



Published in final edited form as:

J Abnorm Psychol. 2016 August ; 125(6): 777–787. doi:10.1037/abn0000164.

Reduced Model-Based Decision-Making in Schizophrenia

Adam J. Culbreth, M.A.¹, Andrew Westbrook, M.A.¹, Nathaniel D. Daw, Ph.D.², Matthew Botvinick, Ph.D.², and Deanna M. Barch, Ph.D.^{1,3}

¹Department of Psychological and Brain Sciences, Washington University in Saint Louis

²Princeton Neuroscience Institute and Department of Psychology, Princeton University

³Department of Psychiatry & Radiology, Washington University in Saint Louis

Abstract

Background—Individuals with schizophrenia have a diminished ability to use reward history to adaptively guide behavior. However, tasks traditionally used to assess such deficits often rely on multiple cognitive and neural processes, leaving etiology unresolved. In the current study, we adopted recent computational formalisms of reinforcement learning to distinguish between model-based and model-free decision-making in hopes of specifying mechanisms associated with reinforcement-learning dysfunction in SZ. Under this framework, decision-making is model-free to the extent that it relies solely on prior reward history, and model-based if it relies on prospective information such as motivational state, future consequences, and the likelihood of obtaining various outcomes.

Methods—Model-based and model-free decision-making was assessed in 33 schizophrenia patients and 30 controls using a 2-stage 2-alternative forced choice task previously demonstrated to discern individual differences in reliance on the two forms of reinforcement-learning.

Results—We show that, compared to controls, schizophrenia patients demonstrate decreased reliance on model-based decision-making. Further, parameter estimates of model-based behavior correlate positively with IQ and working memory measures, suggesting that model-based deficits seen in schizophrenia may be partially explained by higher-order cognitive deficits.

Conclusions—These findings demonstrate specific reinforcement-learning and decision-making deficits and thereby provide valuable insights for understanding disordered behavior in schizophrenia.

Keywords

Decision-making; computational modeling; model-based learning; schizophrenia

Introduction

Schizophrenia (SZ) has long been characterized by deficits in goal-directed decision-making. However, the underlying mechanisms are not well understood. Recent work has

suggested the involvement of specific reward-learning deficits, namely that SZ patients have difficulties creating representations for the value of various actions, and in utilizing such representations to drive behavior (Gold, Waltz, Prentice, Morris, & Heerey, 2008). Given the existence of multiple decision-making processes and systems, a critical next step in this line of research is to identify specific mechanisms underlying aberrant value-learning. In the current study we use a recent reinforcement-learning (RL) framework to identify the contributions of separable value-learning systems to decision-making. This framework formalizes the idea that the value of actions can be learned either solely by considering prior reward history (model-free), or by taking into account the structure of the environment and future consequences of actions (model-based) (Daw, Niv, & Dayan, 2005).

Importantly, such RL algorithms have a number of attractive features. For example, they link data collected at many different levels of analysis (e.g., biological, behavioral), and thus, are well suited for linking symptomology of disorders, such as SZ, back to clinically-relevant, neural and psychological mechanisms. Further, they generate precise quantitative estimates of parameters, which are proposed to govern learning processes (e.g., learning rate). These precise estimates, in turn, afford specific predictions about the neural dynamics and behavioral correlates of reward learning. In addition, RL models are supported by a wealth of human and animal, neurophysiological and behavioral evidence, and they hold great potential for increasing the precision and sophistication of our current understanding of psychiatric disorders (Deserno, Boehme, Heinz, & Schlagenhauf, 2013). Here we leverage a recently developed RL framework to specify aberrant decision-making in SZ in terms of the relative contributions of putative model-free and model-based systems.

Broadly, decision-makers can express either habitual or goal-directed behavior (Balleine, Daw, & O'Doherty, 2008; Dickinson & Balleine, 2002). Habitual behavior is characterized by the repetition of rewarded actions (or avoidance of punished action) (Thorndike, 1927). For example, children may learn to refrain from touching a hot stove that burned them. In contrast, goal-directed behavior reflects decisions made prospectively by weighing actions based on an "internal model" of actions and their probable outcomes (Tolman, 1948). For example, adults may learn to grab the handle of a hot oven if doing so enables them to subsequently extract a delicious cake. Importantly, goal-directed decision-making allows for flexible and exploratory behavior; reductions in this system having been associated with rigid and problematic decision-making in several psychiatric populations (Gillan, Kosinski, Whelan, Phelps, & Daw, 2016; Voon et al., 2014). In the case of SZ, use of inflexible, habitual decision-making systems has been thought to underlie the formation of delusions, where individuals show a bias against disconfirmatory evidence (i.e., alternative explanations) for irrational belief structures (Woodward, Moritz, Cuttler, & Whitman, 2006). Further, use of rigid and inflexible decision-making is characteristic of the negative symptoms of SZ where patients show a reduction in the variety of pleasurable activities they pursue on a daily basis, and a reduced ability to mentally generate such pleasurable options (Hartmann et al., 2015; Strauss, Waltz, & Gold, 2014).

Recently, the post-training distinction between habitual and goal-directed behavior has been proposed to arise from two distinct computational learning rules, model-based and model-free. This account enables the precise quantitative assessment of the relative contributions,

and thereby neural mechanisms of these systems during decision-making (Balleine, et al., 2008; Daw, et al., 2005). The model-free system, for example, has become associated with reward prediction error signaling in midbrain dopamine neurons and, in human fMRI, their targets in areas like the ventral striatum (Daw, Gershman, Seymour, Dayan, & Dolan, 2011; Gläscher, Daw, Dayan, & O'Doherty, 2010; Schultz, Dayan, & Montague, 1997). In contrast, model-based decision-making is reliant on a network of regions needed for representing internal models of the environment to bias value functions including lateral and dorsolateral prefrontal cortex (dlPFC) (Smittenaar, FitzGerald, Romei, Wright, & Dolan, 2013), and ventromedial and orbital prefrontal cortices (vmPFC and OFC) (Lee, Shimojo, & O'Doherty, 2014; McDannald, Lucantonio, Burke, Niv, & Schoenbaum, 2011) (Gläscher, et al., 2010; Lee, et al., 2014), as well as ventral striatum (Daw, et al., 2011; McDannald, et al., 2011). Recent work has also pointed to a critical role of ventral striatal presynaptic dopamine in modulating the relative contributions of these systems (Deserno et al., 2015).

Previous reports have explored mechanisms associated with both habitual and goal-directed control in SZ. For example, SZ patients have impairments in cognitive systems associated with model-based decision-making including proactive cognitive control and working memory (Barch & Ceaser, 2012; Otto, Raio, Chiang, Phelps, & Daw, 2013; Otto, Skatova, Madlon-Kay, & Daw, 2014). SZ patients also have functional and structural abnormalities in brain regions thought to support model-based behavior (dlPFC: (Barch, Csernansky, Conturo, & Snyder, 2002; Minzenberg, Laird, Thelen, Carter, & Glahn, 2009; Semkovska, Bédard, Godbout, Limoge, & Stip, 2004); vmPFC: (Hooker, Bruce, Lincoln, Fisher, & Vinogradov, 2011; Park, Park, Chun, Kim, & Kim, 2008); OFC: (Gur et al., 2000; Pantelis et al., 2003; Plailly, d'Amato, Saoud, & Royet, 2006). SZ patients are also impaired at goal-directed control as assessed using outcome devaluation (Morris, Quail, Griffiths, Green, & Balleine, 2015). In contrast, SZ patients often show intact functioning in apparently model-free RL processes, such as, procedural learning (Kéri et al., 2000; Weickert et al., 2002) and implicit reinforcement-learning (Heerey, Bell-Warren, & Gold, 2008). However, such interpretations are complicated by studies that show impaired ventral striatal prediction error signaling (the key mechanism believed to support model-free learning) in schizophrenia (Juckel, Schlagenhauf, Koslowski, Wüstenberg, et al., 2006; Murray et al., 2007; Schlagenhauf et al., 2014), and are also behaviorally impaired at blocking tasks, a classic test of model-free error-driven learning (Moran, Al-Uzri, Watson, & Reveley, 2003). Predictions of intact model-free learning and impaired model-based learning are consistent with a recent computational study showing that while basic learning of stimulus-response contingencies is intact in schizophrenia, deficits in working memory interact to undermine learning that relies on internal models of the environment (Collins, Brown, Gold, Waltz, & Frank, 2014). However, relative reliance on model-based and model-free RL has never been directly tested in SZ. Applying computational formalisms of these very different forms of learning may allow specific and quantifiable characterization of aberrant value-learning in SZ, and also provide testable hypotheses about specific mechanisms, enabling development of targeted interventions (Deserno, et al., 2013; Montague, Dolan, Friston, & Dayan, 2012).

To investigate relative reliance on model-based versus model-free control in decision-making we utilized a previously validated 2-stage Markov decision task (Daw, et al., 2011; Otto, et al., 2013; Otto, et al., 2014; Smittenaar, et al., 2013; Voon, et al., 2014; Wunderlich,

Smittenaar, & Dolan, 2012). In the first-stage, participants decide between two alternatives, each of which yields a common (70%) or rare (30%) transition to a respective second-stage, where they again choose between two alternatives that are rewarded probabilistically (Figure 1). Reward contingencies of second-stage choices vary trial-by-trial, so that participants must learn from experience to maximize rewarded outcomes. The fact that first-stage choices are rewarded only indirectly, via the second-stage, allows distinguishing model-based from model-free learning strategies, because the former is characterized by evaluating options prospectively in terms of a model of their consequences. Thus, if the likelihood of repeating first-stage choices is based solely on the prior trials being rewarded; choice behavior is considered model-free (i.e., habitual). However, if the likelihood of repeating first-stage-choices depends on the interaction of prior rewards with the transition contingencies they are considered model-based (i.e., goal-directed).

We hypothesized that SZ patients would be more reliant on model-free systems compared to healthy controls. This prediction is consistent with previous reports, which show intact procedural learning and implicit RL in SZ (Bleuler, 1950; Heerey, et al., 2008; Kéri, et al., 2000; Weickert, et al., 2002). We also predicted that SZ patients would show diminished model-based behavior in the Markov decision-making task. This prediction is consistent with literature suggesting that SZ patients have deficits in proactive cognitive control and working memory capacity (Barch & Ceaser, 2012), as well as functional and structural abnormalities in brain regions associated with model-based RL. Furthermore, we hypothesized that patient deficits in model-based RL would be correlated with measures of working memory capacity and premorbid intellectual functioning. This prediction is consistent with the aforementioned literature linking model-based RL to higher order cognitive processes such as cognitive control and working memory capacity (Otto, et al., 2013; Otto, et al., 2014). Finally, we predicted that parameter estimates reflecting diminished model-based RL would be correlated with negative symptoms such that individuals with greater negative symptom severity would demonstrate decreased model-based behavior.

Materials & Methods

Participants

Participants were 33 individuals meeting DSM-IV criteria for SZ or schizoaffective disorder (SZA; N=13), and 30 controls (CN), with no personal or family history of psychosis, from the Saint Louis community. All SZ patients were stable outpatients in the chronic phase of the illness. SZ patients were recruited from outpatient clinics in the Saint Louis community. Controls were recruited through flyers. Exclusion criteria included 1) DSM-IV diagnosis of substance abuse or dependence in the past six months for all participants; 2) DSM-IV diagnosis of major depressive disorder or dysthymia in the past year for controls; 3) changes in medication type or dosage two weeks prior to consent; 4) past head injury with documented neurological sequelae and/or loss of consciousness. We did not exclude participants for current/previous anxiety or personality disorders in either group. One control participant met criteria for a previous depressive episode, but was currently in full remission. We excluded 4 controls (2 for previous head injury and 2 for current substance abuse) and 2

patients (for current substance abuse). The Washington University Institutional Review Board approved the study. Participants provided written informed consent in accordance with Washington University's Human Subject Committee's criteria.

Clinical Assessments

Diagnoses were determined by the Structured Clinical Interview for DSM-IV-TR (First, Spitzer, Gibbon, & Williams, 2001). Negative symptoms were assessed using the Brief Negative Symptom Scale (BNSS) (Kirkpatrick et al., 2011). Raters were master's level clinicians who participated in regular joint interview and rating sessions to ensure reliability. Anhedonia was also assessed using the Snaith-Hamilton Pleasure Scale (Snaith et al., 1995). Premorbid intellectual functioning was estimated using the Wechsler Test of Adult Reading (WTAR) (Wechsler, 2001). The WTAR is a brief reading recognition test, which has shown robust correlations to Wechsler Adult Intelligence Scale. All participants were required to pass a urine drug screen and a Breathalyzer test.

Sequential Learning Task

Participants completed a modified version of a two-stage decision task (Figure 1) (Daw, et al., 2011). Extensive instructions on the task structure and practice trials were completed prior to task administration (see supplement). At the start of each trial, two alternatives were presented. Participants' choices then led, probabilistically, to one of two second-stage "states" comprising two subsequent alternatives. Importantly, each first-stage choice led more frequently (70%) to one of two second-stage states (Figure 1). Choices in the second stage probabilistically resulted in visual feedback of reward or no-reward. In order to ensure learning throughout the task, the probability of receiving a reward for the four second-stage alternatives varied slowly according to Gaussian random walks. Altogether, participants completed 200 trials. Participants were given 3 seconds to make each choice. The interval between first and second-stage stimuli was 1 second, following choice. Feedback following the second-stage choice was presented for 1 second, followed by another 1 second before the next trial began. Participants were informed that they would receive increased bonus money for task accuracy. However, we paid participants a five-dollar bonus for completing the task, regardless of performance. Stimuli on the task consisted of spaceships and aliens, instead of the Tibetan characters in the original report (Daw, et al., 2011), to provide a more intuitive task environment for better engaging the patient population (Decker et al., under review).

N-Back Task

In order to assess the contribution of working memory to model-based and model-free learning estimates, a subset of participants (20 SZ; 20 CN) completed two versions of an N-Back task (1-Back and 2-Back). During the task, participants were instructed to identify letters, presented one at a time on a computer screen, as targets or non-targets. In a given level of the N-back, a letter is a target if the same letter was presented "N" trials before the current trial. Each version (1-Back and 2-Back) consisted of 64 trials and 16 targets per run. Interstimulus intervals were 2 seconds, thus runs were 128 seconds each, regardless of N-back level.

The sensitivity index, d' , was used to quantify N-back performance, controlling for target or non-target response biases. Raw d' values were adjusted by the “log linear” transformation to address extreme false-alarm and hit proportions (Hautus, 1995).

Task Data Analysis

Trial by trial learning was analyzed for signs of model-based vs. model-free updating in two ways (Daw, et al., 2011), first using a hierarchical linear regression examining selection of the same or a different first stage option based on the previous trial’s reward and transition type. This is a simplified approximation to a more detailed RL model, which we also fit. The RL model considers choices in light of the full history of preceding events.

Hierarchical Linear Model (HLM)—We fit a HLM following (Daw, et al., 2011). Analyses were performed using the lme4 linear mixed-effects package in R (Bates & Sarkar, 2007). The dependent variable was the first-stage choice (coded: stay or shift). Predictors included dummy variables indicating whether the previous trial was rewarded or not (1 and -1, respectively), whether the previous trial’s transition was rare or common (-1 and 1, respectively), and their interaction. All coefficients were taken as random effects across participants, and estimates are reported across participants. To observe group differences between model-based and model-free estimates, diagnostic group was dummy coded for controls and patients (0 and 1 respectively). IQ as determined by the Wechsler Test of Adult Reading (WTAR), clinician-rated negative symptoms, the Snaith-Hamilton pleasure scale, and olanzapine equivalent antipsychotic dose (Gardner, Murphy, O’Donnell, Centorrino, & Baldessarini, 2014) were group-centered and interacted as factors with task effects (i.e., reward, rarity, and the reward \times rarity interaction) to determine individual difference relationships.

RL Modeling—We modeled choice behavior as reflecting a combination of model-based and model-free value-learning, using an adaptation of a previously described hybrid RL model (Daw, et al., 2011; Doll, Shohamy, & Daw, 2015). Altogether, the task consisted of 3 states (stage 1. state a; stage 2. state b & c) with two possible actions at each state (a1, a2) (Figure 1). The model predicted choice behavior as reflecting trial-wise RL of the value of all six state-action pairs. Specific parameters were incorporated to quantify the extent to which individuals’ choice patterns reflected reliance on model-based or model-free estimates of state-action values during decision-making.

Model-free component—Model-free RL was computed using a modified SARSA (State-Action-Reward-State-Action) temporal difference-learning algorithm. This algorithm updates, at each stage i of trial t , the action value (Q_{MF}) for each state (s) action (a) pair visited:

$$Q_{MF}(s_{(i,t)}, a_{(i,t)}) = Q_{MF}(s_{(i,t)}, a_{(i,t)}) + \alpha \delta_{(t)}$$

where

$$\delta_{(t)} = r_{(t)} / \alpha - Q_{MF}(s_{(i,t)}, a_{(i,t)})$$

and α is a free learning-rate parameter, which we have set equivalent for both task stages. Following Doll et al. 2014, the reward term is divided by subjects' learning rates. This division does not influence overall choice likelihood. It does, however, rescale beta weights in the softmax choice rule, and reduces the correlation between subjects' beta weights and their learning rates, thereby enhancing parameter estimation.

Note that the update rule for both first and second stages i is in terms of the terminal reward only, and we have omitted subsequent-stage action values from δ . This corresponds to the restriction $\lambda=1$ in the model of (Daw, et al., 2011). We did so because in the second-stage, no subsequent stages are visited, and in the first-stage, we observed (analyses not reported) that choices did not depend on model-free, TD(0), first-stage updates from second-stage action values. Specifically, beta weights (described below) relating choices to model-free updates of first- to second-stage transitions were not reliably different from zero across participants. For this reason, no separate eligibility trace was needed in the calculation of model-free state-action values because there was only a single stage-skipping update for the first-stage.

Model-based component—Model-based action values (Q_{MB}) for first-stage actions are defined prospectively, considering the maximally valued outcomes that one could obtain, given state transitions.

$$Q_{MB}(s_1, a_j) = P(s_{2a}|s_1, a_j) \max_{a' \in a_1, a_2} Q_{MF}(s_{2a}, a') + P(s_{2b}|s_1, a_j) \max_{a' \in a_1, a_2} Q_{MF}(s_{2b}, a')$$

This equation gives the model-based value of action a_j in state s_1 , based on the probability that each stage 1 action would lead to a given second-stage state and the maximally valued actions a' in second-stage states s_{2a} and s_{2b} . We variously modeled transition structure learning.

$$P(s_{2k}|s_1, a_j)$$

by assuming that either 1) participants knew the correct transition structure from the beginning, 2) that they guessed which transition was rare and which was common based upon experienced transition frequencies (Daw, et al., 2011), or 3) that they calculated transition probabilities from a running tally of transitions from each action in each state. Results were the same regardless of the model used for transition structure learning. We present results specifically from the model in which we assume that participants calculated transition probabilities from a running tally of transitions from each action in each state. For example the following equation yields the probability of observing second stage state (s_{2k}) given action (a_j) in the first stage

$$P(s_{2k}|s_1, a_j) = n(s_{2k}|s_1, a_j) / n(s_1, a_j)$$

where n is the number of observed states.

Combining model-based and model-free values to predict choices—Model-based and model-free value estimates were combined to predict first-stage choices using the softmax rule.

$$P(a_{(1,t)}=a|s_{(1,t)}) = \frac{\exp[\beta^{MF} Q_{s_{(1,t)},a}^{MF} + \beta^{MB} Q_{s_{(1,t)},a}^{MB} + \rho \cdot rep(a)]}{\sum_{a'} \exp[\beta^{MF} Q_{s_{(1,t)},a'}^{MF} + \beta^{MB} Q_{s_{(1,t)},a'}^{MB} + \rho \cdot rep(a')]}$$

For first-stage choices, we included free parameters β^{MF} and β^{MB} describing the extent to which model-free and model-based value estimates predict choice behavior. Note that this formulation is algebraically equivalent to that used by Daw et al., 2011 with $\beta^{MB} = w\beta$ and $\beta^{MF} = ((1 - w)\beta)$. Following Daw et al. (2011), we also included a stickiness parameter ρ to account for individual differences in perseveration in first-stage choices from the previous trial; rep is an indicator function set to 1 in case the individual repeats their first-stage choice from the previous trial and 0 otherwise.

Second-stage model-based and model-free values are identical ($Q^{MB} = Q^{MF}$) and the choice is thus weighted by a single parameter β^{stage2} .

$$P(a_{(2,t)}=a|s_{(2,t)}) = \frac{\exp(\beta^{stage2} Q_{s_{(2,t)},a}^{MF})}{\sum_{a'} \exp(\beta^{stage2} Q_{s_{(2,t)},a'}^{MF})}$$

In total the model had 5 free parameters ($\beta^{MF}, \beta^{MB}, \beta^{stage2}, \alpha, \rho$). Parameter values were estimated for each participant individually by likelihood maximization using the MATLAB function `fmincon`, and then subjected to group and individual difference analyses. Given the lack of normality in parameter distributions, non-parametric Wilcoxon ranked-sum tests were used for group differences. We also tested for individual difference correlations between symptom, neurocognitive, medication metrics, and model parameter estimates.

Reaction Time—Finally, we analyzed median second-stage reaction times for additional evidence of group differences in model-based decision-making. We conducted a repeated-measures ANOVA with two factors: diagnostic group and transition type. In order to observe individual difference relationships of second-stage reaction time differences between transition types, partial correlations were conducted between an RT difference score (rare–common) and negative symptoms, neurocognitive measures, and the Snaith-Hamilton pleasure scale holding diagnostic group status constant.

Results

Demographics

Groups did not significantly differ in age, gender, ethnicity, or parental education (Table 1). However, the personal education of the SZ group was lower than the controls. The SZ group self-reported increased levels of anhedonia compared to the controls.

Task Behavior

To assess relative contributions of model-based and model-free learning, we analyzed first-stage choice behavior that, critically, should vary as a function of reward and transition history, depending on model-based or model-free learning. For example, consider a trial where a first-stage choice results in a rare transition leading to an unlikely second-stage state where a reward is obtained (rare-rewarded condition, Figure 2). The pure model-free learner would stay with the same first stage choice, following a reward, ignoring the transition structure. However, the model-based learner, taking into consideration both reward and task structure, would show a reward \times rarity interaction: a decreased likelihood of repeating the same first-stage choice because shifting would increase the probability of encountering the previously rewarded second-stage state. Figure 2 illustrates frequency of staying with previous first-stage choices for each group.

1. Hierarchical Linear Model—We fit a HLM predicting the current first-stage choice as a function of the previous trial's reward, the previous transition rarity, and diagnostic status (Table 2). We found a significant effect of reward, such that the probability of staying with the same first-stage choice increased when the previous choice was rewarded, demonstrating model-free learning. There was also a significant reward \times rarity interaction – a marker of model-based learning – such that participants were less likely to repeat the prior first-stage choice that led to a reward if it was preceded by a rare versus common transition. Furthermore, there was a significant reward \times rarity \times group interaction, supporting that the SZ group demonstrated less model-based learning compared to the CN group. Finally, the group \times reward interaction was not significant, suggesting that groups did not differ on estimates of model-free learning.

We further examined individual difference interactions with model-based and model-free decision-making estimates (Table 3). Contrary to our hypotheses, clinician-rated negative symptoms (BNSS total and BNSS Avolition/Anhedonia Subscale) did not interact with estimates of model-based or model-free decision-making. However, separately, we observed a significant reward \times rarity \times IQ interaction, controlling for group, suggesting that, independent of diagnostic status, individuals with greater intellectual capacity utilized model-based control more than those with reduced premorbid functioning. This relationship was also observed in the CN and SZ groups separately. In another model, we observed a significant Snaithe \times reward interaction suggesting a positive relationship between hedonic capacity and model-free decision-making. However, when fitting the model separately for each group, this relationship remained significant only in the patient group. Olanzapine equivalent dose and age did not interact with model-free and model-based effects.

Finally, given significant effects of IQ and diagnostic group on model-based learning estimates we fit a HLM, which included both factors into the same model to determine if IQ accounted for the diagnostic group effect (supplement). The results indicated a significant effect of IQ (IQ \times rarity \times reward interaction), and a trend level effect of group (group \times rarity \times reward) on model-based learning estimates. These results suggest that while the IQ differences associated with SZ account for part of the effect of diagnostic group on model-based learning, there is still some effect of psychosis above and beyond differences in premorbid intellectual functioning, which has been found to be reduced in association with the development of psychosis (Agnew-Blais et al., 2015; David, Malmberg, Brandt, Allebeck, & Lewis, 1997).

2. Computational Modeling of Learning Processes—The preceding analysis is based on a simplified marker of model-based and model-free learning in terms of experience on the previous trial. To verify that our results were not dependent on this simplification, we also fit a full RL model to choice and reward data to estimate the influence of model-based and model-free learning on decision-making for each subject (Daw, et al., 2011; Doll, et al., 2015). Our model had 5 free parameters: β^{MB} (model-based weighting factor), β^{MF} (model-free weighting factor), β^{stage2} (second-stage inverse temperature), ρ (preservation of the first-stage choice), α (learning rate). Table 4 provides the group level estimates for each parameter. Converging with previous HLM analyses, the model-based weighting factor (β^{MB}) was significantly blunted in the SZ group suggesting decreased reliance on model-based learning in the patient group. However, model-free parameter estimates did not differ between groups. No further group differences were reliable. Note that Table 4 excludes two outlier participants with exceedingly large β^{MF} estimates from the SZ group, but none of our conclusions change when all participants are included.

Correlations of individual difference metrics with parameter estimates from the computational model mirrored the HLM analyses. β^{MB} positively correlated with IQ ($r=0.42, p<0.01$). In order to more fully explore the relationship between β^{MB} and premorbid intellectual functioning we correlated performance on an n-back task, d' , with β^{MB} in a subset of participants and found these variables to be trend-level related ($r=0.30; p=0.06$). Consistent with previous analyses the effect of diagnostic group was estimated with the same sign but did not remain significant after accounting for premorbid intellectual functioning ($p=0.1$). The Snaith-Hamilton pleasure scale positively correlated with β^{MF} ($r=0.375, p<0.01$) (supplemental materials). Correlations between negative symptoms (BNSS total and BNSS Avolition/Anhedonia Subscale), olanzapine equivalents, age, and parameters estimates were not significant. Finally, β^{MB} positively correlated with task accuracy ($r=0.27; p=0.03$) suggesting that those who performed better on the task more readily utilized model-based strategies. This relationship was not observed between β^{MF} and task performance ($r=-0.05; p=0.68$).

3. Reaction Time—We analyzed second-stage median reaction times (RT) as another indicator of learning of the transition model, a necessary component for model-based choice (Figure 3). Importantly, model-based decisions are informed by transition expectations, and to the extent subjects anticipate events according to a learned transition model, second-stage

RTs following rare transitions may be slower than RTs following common transitions (reflecting expectancy violations). Model-free decisions, by contrast, do not utilize transition expectations. Thus, differences in the second-stage RT between transition types are consistent with model-based learning during decision-making (Deserno, et al., 2015). In our data, the model-based weighting factor, β^{MB} , correlated with the subject-level RT differences for rare compared to common transitions ($r=0.615$; $p<0.001$). A repeated-measures ANOVA revealed a significant main effect of transition on RT, where choices following a rare transition were slower than those following common transitions ($F=38.60, p<0.001$). Contrary to our hypotheses, the interaction between transition type and diagnostic group was not significant ($F=1.64, p=0.2$). However, when analyzing the simple effect of RT within each group separately, we found that the size of the transition effect for the CN group ($\eta^2_{\text{partial}}=0.272$) was over twice the size of the SZ group ($\eta^2_{\text{partial}}=0.102$) consistent with lower model-based decision-making in SZ. Finally, a significant partial correlation was found between RT slowing following rare versus common transitions and IQ, holding diagnostic group constant ($r = 0.30; p=0.02$) providing converging evidence for the aforementioned relationship between IQ and model-based control. We also examined the relationship between RT slowing and N-Back performance as another maker of intellectual functioning and found these variables to be significantly related ($r=0.38$; $p<0.02$). Finally, no significant relationship was found between RT slowing and self-reported or clinician-rated negative symptoms (BNSS total and BNSS Avolition/Anhedonia Subscale).

Discussion

The goal of the current experiment was to assess reliance on model-based and model-free learning in SZ. Consistent with our hypotheses, SZ patients demonstrated reduced model-based learning compared to controls, suggesting either diminished motivation or diminished capacity for utilizing internal models of the environment during decision-making. In contrast, model-free estimates did not differ between groups. Supplementary analyses of second-stage reaction times provided evidence that subjects in both groups exhibited knowledge of the transition model. In general, participants took longer to respond for second-stage choices preceded by rare compared to common transitions. That these expectancy effects are detectable even in SZ suggests that these subjects understood the task instructions and model structure, and that their deficits in model-based choice behavior may relate to the use of this information to guide action. However, this effect (though not significantly different) was almost twice as large in the controls, suggesting a more robust effect of model-based learning for controls. Model-based learning correlated positively with measures of intellectual functioning (IQ and N-Back) in both SZ patients and controls. Furthermore, even though the groups did not differ significantly on IQ, supplementary analyses showed that IQ absorbed some, but not all, of the variance associated with diagnostic group. Specifically, group differences in model-based learning were still trend-level when accounting for IQ, suggesting that aspects of psychosis may influence model-based learning estimates above and beyond differences in IQ. These results were similar when controlling for n-back performance. Finally, the hypothesis that model-based and model-free learning estimates would be related to negative symptom severity was not

robustly supported, although some evidence did suggest that reduced model-free learning was associated with greater levels of self-reported anhedonia.

The current result of reduced model-based and intact model-free learning in SZ is consistent with several recent reports that have examined the role of higher-order cognition in reward learning. For example, Collins et al., showed working-memory impairments (a strong correlate of model-based learning) entirely accounted for impairments on a RL task in SZ (Collins, et al., 2014). Similarly, Strauss et al. found that SZ choice-behavior was best-fit by a standard actor-critic model, putatively driven by ventral striatal prediction-error signaling, a conceptually similar mechanism to model-free learning (Strauss et al., 2015). In a recent report by our group (Culbreth, Gold, Cools, & Barch, 2015), we found that choice behavior of SZ patients on a reward-learning task was related to hypoactivation of a fronto-parietal network of brain regions (strongly tied to model-based learning), whereas striatal regions did not show relationships to task behavior. Such prior reports are consistent with our findings of intact model-free but reduced model-based decision-making in SZ, and also consistent with the relationship between model-based behavior and N-Back performance seen in the current study. Further, our results are also consistent with previous reports, which have examined the external correlates of model-based and model-free learning in healthy subjects. For example, as shown for cognitive control, working memory, and IQ (Gillan, et al., 2016; Otto, et al., 2013; Otto, et al., 2014), we also found higher-order cognitive metrics correlating with model-based decision-making. Taken together, such results suggest a critical link between higher-order cognitive deficits and value-based decision-making in SZ.

While the current results demonstrate a clear relationship between reward learning and higher-order cognition in SZ, we failed to see robust correlations between model-based learning and negative symptoms, though we did see a relationship between individual differences in model-free learning and self-reported anhedonia. It is not entirely clear to us why we did not also see a relationship between model-based learning and negative symptoms. However, these results are conceptually consistent with a recent report, which used an identical task in a large non-clinical sample and failed to find correlations between model-based learning and the negative symptom traits of schizotypy (Gillan, et al., 2016). One possible explanation is that the stimulus-response relationships in the current task design rely on working memory and cognitive control in order to learn and leverage the transition matrix during decision-making. Thus, it is possible that these cognitive demands limited our ability to observe a more direct relationship to individual differences in negative symptoms. Consistent with this interpretation, many previous studies that have demonstrated a significant relationship between negative symptoms and reward learning have implemented somewhat simpler task designs where stimulus-response learning might require less involvement of higher-order cognitive processes compared to the current design (Gold et al., 2012; Shurman, Horan, & Nuechterlein, 2005; Waltz, Frank, Wiecki, & Gold, 2011; Waltz & Gold, 2007). Further, several recent reports examining reward-learning in SZ which also used paradigms that might engage cognitive control relatively more strongly have also failed to show correlations between task performance and negative symptoms. This is consistent with the hypothesis that the role of higher order cognitive processes in decision-making deficits may be somewhat independent of negative symptoms (Collins, et al., 2014; Culbreth, et al., 2015; Strauss, et al., 2015).

While no significant correlations between model-based learning and negative symptoms were observed in the current report, we did observe a positive correlation between hedonic capacity and model-free learning estimates, suggesting that those SZ patients with higher levels of anhedonia demonstrate reduced model-free learning. This finding is consistent with previous literature showing that reduced ventral striatal prediction error signaling (a mechanism proposed to underlie model-free learning) is robustly correlated with negative symptom severity (Juckel, Schlagenhauf, Koslowski, Filonov, et al., 2006; Juckel, Schlagenhauf, Koslowski, Wüstenberg, et al., 2006).

Though our results support our hypothesis of a specific deficit in model-based decision-making in SZ, it is worth considering why premorbid IQ and working memory capacity may explain some between-group variance. First, we included such measures into our study design to better understand covariates of model-based learning and provide converging evidence with previous reports, not to suggest that such factors should be controlled for to establish an effect of diagnostic group. Controlling for such variables to establish an independent effect of diagnostic group is ill informed as this would remove schizophrenia-related variability from the data (Meehl, 1971). There are multiple longitudinal studies suggesting that premorbid IQ is associated with conversion to psychosis (Agnew-Blais, et al., 2015; David, et al., 1997). Thus, the interpretation that IQ explains the group differences we observed in model-based learning neglects previous data suggesting that IQ is a causally related to the pathogenesis of SZ. We instead view the findings regarding individual difference relationships between working memory, IQ, and model-based learning estimates to be an attempt to better understand factors that might be contributing to such a group difference. In the case of these data, it appears that the reductions in model-based learning seen in SZ may be due, in part, to reduced functioning of higher-order cognitive processes, such as working memory. One intriguing possibility may be that core deficits in model-based learning contribute to decreased performance on IQ-type assessments; however, further research is needed to explore such hypotheses.

Future Directions

Our study provides precise, computational evidence for diminished model-based learning in SZ. Moreover, our results suggest directly testable hypotheses about dysfunction, in SZ, in neural systems thought to support model-based RL. Previous literature suggests three likely biological targets for such a deficit: abnormalities in presynaptic dopamine in the ventral striatum (Deserno, et al., 2015), lateral prefrontal cortex functioning (Deserno, et al., 2015; Gläscher, et al., 2010), or cortico-striatal connectivity (Deserno, et al., 2015). Studies addressing these biological functions will allow researchers to specify the neural structure of the RL deficit in SZ, and may yield biologically-informed targets for novel interventions.

While the current study provides robust evidence for diminished model-based decision-making in SZ, more work is need to understand why SZ patients show reductions in model-based learning. Importantly, model-based behavior relies on a number of underlying cognitive and affective processes, and further work will be needed to disentangle such processes in order to understand why SZ is associated with reductions in model-based learning. Several possibilities exist, such that SZ patients may: 1) have difficulties generating

complex internal models of task environments 2) generate inaccurate models of task environments 3) generate models more slowly 4) be able to generate internal models but fail to exert the cognitive effort required to produce such models. Thus, follow up studies will be needed to more accurately characterize the nature of reduced model-based decision-making seen in the current report in order to understand differential patterns in healthy and psychiatric populations.

Finally, it remains to be seen how deficits in model-based and model-free learning may manifest across psychiatric disorders (e.g., depression, anxiety disorders). In a recent article, Gillan et al., assessed model-based learning in a large general population sample and collected a number of self-reported symptom measures (including depression, schizotypy, anxiety, etc.) They found that model-based learning did not vary as a function of depressive symptoms or the negative symptoms in schizotypy. In contrast, they found that model-based learning was negatively related to symptoms of compulsion, intrusive thought patterns, and substance abuse, suggesting that disorders characterized by obsessive behavior and intrusive thinking might be the most likely to exhibit deficits in model-based learning. While the Gillan et al study provided a large sample and assessed symptom severity in multiple domains, it was a non-clinical sample and it is possible that that behavior in individuals meeting full criteria for psychiatric disorders might differ. Thus, future research is needed to more thoroughly evaluate model-based learning across diagnostic boundaries in clinical populations.

Limitations

The current study has several limitations. First, we did not explicitly assess whether individuals with schizophrenia understood the task instructions and the transition structure. We believe that the significant second-stage RT differences between rare and common transitions coupled with the trend-level model-based learning effect in a HLM including only the patient group suggests that schizophrenia patients understood the transition function and task instructions (see supplemental materials), but were more specifically impaired at leveraging this information to guide decision behavior. Furthermore, extensive instructions were given pertaining to the nature of the task and the importance of the mapping between first and second stage states (see supplemental materials). Second, we did not collect measures of positive and disorganized symptoms. This was done for two reasons: we were trying to make the session relative short for participants, and our participants were stable outpatients with relatively low levels of positive symptoms and relatively little variation in positive symptoms, reducing the likelihood that they would show meaningful relationships to task variables. Third, while we found a robust relationship between IQ and model-based learning, it should be noted that our IQ measure, WTAR, is concise measure and thus cannot index particular cognitive domain that may be driving this relationship. However, our n-back analyses suggest that working memory may be one cognitive domain particularly important in model-based learning. Finally, the majority of our SZ sample was currently taking anti-psychotic medications, which may have altered reward-related responses. However, no correlations between olanzapine equivalents and task parameter estimates were significant.

Summary

The goal of the current experiment was to assess reliance on model-based and model-free learning in SZ. We found evidence for reduced model-based behavior in SZ patients, suggesting either that individuals with SZ have diminished capacity or diminished motivation to utilize internal models of the environment for goal-directed behavior. These findings are consistent with reports showing that SZ patients have deficits in processing domains that support model-based learning such as cognitive control and working memory capacity. Our results were also specific in that there was no group difference in model-free estimates suggesting that basic, habitual RL is unaltered in SZ. Importantly, our findings motivate investigations into specific cortico-striatal systems that likely mediate diminished model-based behavior in SZ. Such investigations will deepen our understanding of etiology, and facilitate the discovery of targeted, novel interventions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors would like to thank the participants in this study who gave generously of their time. We also thank Catherine Hartley for providing the modified task stimuli utilized in this study. This work was supported by National Institute of Mental Health R01 MH066031.

Financial Disclosures

DMB has served as a consultant for Roche, Amgen, and Pfizer. AJC, AW, NDD, and MMB have no interests to disclose.

References

- Agnew-Blais JC, Buka SL, Fitzmaurice GM, Smoller JW, Goldstein JM, Seidman LJ. Early Childhood IQ Trajectories in Individuals Later Developing Schizophrenia and Affective Psychoses in the New England Family Studies. *Schizophrenia bulletin*. 2015:svb027.
- Balleine BW, Daw ND, O'Doherty JP. Multiple forms of value learning and the function of dopamine. *Neuroeconomics: decision making and the brain*. 2008:367–385.
- Barch DM, Ceaser A. Cognition in schizophrenia: core psychological and neural mechanisms. *Trends in cognitive sciences*. 2012; 16(1):27–34. [PubMed: 22169777]
- Barch DM, Csernansky JG, Conturo T, Snyder AZ. Working and long-term memory deficits in schizophrenia: is there a common prefrontal mechanism? *Journal of abnormal psychology*. 2002; 111(3):478. [PubMed: 12150424]
- Bates, D.; Sarkar, D. lme4: Linear mixed-effects models using Eigen and S4 classes. 'R' package. Version 0.9975–12. 2007. URL <http://CRAN.R-project.org>
- Bleuler, E. *Dementia praecox or the group of schizophrenias*. 1950.
- Collins AGE, Brown JK, Gold JM, Waltz JA, Frank MJ. Working Memory Contributions to Reinforcement Learning Impairments in Schizophrenia. *The Journal of Neuroscience*. 2014; 34(41): 13747–13756. DOI: 10.1523/jneurosci.0989-14.2014 [PubMed: 25297101]
- Culbreth AJ, Gold JM, Cools R, Barch DM. Impaired Activation in Cognitive Control Regions Predicts Reversal Learning in Schizophrenia. *Schizophrenia bulletin*. 2015:svb075.
- David AS, Malmberg A, Brandt L, Allebeck P, Lewis G. IQ and risk for schizophrenia: a population-based cohort study. *Psychological medicine*. 1997; 27(06):1311–1323. [PubMed: 9403903]
- Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ. Model-based influences on humans' choices and striatal prediction errors. *Neuron*. 2011; 69(6):1204–1215. [PubMed: 21435563]

- Daw ND, Niv Y, Dayan P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature neuroscience*. 2005; 8(12):1704–1711. [PubMed: 16286932]
- Deserno L, Boehme R, Heinz A, Schlagenhauf F. Reinforcement learning and dopamine in schizophrenia: dimensions of symptoms or specific features of a disease group? *Frontiers in psychiatry*. 2013; 4
- Deserno L, Huys QJ, Boehme R, Buchert R, Heinze H-J, Grace AA, ... Schlagenhauf F. Ventral striatal dopamine reflects behavioral and neural signatures of model-based control during sequential decision making. *Proceedings of the National Academy of Sciences*. 2015; 201417219
- Dickinson A, Balleine B. The role of learning in the operation of motivational systems. *Stevens' handbook of experimental psychology*. 2002
- Doll BB, Shohamy D, Daw ND. Multiple memory systems as substrates for multiple decision systems. *Neurobiology of learning and memory*. 2015; 117:4–13. [PubMed: 24846190]
- First, MB.; Spitzer, RL.; Gibbon, M.; Williams, JB. Structured Clinical Interview for DSM-IV-TR Axis I Disorders—Patient Edition. New York: Biometrics Research Department, New York State Psychiatric Institute; 2001. SCID-I/P. 2/2001 Revision
- Gardner DM, Murphy AL, O'Donnell H, Centorrino F, Baldessarini RJ. International consensus study of antipsychotic dosing. *Psychopharmacology*. 2014; 12(2):235–243.
- Gillan CM, Kosinski M, Whelan R, Phelps EA, Daw ND. Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *eLife*. 2016; 5:e11305. [PubMed: 26928075]
- Gläscher J, Daw N, Dayan P, O'Doherty JP. States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*. 2010; 66(4):585–595. [PubMed: 20510862]
- Gold JM, Waltz JA, Matveeva TM, Kasanova Z, Strauss GP, Herbener ES, ... Frank MJ. Negative symptoms and the failure to represent the expected reward value of actions: behavioral and computational modeling evidence. *Archives of general psychiatry*. 2012; 69(2):129–138. [PubMed: 22310503]
- Gold JM, Waltz JA, Prentice KJ, Morris SE, Heerey EA. Reward processing in schizophrenia: a deficit in the representation of value. *Schizophrenia bulletin*. 2008; 34(5):835–847. [PubMed: 18591195]
- Gur RE, Cowell PE, Latshaw A, Turetsky BI, Grossman RI, Arnold SE, ... Gur RC. Reduced dorsal and orbital gray matter volumes in schizophrenia. *Archives of general psychiatry*. 2000; 57(8):761–768. [PubMed: 10920464]
- Hartmann MN, Kluge A, Kalis A, Mojzisch A, Tobler PN, Kaiser S. Apathy in schizophrenia as a deficit in the generation of options for action. *Journal of abnormal psychology*. 2015; 124(2):309. [PubMed: 25688424]
- Hautus MJ. Corrections for extreme proportions and their biasing effects on estimated values of d' . *Behavior Research Methods, Instruments, & Computers*. 1995; 27(1):46–51.
- Heerey EA, Bell-Warren KR, Gold JM. Decision-making impairments in the context of intact reward sensitivity in schizophrenia. *Biological psychiatry*. 2008; 64(1):62–69. [PubMed: 18377874]
- Hooker CI, Bruce L, Lincoln SH, Fisher M, Vinogradov S. Theory of mind skills are related to gray matter volume in the ventromedial prefrontal cortex in schizophrenia. *Biological psychiatry*. 2011; 70(12):1169–1178. [PubMed: 21917239]
- Juckel G, Schlagenhauf F, Koslowski M, Filonov D, Wüstenberg T, Villringer A, ... Wrase J. Dysfunction of ventral striatal reward prediction in schizophrenic patients treated with typical, not atypical, neuroleptics. *Psychopharmacology*. 2006; 187(2):222–228. [PubMed: 16721614]
- Juckel G, Schlagenhauf F, Koslowski M, Wüstenberg T, Villringer A, Knutson B, ... Heinz A. Dysfunction of ventral striatal reward prediction in schizophrenia. *Neuroimage*. 2006; 29(2):409–416. [PubMed: 16139525]
- Kéri S, Kelemen O, Szekeres G, Bagoczky N, Erdelyi R, Antal A, ... Janka Z. Schizophrenics know more than they can tell: probabilistic classification learning in schizophrenia. *Psychological medicine*. 2000; 30(01):149–155. [PubMed: 10722185]
- Kirkpatrick B, Strauss GP, Nguyen L, Fischer BA, Daniel DG, Cienfuegos A, Marder SR. The brief negative symptom scale: psychometric properties. *Schizophrenia bulletin*. 2011; 37(2):300–305. [PubMed: 20558531]

- Lee SW, Shimojo S, O'Doherty JP. Neural Computations Underlying Arbitration between Model-Based and Model-free Learning. *Neuron*. 2014; 81(3):687–699. [PubMed: 24507199]
- McDannald MA, Lucantonio F, Burke KA, Niv Y, Schoenbaum G. Ventral striatum and orbitofrontal cortex are both required for model-based, but not model-free, reinforcement learning. *The Journal of Neuroscience*. 2011; 31(7):2700–2705. [PubMed: 21325538]
- Meehl, PE. High school yearbooks: a reply to Schwarz. 1971.
- Minzenberg MJ, Laird AR, Thelen S, Carter CS, Glahn DC. Meta-analysis of 41 functional neuroimaging studies of executive function in schizophrenia. *Archives of general psychiatry*. 2009; 66(8):811–822. [PubMed: 19652121]
- Montague PR, Dolan RJ, Friston KJ, Dayan P. Computational psychiatry. *Trends in cognitive sciences*. 2012; 16(1):72–80. [PubMed: 22177032]
- Moran P, Al-Uzri M, Watson J, Reveley M. Reduced Kamin blocking in non paranoid schizophrenia: associations with schizotypy. *Journal of Psychiatric Research*. 2003; 37(2):155–163. [PubMed: 12842169]
- Morris RW, Quail S, Griffiths KR, Green MJ, Balleine BW. Corticostriatal control of goal-directed action is impaired in schizophrenia. *Biological psychiatry*. 2015; 77(2):187–195. [PubMed: 25062683]
- Murray G, Corlett P, Clark L, Pessiglione M, Blackwell A, Honey G, ... Fletcher P. Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. *Molecular psychiatry*. 2007; 13(3):267–276.
- Otto AR, Raio CM, Chiang A, Phelps EA, Daw ND. Working-memory capacity protects model-based learning from stress. *Proceedings of the National Academy of Sciences*. 2013; 110(52):20941–20946.
- Otto, AR.; Skatova, A.; Madlon-Kay, S.; Daw, ND. Cognitive Control Predicts Use of Model-based Reinforcement Learning. 2014.
- Pantelis C, Velakoulis D, McGorry PD, Wood SJ, Suckling J, Phillips LJ, ... Soulsby B. Neuroanatomical abnormalities before and after onset of psychosis: a cross-sectional and longitudinal MRI comparison. *The Lancet*. 2003; 361(9354):281–288.
- Park IH, Park HJ, Chun JW, Kim EY, Kim JJ. Dysfunctional modulation of emotional interference in the medial prefrontal cortex in patients with schizophrenia. *Neuroscience letters*. 2008; 440(2):119–124. [PubMed: 18562102]
- Plailly J, d'Amato T, Saoud M, Royet JP. Left temporo-limbic and orbital dysfunction in schizophrenia during odor familiarity and hedonicity judgments. *Neuroimage*. 2006; 29(1):302–313. [PubMed: 16099179]
- Schlagenhauf F, Huys QJ, Deserno L, Rapp MA, Beck A, Heinze HJ, ... Heinz A. Striatal dysfunction during reversal learning in unmedicated schizophrenia patients. *Neuroimage*. 2014; 89:171–180. [PubMed: 24291614]
- Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. *Science*. 1997; 275(5306):1593–1599. [PubMed: 9054347]
- Semkovska M, Bédard MA, Godbout L, Limoge F, Stip E. Assessment of executive dysfunction during activities of daily living in schizophrenia. *Schizophrenia research*. 2004; 69(2):289–300. [PubMed: 15469200]
- Shurman B, Horan WP, Nuechterlein KH. Schizophrenia patients demonstrate a distinctive pattern of decision-making impairment on the Iowa Gambling Task. *Schizophrenia research*. 2005; 72(2):215–224. [PubMed: 15560966]
- Smittenaar P, FitzGerald TH, Romei V, Wright ND, Dolan RJ. Disruption of dorsolateral prefrontal cortex decreases model-based in favor of model-free control in humans. *Neuron*. 2013; 80(4):914–919. [PubMed: 24206669]
- Snaith R, Hamilton M, Morley S, Humayan A, Hargreaves D, Trigwell P. A scale for the assessment of hedonic tone the Snaith-Hamilton Pleasure Scale. *The British Journal of Psychiatry*. 1995; 167(1):99–103. [PubMed: 7551619]
- Strauss GP, Thaler NS, Matveeva TM, Vogel SJ, Sutton GP, Lee BG, Allen DN. Predicting Psychosis Across Diagnostic Boundaries: Behavioral and Computational Modeling Evidence for Impaired Reinforcement Learning in Schizophrenia and Bipolar Disorder With a History of Psychosis. 2015

- Strauss GP, Waltz JA, Gold JM. A review of reward processing and motivational impairment in schizophrenia. *Schizophrenia bulletin*. 2014; 40(Suppl 2):S107–S116. [PubMed: 24375459]
- Tolman EC. Cognitive maps in rats and men. *Psychological review*. 1948; 55(4):189. [PubMed: 18870876]
- Voon V, Derbyshire K, Rück C, Irvine M, Worbe Y, Enander J, ... Sahakian B. Disorders of compulsivity: a common bias towards learning habits. *Molecular psychiatry*. 2014
- Waltz JA, Frank MJ, Wiecki TV, Gold JM. Altered probabilistic learning and response biases in schizophrenia: behavioral evidence and neurocomputational modeling. *Neuropsychology*. 2011; 25(1):86. [PubMed: 21090899]
- Waltz JA, Gold JM. Probabilistic reversal learning impairments in schizophrenia: further evidence of orbitofrontal dysfunction. *Schizophrenia research*. 2007; 93(1):296–303. [PubMed: 17482797]
- Wechsler, D. Wechsler Test of Adult Reading: WTAR. Psychological Corporation; 2001.
- Weickert TW, Terrazas A, Bigelow LB, Malley JD, Hyde T, Egan MF, ... Goldberg TE. Habit and skill learning in schizophrenia: evidence of normal striatal processing with abnormal cortical input. *Learning & Memory*. 2002; 9(6):430–442. [PubMed: 12464703]
- Woodward TS, Moritz S, Cuttler C, Whitman JC. The contribution of a cognitive bias against disconfirmatory evidence (BADE) to delusions in schizophrenia. *Journal of Clinical and Experimental Neuropsychology*. 2006; 28(4):605–617. [PubMed: 16624787]
- Wunderlich K, Smittenaar P, Dolan RJ. Dopamine enhances model-based over model-free choice behavior. *Neuron*. 2012; 75(3):418–424. [PubMed: 22884326]

General Scientific Summary

This study uses a recent reinforcement-learning framework to specify deficits in reward learning in schizophrenia. We show that while individuals with schizophrenia display intact slow, habitual learning processes, they have a reduced ability to make decisions using future consequences of actions.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

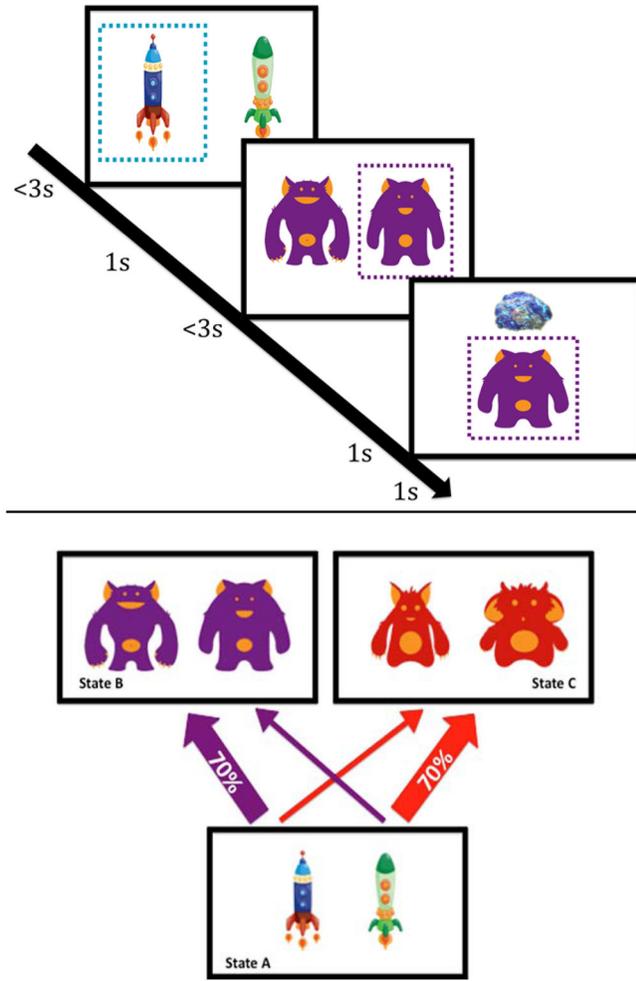


Figure 1. **A.** Sample trial diagram: binary choice at stage one (spaceships) leads probabilistically to one of two second stage states (aliens) each with two choices that probabilistically result in reward/no-reward (treasure). **B.** Task structure: Each stage one choice has a common or rare transition to each of the second stage states.

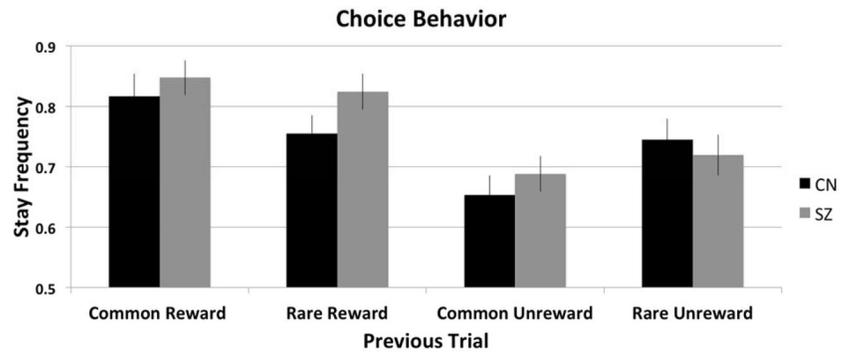


Figure 2. First-stage choice behavior (coded as stay/shift) averaged across individuals within each group.
Note: Error Bars are presented as standard errors.

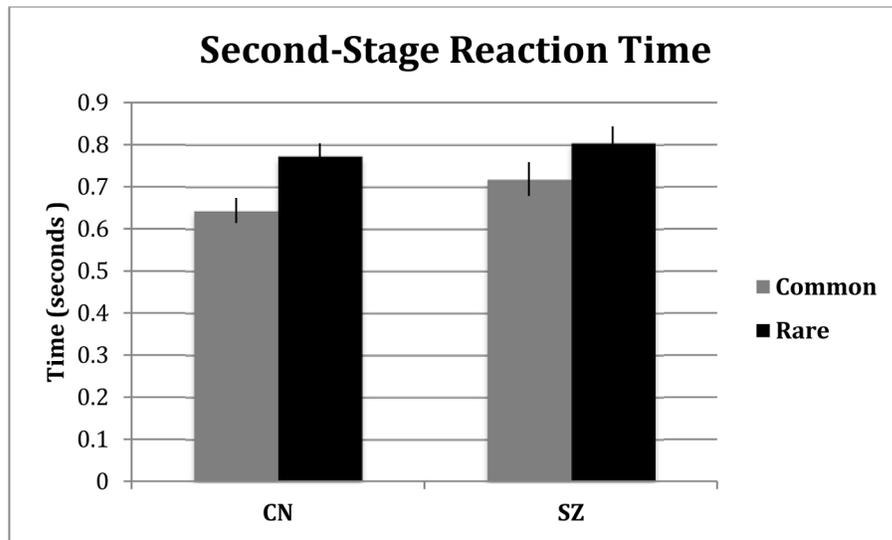


Figure 3.
Second-stage reaction time by transition type for each group
Note: Error Bars are presented as standard errors.

Participant Characteristics

Table 1

Characteristics	Healthy Controls (N=30)		Individuals with Schizophrenia (N=33)		p-value
Demographics	Mean	SD	Mean	SD	
Age (years)	35.9	8.2	36.7	9.25	0.72
Sex (% male)	46.7%		51.5%		0.70
Ethnicity					
<i>African-American</i>	58.0%		56.7%		
<i>Caucasian</i>	42.0%		43.3%		
Personal Education (years)	15.2	2.7	12.8	2.3	<0.001
Parental Education (years)	14.4	2.6	14.4	3.50	0.97
Medication status					
Atypical antipsychotics (%)	NA		55%		
Typical antipsychotics (%)	NA		3%		
Typical and atypical (%)	NA		6%		
Medicated (no antipsychotics)			15%		
Not Medicated (%)	NA		21%		
Olanzapine Equivalent Dose	NA		14.7	9.2	
Clinical ratings					
Brief Negative Symptom Scale					
Total Score	NA		24.5	13.2	
Avolition/Anhedonia Subscale	NA		16.0	8.8	
Self-Report					
Snaith Hamilton Pleasure Scale	50.7	5.9	43.5	10.5	0.001
Neurocognitive Measures					
Wechsler Test of Adult Reading	101.2	14.3	98.3	12.0	0.39

Table 2

Coefficients predicting response repetition from the previous trial outcome, the transition type, and diagnostic group.

Coefficient	Estimate (SE)	p-value
Intercept	1.51 (0.13)	<0.001
Reward	0.49 (0.07)	<0.001
Rarity	0.06 (0.04)	0.097
Group	0.10 (0.12)	0.46
Reward × Rarity	0.26 (0.06)	<0.001
Rarity × Group	-0.01 (0.03)	0.681
Reward × Group	0.09 (0.07)	0.177
Reward × Rarity × Group	-0.11 (0.06)	0.049

Note: Error bars are presented as standard errors

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3

External Correlates of model-based and model-free learning

Brief Negative Symptom Scale (BNSS)		
Coefficient	Estimate (SE)	p-value
Intercept	1.63 (0.17)	<0.001
Reward	0.61 (0.11)	<0.001
Rarity	0.02 (0.06)	0.67
BNSS	-0.03 (0.01)	0.03
Reward × Rarity	0.12 (0.07)	0.09
Reward × BNSS	-0.007 (0.008)	0.40
Rarity × BNSS	0.0003 (0.004)	0.085
Reward × Rarity × BNSS	0.006 (0.005)	0.24

IQ		
Coefficient	Estimate (SE)	p-value
Intercept	1.51 (0.13)	<0.001
Reward	0.50 (0.07)	<0.001
Rarity	0.06 (0.04)	0.10
Group	0.12 (0.11)	0.28
IQ	0.02 (0.009)	0.07
Reward × Rarity	0.25 (0.05)	<0.001
Reward × IQ	0.004 (0.005)	0.49
Rarity × IQ	0.004 (0.003)	0.13
Reward × Rarity × IQ	0.014 (0.004)	<0.001

Snaith-Hamilton Pleasure Scale (Hedonic Capacity)		
Coefficient	Estimate (SE)	p-value
Intercept	1.51 (0.13)	<0.001
Reward	0.50 (0.07)	<0.001
Rarity	0.06 (0.04)	0.09
Group	0.12 (0.11)	0.28
SNAITH	0.03 (0.01)	0.09
Reward × Rarity	0.26 (0.06)	<0.001
Reward × SNAITH	0.02 (0.008)	0.03
Rarity × SNAITH	0.003 (0.004)	0.50
Reward × Rarity × SNAITH	0.01 (0.01)	0.11

Note: Coefficients predicting response repetition from the outcome of the previous trial, the transition type, and negative symptoms (BNSS total score), WTAR estimated full scale IQ, and the Snaith-Hamilton Pleasure Scale. Models including Snaith and IQ were tested across both groups. The model including BNSS was included only the SZ group

Table 4

Parameter Estimates from Hybrid Model

Parameter	Schizophrenia	Control	p-value
β^{MF}	0.79 (0.17)	0.66 (0.10)	0.94
β^{MB}	0.57 (0.34)	1.28 (0.40)	0.04
β^{stage2}	1.03 (0.22)	1.40 (0.20)	0.14
α	0.61 (0.06)	0.65 (0.06)	0.96
ρ	1.08 (0.16)	1.01 (0.16)	0.69
NLL	198.17 (7.77)	195.90 (9.73)	0.95

Note: β^{MF} : model-free parameter; β^{MB} : model-based parameter; β^{stage2} : stage-2 inverse temperature; α : learning rate; ρ : perseveration parameter; NLL: negative log likelihood.