queen Mary Hospital, Pok Fu Lam, Hong Kong; <sup>2</sup>State Key Laboratory of Brain and Cognitive Sciences, The University of Hong Kong, Pok Fu Lam, Hong Kong; <sup>3</sup>Maryland Psychiatric Research Center, Department of Psychiatry, University of Maryland School of Medicine, Baltimore, MD

<sup>4</sup>These authors contributed equally to the article.

\*To whom correspondence should be addressed; Department of Psychiatry, The University of Hong Kong, Queen Mary Hospital, 102 Pok Fu Lam Rd, Pok Fu Lam, Hong Kong; tel: 852-22554486, fax: 852-28551345, e-mail: [changwc@hku.hk](mailto:changwc@hku.hk)

Numerous studies have identified reinforcement learning (RL) deficits in schizophrenia. Most have focused on chronic patients with longstanding antipsychotic treatment, however, and studies of RL in early-illness patients have produced mixed results, particularly regarding gradual/procedural learning. No study has directly contrasted both rapid and gradual RL in first-episode psychosis (FEP) samples. We examined probabilistic RL in 34 FEP patients and 36 controls, using Go/NoGo (GNG) and Gain vs Loss-Avoidance (GLA) paradigms. Our results were mixed, with FEP patients exhibiting greater impairment in the ability to use positive, as opposed to negative, feedback to drive rapid RL on the GLA, but not the GNG. By contrast, patients and controls showed similar improvement across the acquisition. Finally, we found no significant between-group differences in the postacquisition expression of value-based preference in both tasks. Negative symptoms were modestly associated with RL measures, while the overall bias to engage in Go-responding correlated significantly with psychosis severity in FEP patients, consistent with striatal hyperdopaminergia. Taken together, FEP patients demonstrated more circumscribed RL impairments than previous studies have documented in chronic samples, possibly reflecting differential symptom profiles between first-episode and chronic samples. Our finding of relatively preserved gradual/procedural RL, in briefly medicated FEP patients, might suggest spared or restored basal ganglia function. Our findings of preserved abilities to use representations of expected value to guide decision making, and our mixed results regarding rapid RL, may reflect a lesser degree of prefrontal cortical functional impairment in FEP than in chronic samples. Further longitudinal research, in larger samples, is required.

**Key words:** decision making/psychosis/prefrontal cortex/basal ganglia/dopamine/prediction error

### Introduction

Recent years have seen increased interest in mechanisms of reinforcement learning (RL) in schizophrenia (SZ),<sup>1–3</sup> based on evidence that dopamine pathways<sup>4,5</sup> and frontostriatal circuits<sup>6</sup> are both critically involved in RL and implicated in the pathophysiology of SZ.<sup>7–9</sup> Complicating the picture, however, findings point to a *variety* of RL deficits in SZ,<sup>1–3</sup> with different aspects of RL likely drawing on different neural substrates and relating to different symptom dimensions. The capacity for rapid/explicit RL, eg, relies heavily on fronto-hippocampal circuits, perhaps centered on orbitofrontal cortex,<sup>10–12</sup> whereas gradual/procedural learning, and the expression of acquired habits, depends on the ability to learn from repeated response sequences, or from repeated receipt of probabilistic feedback, and is thought to rely, predominantly, on the basal ganglia (BG).<sup>13</sup> Furthermore, both rapid/explicit and gradual/procedural forms of RL can be driven by either rewards or punishments or both.<sup>6,14</sup>

Behavioral studies have revealed impairment in rapid trial-to-trial learning of reinforcement contingencies in SZ,<sup>1</sup> especially in patients with more severe negative symptoms.<sup>15,16</sup> Findings regarding gradual/procedural learning in SZ, using paradigms such as serial reaction time (SRT) tasks<sup>17,18</sup> and probabilistic discrimination learning,<sup>19–22</sup> are more mixed. Several studies indicate that SZ patients may demonstrate compromised reward-driven learning, in the presence of relatively intact punishment-driven learning.<sup>15,16,23,24</sup> Functional neuroimaging studies have confirmed a role for frontostriatal dysfunction in the RL deficits observed in SZ, especially regarding reward anticipation<sup>25–28</sup> and reward prediction error (RPE) signaling.<sup>29–32</sup>

Of note, the majority of previous RL studies in SZ have focused on patients with chronic illness, where it is difficult to address potentially adverse consequences of prolonged treatment exposure and illness chronicity.

However, at least 10 studies have examined RL in first-episode psychosis (FEP). Four studies probed rapid RL using reversal learning or related paradigms, and all found evidence of impairment in FEP patients,<sup>33–36</sup> suggesting deficits in the use of feedback to make rapid behavioral adjustments. Five studies examined gradual/procedural RL using SRT or related tasks, with 3 demonstrating significantly poorer learning in patients, relative to controls.<sup>37–39</sup> Two neuroimaging studies reported a lack of between-group differences in procedural RL,<sup>29,40</sup> but the goal of these studies was likely to determine whether group differences in neural responses would be observed in the presence of similar behavioral performance. Finally, a study using a causal learning paradigm to examine PE signaling<sup>41</sup> also found that patients exhibited intact performance but abnormal neural correlates of learning.

To date, no study has investigated RL with the goal of distinguishing between rapid/explicit and gradual/procedural RL, and between reward- and punishment-driven RL. Furthermore, no FEP study has investigated RL with the goal of distinguishing between the acquisition of reward contingencies and the successful use of value representations, established over the course of learning, in decision making. In the current study, we sought to do this in a representative cohort of Chinese patients with clinically stabilized FEP, using 2 probabilistic RL paradigms: the Go/NoGo (GNG) task and the Gain vs Loss-Avoidance (GLA) task. These tasks confer several advantages over RL paradigms used to date. First, in each task, both rapid and gradual RL can be examined. Second, both tasks allow one to assess learning from both positive and negative feedback. Moreover, these tasks allowed us to investigate 2 additional aspects of RL not previously assessed in FEP. The GNG task enables us to quantify the general tendency of participants to make (vs withhold) responses, irrespective of reward contingencies.<sup>16</sup> This behavioral index (termed “Go-response bias”) has been hypothesized to result from high tonic dopamine levels in the striatum.<sup>42</sup> The GLA task allows one to dissociate contributions that expected value (EV) representations make to RL, from those made by RPE signaling.<sup>43</sup> In this task, selection of the correct item from *some* stimulus-pairs (gains vs neutral outcomes) leads to reward; in other pairs (losses vs neutral outcomes), correct-response selection results in loss-avoidance. Both types of correct selections should lead to positive RPEs, but some selections are associated with positive EV, others with negative EV. Examining participants’ preferences between stimuli with different EVs, but identical frequencies of positive RPEs, in a Test/transfer phase following Acquisition, offers a critical evaluation of impairment in EV representation.

Previous studies have shown evidence of impairment on these tasks in chronic SZ.<sup>16,43</sup> In a study using the GNG task,<sup>16</sup> chronic patients showed attenuations in both rapid punishment-driven RL and gradual reward-driven

learning, despite higher overall levels of Go-responding (a Go-response bias). Furthermore, measures of all 3 of these constructs tracked with negative symptom severity. Our prior study using the GLA task<sup>43</sup> revealed that chronic patients with the highest negative symptoms showed no preference for frequently rewarded stimuli (positive EV) over stimuli frequently associated with loss-avoidance (negative EV). These findings, and others from our group,<sup>15,23,44</sup> have documented the impact of negative symptom severity on aspects of RL and decision making, suggesting that this symptomatic feature of the illness, rather than the diagnosis per se, may be primarily implicated in the extent of RL deficits.

Because many of our previous results have been related to negative symptom severity, it is difficult to make strong predictions about the overall performance of a FEP sample, where we would expect prominent positive symptoms, but less-severe negative symptoms than those observed in prior studies of chronic patients at the Maryland Psychiatric Research Center (MPRC). Thus, we predicted that FEP patients would demonstrate relatively mild deficits, when compared with controls, in overall RL (learning across the Acquisition phase). However, we expected that patients with the most severe negative symptoms would show deficits in rapid RL, and in EV representation. Based on computational models, as well as evidence indicating that acute psychosis is accompanied by striatal hyperdopaminergia,<sup>7</sup> we predicted that a Go-response bias on the GNG task would be characteristic of FEP patients with the most severe *positive* symptoms.

## Methods

### Participants

Thirty-four patients in their first psychotic episode, aged 15–40 years, were recruited from the outpatient unit of a specialized early intervention service for FEP in Hong Kong (HK). Diagnosis was ascertained at intake using the Chinese-bilingual Structured Clinical Interview for DSM-IV (CB-SCID-I/P)<sup>45</sup> and revisited after clinical stabilization (28 received a DSM-IV<sup>46</sup> diagnosis of SZ and 6 received a diagnosis of schizophreniform disorder). Study assessments were administered to patients 3–7 months following antipsychotic initiation (mean: 161.9 d; SD: 34.2 d), at which point all patients had been on stable antipsychotic regimens for at least 4 weeks (most with second-generation antipsychotic (SGA) monotherapy; see [table 1](#) and [supplementary table S1](#) for details).

Thirty-six healthy controls were recruited from the community via advertisements and word-of-mouth among recruited participants. Patients and controls were matched for age, gender, and educational level. Controls were screened to confirm that they had no psychiatric diagnosis (by CB-SCID-I/P), family history of psychotic disorder, and were not taking any psychotropic medications.

**Table 1.** Demographics, Cognitive Functions, Clinical and Treatment Characteristics of Patients and Controls

Variables <sup>a</sup>	GNG Task		<i>P</i>	GLA Task		<i>P</i>
	Patients ( <i>n</i> = 31)	Controls ( <i>n</i> = 33)		Patients ( <i>n</i> = 31)	Controls ( <i>n</i> = 33)	
<b>Demographics</b>						
Age	24.8 (7.4)	23.7 (7.5)	.54	24.1 (7.5)	23.7 (7.5)	.83
Male gender, <i>n</i> (%) <sup>b</sup>	13 (41.9)	16 (48.5)	.68	12 (38.7)	14 (42.4)	.76
Nicotine-dependent, <i>n</i> (%)	4 (12.9)	3 (9.1)	.70	4 (12.9)	3 (9.1)	.70
Years of education	12.5 (2.7)	12.7 (2.7)	.75	12.2 (2.6)	12.8 (2.6)	.43
<b>Cognitive function</b>						
Letter number span	12.9 (3.7)	15.6 (4.4)	<.01	12.9 (3.7)	16.0 (3.0)	<.01
Visual pattern	19.7 (3.9)	24.7 (4.7)	<.01	20.4 (16.3)	24.9 (4.7)	<.01
Logical memory	10.9 (4.4)	12.6 (4.4)	.12	10.9 (4.2)	12.6 (4.4)	.11
Visual reproduction	20.9 (1.2)	21.2 (2.0)	.52	21.0 (1.2)	21.1 (2.1)	.69
Monotone counting	11.4 (1.8)	12.0 (0.2)	.08	11.4 (1.8)	12.0 (0.2)	.08
Intelligence estimate	96.8 (12.1)	114.3 (11.9)	<.001	96.4 (11.2)	114.4 (11.3)	<.001
<b>Clinical characteristics</b>						
Age at onset	23.7 (7.2)	N/A		22.9 (7.3)	N/A	
DUP in days (median) <sup>c</sup>	123	N/A		123	N/A	
Diagnoses	25 SZ, 6 SP	N/A		25 SZ, 6 SP	N/A	
PANSS total score	41.6 (8.7)	N/A		41.9 (8.5)	N/A	
SANS total score	14.2 (13.1)	N/A		11.2 (12.4)	N/A	
CDS total score	1.4 (2.6)	N/A		1.6 (2.8)	N/A	
<b>Treatment characteristics</b>						
Haloperidol equivalents <sup>d</sup>	3.7 (2.5)	N/A		3.7 (2.6)	N/A	
Antipsychotic regimen, <i>n</i> (%)						
FGA monotherapy	2 (6.5) <sup>e</sup>	N/A		2 (6.5) <sup>e</sup>	N/A	
SGA monotherapy	25 (80.6) <sup>f</sup>	N/A		25 (80.6) <sup>g</sup>	N/A	
Combined antipsychotics	4 (12.9) <sup>h</sup>	N/A		4 (12.9) <sup>h</sup>	N/A	

*Note:* CDS, Calgary Depression Scale; DUP, duration of untreated psychosis; FGA, first-generation antipsychotic; GNG, Go/NoGo; GLA, Gain vs Loss-Avoidance; PANSS, Positive and Negative Syndrome Scale; SANS, Scale for the Assessment of Negative Symptoms; SGA, second-generation antipsychotic; SP, schizophreniform disorder; SZ, schizophrenia.

<sup>a</sup>Variables were presented in mean and SD except gender, antipsychotic regimen, and DUP. Gender and antipsychotic regimen were presented in number and percentage, while DUP was presented in median due to its skewed distribution.

<sup>b</sup>Independent-samples *t*-tests were performed for patient-control comparison in demographics and cognitive functions except in the case of assessing between-group differences in gender breakdown and smoker/nonsmoker breakdown. These differences were examined using Chi-square tests.

<sup>c</sup>DUP was measured by the Interview for Retrospective Assessment of the Onset of Schizophrenia (IRAOS).<sup>64</sup>

<sup>d</sup>Haloperidol equivalents were computed according to the method of Andreasen et al (2010).<sup>56</sup>

<sup>e</sup>1 on haloperidol and 1 on flupenthixol depot injection.

<sup>f</sup>14 on risperidone, 4 on amisulpride, 4 on olanzapine, 2 on quetiapine, and 1 on aripiprazole.

<sup>g</sup>12 on risperidone, 5 on quetiapine, 4 on amisulpride, 3 on olanzapine, and 1 on aripiprazole.

<sup>h</sup>2 on risperidone and quetiapine, and another 2 on risperidone and aripiprazole.

The study was approved by the local institutional review boards, and all participants provided written informed consent. Any individual showing evidence of substance abuse (according to the Alcohol Use Scale and the Drug Use Scale<sup>47</sup>), intellectual disability, or neurological disease was excluded from participation.

### Clinical and Cognitive Assessments

Positive and disorganization symptoms were assessed using the Positive and Negative Syndrome Scale (PANSS).<sup>48</sup> Based on our previous work, negative symptom severity was quantified using the Scale for the Assessment of Negative Symptoms (SANS).<sup>49</sup> As

evidence suggested that the amotivation subdomain of negative symptoms may be specifically linked to RL impairment,<sup>3</sup> we computed an amotivation score by summing all items from the Avolition/Apathy and Anhedonia/Asociality subscales of the SANS, excluding global items. Depression was evaluated using the Calgary Depression Scale (CDS).<sup>50</sup> Cognitive assessments, comprising letter-number span (working memory measure),<sup>51</sup> visual pattern test,<sup>52</sup> monotone counting,<sup>53</sup> and logical memory and visual reproduction subtests from the Wechsler Memory Scale-Revised (WMS-R)<sup>54</sup> were administered to all participants. The 3-subtest short-form<sup>55</sup> of the WAIS-III was used to generate IQ estimates. Controls had higher IQs and performed better than patients on the letter-number

span and visual pattern test (table 1). No group differences were observed on other cognitive tests.

*RL Tasks*

Two computerized probabilistic RL paradigms were administered to each participant. Details of these RL tasks have been described in previous reports,<sup>16,43</sup> and the same task parameters were used in the current study, in the case of each paradigm (see [supplementary materials](#)). Participants were compensated HK\$100 (US\$13) for completion of the study and could earn bonuses of up to HK\$200 (US\$26), based on RL task performance. The range of compensation for participation was roughly on par with that of previous studies conducted at the MPRC,<sup>16,43</sup> representing about 2% of the median monthly income in HK.

*Outcome Measures and Statistical Analysis*

In each of the 2 experiments, data from 3 patients and 3 controls who did not appear to understand or engage with the task according to the tester, and responded with random button-pressing, were removed from analysis.

We operationalized 4 constructs, in the context of our RL tasks: (1) an overall bias to make Go-responses on the GNG task, (2) rapid/explicit RL, (3) overall RL (thought to involve a combination of explicit and procedural mechanisms), and (4) the expression of postacquisition value-based preference in the absence of feedback (see [table 2](#)). In the context of the GLA task, we also assessed the relative benefit of positive EV vs negative EV during Acquisition by computing gain-pair vs loss-avoidance-pair difference scores for each individual. Computation of RL measures and statistical analyses conducted in the

**Table 2.** Hypotheses, Reinforcement Learning Constructs and Corresponding Dependent Variables Based on the GNG and the GLA Tasks Studied in First-Episode Psychosis Patients

	Reinforcement Learning Constructs and Hypotheses	Results from the Current FEP sample
Construct 1 Hypothesis	Go-Response Bias <sup>16</sup> Will correlate with positive symptom severity	- No Go-response bias was observed in the overall patient sample. - Overall Go-rate correlated with PANSS positive symptom score.
Measure	Overall Go-rate (GNG)	
Construct 2 Hypothesis	Rapid Reinforcement Learning <sup>15,16</sup> Will be disrupted in FEP patients	- Patients had lower win-stay rate on the GLA task than controls. - GLA win-stay rate correlated with SANS total score. - No rapid RL impairment was observed on the GNG task.
Measures	Will correlate with negative symptom severity Lose-shift rate (GNG & GLA) Win-stay rate (GNG & GLA)	
Construct 3 Hypothesis	Overall Reinforcement Learning <sup>15,16,43</sup> Will be mildly impaired in FEP patients	- Patients performed worse than controls on the positive stimuli in the GNG task and both positive (gain) and negative (loss-avoidance) stimuli in the GLA task. - Patients showed longer response latencies to positive stimuli than controls in the GNG task, but no deficit in RT acceleration to positive stimuli. - Overall RL measures did not correlate with negative symptoms
Measures	Will correlate with negative symptom severity Acquisition RT acceleration to positive stimuli (GNG) Learning across Acquisition blocks (GNG & GLA)	
Construct 4 Hypothesis	Value-guided Choice <sup>16,43</sup> Will be mildly impaired in FEP patients	- Comparable performance between patients and controls on value-guided decision-making in both GNG and GLA tasks. - Patients' [FP – NNeu] contrast values correlated with SANS amotivation scores.
Measures	Will correlate with negative symptom severity [FP – NNeu] contrast and [NP – NNeu] contrast at Test/transfer phase (GNG) [FW – FLA] preference at Test/transfer phase (GLA)	

*Note:* FLA, frequent loss-avoider; FP, familiar-positive; FW, frequent winner; GLA, gain vs loss-avoidance; GNG, Go/NoGo; NNeu, novel-neutral; NP, novel-positive; PANSS, Positive and Negative Syndrome Scale; POS, positive symptom dimension score; RL, reinforcement learning; RT, response time; SANS, Scale of the Assessment of Negative Symptoms.

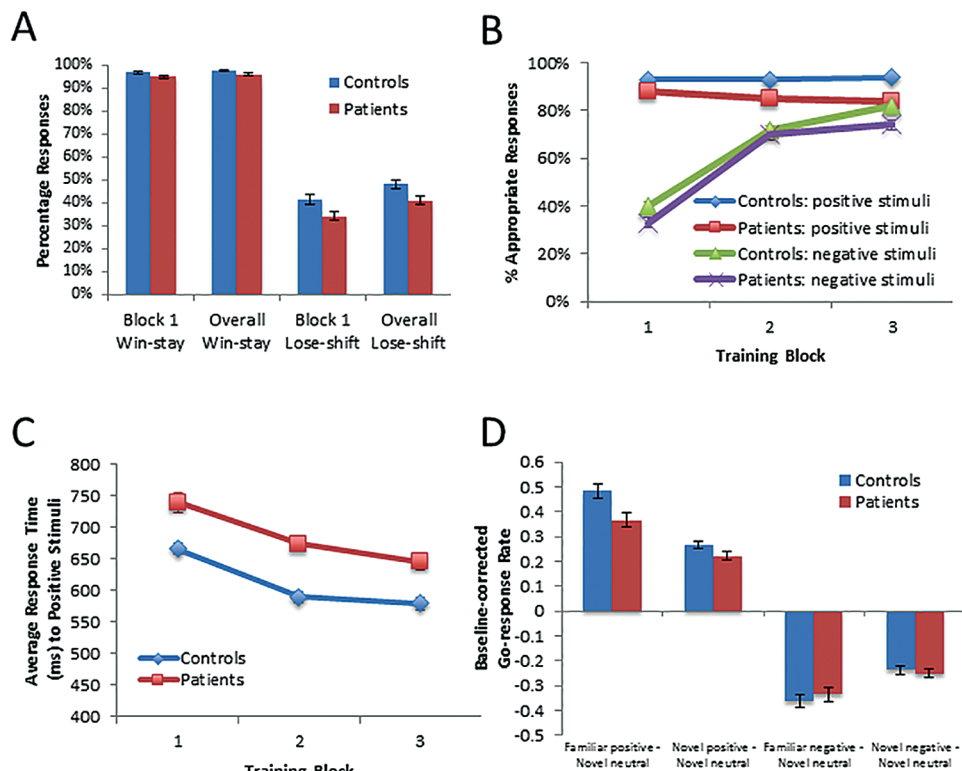
current study followed the methods adopted by previous reports<sup>16,43</sup> (see [supplementary materials](#)). For the hypothesized between-group comparisons (eg, win-stay and lose-shift scores in both tasks, and overall Go-response rates in the GNG task), we report *t*-values, *P*-values, estimates of effect sizes (ES), and Bayes factors which indicate the likelihood of accepting the null hypothesis of no group difference (see [supplementary table S2](#)). In light of our small sample size, which limited power to detect interaction effects, we also did post hoc *t*-tests and calculated ES for variables of key theoretical interest, for descriptive purposes. Our overall interpretations were constrained by the ANOVA results. Correlation analyses were conducted to assess the relationships of RL performance with measures of symptom severity, cognitive functions, and antipsychotic dose.<sup>56</sup> We did not apply multiple-comparison correction to correlations that were hypothesis-driven ([table 2](#)), but did so in the case of correlations that were not hypothesis-driven.

**Results**

*Performance on the GNG Task*

We observed no significant between-group difference in the “Go-response bias”, as assessed by participants’ overall rate of Go-responding ( $t_{44.8} = 0.04, P = .97,$

$D = 0.01$ ). Concerning rapid RL measures, there were no significant between-group differences in either block 1 win-stay ( $t_{62} = -1.2, P = .25, D = 0.20$ ) or lose-shift scores ( $t_{62} = -1.3, P = .21, D = 0.35$ ), but tests of group-differences in win-stay and lose-shift scores across the Acquisition phase revealed small-to-medium ES ( $t_{62} = -1.5, P = .14, D = 0.40$  for win-stay rates;  $t_{62} = -1.5, P = .13, D = 0.40$  for lose-shift rates; [figure 1A](#); [supplementary table S2](#)). Regarding overall RL, patients exhibited significantly lower accuracy rates than controls during Acquisition ([figure 1B](#)). This was confirmed by a repeated-measures ANOVA, which revealed significant main effects of group ( $F_{1,62} = 9.7, P = .003$ ; controls performing better than patients), block ( $F_{1.7,107.3} = 86.3, P < .001$ ; performance improving over time), and valence ( $F_{1,62} = 115.6, P < .001$ ; performance better with positive than negative stimuli). A significant valence  $\times$  block interaction was also observed ( $F_{2,124} = 64.3, P < .001$ ). This is the result of accuracy for negative stimuli rapidly increasing after block 1, whereas performance for positive stimuli was largely stable across blocks. None of the other interaction effects approached significance (all  $F$ s  $< 1$ ). To better understand the nature of the main effect of group, we examined overall performance on the gain and loss stimuli. Controls performed much better than patients on the gain stimuli ( $t_{62} = -2.8, P = .01, D = 0.80,$



**Fig. 1.** Performance of patients and controls on the Go/NoGo (GNG) task. (A) Win-stay and lose-shift scores from the first block and across Acquisition; (B) Performance across the Acquisition phase, expressed in percentages of appropriate responses; (C) Average response-time to positive stimuli across Acquisition; (D) Go-response rates to familiar and novel stimuli in the Test/transfer phase, controlling for Go-response rates to neutral stimuli.

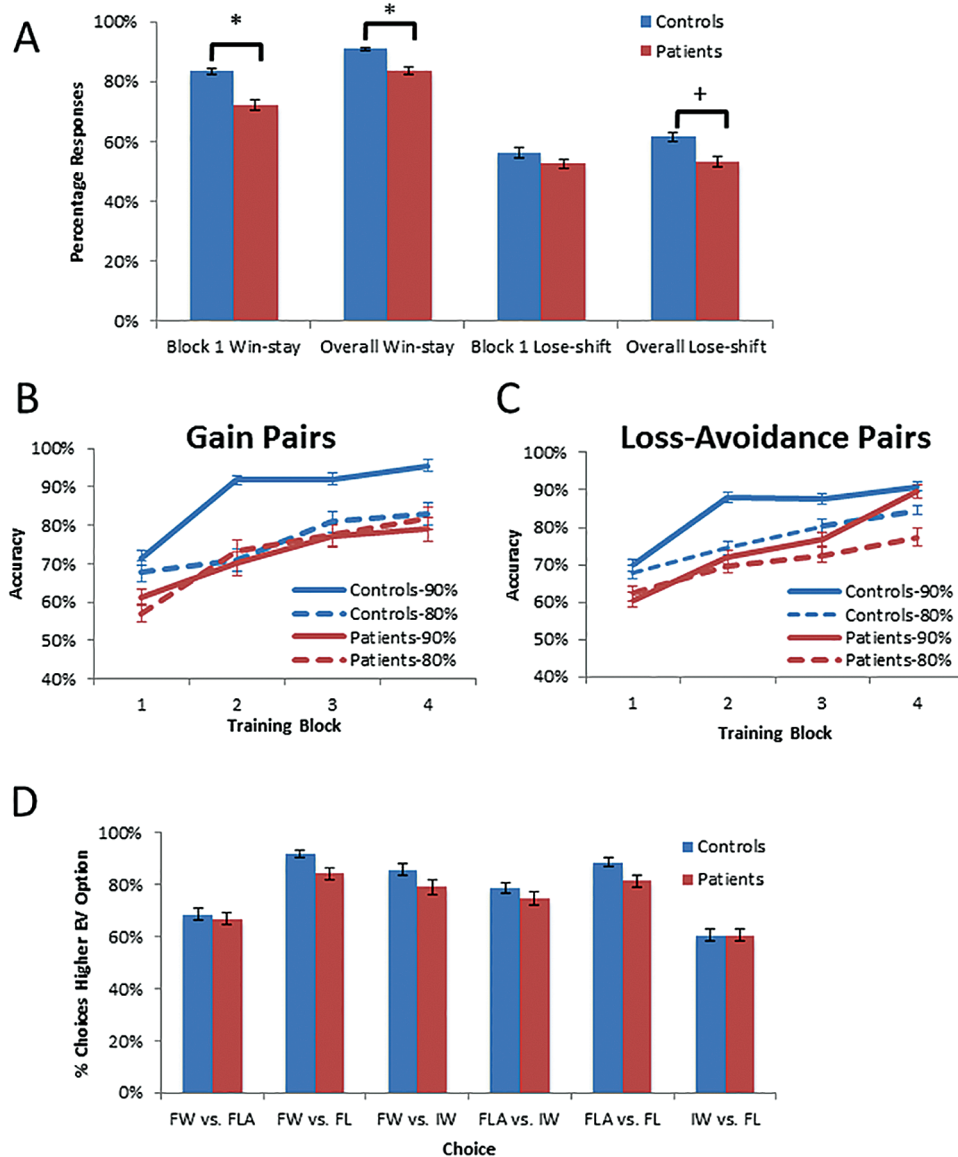
and numerically better than patients on the loss stimuli ( $t_{62} = -1.4, P = .16, D = 0.38$ ). While these ES appear to be of somewhat different magnitudes, this difference was not substantial enough to yield a significant group  $\times$  valence interaction. Regarding Go-response latencies to positive stimuli, we observed a main effect of group ( $F_{1,62} = 8.4, P = .005$ ; patients slower than controls) and a significant main effect of block ( $F_{2,124} = 17.5, P < .001$ ; figure 1C), but we found no significant group  $\times$  block interaction ( $F_{2,124} = 0.2, P = .86$ ). There was no group-difference on response-time acceleration to positive stimuli ( $t_{62} = 0.52, P = .61$ ).

Finally, concerning the postacquisition of value-based preferences (as assessed by Go-response rates to Test/transfer stimuli), a 2-way ANOVA revealed a main effect

of trial-type ( $F_{3,186} = 321.6, P < .001$ ) and a trend toward group  $\times$  trial-type interaction ( $F_{3,186} = 2.2, P = .088$ ), but no main effect of group ( $F_{1,62} = 0.6, P = .44$ ; figure 1D). These results thus indicated comparable performance between patients and controls on value-guided decision making (see [supplementary materials](#) for additional exploratory analyses).

#### Performance on the GLA Task

On the GLA tasks, patients showed evidence of impaired rapid RL, in that they displayed significantly lower win-stay rates in both block 1 and across the Acquisition phase than controls (both with  $t_{62} = -2.9, P = .01, D = 0.70$ ; figure 2A). Group-differences in lose-shift rates were



**Fig. 2.** Performance of patients and controls on the gain vs loss-avoidance (GLA) task. (A) Win-stay and lose-shift scores from the first block and across Acquisition; (B) Performance across Acquisition blocks: Gain/Miss pairs; (C) Performance across Acquisition blocks: Loss/Avoid pairs; (D) Rates of choosing the option with higher EV from novel stimulus-pairings in the Test/transfer phase. \*  $P < .05$ .

small-to-medium ( $t_{62} = -0.8, P = .42, D = 0.15$  for block 1;  $t_{62} = -1.8, P = 0.07, D = 0.45$  for the Acquisition phase). Patients again showed evidence of a deficit in overall RL, as a repeated-measures ANOVA revealed significant main effects of group ( $F_{1,62} = 7.7, P = .007$ ; patients performing worse than controls), block ( $F_{2,6,158.9} = 66.7, P < .001$ ; performance improving over time), and probability ( $F_{1,62} = 5.9, P = .02$ ; better performance with 90% than 80% pairs; see [figure 2](#), Panels B and C, for overall RL patterns across blocks). The main effect of valence was not significant ( $F_{1,62} = 0.03, P = .86$ ), and there were no significant higher-order interactions, although group  $\times$  probability interaction trended toward significance ( $F_{1,62} = 2.8, P = .10$ ). Patients showed somewhat more substantial impairment with the 90% stimuli ( $t_{39,6} = -3.5, P = .001, D = 0.89$ ) than the 80% stimuli ( $t_{62} = -1.1, P = .27, D = 0.29$ ; [figure 2](#)), an effect we observed in our prior study of chronic patients. Again, to better understand the nature of the main effect of group, we examined overall performance on the gain and loss-avoidance stimuli. There was no evidence of differential impairment, as controls performed better than patients on both the gain ( $t = -2.1, P = .04, D = -0.63$ ), and the loss-avoidance stimuli ( $t = -2.6, P = .01, D = 0.70$ ). Comparable GLA difference scores during Acquisition were observed between patients and controls ( $t_{62} = 0.41, P = .68$ ). We also failed to find significant group differences regarding stimulus preference in the Test/transfer phase ([figure 2D](#)), including between the gain and loss-avoidance stimuli that are critical for evaluating the role of EV in guiding decisions. Thus, despite impaired performance during Acquisition, patients were able to use their experience with the stimulus items to develop similar preferences, relative to controls, when faced with novel stimulus-pairings.

#### Relationships Between RL Measures and Clinical Variables

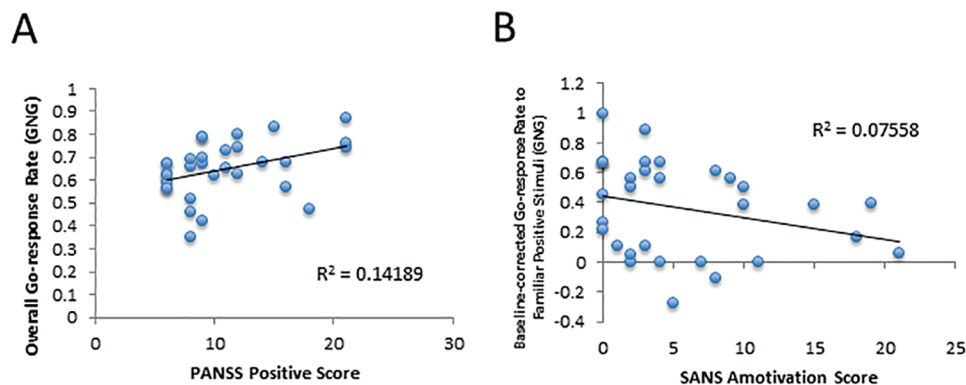
We observed a significant positive correlation between PANSS positive symptom scores and overall Go-response rates from the GNG task ([supplementary table S4](#); [supplementary figure S1](#)). Significant correlations were observed

between SANS amotivation scores and Go-response rates to Familiar-Positive stimuli in the Test/transfer phase of the GNG task, and between SANS total scores and win-stay rates across the Acquisition phase of the GLA task ([supplementary table S3](#); [figure 3](#)). There were no significant correlations between RL measures and ratings of disorganization or depressive symptoms. No significant correlations were observed between RL variables and anti-psychotic dose when corrected for multiple-comparisons ([supplementary table S4](#)). Furthermore, no significant correlations were observed between RL variables and measures of IQ, working memory, episodic memory, or vigilance in FEP patients after correction for multiple comparisons was applied (see [supplementary tables S5 and S6](#)).

#### Discussion

Across both experimental paradigms, there was evidence for mild RL impairment in medicated FEP patients as seen in significant main effects of group on overall performance during the Acquisition phase of both tasks. This Acquisition deficit was not limited to positively valenced items on either task, however. On both tasks, FEP patients exhibited reliable deficits in the loss-avoidance condition (ES of group-difference = 0.70 for the GLA; ES of group-difference = 0.38 for the GNG). Contrary to our hypothesis, FEP patients exhibited comparable performance in value-guided decision making in the Test/transfer phase of the GNG task, as well as a normative preference for gain stimuli, relative to loss-avoidance stimuli, on the GLA task. Finally, as hypothesized, we found associations between negative symptoms scores and reward-driven learning measures from both tasks, and we observed that patients' overall Go rates from the GNG task correlated with positive symptom scores (see [table 2](#)).

In short, RL impairments in FEP appear to be less severe and more circumscribed than those observed in chronic SZ tested on the same tasks.<sup>16,43</sup> We found *some* evidence that FEP patients exhibited a deficit in using positive RPEs to modify value representations, similar to chronically ill patients.<sup>15,16,23,24</sup> However, given the fact that



**Fig. 3.** Scatter plots illustrating relationships between experimental variables and (A) positive symptoms and (B) negative symptoms.

our FEP patients had much less severe negative symptoms (mean overall SANS = 14.2 for patients performing GNG; mean overall SANS = 11.2 for patients performing GLA) than the MPRC samples (mean overall SANS = 32.8 for patients performing GNG<sup>16</sup>; mean overall SANS = 37.0 for high-negative-symptom patients performing GLA<sup>43</sup>; mean overall SANS = 22.9 for low-negative-symptom patients performing GLA<sup>43</sup>), we expected some of the deficits most closely associated with severe negative symptoms to be absent in the current FEP cohort. This was, in fact, that case. Specifically, our finding of marginal group-differences in lose-shift rates, with discrepant ESs across tasks, contrasts with the robust deficits in lose-shift behavior observed in chronic patients in 2 previous studies.<sup>15,16</sup> Results from the current study may suggest that the ability to use negative RPEs to rapidly update value representations is less impaired in clinically stable FEP patients. Additionally, unlike chronic patients with severe negative symptoms,<sup>43</sup> FEP patients and controls showed similar preferences for frequently rewarded stimuli (positive EV) over stimuli frequently associated with loss-avoidance (negative EV), on the GLA task, suggesting that FEP patients were using EV representations to guide decision making. The fact that EV representation was intact in a FEP sample with low levels of negative symptoms and in chronic patients with low levels of negative symptoms, may be seen as additional evidence for the specificity of the relationship between negative symptom severity and the ability to use EV to guide behavior.

Importantly, we still saw modest associations between negative symptoms scores and reward-driven learning measures despite the reduced dynamic range of negative symptoms in the current FEP sample. We also observed a significant correlation between positive symptom scores and overall Go-response rates from the GNG task, thought to reflect tonic striatal dopamine levels. Our finding that the general tendency to execute Go-responses was characteristic of FEP patients with the most severe positive symptoms accords with the notion that Go-response rates and psychotic symptoms may be of similar origin.<sup>42,57</sup>

We interpret our findings of a relatively spared ability to use acquired value representations to guide subsequent responding, in FEP patients, to be suggestive of both an intact or restored (BG-dependent) habit-learning system, and preserved capacity to draw on ventromedial prefrontal cortex value representations in making choices.<sup>6,13</sup> Rapid/explicit RL, on the other hand, is believed to rely most heavily on lateral prefrontal and hippocampal systems.<sup>10–12,58</sup> It is possible that FEP patients *rely* more heavily on BG function, during the Acquisition, in order to learn across trial blocks and achieve similar performance levels in the Test/transfer phases of our tasks. That is, patients and controls may be using different learning strategies during the Acquisition phases.

How do we explain differences in findings across tasks? Several factors can be speculatively considered. In the

GNG task, Pavlovian biases (respond for reward, withhold response for punishment-avoidance) may serve to support performance, whereas, in the GLA task, participants must respond both to gain rewards and avoid punishments. In the GNG paradigm, negative feedback is provided for choosing a response to single stimulus, whereas, in the GLA paradigm, negative feedback for the choice of one stimulus (from a stimulus-pair) must be used to learn that the other stimulus is likely more optimal—perhaps involving demands on working memory. By contrast, it is possible that response-selection (as is required for the GNG task) places greater demands of BG-driven procedural learning capacities.

In addition to differences in the symptom profiles of the patient samples in the respective studies, several other factors may account for the discrepant findings between the current FEP sample and the chronic patients examined at the MPRC.<sup>16,43</sup> First, the current FEP cohort (mean age: 24.1 y for the GNG task, 24.8 y for the GLA task) was much younger than the MPRC samples (average age: 40 y),<sup>16,43</sup> and there is evidence of an effect of age on RL performance.<sup>59</sup> Second, patients in the current study had much shorter illness duration than those in our previous studies.<sup>16,43</sup> Thus, it is not possible for us to distinguish between an effect of negative symptom severity and an effect of illness chronicity. Examining the longitudinal course of RL impairments in first-episode cohorts would help answer this question. Third, our FEP cohort showed evidence of relatively preserved learning and memory from several standardized measures. While these measures were not correlated with RL performance, the fact that this cohort of patients failed to demonstrate deficits that are common in SZ samples may raise the question of whether this was an atypical sample. Finally, roughly half of patients studied at the MPRC<sup>16,43</sup> were treatment-resistant with clozapine being prescribed. In contrast, the current sample had been clinically stabilized with treatment and the majority (74.2%) achieved positive symptom remission with SGAs other than clozapine (according to criteria from Andreasen et al.<sup>60</sup>). The limited variance in psychosis severity ratings might explain the absence of Go-response bias in the entire sample of FEP patients.

Several study limitations need to be acknowledged. Our sample size is modest, and our power to detect small effects was limited. Also, because all patients in the current study were receiving antipsychotics at the time of testing, we cannot rule out an effect of dopamine D2-receptor antagonists on RL.<sup>25,26,61–63</sup> Prospective investigation of RL prior to and following antipsychotic treatment in FEP patients is required to differentiate the effects of illness and medication on RL.

### Supplementary Material

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



## Funding

This study was funded by the HK Research Grants Council (GRF: 17124715). Support for J.A.W. and J.M.G. was provided by National Institutes of Health Grant R01MH080066.

## Acknowledgments

We thank all the coordinating clinical staff. The authors declared no conflicts of interest regarding the subject of this study.

## References

- 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.
- Ziauddeen H, Murray GK. The relevance of reward pathways for schizophrenia. *Curr Opin Psychiatr.* 2010;23:91–96.
- Strauss GP, Waltz JA, Gold JM. A review of reward processing and motivational impairment in schizophrenia. *Schizophr Bull.* 2014;40(suppl 2):S107–S116.
- Glimcher PW. Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. *Proc Natl Acad Sci USA.* 2011;108(suppl 3):15647–15654.
- Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. *Science.* 1997;275:1593–1599.
- Frank MJ, Claus ED. Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. *Psychol Rev.* 2006;113:300–326.
- Howes OD, Kambitz J, Kim E, et al. The nature of dopamine dysfunction in schizophrenia and what this means for treatment. *Arch Gen Psychiatry.* 2012;69:776–786.
- Pantelis C, Barnes TR, Nelson HE, et al. Frontal-striatal cognitive deficits in patients with chronic schizophrenia. *Brain.* 1997;120(Pt 10):1823–1843.
- Fornito A, Harrison BJ, Goodby E, et al. Functional dysconnectivity of corticostriatal circuitry as a risk phenotype for psychosis. *JAMA Psychiatry.* 2013;70:1143–1151.
- Monchi O, Petrides M, Petre V, Worsley K, Dagher A. Wisconsin card sorting revisited: distinct neural circuits participating in different stages of the task identified by event-related functional magnetic resonance imaging. *J Neurosci.* 2001;21:7733–7741.
- Cools R, Clark L, Owen AM, Robbins TW. Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. *J Neurosci.* 2002;22:4563–4567.
- McClelland JL, McNaughton BL, O'Reilly RC. Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol Rev.* 1995;102:419–457.
- Cohen MX, Frank MJ. Neurocomputational models of basal ganglia function in learning, memory and choice. *Behav Brain Res.* 2009;199:141–156.
- Frank MJ, Seeberger LC, O'Reilly RC. By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science.* 2004;306:1940–1943.
- 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.
- Waltz JA, Frank MJ, Wiecki TV, Gold JM. Altered probabilistic learning and response biases in schizophrenia: behavioral evidence and neurocomputational modeling. *Neuropsychology.* 2011;25:86–97.
- Kern RS, Green MF, Wallace CJ. Declarative and procedural learning in schizophrenia: a test of the integrity of divergent memory systems. *Cogn Neuropsychiatry.* 1997;2:39–50.
- Green MF, Kern RS, Williams O, McGurk S, Kee K. Procedural learning in schizophrenia: evidence from serial reaction time. *Cogn Neuropsychiatry.* 1997;2:123–134.
- Weickert TW, Terrazas A, Bigelow LB, et al. Habit and skill learning in schizophrenia: evidence of normal striatal processing with abnormal cortical input. *Learn Mem.* 2002;9:430–442.
- Kéri S, Nagy O, Kelemen O, Myers CE, Gluck MA. Dissociation between medial temporal lobe and basal ganglia memory systems in schizophrenia. *Schizophr Res.* 2005;77:321–328.
- Gomar JJ, Pomarol-Clotet E, Sarró S, Salvador R, Myers CE, McKenna PJ. Procedural learning in schizophrenia: reconciling the discrepant findings. *Biol Psychiatry.* 2011;69:49–54.
- Somlai Z, Moustafa AA, Kéri S, Myers CE, Gluck MA. General functioning predicts reward and punishment learning in schizophrenia. *Schizophr Res.* 2011;127:131–136.
- Strauss GP, Frank MJ, Waltz JA, Kasanova Z, Herbener ES, Gold JM. Deficits in positive reinforcement learning and uncertainty-driven exploration are associated with distinct aspects of negative symptoms in schizophrenia. *Biol Psychiatry.* 2011;69:424–431.
- Reinen J, Smith EE, Insel C, et al. Patients with schizophrenia are impaired when learning in the context of pursuing rewards. *Schizophr Res.* 2014;152:309–310.
- Juckel G, Schlagenhauf F, Koslowski M, et al. Dysfunction of ventral striatal reward prediction in schizophrenic patients treated with typical, not atypical, neuroleptics. *Psychopharmacology (Berl).* 2006;187:222–228.
- Schlagenhauf F, Juckel G, Koslowski M, et al. Reward system activation in schizophrenic patients switched from typical neuroleptics to olanzapine. *Psychopharmacology (Berl).* 2008;196:673–684.
- Simon JJ, Biller A, Walther S, et al. Neural correlates of reward processing in schizophrenia—relationship to apathy and depression. *Schizophr Res.* 2010;118:154–161.
- Nielsen MØ, Rostrup E, Wulff S, et al. Alterations of the brain reward system in antipsychotic naïve schizophrenia patients. *Biol Psychiatry.* 2012;71:898–905.
- Murray GK, Corlett PR, Clark L, et al. Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. *Mol Psychiatry.* 2008;13:239, 267–276.
- Waltz JA, Schweitzer JB, Gold JM, et al. Patients with schizophrenia have a reduced neural response to both unpredictable and predictable primary reinforcers. *Neuropsychopharmacology.* 2009;34:1567–1577.
- Gradin VB, Kumar P, Waiter G, et al. Expected value and prediction error abnormalities in depression and schizophrenia. *Brain.* 2011;134:1751–1764.
- Morris RW, Vercammen A, Lenroot R, et al. Disambiguating ventral striatum fMRI-related BOLD signal during reward prediction in schizophrenia. *Mol Psychiatry.* 2012;17:280–289.

33. Murray GK, Cheng F, Clark L, et al. Reinforcement and reversal learning in first-episode psychosis. *Schizophr Bull.* 2008;34:848–855.
34. Leeson VC, Robbins TW, Matheson E, et al. Discrimination learning, reversal, and set-shifting in first-episode schizophrenia: stability over six years and specific associations with medication type and disorganization syndrome. *Biol Psychiatry.* 2009;66:586–593.
35. Huddy VC, Hodgson TL, Ron MA, Barnes TR, Joyce EM. Abnormal negative feedback processing in first episode schizophrenia: evidence from an oculomotor rule switching task. *Psychol Med.* 2011;41:1805–1814.
36. Schlagenhauf F, Huys QJ, Deserno L, et al. Striatal dysfunction during reversal learning in unmedicated schizophrenia patients. *Neuroimage.* 2014;89:171–180.
37. Exner C, Boucsein K, Degner D, Irle E. State-dependent implicit learning deficit in schizophrenia: evidence from 20-month follow-up. *Psychiatry Res.* 2006;142:39–52.
38. Pedersen A, Siegmund A, Ohrmann P, et al. Reduced implicit and explicit sequence learning in first-episode schizophrenia. *Neuropsychologia.* 2008;46:186–195.
39. Murray GK, Clark L, Corlett PR, et al. Incentive motivation in first-episode psychosis: a behavioural study. *BMC Psychiatry.* 2008;8:34.
40. Purdon SE, Waldie B, Woodward ND, Wilman AH, Tibbo PG. Procedural learning in first episode schizophrenia investigated with functional magnetic resonance imaging. *Neuropsychology.* 2011;25:147–158.
41. Corlett PR, Murray GK, Honey GD, et al. Disrupted prediction-error signal in psychosis: evidence for an associative account of delusions. *Brain.* 2007;130:2387–2400.
42. Frank MJ, O'Reilly RC. A mechanistic account of striatal dopamine function in human cognition: psychopharmacological studies with cabergoline and haloperidol. *Behav Neurosci.* 2006;120:497–517.
43. Gold JM, Waltz JA, Matveeva TM, et al. Negative symptoms and the failure to represent the expected reward value of actions: behavioral and computational modeling evidence. *Arch Gen Psychiatry.* 2012;69:129–138.
44. Gold JM, Strauss GP, Waltz JA, Robinson BM, Brown JK, Frank MJ. Negative symptoms of schizophrenia are associated with abnormal effort-cost computations. *Biol Psychiatry.* 2013;74:130–136.
45. So E, Kam I, Leung CM, Chung D, Liu Z, Fong S. The Chinese-bilingual SCID-I/P project: stage 1: reliability for mood disorders and schizophrenia. *HK J Psychiatry.* 2003;13:7–18.
46. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders.* 4th ed. Washington, DC: American Psychiatric Association; 1994.
47. Drake RE, Mueser KT, McHugo GJ. Clinician rating scales: alcohol use scale (AUS), drug use scale (DUS), and substance abuse treatment scale (SATS). In: Sederer L, Dickey B, eds. *Outcomes Assessment in Clinical Practice.* Baltimore, MD: Williams & Wilkins; 1996.
48. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull.* 1987;13:261–276.
49. Andreasen NC. *The Scale for the Assessment of Negative Symptoms (SANS).* Iowa City: University of Iowa; 1984.
50. 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.
51. 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.
52. Della Sala S, Gray C, Baddeley AD, Wilson L. *The Visual Pattern Test: A New Test of Short-Term Visual Recall.* Bury St Edmunds, UK: Thames Valley Test Company; 1997.
53. Wilkins AJ, Shallice T, McCarthy R. Frontal lesions and sustained attention. *Neuropsychologia.* 1987;25:359–365.
54. Wechsler D. *Wechsler Memory Scale.* 3rd ed. San Antonio, TX: The Psychological Corporation; 1997.
55. Chan ELS, Chen EYH, Chan RCK. Three-subtest short form of the Wechsler Adult Intelligence Scale—III for patients with psychotic disorders: a preliminary report. *HK J Psychiatry.* 2005;15:39–42.
56. Andreasen NC, Pressler M, Nopoulos P, Miller D, Ho BC. Antipsychotic dose equivalents and dose-years: a standardized method for comparing exposure to different drugs. *Biol Psychiatry.* 2010;67:255–262.
57. 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.
58. Jocham G, Klein TA, Ullsperger M. Dopamine-mediated reinforcement learning signals in the striatum and ventromedial prefrontal cortex underlie value-based choices. *J Neurosci.* 2011;31:1606–1613.
59. Frank MJ, Kong L. Learning to avoid in older age. *Psychol Aging.* 2008;23:392–398.
60. Andreasen NC, Carpenter WT Jr, Kane JM, Lasser RA, Marder SR, Weinberger DR. Remission in schizophrenia: proposed criteria and rationale for consensus. *Am J Psychiatry.* 2005;162:441–449.
61. Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature.* 2006;442:1042–1045.
62. Menon M, Jensen J, Vitcu I, et al. Temporal difference modeling of the blood-oxygen level dependent response during aversive conditioning in humans: effects of dopaminergic modulation. *Biol Psychiatry.* 2007;62:765–772.
63. Nielsen MO, Rostrup E, Wulff S, et al. Improvement of brain reward abnormalities by antipsychotic monotherapy in schizophrenia. *Arch Gen Psychiatry.* 2012;69:1195–1204.
64. Häfner H, Riecher-Rössler A, Hambrecht M, et al. IRAOS: an instrument for the assessment of onset and early course of schizophrenia. *Schizophr Res.* 1992;6:209–223.