

# Reduced susceptibility to confirmation bias in schizophrenia

Bradley B. Doll · James A. Waltz · Jeffrey Cockburn ·  
Jaime K. Brown · Michael J. Frank · James M. Gold

© Psychonomic Society, Inc. 2014

**Abstract** Patients with schizophrenia (SZ) show cognitive impairments on a wide range of tasks, with clear deficiencies in tasks reliant on prefrontal cortex function and less consistently observed impairments in tasks recruiting the striatum. This study leverages tasks hypothesized to differentially recruit these neural structures to assess relative deficiencies of each. Forty-eight patients and 38 controls completed two reinforcement learning tasks hypothesized to interrogate prefrontal and striatal functions and their interaction. In each task, participants learned reward discriminations by trial and error and were tested on novel stimulus combinations to assess learned values. In the task putatively assessing fronto-striatal interaction, participants were (inaccurately) instructed that one of the stimuli was valuable. Consistent with prior reports and a model of confirmation bias, this manipulation resulted in overvaluation of the instructed stimulus after its true value had been experienced. Patients showed less susceptibility to this confirmation bias effect than did controls. In the choice bias task hypothesized to more purely assess striatal function, biases in endogenously and exogenously chosen actions were

assessed. No group differences were observed. In the subset of participants who showed learning in both tasks, larger group differences were observed in the confirmation bias task than in the choice bias task. In the confirmation bias task, patients also showed impairment in the task conditions with no prior instruction. This deficit was most readily observed on the most deterministic discriminations. Taken together, these results suggest impairments in fronto-striatal interaction in SZ, rather than in striatal function per se.

**Keywords** Schizophrenia · Reward · Decision-making

## Introduction

The psychological impairments observed in schizophrenia (SZ) involve most, but not all, cognitive functions (Gold, Hahn, Strauss, & Waltz, 2009). Deficits in tasks thought to rely on the integrity of the frontal lobes, such as working memory and executive control (Barch & Ceaser, 2012; Lesh, Niendam, Minzenberg & Carter 2011), have been repeatedly observed and modeled (Cohen & Servan-Schreiber, 1992). Episodic memory deficits, likely related to both prefrontal and medial temporal lobe abnormalities, are also prominent (Ragland et al., 2009). In contrast, deficits in reinforcement learning and reward-based decision making, dependent on both the basal ganglia and prefrontal cortex, have been observed in some cases, but not in others (Averbeck, Evans, Chouhan, Bristow & Shergill 2011; Gold et al., 2009; Somlai, Moustafa, Kéri, Myers & Gluck 2011; Waltz, Frank, Robinson, & Gold, 2007; Weickert et al., 2002).

When studying patient impairments in cognitive function, it is important to distinguish between *global* performance deficits and *selective* alterations in particular processes putatively related to underlying neural mechanisms. Notably, several reinforcement learning studies have reported that patients

---

B. B. Doll (✉)  
Center for Neural Science, New York University, 4 Washington Pl,  
Room 873B, New York, NY 10003, USA  
e-mail: bradley.doll@nyu.edu

B. B. Doll  
Department of Psychology, Columbia University, New York, NY,  
USA

J. A. Waltz · J. K. Brown · J. M. Gold  
Maryland Psychiatric Research Center, University of Maryland  
School of Medicine, Baltimore, MD, USA

J. Cockburn · M. J. Frank  
Department of Cognitive, Linguistic, and Psychological Sciences,  
Brown University, Providence, RI, USA

with SZ exhibit selective deficits in learning from probabilistic rewards/gains, and not punishments/losses (Gold et al., 2012; Strauss et al., 2011; Waltz et al., 2007; Waltz, Frank, Wiecki, & Gold, 2011; Waltz et al., 2010). While initial reports attributed this pattern to deficiencies in striatal dopaminergic function similar to those observed in Parkinson's disease (Waltz et al., 2007; Waltz et al., 2011), subsequent studies have implied that the source of dysfunction may lie in reward-processing areas of the prefrontal cortex and fronto-striatal communication (Gold et al., 2012; Strauss et al., 2011; Waltz et al., 2010). Here, we employed tasks designed to assess the function and interaction of the prefrontal cortex and striatum in an attempt to assess their relative perturbation in SZ. These tasks were chosen to assess biases in learning and decision-making processes that arise in healthy individuals—biases that are thought to reflect the signatures of striatal function or prefrontal–striatal communication. If patients exhibit selective alterations in these neural systems, we might expect to observe a failure to exhibit one of these biases and, hence, a paradoxical advantage in the relevant task.

We utilized two variants of a well-characterized probabilistic reinforcement learning task (Frank, Moustafa, Haughey, Curran, & Hutchison, 2007; Frank, Seeberger, & O'Reilly, 2004), which assesses the striatal-dependent capacity to learn from reinforcement and produce extrinsically motivated actions, in order to investigate biases in learning. The first variant focuses on the interaction of cognition and motivation in a task where participants combine experimenter-given prior information with trial-to-trial reinforcement. The second variant focuses more strictly on the learning of actions in a task with endogenously and exogenously motivated actions.

The first task assesses how striatal learning mechanisms are altered as a function of explicit prior information about task contingencies given by verbal instruction. In particular, prior modeling and empirical studies have suggested that prefrontal mechanisms representing explicit task rules provide top-down input to modify striatal learning (Biele, Rieskamp, Krugel, & Heekeren, 2011; Doll, Hutchison, & Frank, 2011; Doll, Jacobs, Sanfey, & Frank, 2009). Notably, this effect produces a *confirmation bias* (Nickerson, 1998) in healthy individuals, such that participants express higher learned reward value for choices that had been previously instructed than for those with similar or even higher objective reward probabilities (Doll et al., 2011; Doll et al., 2009; Staudinger & Büchel, 2013). This effect is best captured by computational models (Biele, Rieskamp, & Gonzalez, 2009; Doll et al., 2009) that bias the interpretation of outcomes following instructed stimulus selection. Specifically, our model (Doll et al., 2009) posits that prefrontal cortical instruction representations increase the activation of the striatal neuronal populations representing the value of the instructed stimulus. Subsequent dopaminergic prediction errors following instructed stimulus selection

ingrain this distortion by further exaggerating activation and activity-dependent plasticity. This bias increases the impact of instruction-consistent gains and dampens the impact of instruction-inconsistent losses, causing the brain to overvalue the instructed stimulus (hence, a confirmation bias). Functional imaging studies have further provided evidence for this type of bias during learning (Staudinger & Büchel, 2013), showing the exaggerated positive and blunted negative prediction errors posited to drive this effect (Biele et al., 2011). Furthermore, individual differences in susceptibility to such confirmation biases are related to variations in genetic function linked to fronto-striatal processing. In particular, participants with increased dopamine levels in the prefrontal cortex as indexed by COMT val158met genotype (Gogos et al., 1998; Slifstein et al., 2008), which is frequently associated with a more robust working memory (de Frias et al., 2010; Egan et al., 2001; Tunbridge, Bannerman, Sharp, & Harrison, 2004), showed increased adherence to, and valuation of, the instructed stimulus (Doll et al., 2011), thereby exhibiting greater confirmation bias. Moreover, striatal dopaminergic genes normally associated with enhanced reinforcement learning were, in the context of this task, predictive of the extent to which such learning was distorted by instructed priors (Doll et al., 2011; Frank et al., 2007). Thus, in this task environment, genetic variants that typically confer a cognitive advantage are associated with less veridical performance.

Cortical pathways can accomplish action selection on their own when external events trigger responses, whereas volitional behavior requires additional recruitment of the basal ganglia to inform action selection (Brown & Marsden, 1988; François-Brosseau et al., 2009). A second task leveraged this differential recruitment of the basal ganglia to putatively probe striatal action value learning as a tendency for participants to learn relatively higher reinforcement values for actions they had freely chosen, relative to those that were chosen for them. This type of *choice bias* has been related to striatal processing and dopaminergic signaling that is preferentially engaged when endogenous choice is required (Sharot, De Martino, & Dolan, 2009). That is, the bias of learning more from the outcomes of endogenously chosen actions versus exogenously chosen ones is hypothesized to arise from striatal mechanisms that amplify the impact of rewards following active choices.

We thus reasoned that if the source of reward learning deficits in SZ lies in striatal dopaminergic mechanisms, patients would show reduced evidence for a choice bias. In contrast, if patients exhibit reduced prefrontal–striatal communication, patients would be expected to exhibit reduced confirmation bias. In this task setting, poor prefrontal–striatal communication should lead to a paradoxical improvement in patient performance, with patients more able than controls to report task contingencies veridically.

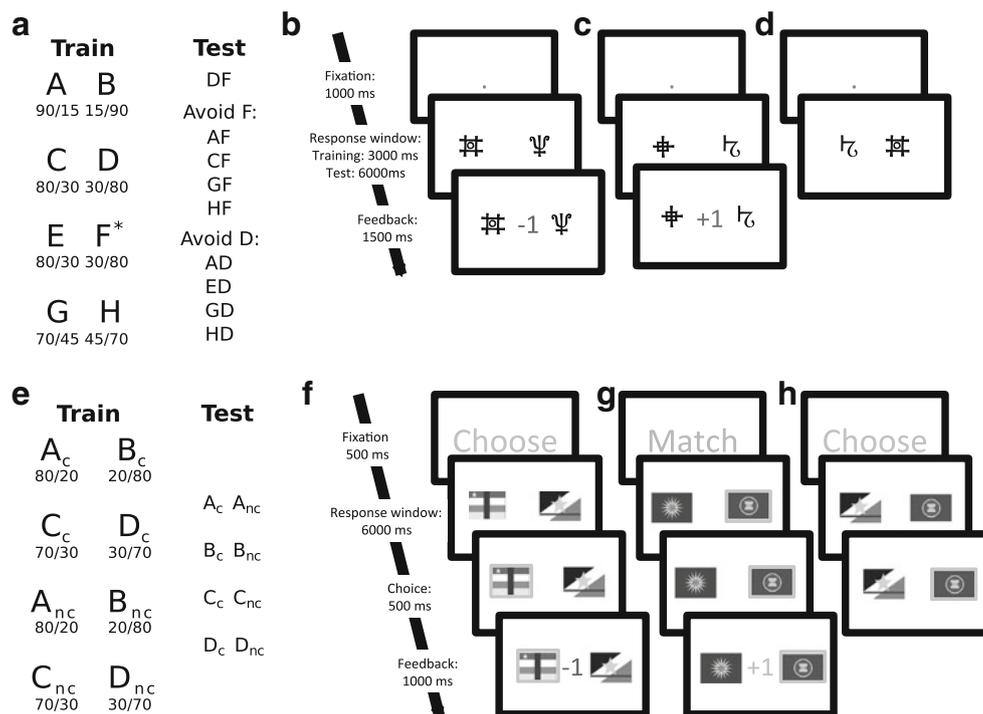
## Method

### Experimental tasks

In the instructed version of the probabilistic selection (PS) task, participants learned four stimulus discriminations (AB, CD, EF, and GH, each stimulus represented in the task by a unique character in the Agathadaimon font set) by trial and error (Fig. 1a–d). One stimulus in each training pair was more likely to produce reward (+1 vs -1 point) than was its pair partner (independent probabilities of +1 point: A/B, .9/.15; C/D, .8/.3; E/F, .8/.3; G/H, .7/.45). Before beginning the training phase, participants read the task instructions on a computer screen. Participants were told that the aim of the task was to win as many points as possible and that, in each stimulus pair, one stimulus would be “better” than the other, although there was no absolutely correct answer. Additionally, participants were inaccurately instructed about the value of stimulus F. Specifically, the text “the following stimulus will be good” was presented above the F stimulus. Participants

were tested on comprehension of the instructions before the task begin and were reinstructed until performance on the comprehension test was perfect. Participants completed from two to four 80-trial training blocks (20 of each stimulus pair, randomly interleaved) and advanced to the test phase when satisfying training accuracy criteria (A/B  $\geq$  70 % A choices, C/D  $\geq$  65 % C choices, G/H  $\geq$  60 % G choices) or completing 4 training blocks.

In the test phase, participants were serially presented with all combinations of stimuli from the training phase, with each combination repeating 4 times, randomly interleaved. Participants were instructed that they would not receive feedback for responses during this phase and that they should pick the symbol they felt was correct more often on the basis of what they had learned during training. After the test phase, participants were probed for their memory of the instructions. In the memory probe, each stimulus was shown (in random order), and participants were asked to indicate which one they were told would be good at the beginning of the experiment. This memory test was administered to assess whether any



**Fig. 1** Experimental tasks. **a** Instructed probabilistic selection task. In the training phase, participants repeatedly chose between two stimuli from each of four randomly presented stimulus pairs (AB, CD, EF, GH; stimuli represented in task by Agathadaimon font characters) and received positive/negative (+1/-1) feedback in accordance with stimulus reward probabilities. Prior to training, participants were given erroneous instruction about the value of stimulus F. In the test phase, participants chose between novel combinations of stimuli and received no feedback. Test trials permitting inference of the learned valuation of the instructed stimulus are those featuring stimulus F and stimulus D, which has an identical reward probability but no instruction. **b, c** Examples of training-phase trials with different stimulus pairings. **d** Example of a test-phase trial,

involving a novel stimulus pairing. **e** Choice bias task. In the training phase, participants volitionally chose between “choice” ( $A_c B_c$ ,  $C_c D_c$ ) pairs, which were yoked to “no-choice” pairs ( $A_{nc} B_{nc}$ ,  $C_{nc} D_{nc}$ ), and received feedback in accordance with the designated reward probabilities. In the test phase, participants selected stimuli from novel combinations of training stimuli and did not receive feedback. Test trials permitting inference of choice bias were pairs with identical reward probability where one stimulus was freely chosen in the training phase and the other was not. **f** A “free-choice” trial from the training phase. **g** A “forced-choice” trial from the training phase (directed choice during response window). **h** Example of a test-phase trial, involving a novel stimulus pairing

reduction in susceptibility to instructions could be attributed to simple forgetting. Finally, participants were again shown all of the stimuli from the task and were asked to estimate how frequently each would win if selected 100 times.

In the choice bias version of the PS task, participants again learned four stimulus discriminations by trial and error (Fig. 1e–h). We refer to these eight stimuli, each represented in the task by a unique flag picture, as  $A_c B_c$ ,  $C_c D_c$ ,  $A_{nc} B_{nc}$ , and  $C_{nc} D_{nc}$ . One stimulus in each training pair was more likely to produce a reward (+1 vs. -1 point; probabilities of +1:  $A_c/B_c$ , .8/.2;  $C_c/D_c$ , .7/.3;  $A_{nc}/B_{nc}$ , .8/.2;  $C_{nc}/D_{nc}$ , .7/.3). When presented with  $A_c B_c$  and  $C_c D_c$  pairs, participants were free to choose either stimulus (choice condition), but when presented with  $A_{nc} B_{nc}$  and  $C_{nc} D_{nc}$  pairs, participants were forced to pick a preselected stimulus (no-choice condition). Critically, “no-choice” trials were yoked to choice trials to ensure identical sampling and feedback between conditions. For example, if a participant was presented with  $A_c B_c$  on a choice trial, selected  $A_c$ , and received -1 as feedback, sometime in the near future he or she would be presented with an  $A_{nc} B_{nc}$  pair in a no-choice trial, would be forced to select  $A_{nc}$ , and would receive -1 as feedback. This design enables us to precisely control the trial-by-trial sequence of reinforcement history for each of the choice and no-choice symbols, permitting us to later assess whether a bias to value the choice stimuli was present and whether any such biases interacted with the values of the stimuli.

Participants completed at minimum two and at maximum four training blocks and advanced to the test phase after satisfying training accuracy criteria ( $A_c B_c > 65\%$   $A_c$  choices and  $C_c D_c > 55\%$   $C_c$  choices) or after completing the maximum permitted blocks. At test, participants were presented with all possible stimulus pairings to evaluate what they had learned. The test phase included four repetitions of each of the noncritical stimulus pairings and eight repetitions of the critical choice bias trials ( $A_c A_{nc}$ ,  $B_c B_{nc}$ ,  $C_c C_{nc}$ ,  $D_c D_{nc}$ ). During the test phase, participants were free to choose on all trials, but to prevent any additional learning, they were no longer given feedback. Importantly, participants encountered trials on which they had to choose between choice and no-choice stimuli with identical reward value ( $A_c A_{nc}$ ,  $B_c B_{nc}$ ,  $C_c C_{nc}$ ,  $D_c D_{nc}$ ). These trials, where stimuli were equated for sampling and feedback, isolated the value associated with choice across stimuli with a range of reward probabilities.

## Participants

After explanation of study procedures, all participants provided written informed consent. All participants were compensated for study participation. Participants completed standard cognitive batteries, including the MATRICS battery (Green et al., 2004; MATRICS, 1996), the Wechsler Abbreviated

Scale of Intelligence (WASI) (Wechsler, 1999), and the Wechsler Test of Adult Reading (WTAR; Wechsler, 2001). Overall symptom severity in patients was characterized using the Brief Psychiatric Ratings Scale (BPRS; Overall & Gorman, 1962), and negative symptom severity was quantified using the Scale for the Assessment of Negative Symptoms (SANS; Andreasen, 1983) and the Brief Negative Symptom Scale (BNSS; Kirkpatrick et al., 2011).

A total of 48 individuals meeting DSM-IV (DSM, 1994) criteria for SZ or schizoaffective disorder and 38 healthy controls (HCs) participated in the study. All patients were on stable doses of medication for at least 4 weeks at time of testing and were considered to be clinically stable by treatment providers. The outpatients were recruited from the Maryland Psychiatric Research Center outpatient clinics and other local clinics. HCs were recruited from the community via random digit dialing, word of mouth among participants, and newspaper advertisements. HCs had no current Axis I or II diagnoses as established by the SCID (First, Spitzer, Gibbon, & Williams, 1997) and SID-P (Pfohl, Blum, & Zimmerman, 1997), had no family history of psychosis, and were not taking psychotropic medications. All participants denied a history of significant neurological injury or disease and significant medical or substance use disorders within the last 6 months. All participants provided informed consent for a protocol approved by the University of Maryland School of Medicine Institutional Review Board.

As was noted above, our main analyses are focused on the posttraining test/transfer phase of both tasks. In order to interpret transfer phase results, it is necessary to exclude from analysis participants who failed to acquire the task during the training phase. Thus, we excluded participants who performed below chance (i.e., less than 50% selection of the stimulus with higher reward probability) on the easiest test trial discriminations (training pairs A/B and C/D in the test phase) or at chance on both of these discriminations. This eliminated 13 participants (10 from the patient group) from the instructed task and 18 participants (13 from the patient group) from the choice bias task. Demographic, cognitive, and clinical information for the 45 patients and 37 controls exhibiting above-chance performance on at least one of the experimental tasks is shown in Table 1. No group differences were observed in age, gender, race, or parental education. Due to the exclusion of participants who performed below chance on the easiest test trials from each task, overlapping, but not identical, subsets of patients and controls were included in the analyses of data from each task. On the instructed PS task, 38 patients and 35 controls showed above-chance performance. On the choice bias PS task, 35 patients showed above-chance performance, along with 33 controls. The inclusion of different subsets of patients across analyses did not affect the matching of the groups on demographic variables.

**Table 1** Characterizing information for participants

	SZs ( <i>n</i> = 45) Mean ( <i>SD</i> )	HCS ( <i>n</i> = 37) Mean ( <i>SD</i> )	<i>p</i>
<b>Demographic Info</b>			
Age	38.16 (10.84)	37.00 (12.79)	.65
Education level	13.40 (1.96)	15.00 (2.01)	<.001*
Father's education level	13.48 (3.64)	14.36 (3.02)	.247
Sex (number females)	15	12	.93
Race:			.47
Number caucasian	25	21	
Number Black	16	14	
Number Asian	0	1	
Number other/mixed	4	1	
<b>Neurocognition</b>			
WTAR-scaled score	99.36 (16.25)	112.49 (10.57)	<.001*
Estimated IQ (WASI)	102.27 (13.33)	117.7 (8.67)	<.001*
MATRICES-WM Domain	40.067 (10.70)	51.78 (8.93)	<.001*
MATRICES composite	33.58 (12.70)	54.49 (8.04)	<.001*
<b>Clinical</b>			
BPRS total	33.25 (1.01)		
SANS total	34.08 (2.54)		
SANS Avol. + Anhed.	21.50 (1.47)		
Antipsychotics:			
FGA monotherapy	6		
SGA monotherapy	30		
FGA + SGA	5		
Multiple SGAs	4		
Oral-halop. Dose equiv.	11.65 (1.24)		

**Note.** Daily antipsychotic doses converted into oral-haloperidol dose equivalents on the basis of Andreasen, Pressler, Nopoulos, Miller, and Ho (2010). SZs, patients with schizophrenia; HCs, healthy controls; WTAR, Wechsler Test of Adult Reading; WM, working memory; BPRS, Brief Psychiatric Rating Scale; SANS, Scale for the Assessment of Negative Symptoms; Avol., avolition; Anhed., anhedonia; FGA, first-generation antipsychotic; SGA, second-generation antipsychotic; halop., haloperidol; equiv., equivalent.

Exclusion of participants does, however, potentially impact the generalizability of any findings to the broader population of patients with SZ. To address this possibility, we compared characterizing variables of patients who were excluded on at least one task ( $n = 16$ ) with those of patients who were excluded on neither ( $n = 32$ ; Table 2). While we observed no differences in demographic or clinical variables, excluded patients performed more poorly than not excluded patients on tests of cognitive ability. No differences were observed between excluded ( $n = 7$ ) and never excluded ( $n = 31$ ) control participants ( $ps > .29$ ). Below (see the Results section), we present the key analyses both with and without participants excluded (see also the Discussion section).

**Table 2** Characterizing information for patients who showed greater than chance learning in both tasks and those who failed to acquire the contingencies in at least one task

	SZs Inc ( <i>n</i> = 28) Mean ( <i>SD</i> )	SZs Ex ( <i>n</i> = 20) Mean ( <i>SD</i> )	<i>p</i>
<b>Demographic Info</b>			
Age	36.43 (10.4)	41.2 (10.4)	$p = .71$
Education Level	13.71 (1.9)	12.4 (2.3)	$p = .75$
Father's education level	13.4 (4.1)	13.3 (2.7)	$p = .15$
Sex (number females)	10	5	$p = .63$
Race:			
Number Caucasian	16	9	
Number Black	10	6	
Number Asian	0	0	
Number other/mixed	2	5	
<b>Neurocognition</b>			
WTAR-scaled score	102 (16.6)	92.11 (16.4)	$p = .065$
Estimated IQ (WASI)	104.88 (14.1)	95.59 (12.5)	$p = .033*$
MATRICES-WM Domain	42.61 (11.8)	34.18 (6.5)	$p = .01*$
MATRICES composite	38.08 (12.9)	24.53 (7.8)	$p = .0004*$
<b>Clinical</b>			
BPRS total	33.07 (7)	33.11 (6.5)	$p = .98$
SANS total	33.28 (14.3)	33.77 (18.8)	$p = .93$
SANS Avol. + Anhed.	21.25 (10.1)	21.39 (9.3)	$p = .96$
Antipsychotics:			
FGA monotherapy	5	2	
SGA monotherapy	19	12	
FGA + SGA	2	4	
Multiple SGAs	2	2	
Oral-halop. dose equiv.	12.98 (9.9)	10.27 (6.4)	$p = .29$

**Note.** Daily antipsychotic doses converted into oral-haloperidol dose equivalents on the basis of Andreasen et al. (2010). SZs, patients with schizophrenia; Inc, included in analysis of both tasks; Ex, excluded from analysis of at least one task; WTAR, Wechsler Test of Adult Reading; WM, working memory; BPRS, Brief Psychiatric Rating Scale; SANS, Scale for the Assessment of Negative Symptoms; Avol., avolition; Anhed., anhedonia; FGA, first-generation antipsychotic; SGA, second-generation antipsychotic; halop., haloperidol; equiv., equivalent.

### Statistical analysis

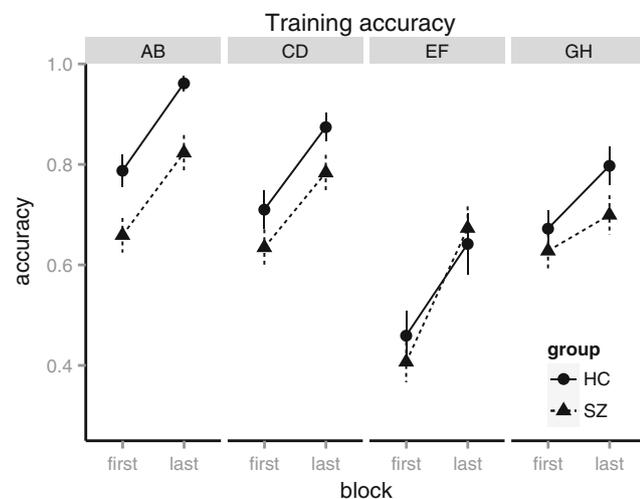
We modeled binomial choice data with multilevel logistic regression (lme4 linear mixed effects package for R; Bates, Maechler, & Bolker, 2011), in which participant accuracy (selection of the statistically superior stimulus in a pair) was the dependent variable. In these models, we additionally entered all within-subjects effects as random by participant (Schielzeth & Forstmeier, 2009). In the cross-task analysis, multilevel linear regression was used to model the continuous outcome variable ( $z$ -scored bias performance). In this model

and in the logistic models for factors with greater than two levels,  $\chi^2$  and  $p$  values were derived for the estimates from type III analysis of variance tables from the ANOVA function in the car package for R (Fox & Weisberg, 2011). In most cases, effect coding (1, -1) was used for regression factors to assess main effects, but where simple effects were preferred, dummy coding (0, 1) was utilized, as is noted in the Results text. We assessed the relationship between negative symptoms and posttask reinforcement frequency estimates in the instructed task with Pearson correlations.

## Results

### Instructed probabilistic selection task: Training phase

We first assessed accuracy on the uninstructed trials in the training phase, entering training block (first and last), stimulus condition (AB, CD, GH), and group (HC, SZ) as independent variables into a multilevel logistic regression (all factors were effect-coded: +1, -1; Fig. 2). Both intercept and training block terms were significantly positive, demonstrating that average performance in the task exceeded chance and increased over training, respectively,  $z_s > 9$ ,  $ps < .0001$ . The condition coefficients and their interactions with block reflect the relative difficulty of the probabilistic discriminations, with AB (reward probability: .9/.15) performance deviating positively,  $z_s > 4.6$ ,  $ps < .0001$ , and GH (.7/.45) negatively from average,  $z_s < 1.9$ ,  $ps < .054$ .



**Fig. 2** Accuracy by group across the training phase of the instructed probabilistic selection task. Patients (SZ) performed less well than healthy controls (HC) on average and had shallower acquisition curves. Notably, groups did not differ on the instructed condition (EF), although both groups were impaired in accuracy on this condition, relative to the GH condition, which had identical reward probabilities. Probabilities of +1 point: A/B, .9/.15; C/D, .8/.3; E/F, .8/.3; G/H, .7/.45. Error bars reflect standard errors of the means

We additionally observed a main effect of group and a group  $\times$  block interaction, indicating overall reduced performance and flatter learning curves in patients than in controls, respectively,  $|z|_s > 2.8$ ,  $ps < .005$ . A marginal interaction of group and stimulus condition,  $\chi^2 = 5.9$ ,  $p = .052$ , and a trend for a group  $\times$  stimulus condition  $\times$  block interaction,  $\chi^2 = 4.89$ ,  $p = .087$ , were also observed. Contrasts on the relevant interactions of these terms with stimulus condition showed that these effects were driven by poorer patient than control performance in the AB (.9/.15) discrimination (Table 3).

Next, we inspected accuracy on the instructed trials in the EF condition and compared it with that in the CD condition, which has identical reinforcement contingencies but no prior instruction. We utilized dummy coding (0, 1) for the stimulus condition (group and training block terms effect-coded: -1, +1), with EF serving as the reference group to assess accuracy on this instructed condition alone (intercept), as well as differences from the CD condition (condition term). The lack of significance of the intercept term in this model reveals that performance on the EF pair averaged across training did not differ from chance,  $z = 1.27$ ,  $p = .2$ . EF performance was significantly worse than that on the uninstructed CD pair (effect of condition,  $z = 5.38$ ,  $p < .0001$ ). Despite poor initial and overall performance on EF induced by the inaccurate instructions, participants learned about the true contingencies over the course of the training phase, as indicated by the significant positive effect of block,  $z = 6.44$ ,  $p < .0001$ . Groups did not differ in EF performance overall (no effect of group,  $z = 0.06$ ,  $p = .95$ ), nor did group interact with any other terms in the model,  $|z|_s < 1.27$ ,  $ps > .2$ .

**Table 3** Logistic regression coefficients from model of control trials (AB, CD, GH; CD as reference condition) in training phase of instructed probabilistic selection task

Coefficient	Estimate	SE	Z	p
Intercept	1.92	0.16	12.05	<.0001*
Last block	0.91	0.10	9.20	<.0001*
Condition AB	0.49	0.10	4.62	<.0001*
Condition GH	-0.19	0.09	-2.08	.038*
SZ group	-0.55	0.16	-3.43	.0006*
Last block $\times$ condition AB	0.34	0.07	5.00	<.0001*
Last block $\times$ condition GH	-0.11	0.06	-1.92	.054
Last block $\times$ SZ group	-0.28	0.10	-2.85	.0043*
Condition AB $\times$ SZ group	-0.24	0.10	-2.29	.022*
Condition GH $\times$ SZ group	-0.01	0.09	-0.11	.91
Last block $\times$ condition AB $\times$ SZ group	-0.15	0.07	-2.20	.028*
Last block $\times$ condition GH $\times$ SZ group	-0.05	0.06	-0.79	.43

*Note.* SE indicates standard error. All terms effect coded: -1, 1. Named levels coded 1. SZ, schizophrenia.

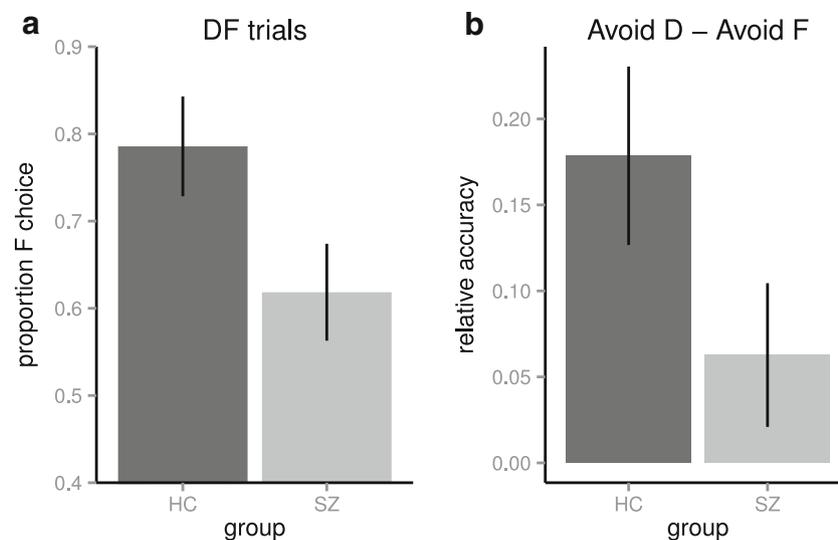
## Instructed probabilistic selection task: Test phase

In analyzing test phase performance, we first excluded all instructed trials to investigate differences in standard reinforcement learning. Participants' overall ability to discriminate the value of rewarding and punishing stimuli was assessed by measuring test phase accuracy on novel combinations of stimuli featuring A, the statistically best stimulus (choose A condition; trials: AC, AD, AG, AH), and those featuring B, the statistically worst stimulus (avoid B condition; trials: BC, BD, BG, BH). We entered group (dummy coded: 0, 1) and test condition as independent variables into a multilevel logistic regression predicting test accuracy. The intercept showed greater than chance accuracy for patients,  $z = 6.67$ ,  $p < .0001$ , and significantly greater performance for controls (effect of group,  $z = 4.16$ ,  $p < .0001$ ). There was neither an effect of condition nor a condition  $\times$  group interaction ( $ps > .2$ ).

We next assessed participants' choices in test trials involving the F stimulus paired with the D stimulus. These trials compared instructed and uninstructed stimuli that had identical reward probability (30 %) in the training phase and served as a measure of confirmation bias. Absent any bias, participants should show no preference for either stimulus. We assessed group differences in choice on these trials in a multilevel logistic model (dependent variable coded, D = 1, F = 0; group independent variable effect coded, SZ = +1, HC = -1). The intercept was significantly negative,  $z = -5.11$ ,  $p < .0001$ , showing that participants across groups preferred the instructed stimulus on these trials and replicating previous

demonstrations of confirmation bias in this task. Additionally, there was an effect of group, with patients less likely to choose the instructed stimulus,  $z = 2.4$ ,  $p = .015$  (Fig. 3a) than were controls. Thus, this pattern is consistent with the prediction that patients exhibit reduced susceptibility to confirmation bias during learning.

Because the uninstructed stimulus D has a low reward probability relative to the paired training stimulus C, participants may prefer the F stimulus simply because they did not choose D enough to learn its value. Patients might thus show differences from controls not because of an underlying difference in confirmation bias, but because they took longer to complete the training phase [average blocks (standard error): 3.1 (0.15) SZ, 2.6 (0.15) HC;  $p = .04$ ] and, with more exposure to the true contingencies, they were possibly better able to learn that the instructions were inaccurate. This interpretation is unlikely to be correct, given that the group effect remained significant after covarying either the total number of D choices during training (group effect,  $z = 2.04$ ,  $p = .04$ ; D choice effect,  $z = 0.97$ ,  $p = .33$ ) or the total number of training trials (group effect,  $z = 2.07$ ,  $p = .038$ ; training trial effect,  $z = 1.5$ ,  $p = .13$ ). Furthermore, the difference in bias between groups cannot be attributed to differences in memory for the instructed stimulus (F) at test. In the posttask memory probe, participants were presented with all stimuli and were asked to indicate the stimulus they were told would be good at the beginning of the experiment. All participants correctly identified the F stimulus as having been instructed (i.e., accuracy score in all participants was 100 %).



**Fig. 3** Instructed probabilistic selection task test phase results. Instructions differentially affect patients (SZ) and healthy controls (HC) at test. **a** Performance on the DF pair at test, which pairs the instructed stimulus F against uninstructed stimulus D, both of which had the identical reward probability of 30 % during the training phase. Although both groups choose the instructed stimulus on average, controls show a more extreme

preference. **b** Controls show a greater impact of instruction on learning than do patients. Difference score compares tendency in the test phase to avoid the uninstructed (D) and instructed (F) stimuli when paired with stimuli learned to be statistically better during training. Error bars reflect standard errors of the means

We next assessed whether instruction introduced a further bias to select the instructed stimulus (F) not only over the equivalently rewarded one (D), but even over statistically superior stimuli. We modeled participants' accuracy in avoidance of either of the 30 % rewarding stimuli (F and D) when paired with those with higher reward probability (avoid F: AF, CF, GF, HF; avoid D: AD, ED, GD, HD). Condition (effect coded avoid F = +1, avoid D = -1) and group (effect coded SZ = +1, HC = -1) were entered into a multilevel logistic regression as independent variables predicting choice accuracy. Note that the test condition term (avoid F, avoid D) in this model estimates confirmation bias by subtracting out the effects of learning common to both the instructed and uninstructed conditions. The model intercept was significantly positive,  $z = 7.9$ ,  $p < .0001$ , revealing that participants performed overall better than chance on these trials, although this effect was carried by significantly higher accuracy on the uninstructed avoid D condition than on avoid F (relative effect of avoid F condition,  $z = -3.3$ ,  $p = .001$ ). We observed no overall differences in accuracy (no effect of group,  $p = .69$ ). However, the group  $\times$  condition interaction was significant,  $z = 2.03$ ,  $p = .043$  (Fig. 3b), with patients showing greater accuracy than controls in avoiding the instructed stimulus. The specificity of this result suggests that group differences in confirmation bias are not explained by group differences in uninstructed learning (in which case, a main effect of group indicating an overall accuracy difference would be expected).

Again, because patients had more training trials on average than did controls, we controlled for the possibility that training, and not group membership, accounted for the effects by entering training duration (number of trials) and the interaction of condition and duration into the model. Neither of these added factors significantly predicted choice,  $|z|_s < 0.66$ ,  $ps > .53$ , and, despite the correlation of group and training duration, a trend for the group  $\times$  condition interaction persisted,  $z = 1.79$ ,  $p = .073$ .

We additionally tested the possibility that differences in confirmation bias could be explained by the training phase differences in learning between groups. We refit the test model, adding average training accuracy and training accuracy  $\times$  test condition as covariates. A main effect of training accuracy was observed,  $z = 4.1$ ,  $p < .0001$ , indicating that better performance in the training phase predicted better performance at test. This effect did not differ across test conditions (condition  $\times$  training accuracy interaction:  $z = 1.35$ ,  $p = .17$ ). Critically, the interaction of group and test condition remained significant,  $z = 2.3$ ,  $p = .02$ . We additionally entered the training accuracy covariate into the model of D versus F trials described above. There was no accuracy effect on DF choice,  $z = 0.78$ ,  $p = .43$ , and the effect of group persisted,  $z = 2.5$ ,  $p = .01$ . Thus, the reduced confirmation bias seen in patients relative to controls is not explained by a general SZ learning impairment.

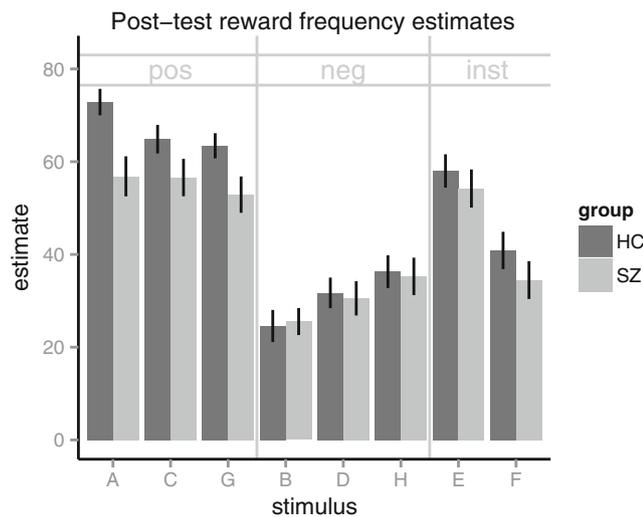
We next assessed whether any of the test phase results could be explained by medication effects. We found no effects of haldol equivalent dosage on uninstructed test trial accuracy across conditions (main effect of condition [choose A, avoid B]:  $z = -.37$ ,  $p = .71$ ) or differentially on the conditions (condition by dosage interaction:  $z = 1.02$ ,  $p = .31$ ). There were also no effects of medication on either of the confirmation bias measures (DF trial medication effect,  $z = 0.56$ ,  $p = .58$ ; avoid F vs. avoid D medication effect,  $z = 1.56$ ,  $p = .12$ ). We note that these null effects are potentially attributable to insufficient variability of medication dosage across participants in the sample (Table 1).

Our model of the confirmation bias effect postulates that prefrontal instruction representations bias striatal activation. As such, individual differences in prefrontal efficacy of instruction maintenance might predict the size of the effect, with greater representational capacity enhancing confirmation bias. We repeated our analyses of the DF trials and avoid F versus avoid D conditions, substituting a proxy for this capacity (MATRICS working memory score) for group membership. For DF trials, a trend toward an effect of working memory was observed,  $z = -0.48$ ,  $p = .088$ , reflecting an association of greater working memory with greater confirmation bias. For avoid D versus avoid F, we found no such relationship (no condition  $\times$  working memory interaction:  $z = -0.057$ ,  $p = .55$ ); nor was there a main effect of working memory,  $z = 0.13$ ,  $p = .36$ . We additionally entered working memory score as a covariate in addition to group in the two models. In neither case were effects of working memory observed (DF trials: no working memory effect,  $z = -0.17$ ,  $p = .58$ ; avoid D vs. avoid F model, no test condition  $\times$  working memory interaction,  $z = 0.03$ ,  $p = .77$ ). The effect of group in the DF trial model remained significant,  $z = 0.73$ ,  $p = .019$ , and a trend was observed for the group  $\times$  condition interaction in the avoid D versus avoid F model,  $z = 0.19$ ,  $p = .086$ .

Finally, we considered the generalizability of the observed effects, given the subset of participants who were unable to acquire the task contingencies (see the Method section). We refit our models of confirmation bias to all 86 participants. The observed pattern of effects persisted, with a significant group difference in DF trials,  $z = 2.37$ ,  $p = .018$ , and a marginal difference in avoid D versus avoid F trials,  $z = 0.17$ ,  $p = .066$ .

#### Instructed probabilistic selection task: Posttest reward frequency estimates

We utilized a multilevel linear regression to test for group differences in posttest estimates of stimulus reward frequencies (Fig. 4). The frequently rewarded stimuli and the frequently punished stimuli were grouped into a two-level factor (positive, A, C, G; negative, B, D, H; instructed pair excluded) and were entered, along with participant group, as effect-coded (SZ = +1, HC = -1) independent variables into a model



**Fig. 4** Posttest reward frequency estimates for instructed probabilistic selection task. After completing the task, participants were shown each stimulus and were asked to estimate the number of times each would pay +1 point if selected 100 times. Relative to healthy controls (HC), patients (SZ) underestimated the value of the more frequently rewarding (positive) stimuli (pos: A, C, G). The groups did not differ in estimation of the value of the infrequently rewarding (negative) stimuli (neg: B, D, H) or the instructed stimulus (F) or its pair (E). Error bars reflect standard errors of the means

predicting posttest reward frequency estimates. We observed main effects of both group,  $t = -2.29$ ,  $p = .022$ , and reward frequency,  $t = 9.47$ ,  $p < .0001$ , that were moderated by a significant group  $\times$  frequency interaction,  $t = -2.11$ ,  $p = .035$ . This interaction was driven by underestimation of the reinforcement frequencies of positive stimuli,  $t(71) = 3.291$ ,  $p = .002$ , in patients (mean difference from expected value =  $-24.6$ ,  $SD = 17.8$ ), relative to HCs (mean =  $-13.0$ ,  $SD = 12.0$ ). By contrast, patients and controls did not differ in the extent to which their estimations of reward frequency for frequently punished stimuli differed from actual expected value,  $t(71) = 0.108$  (both groups' estimates deviated by  $< 1\%$ ). Patients and controls also did not differ in the extent to which their estimations of reward frequency for the instructed stimulus [F; SZ mean =  $4.5$ ; HC mean =  $10.9$ ;  $t(71) = 1.112$ ] or its counterpart [E; SZ mean =  $-15.8$ ; HC mean =  $-12.0$ ;  $t(71) = 0.689$ ,  $ps > .2$ ] differed from actual expected value.

#### Choice bias probabilistic selection task: Training phase

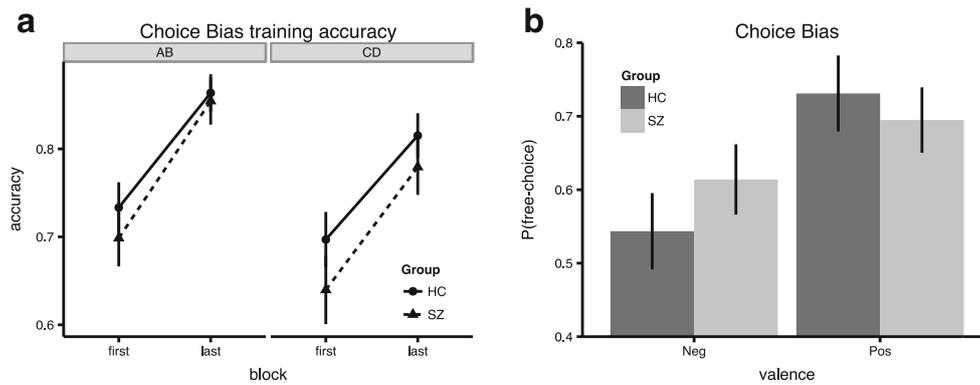
We assessed choice accuracy in the training phase of the choice bias task by entering training block (first, last), stimulus condition ( $A_c B_c$ ,  $C_c D_c$ ), and group (HC, SZ) as independent variables into a multilevel logistic regression (all factors effect coded: +1, -1). Both intercept and training block terms were significantly positive, demonstrating, respectively, that average task performance was better than chance and that performance improved during training,  $zs > 8$ ,  $p < .0001$ . The

relative difficulty of the probabilistic discrimination is reflected by a significantly positive stimulus condition term, with performance on  $A_c B_c$  being better overall,  $z = 3.3$ ,  $p = .001$ , and showing the most improvement across training blocks (condition by block interaction:  $z = 2.1$ ,  $p = .03$ ). We found no main effect of group on training accuracy and no interaction, all  $zs < 1$ , all  $ps > .4$  (Fig. 5a), nor did the groups differ in number of training trials required to meet criteria (mean HC =  $2.3$ , SZ =  $2.5$ ,  $p = .3$ ).

#### Choice bias probabilistic selection task: Test phase

First, we assessed participants' ability to discriminate among the stimuli they had learned about during the training phase. This was done by determining how reliably participants chose the statistically best (.8 reward probability:  $A_c$  and  $A_{nc}$ ) and avoided the worst (.2 reward probability:  $B_c$  and  $B_{nc}$ ) stimuli. We subdivided our analysis further according to whether participants were free or forced to select those stimuli. As such, we entered group (HC, SZ), test condition (choose A, avoid B), and choice (choice, no choice) as independent variables into a multilevel logistic regression (choose  $A_c$ ,  $A_{nc}$ ,  $C_c$ ,  $A_c D_c$ ; avoid  $B_c$ ,  $B_{nc}$ ,  $C_c$ ,  $B_c D_c$ ; choose  $A_{nc}$ ,  $A_{nc} C_{nc}$ ,  $A_{nc} D_{nc}$ ; avoid  $B_{nc}$ ,  $B_{nc} C_{nc}$ ,  $B_{nc} D_{nc}$ ). A significant positive intercept,  $z = 10.6$ ,  $p < .0001$ , shows that performance at test was above chance. There was also a significant condition  $\times$  choice interaction,  $z = -2.4$ ,  $p = .02$ , which was driven primarily by poor performance avoiding  $B_{nc}$ . Again, we found no effect of group alone or for any of the factors, all  $zs < 1$ , all  $ps > .5$ .

We assessed choice bias directly as a preference for selecting stimuli that had been freely selected during the training phase. To do so, we entered group (HC, SZ) and value (A, B, C, D) into a multilevel logistic regression (all factors effect coded: +1, -1), to assess preference for choice stimuli on  $A_c A_{nc}$ ,  $B_c B_{nc}$ ,  $C_c C_{nc}$ , and  $D_c D_{nc}$  trials. A significantly positive intercept,  $z = 5.5$ ,  $p < .001$ , indicates that participants exhibited an overall preference for stimuli they had chosen freely. There was a significant effect of value,  $\chi^2 > 25$ ,  $p < .001$ , indicating that this preference varied with the learned reward value of the stimuli. Critically, there was no main effect of group,  $z = 0.23$ ,  $p = .8$ , and no interaction with value condition,  $\chi^2 = 5.4$ ,  $p > .1$ . To investigate this effect of value further, we contrasted the choice bias for good (A and C) versus bad (B and D) stimuli. This analysis revealed that participants show a greater choice bias for good stimuli,  $z = 3.8$ ,  $p < .001$ , but no main effect of or interaction with group, all  $zs < 1.3$ , all  $ps > .1$  (Fig. 5b). Next, we probed for any influence of medication (to the extent afforded by the sampled dosage variance) but found no effects of haldol equivalent dosage on choice bias (main effect of dosage,  $z = 0.17$ ,  $p = .86$ ; dosage  $\times$  stimulus pair interaction,  $\chi^2 = 1.4$ ,  $p = .69$ ).



**Fig. 5** Choice bias task results. **a** In the training phase, patients (SZ) did not differ from healthy controls (HCs) in learning task contingencies. Probabilities of +1: A/B, .8/.2; C/D, .7/.3. **b** In the test phase, greater choice bias was observed for the positive stimuli (i.e., highest reward

value; pos: A and C) than for the negative stimuli (neg: B and E), but no difference was observed between patients and controls. Error bars reflect standard errors of the means

Finally, we assessed whether the effects were robust to the inclusion participants who were unable to acquire the reward contingencies. Neither the sign of any of the estimates nor their significance differed when all participants were included (no group effects or interactions:  $z_s < 1.61$ ,  $p_s > .1$ ).

#### Cross-task relationship

The confirmation bias and choice bias tasks were designed to assess prefrontal–striatal communication and striatal function, respectively. The discrepant group differences between tasks suggest that SZ patients have more pronounced impairments in the former than in the latter. To test this impression, we tested whether the difference between patients and controls was larger in the confirmation bias than in the choice bias task. We  $z$ -transformed bias measures separately for the two different tasks ( $z$ -score of the choice bias measure and average of the  $z$ -scores for the two confirmation bias measures) and modeled this outcome variable in a multilevel linear model with task and group as independent variables and random effects of intercept and task by participant. In the 59 participants completing both tasks, we observed a significant interaction of task and group,  $\chi^2 = 4.32$ ,  $p = .038$ , indicating a greater difference between groups in confirmation bias than in choice bias. This effect remained when controlling for the within-subjects difference in training blocks between tasks (task  $\times$  group interaction,  $\chi^2 = 4.03$ ,  $p = .045$ ; no task  $\times$  block difference interaction,  $\chi^2 = 0.99$ ,  $p = .31$ ). We additionally assessed this effect in all participants, although it failed to reach significance,  $\chi^2 = 2.27$ ,  $p = .13$ .

#### Cognitive correlates of negative symptoms

Prior work suggests that reinforcement learning deficits observed in SZ are associated with negative symptoms (Gold et al., 2012; Strauss et al., 2011; Waltz et al., 2007; Waltz et al., 2011; Waltz et al., 2010). We looked for evidence of this

relationship in the present data set by reestimating the models above in the patient group alone and adding summed scores on the SANS avolition and anhedonia subscales as a covariate (utilization of SANS total score as a covariate produced a similar pattern of results). Although the estimated effect of negative symptom severity on acquisition performance was indicative of a negative relationship in all cases, we observed no significant effects of negative symptoms or interactions between negative symptoms and performance measures from the experimental tasks (all  $p_s > .13$ ). We also examined the role of negative symptoms on choose A and avoid B performance, using a categorical approach based on median splits on SANS total score and the avolition and anhedonia scales (as used in Gold et al., 2012), but did not observe any significant effects ( $p_s > .13$ ).

When we examined relationships between negative symptom severity and the accuracy of reinforcement frequency estimates made following the test phase of the confirmation bias task, we found that total scores from the SANS correlated inversely with the accuracy of reward probability estimates from frequently rewarded stimuli,  $r = -.403$ ,  $p = .022$ . This suggests that SZ patients with the most severe negative symptoms showed the greatest tendency to underestimate the reward probabilities of positive stimuli, consistent with findings from several of our previous studies (Gold et al., 2012; Strauss et al., 2011; Waltz et al., 2011).

#### Discussion

Here, we employed learning tasks hypothesized to probe the function and interaction of the basal ganglia and prefrontal cortex in an attempt to assess how these regions are compromised by SZ. The results suggest relatively greater impairments in fronto-striatal communication than in more purely striatal function. The tasks used are variants of a widely studied reinforcement learning task, which has been

repeatedly shown to assay dopaminergic effects on learning (Frank et al., 2007; Frank et al., 2004), consistent with a computational model of the basal ganglia (Frank, 2005) and BOLD activation of this region in fMRI studies utilizing this task (Jocham, Klein, & Ullsperger, 2011; Shiner et al., 2012). The instructed PS task utilized here is hypothesized to interrogate the interaction of the prefrontal cortex and striatum. Indeed, our prior work has supported a model of this task in which instruction representations in the prefrontal cortex bias activation levels in the striatum, causing a confirmation bias, whereby instructed stimuli are overvalued relative to their experienced value (Doll et al., 2011; Doll et al., 2009). The choice bias task utilized here is hypothesized to assess how actions learned through volitional choice are evaluated, in comparison with actions learned through exogenous choice. Learning in both cases is hypothesized to be dependent on the striatum, but the act of volitionally choosing is expected to confer a “boost” to the value of the endogenously chosen stimulus over the exogenously chosen one. Indeed, recent (although limited) evidence indicates that individual differences in choice bias in this task are related to striatal genotypes (Cockburn, Collins, & Frank, 2014), and imaging studies indicate that this type of bias is related to striatal value coding (Sharot et al., 2009).

Consistent with our predictions, patients were less susceptible to the confirmation bias effect in the instructed task, showing better learning of the veridical contingencies when presented with the instructed stimulus in a postlearning test phase, an unusual instance of an impairment in patients conferring a paradoxical performance advantage, relative to controls. Patients who learned both tasks showed larger differences from controls in this putative measure of prefrontal–striatal communication than in the choice bias task, which we hypothesize measures striatal function more exclusively. Moreover, with the exception of the most deterministic discrimination in the instructed task, patients showed relatively intact reinforcement learning. They showed no significant deficits in the choice bias task, in either the endogenous or the exogenous component. While patients in the instructed task were, overall, worse than controls at the uninstructed conditions, these effects were most apparent in the most deterministic learning condition (A/B: .9/.15 reward probability). This somewhat surprising result is consistent with previous observations in patients (Gold et al., 2012) and may be explained by partially dissociable neural substrates in deterministic, as compared with probabilistic, discriminations. Such nearly-deterministic discriminations may rely on rule (Bunge, Kahn, Wallis, Miller, & Wagner, 2003) or value (Frank & Claus, 2006) representation in the prefrontal cortex, in addition to striatal function. A previous report tested predictions of a model of the orbitofrontal cortex and basal ganglia, with results suggesting impaired processing in this frontal region in SZ (Gold et al., 2012). In agreement,

orbitofrontal damage has been associated with impairments in maximizing reward in discriminations with low (Tsuchida, Doll, & Fellows, 2010), rather than high, stochasticity (Noonan et al., 2010; Riceberg & Shapiro, 2012). Thus, patient impairment at the easiest of reinforcement contingencies may be consistent with their relative immunity from the confirmation bias effect, in that both rely on prefrontal integrity.

A possible limiting factor of the present findings is that not all participants were able to acquire the task contingencies. Although the key results within each task (although not between tasks) were robust to inclusion of all participants, we focused on the results in participants demonstrating task learning. This decision was motivated by the fact that, absent any learning, we should also expect no bias—not because of a reduction in bias itself but, rather, because of near-chance performance overall. Thus, the conclusions here strictly apply to the included participants. Exclusion may indeed constrain the generalizability of these results to SZ patients. The excluded participants might, for example, represent a subgroup with possibly more pronounced striatal dysfunction that limited task performance, perhaps in addition to fronto-striatal dysfunction. Of the measured characterizing variables, excluded participants performed worse on measures of cognitive ability. However, it is worth noting that the vast majority of patients (45) showed learning in at least one task, suggesting greater generalizability than the number of excluded patients from either task alone.

Another potential limitation of the results described here is that medication may have had unexpected effects on the behaviors hypothesized to rely on striatal and frontal function. Although we found no medication effects on task variables across patients, the present design is not ideal for uncovering these effects. Medication type and dosage were not randomly assigned here, confounding medication effects with treatment responsiveness and overall clinical severity of illness. Thus, the interpretation of correlations (or lack thereof) between behavioral performance and antipsychotic dose is not straightforward. Future work assessing these behavioral effects in medication-naïve patients would help to resolve this matter.

As was described above, our predictions about, and interpretations of, the results were based on formal modeling efforts and (to a greater extent in the confirmation bias task) supporting results across a number of methodologies. However, we collected no neural measurements in the present sample. The absence of such measurements precludes a strong neural interpretation of the data, which awaits further research. One alternative account of the present data is that the reduced confirmation bias in patients is owing to dysfunction in dopaminergic reward prediction error signaling to the striatum, rather than dysfunction in prefrontal–striatal communication. Indeed, this interpretation has previously been applied to impairments observed in patients with SZ (Waltz et al.,

2007; Waltz et al., 2011; but see Gold et al., 2012, which suggests that these asymmetric learning effects are of frontal, rather than striatal, origin), consistent with observed abnormalities in striatal presynaptic dopamine (Fusar-Poli & Meyer-Lindenberg, 2013). Furthermore, we observed an uninstructed reward learning deficit in the present data. We do not favor this alternative hypothesis, given our prior theory and empirical work, as well as the statistical controls utilized here. In particular, we controlled for the possibility that uninstructed learning deficits produced the confirmation bias effect and found the evidence for bias to be robust to these controls. Moreover, the type of learning deficits observed in the present task seem inconsistent with the view that a striatal reward learning deficit accounts for patient behavior. As was discussed above, learning deficits between groups in the uninstructed case are most clearly observed on the most deterministic choice discriminations. A pure deficit in striatal-dependent learning would predict that these impairments would be most readily observed in the most stochastic choice conditions, where the long-term integration of reward prediction errors over repeated experience is an adaptive strategy for establishing a response policy. For more deterministic discriminations, a working-memory-reliant strategy of remembering the previous outcome and adjusting choice accordingly may suffice (Collins & Frank, 2012). Further research should seek to disentangle these possibilities.

Prior work that motivated the present study has suggested that instructions bias striatal mechanisms via excitatory input from the prefrontal cortex (Doll et al., 2011; Doll et al., 2009), which amplifies positive prediction errors and diminishes negative ones. Although this view has been supported by evidence from neuroimaging and behavioral genetics (Biele et al., 2011; Doll et al., 2011; Staudinger & Büchel, 2013), several other reports appear partially inconsistent (Li, Delgado, & Phelps, 2011; Walsh & Anderson, 2011). These reports are more in line with a model in which prefrontal instruction representations override the striatum for control of behavior (see override model, Doll et al., 2009). Li et al., for example, found diminished negative and also diminished positive reward prediction error signaling in the striatum for instructed, relative to uninstructed, probabilistic discriminations. Consistent with the view advanced here, activation of the prefrontal cortex was also found in the instructed condition, although this activity predicted prediction error blunting in general, instead of the asymmetric enhancement/blunting predicted by our model (but see Biele et al., 2011). In a similar design, Walsh and Anderson (2011) found prediction-error-like learning signals (albeit utilizing EEG instead of fMRI) that did not differ across instructed and uninstructed conditions, although in the instructed case, they became uncoupled from behavior. In these studies, explicit, accurate reward probabilities served as instruction, and this written information (or memorized information-associated cues) was

presented on every trial, perhaps encouraging an alternative, reasoning-based task solution. While a complete understanding of the neural mechanisms by which instructions exert control on behavior is lacking, fronto-striatal coordination is commonly implicated in the extant literature (regardless of whether these regions cooperate or compete). As such, the reduced susceptibility to inaccurate instruction observed here in patients with SZ appears most suggestive of fronto-striatal impairment.

Our model of the confirmation bias effect posits a role for the working memory function of the prefrontal cortex in maintaining the instructions (which permits the biasing of striatal learning). In support of this view, in previous work (Doll et al., 2011), we found that individual differences in a polymorphism of the COMT val158met genotype predicted adherence to inaccurate instruction during training, as well as a confirmation bias measured at test. In particular, Met alleles of this genotype, which are associated with enhanced working memory (de Frias et al., 2010; Egan et al., 2001; Tunbridge et al., 2004), putatively due to higher levels of dopamine in the prefrontal cortex (Gogos et al., 1998; Slifstein et al., 2008), predicted the effect. In the present sample, we found only a weakly suggestive relationship between working memory (as measured by MATRICS-working memory domain) and confirmation bias (as measured by preference on DF test trials). One possibility is that the working memory measure used here insufficiently probes the relevant cognitive function. Future work seeking to measure dissociable components of working memory should be better suited for the measurement and interpretation of associations with confirmation bias.

One motivation for having the same participants complete both tasks was to test our hypothesis that negative symptoms reflect failures to represent positive expected reward value of choices in the prefrontal cortex, rather than dysfunction at the level of prediction error signaling in the striatum. Thus, we expected that patients with high levels of negative symptoms would demonstrate the greatest reduction in confirmation bias, as well as the greatest impairment in uninstructed positive reward learning. In contrast, we expected that choice bias would not be related to negative symptoms. That is, if choice bias occurs on the basis of purely striatal mechanisms, we would not expect a symptom effect. Our results provide mixed support for this view. We did not see a negative symptom effect on the confirmation bias measure. However, we did observe the expected undervaluation of positive reward value, coupled with intact estimation of the value of relatively aversive outcomes, replicating prior results (Gold et al., 2012; Strauss et al., 2011; Waltz et al., 2011). This impairment correlated with negative symptom scores across patients. It is possible that our failure to observe a negative symptom effect on confirmation bias reflects the fact that bias impairment is characteristic of patients with varying symptom



- François-Brosseau, F.E., Martinu, K., Strafella, A.P., Petrides, M., Simard, F., Monchi, O., 2009. Basal ganglia and frontal involvement in self-generated and externally-triggered finger movements in the dominant and non-dominant hand. *Eur J Neurosci* 29, 1277–1286.
- Fusar-Poli, P., Meyer-Lindenberg, A., 2013. Striatal presynaptic dopamine in schizophrenia, part ii: Meta-analysis of [(18)f/(11)c]-dopa pet studies. *Schizophr Bull* 39, 33–42.
- Gogos, J.A., Morgan, M., Luine, V., Santha, M., Ogawa, S., Pfaff, D., Karayiorgou, M., 1998. Catechol-o-methyltransferase-deficient mice exhibit sexually dimorphic changes in catecholamine levels and behavior. *Proc Natl Acad Sci U S A* 95, 9991–9996.
- Gold, J.M., Hahn, B., Strauss, G.P., Waltz, J.A., 2009. Turning it upside down: areas of preserved cognitive function in schizophrenia. *Neuropsychol Rev* 19, 294–311.
- Gold, J.M., Waltz, J.A., Matveeva, T.M., Kasanova, Z., Strauss, G.P., Herbener, E.S., Collins, A.G.E., Frank, M.J., 2012. Negative symptoms and the failure to represent the expected reward value of actions: behavioral and computational modeling evidence. *Arch Gen Psychiatry* 69, 129–138.
- Green, M.F., Nuechterlein, K.H., Gold, J.M., Barch, D.M., Cohen, J., Essock, S., Fenton, W.S., Frese, F., Goldberg, T.E., Heaton, R.K., Keefe, R.S.E., Kern, R.S., Kraemer, H., Stover, E., Weinberger, D.R., Zalcman, S., Marder, S.R., 2004. Approaching a consensus cognitive battery for clinical trials in schizophrenia: The nimh-matrix conference to select cognitive domains and test criteria. *Biol Psychiatry* 56, 301–307.
- Jocham, G., Klein, T.A., Ullsperger, M., 2011. Dopamine-mediated reinforcement learning signals in the striatum and ventromedial prefrontal cortex underlie value-based choices. *J Neurosci* 31, 1606–1613.
- Kirkpatrick, B., Strauss, G.P., Nguyen, L., Fischer, B.A., Daniel, D.G., Cienfuegos, A., Marder, S.R., 2011. The brief negative symptom scale: psychometric properties. *Schizophr Bull* 37, 300–305.
- Lesh, T.A., Niendam, T.A., Minzenberg, M.J., Carter, C.S., 2011. Cognitive control deficits in schizophrenia: Mechanisms and meaning. *Neuropsychopharmacology* 36, 316–338.
- Li, J., Delgado, M.R., Phelps, E.A., 2011. How instructed knowledge modulates the neural systems of reward learning. *Proc Natl Acad Sci U S A* 108, 55–60.
- MATRICES, 1996. MATRICES Consensus Cognitive Battery Administration Manual. Matrics Assessment Inc. Los Angeles, CA.
- Nickerson, R.S., 1998. Confirmation bias: a ubiquitous phenomenon in many guises. *Review of General Psychology* 2, 175–220.
- Noonan, M.P., Walton, M.E., Behrens, T.E.J., Sallet, J., Buckley, M.J., Rushworth, M.F.S., 2010. Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. *Proc Natl Acad Sci U S A* 107, 20547–20552.
- Overall, J.E., Gorman, D.R., 1962. The brief psychiatric rating scale. *Psychological Reports* 10, 899–912.
- Pfohl, B., Blum, N., Zimmerman, M., 1997. Structured Interview for DSM-IV personality (SID-P). Washington, DC: American Psychiatric Press.
- Ragland, J.D., Laird, A.R., Ranganath, C., Blumenfeld, R.S., Gonzales, S.M., Glahn, D.C., 2009. Prefrontal activation deficits during episodic memory in schizophrenia. *Am J Psychiatry* 166, 863–874.
- Riceberg, J.S., Shapiro, M.L., 2012. Reward stability determines the contribution of orbitofrontal cortex to adaptive behavior. *J Neurosci* 32, 16402–16409.
- Schielzeth, H., Forstmeier, W., 2009. Conclusions beyond support: Overconfident estimates in mixed models. *Behav Ecol* 20, 416–420.
- Sharot, T., De Martino, B., Dolan, R.J., 2009. How choice reveals and shapes expected hedonic outcome. *J Neurosci* 29, 3760–3765.
- Shiner, T., Seymour, B., Wunderlich, K., Hill, C., Bhatia, K.P., Dayan, P., Dolan, R.J., 2012. Dopamine and performance in a reinforcement learning task: Evidence from parkinson's disease. *Brain* 135, 1871–1883.
- Slifstein, M., Kolachana, B., Simpson, E.H., Tabares, P., Cheng, B., Duvall, M., Frankle, W.G., Weinberger, D.R., Laruelle, M., Abi-Dargham, A., 2008. Comt genotype predicts cortical-limbic d1 receptor availability measured with [(11)c]nnc112 and pet. *Mol Psychiatry* 13, 821–827.
- Somlai, Z., Moustafa, A.A., Kéri, S., Myers, C.E., Gluck, M.A., 2011. General functioning predicts reward and punishment learning in schizophrenia. *Schizophr Res* 127, 131–136.
- Staudinger, M.R., Büchel, C., 2013. How initial confirmatory experience potentiates the detrimental influence of bad advice. *Neuroimage* 76, 125–133.
- Strauss, G.P., Frank, M.J., Waltz, J.A., Kasanova, Z., Herbener, E.S., Gold, J.M., 2011. Deficits in positive reinforcement learning and uncertainty-driven exploration are associated with distinct aspects of negative symptoms in schizophrenia. *Biol Psychiatry* 69, 424–431.
- Tsuchida, A., Doll, B.B., Fellows, L.K., 2010. Beyond reversal: A critical role for human orbitofrontal cortex in flexible learning from probabilistic feedback. *J Neurosci* 30, 16868–16875.
- Tunbridge, E.M., Bannerman, D.M., Sharp, T., Harrison, P.J., 2004. Catechol-o-methyltransferase inhibition improves set-shifting performance and elevates stimulated dopamine release in the rat prefrontal cortex. *J Neurosci* 24, 5331–5335.
- Walsh, M.M., Anderson, J.R., 2011. Modulation of the feedback-related negativity by instruction and experience. *Proc Natl Acad Sci U S A* 108, 19048–19053.
- Waltz, J.A., Frank, M.J., Robinson, B.M., Gold, J.M., 2007. Selective reinforcement learning deficits in schizophrenia support predictions from computational models of striatal-cortical dysfunction. *Biol Psychiatry* 62, 756–764.
- Waltz, J.A., Frank, M.J., Wiecki, T.V., Gold, J.M., 2011. Altered probabilistic learning and response biases in schizophrenia: Behavioral evidence and neurocomputational modeling. *Neuropsychology* 25, 86–97.
- Waltz, J.A., Schweitzer, J.B., Ross, T.J., Kurup, P.K., Salmeron, B.J., Rose, E.J., Gold, J.M., Stein, E.A., 2010. Abnormal responses to monetary outcomes in cortex, but not in the basal ganglia, in schizophrenia. *Neuropsychopharmacology* 35, 2427–2439.
- Wechsler, D., 1999. Wechsler Abbreviated Scale of Intelligence (WASI). San Antonio TX: The Psychological Corporation.
- Wechsler, D., 2001. Wechsler Test of Adult Reading (WTAR). San Antonio, TX: The Psychological Corporation.
- Weickert, T.W., Terrazas, A., Bigelow, L.B., Malley, J.D., Hyde, T., Egan, M.F., Weinberger, D.R., Goldberg, T.E., 2002. Habit and skill learning in schizophrenia: Evidence of normal striatal processing with abnormal cortical input. *Learn Mem* 9, 430–442.