Schizophrenia Research 127 (2011) 215–222

Contents lists available at ScienceDirect

Schizophrenia Research

journal homepage: www.elsevier.com/locate/schres

Optimizing vs. Matching: Response Strategy in a Probabilistic Learning Task is associated with Negative Symptoms of Schizophrenia

Zuzana Kasanova [a,*], James A. Waltz [a], Gregory P. Strauss [a], Michael J. Frank [b,c], James M. Gold [a]

[a] Department of Psychiatry, Maryland Psychiatric Research Center, University of Maryland School of Medicine, PO Box 21247, Baltimore, MD 21228
[b] Department of Cognitive, Linguistic & Psychological Sciences, Providence, RI
[c] Department of Psychiatry and Human Behavior, Brown University, Providence, RI

ARTICLE INFO

Article history:
Received 14 July 2010
Received in revised form 3 December 2010
Accepted 6 December 2010
Available online 15 January 2011

Keywords:
Schizophrenia
Reinforcement Learning
Probability Learning Task
Negative Symptoms
Matching
Exploration

ABSTRACT

Previous research indicates that behavioral performance in simple probability learning tasks can be organized into response strategy classifications that are thought to predict important personal characteristics and individual differences. Typically, relatively small proportion of subjects can be identified as optimizers for effectively exploiting the environment and choosing the more rewarding stimulus nearly all of the time. In contrast, the vast majority of subjects behaves sub-optimally and adopts the matching or super-matching strategy, apportioning their responses in a way that matches or slightly exceeds the probabilities of reinforcement. In the present study, we administered a two-choice probability learning paradigm to 51 individuals with schizophrenia (SZ) and 29 healthy controls (NC) to examine whether there are differences in the proportion of subjects falling into these response strategy classifications, and to determine whether task performance is differentially associated with symptom severity and neuropsychological functioning. Although the sample of SZ patients did not differ from NC in overall rate of learning or end performance, significant clinical differences emerged when patients were divided into optimizing, super-matching and matching subgroups based upon task performance. Patients classified as optimizers, who adopted the most advantageous learning strategy, exhibited higher levels of positive and negative symptoms than their matching and super-matching counterparts. Importantly, when both positive and negative symptoms were considered together, only negative symptom severity was a significant predictor of whether a subject would behave optimally, with each one standard deviation increase in negative symptoms increasing the odds of a patient being an optimizer by as much as 80%. These data provide a rare example of a greater clinical impairment being associated with better behavioral performance.

Patients with schizophrenia demonstrate a range of cognitive, motivational and affective deficits that limit their adaptive functioning. In the recent literature there has been a renewed focus on the role of basic reward processing mechanisms that could theoretically be related to both cognitive and motivational impairments. Of particular interest is the finding that abnormal reward processing is associated with greater severity of both positive and negative symptoms. For example, Corlett et al. (2007, 2010)) and Murray et al. (2008) have found that abnormal processing of positive feedback may be related to the severity of positive symptoms, a finding that fits with the predictions that emerge from Kapur's notion that abnormal dopamine release would lead to context-inappropriate attributions of salience (Kapur, 2003; Jensen et al., 2008). In contrast, we and others found that abnormalities in reinforcement learning and decision-making (Waltz & Gold, 2007; Farkas et al., 2008; Polgar et al., 2008; Strauss et al., 2011), and associated neural signals (Waltz et al., 2009, 2010), appear to be linked to negative symptoms.

* Corresponding author.
E-mail address: zkasanova@mprc.umaryland.edu (Z. Kasanova).

0920-9964/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.schres.2010.12.003

One problem of the available behavioral evidence is that most of the experiments have involved somewhat complex tasks, and it is possible that non-reward-related cognitive impairments may have affected performance. To address that limitation, we used a very simple two-choice probability learning task in which one choice was rewarded 70% of the time, and the alternative was reinforced only 30% of the time. One interesting feature of such simple tasks is that they tend to elicit non-optimal decision making, in that people often allocate their response choices to match the probability levels of the more frequently rewarded stimulus. That is, rather than choosing the stimulus that on any given trial has the highest expected value in order to maximize overall payoff, people often allocate approximately 70% of their responses to this stimulus, a phenomenon described first by Herrnstein and termed "matching" (Herrnstein, 1970, 1982, 1997). It is noteworthy that there is evidence that many non-human animals (rats, birds, monkeys) reliably demonstrate this behavior, suggesting that higher order cognitive processes are not an essential factor contributing to the widely observed and arguably universal sub-optimal performance (Sugrue et al., 2004; Houston et al., 2007; Herrnstein, 1990; Sakai & Fukai, 2008; Hinson & Staddon, 1983).

This issue of response strategy has been extensively studied in humans, with evidence indicating that healthy individuals most frequently rely on a matching strategy (Estes & Strughan, 1954; Gardner, 1958; Neimark & Shufod, 1959; Healy & Kubovy, 1981; Jarvik, 1951) and to a lesser degree on a super-matching strategy (Myers & Cruse, 1968; Myers et al., 1963; Edwards, 1956; Bereby-Meyer & Erev, 1997), in which rates of choosing the optimal response overshoot the reinforcement rate of that response. Although these are the most common response strategies employed in two-choice probability learning tasks, behavior ranging from chance (50% allocation to each alternative) to maximization (100% allocation to the rewarding alternative) has also been observed (Vulkan, 2000; Baum, 1974; Friedman & Massaro, 1998; Shanks et al., 2002). To explore factors associated with the formulation of these different response strategies, Shanks Shanks et al. (2002) conducted a series of probabilistic learning experiments in which they manipulated variables such as number of trials, frequency and nature of the feedback and monetary payoff. Their results suggest that about 75% of subjects can achieve maximization if provided with monetary incentive and other meaningful feedback about their performance, and the task has a large enough number of learning trials. The authors hypothesized that the behavior of the remaining 25% of subjects (i.e., those who did not reach a level characteristic of maximization and were thus immune to these task manipulations) could be explained by internal factors and individual differences such as sensitivity to feedback, cognitive functioning, proneness to boredom, risk-aversion, and utility representation. However, factors underlying suboptimal performance on these two-choice probability learning tasks remain unresolved, as they have yet to be systematically examined in an empirical study.

Contributing factors affecting subjects' performance can be expected to be population-specific, and in the case of individuals with schizophrenia consist of any number of the core illness dimensions associated with the disease (e.g., positive symptoms, negative symptoms and cognitive impairment; van Os & Kapur, 2009). Our observations that patients with more severe negative symptoms show impairments in learning from positive feedback (Waltz et al., 2007; Strauss et al., 2011) might lead one to predict that such patients would perform sub-optimally on a two-choice probability learning task. On the other hand, computational modeling evidence from our group (Strauss et al., 2011) suggests an association between greater negative symptom severity and reduction in meaningful exploration of the environment that leads to perseveration. Thus, in the context of the two-choice task environment, one might be tempted to predict that higher levels of negative symptoms would be related to paradoxically superior performance. That is, a reduction in exploration would lead the high negative symptom patients to stick with a winning response, resulting in higher overall earnings. It is difficult, a priori, to adjudicate between these two predictions. Thus, we designed a simple probability learning task to directly test these competing hypotheses.

1. Methods

1.1. Participants

Fifty-one patients meeting DSM-IV criteria (First et al., 2001) for schizophrenia or schizoaffective disorder (SZ), and twenty-nine healthy control (NC) subjects volunteered to participate in this study, which was approved by the University of Maryland School of Medicine Institutional Review Board. All participants provided informed consent and received monetary compensation for their participation in the study.

Individuals with SZ were clinically and medically stable outpatients of the Maryland Psychiatric Research Center, as determined by their psychiatrist, therapist and clinical documentation. All patients were receiving antipsychotic medication and were on a stable regimen for a minimum of four weeks prior to entering the study. Almost all patients were being treated with second-generation antipsychotics (see Table 1 for subject demographic, clinical, and neuropsychological assessment data).

All SZ patients were rated for clinical symptoms based on interviews conducted by trained case-workers, using the following measures: the Brief Psychiatric Rating Scale (BPRS; Overall and Gorman, 1962), the Scale for the Assessment of Negative Symptoms (SANS; Andreasen, 1984), and the Calgary Depression Scale (CDS; Addington et al., 1992). Negative symptom ratings from the SANS were used to divide patients into a high negative (HN) symptom group and a low negative (LN) symptom group. In order to do this, we determined the median SANS total score for the entire patient sample (28). All patients with a SANS total score lower than 28 were assigned to the LN group (N = 25) and all patients with a SANS total score greater than, or equal to, 28 were assigned to the HN group (N = 26; subjects whose SANS total scores fell at the median were added to the smaller group).

Normal control (NC) participants were recruited from the community via random digit dialing and word of mouth (from those recruited by random digit dialing). All NC participants had no current Axis I or II diagnoses, as determined by the Structured Clinical Interview for DSM-IV

Table 1
Demographic information and cognitive assessments for patients (N = 51) and controls (N = 29).

Measure	Control M (SD)	LN Patient M (SD)	HN Patient M (SD)	p-value
Age	44.81 (10.49)	43.79 (8.67)	43.88 (10.64	0.917
Education (years)	15.03 (2.23)	13.08 (2.14)	12.56 (2.45)	<0.001
Paternal Education (years)	12.82 (3.43)	14.46 (3.67)	13.16 (3.72)	0.245
Gender (M: F)	21:8	17:8	19:6	0.853
Race				0.579
African American	10	8	9	
Caucasian	19	12	16	
Other	0	5	0	
Standard Neuropsychology				
WRAT	101.03 (17.25)	94.44 (15.11)	90.52 (12.51)	0.041
WTAR	103.52 (17.15)	96.56 (17.29)	93.36 (17.36)	0.091
WASI	113.52 (13.40)	97.04 (15.45)	95.24 (12.15)	<0.001
MATRICS battery	48.33 (15.24)	31.80 (14.12)	29.32 (13.17)	<0.001
Antipsychotic Medication Regimen				
Haloperidol or Fluphenazine only	-	0	1	
Clozapine only	-	6	9	
Other second-generation only	-	8	11	
Clozapine + another antipsychotic	-	5	3	
First -generation + second-generation antipsychotic	-	0	1	
Clinical Ratings				
BPRS total score	-	34.72 (7.15)	37.56 (9.26)	0.231
SANS total score	-	19.16 (8.72)	35.60 (9.06)	<0.001
Calgary Depression Scale	-	1.84 (1.97)	2.56 (2.71)	0.288

(SCID; First et al., 2001), no family history of psychosis, and were not taking any psychotropic medications. In addition, all study participants denied substance abuse within the past 6 months and had no lifetime history of neurological disorder.

Patients and controls were matched on age [$t(68) = 0.246$], parental education [$t(61) = 1.132$], ethnicity [$\chi^2(4) = 0.855$], and gender [$\chi^2(1) = 0.797$]. All participants completed a standard battery of neuropsychological tests, symptom interviews, and computerized reward learning tasks (including the two-choice probability learning task). Neuropsychological tests included the MATRICS battery (Green et al., 2004), Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999), Wide Range Achievement Test (WRAT; Wilkinson, 1993), and Wechsler Test of Adult Reading (WTAR; Wechsler, 2001). Patients and controls did not differ significantly in WRAT [$t(67) = 1.779$] and WTAR [$t(68) = 1.878$]. However, patients had lower WASI estimated IQs [$t(68) = 5.053$] and MATRICS Overall scores [$t(65) = 4.602$].

1.2. Two-Choice Probability Learning Task

Participants performed a simple probability learning task in which they were presented with a pair of identical items centered vertically on either side of a computer screen. The stimuli were treasure chest boxes presented on a black-colored background. The participants were asked to press a button on a response pad to select one of the two treasure chests. The left button corresponded to the left treasure box and the right button to the right treasure box. The selection of the treasure chest on one side was reinforced on 70% of trials and the selection of the treasure chest on the other side was reinforced on 30% of the trials. The location of the more frequently reinforced treasure box was counterbalanced across all participants and once assigned, it remained constant throughout the task. If the participant selected the winning item, the chosen treasure box was replaced by a nickel, coupled with the word "win" and a cash register sound. If the selected item did not win, the treasure chest remained on the screen and the words "Not a winner, better luck next time" were displayed (without any accompanying sound). The participants were not informed of the actual probabilities of reinforcement, and the instructions indicated that there was no cue, pattern or system that could be used to earn a coin on each trial. However, in order to help the participants decide on the best strategy, they were advised to sample both of their options sufficiently, pay attention to the outcome of their choices, and learn from experience. All participants completed a brief practice session consisting of 5 trials to ensure that they understood the instructions and had an opportunity to ask questions. Subsequently, a total of 300 trials were administered in one session, divided into 6 blocks of 50 trials. All trials were response terminated and the task took approximately twenty minutes to complete, with short breaks between blocks. Participants were able to view their running tally of money earned during the task via a display box located in the left corner of the computer screen.

1.3. Data Analysis

The classification of response strategies was based upon previous research using similar two-choice reinforcement learning tasks (Bogacz et al., 2007; Cohen et al., 2007; Miller et al., 2005; Shanks et al., 2002) that have generally classified performance according to 4 categories thought to reflect different strategies: 1) random chance, possibly reflecting failure to learn; 2) matching the reward probability 3) super-matching, overshooting the reward frequency of the best choice or, 4) optimizing, a strategy of almost always selecting the best response.

We used a five-stage iterative procedure to assign subjects to performance classes. Because we did not know the true probabilities of choosing the best response corresponding to the super-matcher and optimizer response strategies, we made initial classifications of participants using binomial expansion of our initial estimates of the probabilities corresponding to each classification, followed by maximum likelihood estimation to assign subjects to categories based on performance during the middle 100 trials. We then determined the mean probability of optimal choice associated with each performance class and performed binomial expansions of those probabilities, before re-classifying subjects based on maximum likelihood estimation (see Supplementary Materials for details on the classification procedure).

Since the four performance groups can be ranked in order of approach to the optimal strategy (chance<matchers<super-matchers<optimizers), we compared the two groups on degree of optimality in their strategy using a Mantel-Haenszel chi-square test (Mantel, 1963) difference in average rank order of strategies. After classifying subjects into the four performance groups, we excluded the subjects who performed at chance levels from further analyses [n = 2/29 controls, 8/51 patients]. Subsequently, we employed a two-way analysis of variance test (ANOVA) to determine whether there were differences in proportions of patients and controls in each class. Additionally, we proceeded to perform a series of t-tests and one-way ANOVAs on neuropsychological and clinical measures using performance classification as a between-subjects factor. We then analyzed correlations between experimental performance measures and clinical and neuropsychological functioning of both patient and control groups. Subsequently, we standardized the BPRS and SANS scales by computing z-scores in order to perform binomial regression analyses, which allowed us to examine the extent to which positive and negative symptoms predict behavioral task performance.

2. Results

2.1. Comparison of Overall Learning between Patients and Controls

There was no significant difference between normal controls and patients in the proportion of trials in which they chose the optimal side (Figs. 1 and 2), with no significant difference in overall performance [control mean = 0.78, SD = 0.10; patient mean = 0.74, SD = 0.10; $t(68) = 1.595$, $p = 0.115$], early learning [first hundred trials; NCs: 0.69 ± 0.13; SZs: 0.65 ± 0.10; $t(68) = 1.546$, $p = 0.127$], or end performance [last hundred trials; NCs: 0.83 ± 0.11; SZs: 0.80 ± 0.12; $t(68) = 1.254$, $p = 0.214$]. Patients and controls also showed similar increases in their proportions of choices of the optimal side from the first to the last hundred trials [t(68) = 0.195].

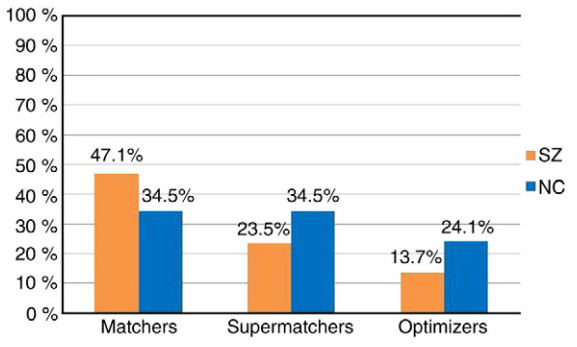

Fig. 1. Proportion of SZ patients and controls in each performance group.

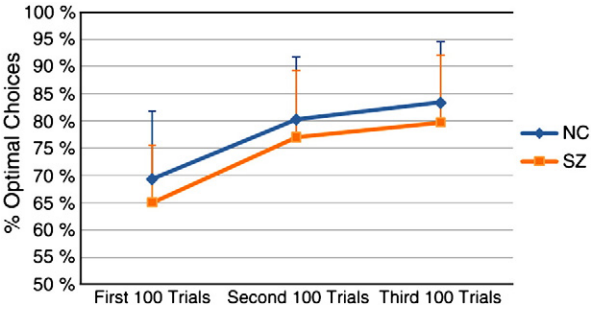

Fig. 2. Average percentage of optimal choices of SZ patients and controls in all three blocks.

2.2. Proportion of Subjects meeting Criteria for each Response Style

We then classified participants into the four response style categories (random chance, matchers, super-matchers, optimizers), and examined differences in the proportion of SZ and NC participants who fell into these classifications. Overall, the NC group showed a trend toward a greater proportion of subjects meeting criteria for the better response strategy classifications (supermatchers, optimizers) than patients [Mantel-Haenszel $\chi^2(1) = 3.53$, p = 0.06].

2.3. Comparison of SZ and NC Response Style Groups on Neuropsychological Functioning and Symptom Severity

As noted in the introduction, one can be led to make predictions about learning performance as a function of symptom status. To address this issue, we worked "backwards" by using end performance (last hundred trials) to form distinct groupings, and examined the symptom differences in these classes. As seen in Table 2, the optimizing group of patients had higher levels of negative symptoms. A one-way ANOVA indicated that the matchers, super-matchers, and optimizers significantly differed in the severity of SANS total negative symptoms [$F(2,42) = 7.30$, $p = 0.002$]. Follow-up, LSD post hoc comparisons indicated that patients within the optimizing group ($M = 39.14$, $SD = 15.41$) had a significantly greater severity of negative symptoms than those classified as matchers ($M = 27.87$, $SD = 8.49$) and super-matchers ($M = 19.67$, $SD = 11.46$). Thus, more severe negative symptoms were associated with greater optimizing, which results in greater overall earnings in the task, a rare instance of better performance being associated with greater symptom severity. Additional analyses revealed that, when classified according to performance, patients differed in the severity of some negative symptoms, but not others (Figure S1). Specifically, differences between performance subgroups were observe for Affective Blunting [F(2,39) = 4.54, p = 0.02], Avolition [F(2,39) = 3.38,

Table 2
Demographic information and cognitive assessments for performance groups.

Measure	Matchers M (SD)	Super-matchers M (SD)	Optimizers M (SD)	F / χ2	p-value
Patients					
Age	44.48 (9.44)	44.97 (9.09)	43.81 (10.19)	0.033	0.967
Education (years)	12.67 (2.70)	13.08 (2.02)	13.14 (2.04)	0.173	0.841
Paternal Education (years)	13.32 (3.88)	14.45 (4.41)	13.71 (3.15)	0.292	0.749
Gender (M: F)	13:11	9:03	7:00	5.618	0.444
Race				1.623	0.579
African American	9	5	1		
Caucasian	14	7	5		
Other	1	0	1		
Standard Neuropsychology					
WRAT	90.26 (13.68)	93.08 (18.54)	101.43 (8.98)	1.559	0.223
WTAR	90.83 (17.26)	94.67 (20.82)	106.29 (12.41)	2.066	0.14
WASI	93.54 (12.84)	101.33 (15.87)	111.86 (12.13)	1.819	0.175
MATRICS battery composite score	31.96 (14.63)	30.92 (16.93)	29.71 (8.80)	0.069	0.934
Clinical Ratings					
BPRS Total	32.44 (6.38)	36.83 (6.74)	43.86 (12.68)	5.998	0.005
BPRS Positive	1.86 (0.88)	2.63 (1.03)	3.04 (1.51)	4.352	0.020
BPRS Negative	1.64 (0.66)	1.56 (0.61)	2.36 (0.92)	3.381	0.044
BPRS Disorganized	1.24 (0.40)	1.55 (0.48)	1.46 (0.74)	1.690	0.198
SANS Total	27.87 (8.49)	19.67 (11.67)	39.14 (15.41)	7.305	0.002
Calgary Depression Scale	1.74 (1.69)	2.08 (2.35)	3.71 (3.95)	1.894	0.164
Antipsychotic Medication Regimen					
Haloperidol or Fluphenazine only	1	0	0		
Clozapine only	10	4	4		
Other second-generation only	10	6	1		
Clozapine + another antipsychotic	3	2	2		
First -generation + second-generation antipsychotic	0	0	1		
Controls					
Age	46.67 (9.86)	43.01 (10.46)	41.34 (11.26)	0.597	0.558
Education (years)	15.50 (1.96)	15.10 (2.23)	13.71 (1.80)	1.680	0.208
Paternal Education (years)	13.80 (3.23)	11.80 (4.19)	12.17 (2.56)	0.885	0.426
Gender (M: F)	7:03	6:04	6:01		
Race					
African American	2	4	4		
Caucasian	8	6	3		
Other	0	0	0		
Standard Neuropsychology					
WRAT	104.50 (15.82)	97.80 (17.11)	96.14 (19.93)	0.585	0.565
WTAR	107.20 (12.84)	99.90 (19.26)	100.14 (21.51)	0.516	0.603
WASI	117.50 (10.47)	108.70 (12.58)	110.71 (17.16)	1.197	0.320
MATRICS battery composite score	55.56 (12.31)	42.78 (14.85)	46.86 (19.13)	1.620	0.221

$p = 0.04$], and Anhedonia/Asociality [$F(2,39) = 4.99$, $p = 0.01$], but not Alogia [$F(2,39) = 1.93$, $p > 0.10$].

However, as seen in Table 2, a parallel result was observed with positive symptoms. The analysis of variance revealed significant differences in BPRS total score, [$F(2,42) = 5.998$, $p = 0.005$], as well as the positive cluster [$F(2,42) = 4.352$, $p = 0.020$] and negative cluster [$F(2,42) = 3.381$, $p = 0.044$] scores. No group differences were observed for disorganization cluster scores [$F(2,42) = 1.69$, $p = 0.198$]. Positive symptom scores in optimizers ($M = 3.04$, $SD = 1.51$) and super-matchers ($M = 2.63$, $SD = 1.03$) did not differ significantly ($p = 0.414$), and both groups had higher scores than matchers ($M = 1.86$, $SD = 0.89$; p's = 0.013 and 0.046, respectively). That is, higher levels of positive symptoms were associated with more behavioral optimizing.

Because both positive and negative symptoms appear to be related to being classified as an optimizer, we examined the role of both symptom types in a binomial logistic regression analysis. In this approach, the specific contribution of each symptom type is examined using standardized BPRS and SANS scores, and yields a test for statistical significance, as well as an odds ratio and a confidence interval that is more easily interpreted. That is, one can determine how much influence a one standard deviation increase in either BPRS positive or SANS negative symptoms makes in establishing the classification as an optimizer versus all other behavior patterns. The results suggest that the odds of being an optimizer increase with negative symptoms [$Exp(B) = 3.918$, CI = 1.207-12.722, $p = 0.023$], whereas the effect of positive symptoms is not significant [$Exp(B) = 2.024$, CI = 0.803-5.101, $p = 0.135$]. This essentially means that with each one standard deviation increase in negative symptom severity, the chances of being classified as an optimizer increases by almost 80%.

We also examined whether performance on the WASI, WTAR, WRAT and MATRICS battery composite score differed among patients meeting the various response strategy classifications. As seen in Table 2, no significant differences in performance on these neuropsychological measures emerged among the three response strategy groups for either

patients or controls. It is noteworthy that, despite having higher levels of positive and negative symptoms, the patient optimizers performed best on cognitive tests reflecting premorbid functioning (WTAR, WRAT) and were also found to have the highest current IQ score as assessed by the WASI (M = 111.86, SD = 12.13), when compared to matchers (M = 93.54, SD = 12.84) and super-matchers (M = 101.33, SD = 15.87).

3. Discussion

Consistent with numerous studies from the literature on reward learning in healthy individuals, we observed that, when presented with a two-choice probability learning task, the majority of subjects show suboptimal levels of performance in that they allocate choices according to their relative expected values (in this case, 70% and 30%). In examining variability in performance based on the underlying clinical characteristics of patients with schizophrenia, we found that the most adaptive and profitable strategy was adopted by a subgroup of patients exhibiting the most severe negative and positive symptoms. That is, the patients who chose the optimal side on nearly all of the trials exhibited higher levels of negative symptoms than the ones who matched or super-matched the probabilities of reinforcement. Although we observed that patients who showed behavioral optimizing in this task also exhibited the greatest positive symptoms, we found that the severity of negative rather than positive symptoms had an impact on predicting which behavior pattern a patient will adopt.

Importantly, the group of patients showing optimal behavior did not differ significantly from patients in other performance classifications in age, educational level, parental educational level, or racial/gender make-up. Furthermore, subgroups of patients identified by task performance did not differ in disease duration, medication status, or measures of neuropsychological functioning, such as working memory capacity, speed of processing, or hypothesis testing. Thus, it seems unlikely that reward maximization was the result of general intellectual impairment. One potential explanation for these results is that patients classified as optimizers have a tendency toward reduced exploration of response options under conditions of uncertainty. Supporting this interpretation are recent computational modeling results from our group, which indicate that higher levels of negative symptoms are associated with reduced uncertainty-driven exploration on a behavioral reward learning task (Strauss et al., 2011). In the current task environment, patients with the highest ratings for negative symptoms showed a reduced tendency to explore response alternatives defined by spatial locations. Based on additional work from our group (e.g., Strauss et al, 2011) we suspect that this reduced tendency to explore response alternatives would extend to tasks in which optimal responses are defined by other stimulus features.

We acknowledge that, in static environments such as those in the current task, it is plausible to expect that a tendency toward reduced exploration could lead to significantly fewer shifts from the optimal response and therefore result in earning a higher percentage of rewards. That is, sticking to a winning choice may make sense, provided that the subject is certain that the reward frequencies and magnitudes are, and will remain, constant. Such behavior would be less than optimal if the previous poor choice had changed in value and was now more desirable. Thus, consistently selecting a previously-rewarding choice comes at the cost of not knowing whether the environment has changed and potentially leading an individual to avoid certain responses. Such a learning pattern could be considered a viable contributor to symptoms of avolition and reduced reward seeking in people with schizophrenia.

Numerous brain systems have been linked to exploratory behavior, including the dopaminergic, cholinergic, and noradrenergic systems, especially through their targets in the prefrontal cortex (PFC; Yoshida & Ishii, 2006; Cohen et al., 2007; Bogacz et al., 2007; Padoa-Schioppa & Assad, 2006, Daw et al., 2006). Importantly, PFC is known to play a critical role in tracking reinforcement, computing its magnitude and representing value (Frank & Claus, 2006; Paulus et al., 2004). Furthermore, recent neuroimaging work (Miller et al., 2005) suggests that probability matching relies on PFC function, possibly in the service of explicitly representing reinforcement histories. If probability matching is, in fact, a phenomenon caused by PFC-dependent feedback sensitivity, we would expect patients with schizophrenia to perform poorly in such environments. The association of negative symptoms with decreased sensitivity to feedback could result in less outcome-driven and more internally-generated behavior. Using a similar two choice guessing task, Paulus and colleagues (1999) have demonstrated that long histories of previous responses (aside from external cues) exert a greater influence on the choices of individuals with schizophrenia relative to normal controls. In our study, patients with the most severe negative symptoms also may have been driven inordinately by response histories and habits, rather than recent feedback.

Paradoxically, our task environment provided an opportunity for the failure to engage in meaningful exploration and relative insensitivity to probabilistic, and occasionally misleading, feedback to be advantageous and result in maximum payoff. However, in volatile, non-stationary learning environments, as most real-world environments are, this quality would be extremely maladaptive, and would clinically manifest in a very limited, perseverative behavioral repertoire, where responses are not sensitive to changes in context. One would expect that the inevitable failures that would occur would lead to further withdrawal from novel or challenging situations and result in a form of inertia. Although our findings can be interpreted in a way that supports this notion, our task design did not allow us to study the effect of the volatility of the environment on exploratory behavior since the respective probabilities as well as the magnitude of reinforcement were kept constant. The impact of symptoms on a patient's tendency to exploit versus explore in non-stationary environments is the subject of ongoing work in our group.

The fact that our group-wise analyses found no significant differences in global percentage of trials on which the optimal choice is made or learning rate between patient and normal controls should not be seen as evidence of intact reinforcement learning in schizophrenia: the current task was very simple and there is reliable evidence of impairment when more challenging tasks are used (Waltz et al., 2007; Strauss et al., 2011). The current data suggest that a more fine-grained understanding of

reward processing and learning deficits may be obtained by dividing the larger population of patients with schizophrenia into more homogeneous subgroups, possibly using variables such as negative symptoms severity as a classifying factor. Given the heterogeneity of the illness, it appears highly likely that the mapping from cognitive process to neural mechanism to behavior will be much more successful using a symptomatic endpoint rather than a broad diagnostic class.

Role of funding source

This project was supported by NIMH grant R01 MH080066-01. Study sponsors had no role in paradigm formation, task design, participant recruitment, data collection, analysis and interpretation, nor were they involved in writing of the report; and in the decision to submit the paper for publication.

Contributors

Drs. Waltz, Gold and Frank formulated the paradigm, designed the study, and wrote the IRB protocol. Data collection was performed by Rebecca Wilbur, Sharon August and Zuzana Kasanova at the Maryland Psychiatric Research Center. Dr. Waltz and Zuzana Kasanova performed statistical analyses. Zuzana Kasanova wrote the first draft of the manuscript and all subsequent drafts were edited by Drs. Gold, Waltz, Strauss and Frank. All authors contributed to and have approved the final manuscript.

Conflict of interest

The authors have no conflicts of interest.

Acknowledgements

We would like to thank Benjamin Robinson for his assistance in task development and to Rebecca Wilbur and Sharon August for their assistance in data collection for this project. We would also like to thank Robert McMahon for his help with data analyses, and other MPRC staff members who made subject participation and study completion possible. And finally, we would like to thank the volunteers with schizophrenia and healthy controls for taking part in the study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at doi:10.1016/j.schres.2010.12.003. Additional details on methods and analyses. Includes two supplementary tables and one supplementary figure.

References

Addington, D., Addington, J., Maticka-Tyndale, E., Joyce, J., 1992. Reliability and validity of a depression rating scale for schizophrenics. Schizophr. Res. 6, 201–208.

Andreasen, N.C., 1984. The Scale for the Assessment of Negative Symptoms (SANS). University of Iowa, Iowa City, IA.

Baum, W.M., 1974. On two types of deviation from the matching law: bias and undermatching. J. Exp. Anal. Behav. 22.

Bereby-Meyer, Y., Erev, I., 1997. On learning to become successful loser: A comparison of alternative abstractions of learning processes in the loss domain. Technion- Israel Institute of Technology, Mimeo.

Bogacz, R., McClure, S.M., Li, J., Cohen, J.D., Montague, P.R., 2007. Short-term memory traces for action bias in human reinforcement learning. Brain Res. 1153, 111–121.

Cohen, J.D., McClure, S.M., Yu, A.J., 2007. Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. Philos. T. Roy. Soc. 362, 933–942.

Corlett, P.R., Murray, G.K., Honey, G.D., Aitken, M.R., Shanks, D.R., Robbins, T. W., Bullmore, E.T., Dickinson, A., Fletcher, P.C., 2007. Disrupted prediction-error signal in psychosis: evidence for an associative account of delusions. Brain 130, 2387–2400.

Corlett, P.R., Taylor, J.R., Wang, X.J., Fletcher, P.C., Krystal, J.H., 2010. Toward a neurobiology of delusions. Prog. Neurobiol. 92, 345–369.

Daw, N.D., O'Doherty, J.P., Dayan, P., Seymour, B., Dolan, R.J., 2006. Cortical substrates for exploratory decisions in humans. Nature 441, 876–879.

Edwards, W., 1956. Reward probability, amount, and information as determiners of sequential two-alternative decision. J. Exp. Psychol. 52, 177–188.

Estes, W.K., Strughan, J.H., 1954. Analysis of a verbal conditioning situation in terms of statistical learning theory. J. Exp. Psychol. 47, 225–234.

Farkas, M., Polgar, P., Kelemen, O., Rethelyi, J., Bitter, I., Myers, C.E., Gluck, M. A., Keri, S., 2008. Associative learning in deficit and nondeficit schizophrenia. NeuroReport 19, 55–58.

First, M.B., Spitzer, R.L., Gibbon, M., Williams, J.B.W., 2001. Structured Clinical Interview for DSM-IV-TR Axis I Disorders-Patient Edition (SCID-I/P 2/ 2001 Revision). Biometrics Research Department, New York State Psychiatric Institute, New York.

Frank, M.J., Claus, E.D., 2006. Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. Psychol. Rev. 113, 300–326.

Friedman, D., Massaro, D.W., 1998. Understanding variability in binary and continuous choice. Psychon. B. Rev. 5, 370–389.

Gardner, R., 1958. Multi-choice decision behavior. Am. J. Psychol. 71.

Green, M.F., Nuechterlein, K.H., Gold, J.M., et al., 2004. Approaching a consensus cognitive battery for clinical trials in schizophrenia: the NIMH-MATRICS conference to select cognitive domains and test criteria. Biol. Psychiat. 56, 301–307.

Healy, A.F., Kubovy, M., 1981. Probability matching and the formation of conservative decision rules in a numerical analog of signal detection. J. Exp. Psychol Hum. L. 7, 344–354.

Herrnstein, R.J., 1970. On the law of effect. J. Exp. Anal. Behav. 13, 243–266.

Herrnstein, R.J., 1982. Melioration as behavioral dynamism. In: Commons, M. L., Herrnstein, R.J., Rachlin, H. (Eds.), Quantitative analyses of behavior. Vol. II: Matching and maximizing accounts. Ballinger; Cambridge, MA.

Herrnstein, R.J., 1990. Rational choice theory: necessary but not sufficient. Am. Psychol. 45, 356–367.

Herrnstein, R.J., 1997. In: Rachlin, H., Laibson, D.I. (Eds.), The Matching Law: Papers in Psychology and Economics. Harvard University Press, Cambridge, MA.

Hinson, J.M., Staddon, J.E.R., 1983. Matching, maximizing, and hill-climbing. J. Exp. Anal. Behav. 40, 321–331.

Houston, A.I., McNamara, J.M., Steer, M.D., 2007. Do we expect natural selection to produce rational behavior? Philos. T. Roy. Soc. 362, 1531–1543.

Jarvik, M.E., 1951. Probability learning and a negative recency effect in a serial anticipation of alternative symbols. J. Exp. Psychol. 41, 291–297.

Jensen, J., Willeit, M., Zipursky, R.B., Savin, I., Smith, A.J., Menon, M., Crawley, A.P., Kapur, S., 2008. The formation of abnormal associations in schizophrenia: neural and behavioral evidence. Neuropsychopharmacol 33, 473–479.

Kapur, S., 2003. Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. Am. J. Psychiat. 160, 13–23.

Mantel, N., 1963. Chi-square Tests with One Degree of Freedom: Extensions of the Mantel-Haenszel Procedure. J. Am. Stat. Asoc. 58, 690–700.

Miller, M.B., Valsangkar-Smyth, M., Newman, S., Dumont, H., Wolford, G., 2005. Brain activations associated with probability matching. Neuropsychologia 43, 1598–1608.

Murray, G.K., Cheng, F., Clark, L., Barnett, J.H., Blackwell, A.D., Fletcher, P.C., Robbins, T.W., Bullmore, E.T., Jones, P.B., 2008. Reinforcement and reversal learning in first-episode psychosis. Schizophr. Bull. 34, 848–855.

Myers, J.L., Cruse, D., 1968. Two-choice discrimination learning as a function of stimulus and event probabilities. J. Exp. Psychol. 77, 453–459.

Myers, J.L., Fort, J.G., Katz, L., Suydam, M., 1963. Differential memory gains and losses and event probability in a two-choice situation. J. Exp. Psychol. 66, 521–522.

Neimark, E., Shufod, E., 1959. Comparison of predictions and estimates in a probability learning situation. J. Exp. Psychol. 57, 294–298.

Overall, J.E., Gorman, D.R., 1962. The Brief Psychiatric Rating Scale. Psychol. Rep. 10, 799–812.

Padoa-Schioppa, C., Assad, J.A., 2006. Neurons in the orbitofrontal cortex encode economic value. Nature 441, 223–226 (Electronic publication ahead of print 2006 Apr 23).

Paulus, M.P., Geyer, M.A., Braff, D.L., 1999. Long-range correlations in choice sequences of schizophrenic patients. Schizophr. Res. 35, 69–75.

Paulus, M.P., Feinstein, J.S., Tapert, S.F., Liu, T.T., 2004. Trend detection via temporal difference model predicts inferior prefrontal cortex activation during acquisition of advantageous action selection. Neuroimage 21, 733–743.

Polgar, P., Farkas, M., Nagy, O., Kelemen, O., Rethelyi, J., Bitter, I., Myers, C.E., Gluck, M.A., Keri, S., 2008. How to find the way out from four rooms? The learning of "chaining" associations may shed light on the neuropsychology of the deficit syndrome of schizophrenia. Schizophr. Res. 99 (1–3), 200–207.

Sakai, Y., Fukai, T., 2008. When does reward maximization lead to matching law? PLoS ONE 3 (11).

Shanks, D.R., Tunney, R.J., McCarthy, J.D., 2002. A re-examination of probability matching and rational choice. J. Behav. Decis. Making. 15, 233–250.
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. Psychiat. 69 (5), 424–431.
Sugrue, L.P., Corrado, G.S., Newsome, W.L., 2004. Matching behavior and the representation of value in the parietal cortex. Science 304, 1782–1786.
van Os, J., Kapur, S., 2009. Schizophrenia. Lancet 374, 635–645.
Vulkan, N., 2000. An economist's perspective on probability matching. J Econ Surv. 14, 101–118.
Waltz, J.A., Gold, J.M., 2007. Probabilistic reversal learning impairments in schizophrenia: Further evidence of orbitofrontal dysfunction. Schizophr. Res. 93, 296–303.
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 Psychiat. 62, 756–764.
Waltz, J.A., Schweitzer, J.B., Gold, J.M., Kurup, P.K., Ross, T.J., Salmeron, B.J., Rose, E.J., McClure, S.M., Stein, E.A., 2009. Patients with schizophrenia have a reduced neural response to both unpredictable and predictable primary reinforcers. Neuropsychopharmacology 6, 1567–1577. (Electronic publication ahead of print 2008 Dec 3. PubMed PMID: 19052540).
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 12, 2427–2439. (Electronic publication ahead of print 2010 Aug 18. PubMed PMID: 20720534; PubMed Central PMCID: PMC2955756).
Wechsler, D., 1999. Wechsler Abbreviated Scale of Intelligence. The Psychological Corporation, San Antonio, TX.
Wechsler, D., 2001. Wechsler Test of Adult Reading. The Psychological Corporation, San Antonio, TX.
Wilkinson, G.S., 1993. Wide Range Achievement test. 3 rd PAR/Psychological Assessment Resources, Inc, Lutz, FL.
Yoshida, W., Ishii, S., 2006. Resolution of uncertainty in prefrontal cortex. Neuron 50, 781–789.