



Published in final edited form as:

Biol Psychiatry. 2011 March 1; 69(5): 424–431. doi:10.1016/j.biopsych.2010.10.015.

Deficits in Positive Reinforcement Learning and Uncertainty-Driven Exploration are Associated with Distinct Aspects of Negative Symptoms in Schizophrenia

Gregory P. Strauss, Ph.D.^{1,*}, Michael J. Frank, Ph.D.^{2,*}, James A. Waltz, Ph.D.¹, Zuzana Kasanova, B.A.¹, Ellen S. Herbener, Ph.D.³, and James M. Gold, Ph.D.^{1,†}

¹University of Maryland School of Medicine, Department of Psychiatry; Maryland Psychiatric Research Center

²Brown University, Departments of Cognitive & Linguistic Sciences and Psychology

³University of Illinois, Chicago, Departments of Psychiatry and Psychology

Abstract

Background—Negative symptoms are core features of schizophrenia; however, the cognitive and neural basis for individual negative symptom domains remains unclear. Converging evidence suggests a role for striatal and prefrontal dopamine in reward learning and the exploration of actions that might produce outcomes that are better than the status quo. The current study examines whether deficits in reinforcement learning and uncertainty-driven exploration predict specific negative symptoms domains.

Methods—We administered a temporal decision making task, which required trial-by-trial adjustment of reaction time (RT) to maximize reward receipt, to 51 patients with schizophrenia and 39 age-matched healthy controls. Task conditions were designed such that expected value (probability * magnitude) increased (IEV), decreased (DEV), or remained constant (CEV) with increasing response times. Computational analyses were applied to estimate the degree to which trial-by-trial responses are influenced by reinforcement history.

Results—Individuals with schizophrenia showed impaired Go learning, but intact NoGo learning relative to controls. These effects were pronounced as a function of global measures of negative symptom. Uncertainty-based exploration was substantially reduced in individuals with schizophrenia, and selectively correlated with clinical ratings of anhedonia.

Conclusions—Schizophrenia patients, particularly those with high negative symptoms, failed to speed RT's to increase positive outcomes and showed reduced tendency to explore when alternative actions could lead to better outcomes than the status quo. Results are interpreted in the

© 2010 Society of Biological Psychiatry. Published by Elsevier Inc. All rights reserved.

†Correspondence concerning this article should be addressed to Gregory P. Strauss, Ph.D., gstrauss@mprc.umaryland.edu. Phone: +1-410-402-6053. Fax: +1-410-402-7198. University of Maryland School of Medicine, Maryland Psychiatric Research Center, P.O. Box 21247, Baltimore, MD, 21228 USA.

*These authors contributed equally to this work.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Financial Disclosures

The authors have no conflicts of interest, biomedical, financial, or otherwise, to declare.

context of current computational, genetic, and pharmacological data supporting the roles of striatal and prefrontal dopamine in these processes.

Keywords

Schizophrenia; Reward; Reinforcement Learning; Negative Symptoms; Dopamine; Computational Model

Introduction

Dopaminergic (DA) signaling plays a key role in the detection, evaluation, and prediction of rewards. Several structures that receive DA input are differentially involved in specific aspects of reward learning. For example, the striatum and orbitofrontal cortex (OFC) have been found to be involved in reward prediction and reward based decision making, with the OFC being particularly responsive to reward magnitudes (1–6). In reinforcement learning models of corticostriatal circuitry (7,8), phasic DA signals are proposed to modify synaptic plasticity in the corticostriatal pathway (9,10), and subsequently reinforce “Go” (learning to pursue actions that have high reward probability) and “NoGo” learning (learning to avoid actions with low reward probabilities) (7,11). Specifically, increases in phasic striatal DA support Go learning from positive feedback via D1 receptor stimulation, whereas decreases in phasic striatal DA support avoidance learning from negative feedback via D2 receptor disinhibition.

This model has been applied to understand patterns of reward learning in Parkinson’s Disease (PD) patients (7), who have depleted striatal DA levels as a result of the disease, but increased striatal DA levels following DA medication (12,13). Supporting the models, it has been found that PD patients on medication learn better from positive than negative decision outcomes, but patients off medication showed the opposite bias (8,14–16). Imaging studies show that these biases are accompanied by medication-induced increased sensitivity to positive prediction errors and reduced sensitivity to negative prediction errors in the ventral and dorsolateral striatum (17). Behaviorally, these medication-induced effects have been primarily observed in tasks where participants learned stimulus-response relationships as a function of reinforcement (e.g., probabilistic learning).

Recent studies have found that SZ patients exhibit reinforcement learning abnormalities, specifically in learning to integrate the history of probabilistic positive decisions across trials (18–20). These deficits have been attributed to deficient phasic striatal dopaminergic signals and D1 receptor functionality, leading to poor Go learning. A deficit in learning to repeat those actions most likely to yield positive reinforcement may provide an intuitively appealing account that could explain negative symptoms. Several studies show that negative symptoms are associated with impairments in rapid, trial-to-trial behavioral adaptation in response to recent changes in reinforcement values, particularly during early phases of learning. We have argued that this deficit in rapid acquisition associated with negative symptoms is likely to stem from prefrontal cortical dysfunction (19,20). Similar patterns of early learning deficits are seen in patients with orbitofrontal damage (21) and in healthy participants with the val/val genotype of the COMT gene (15), who have reduced prefrontal (and particularly, orbitofrontal) dopamine levels(22). However, relationships have also been reported between negative symptoms and deficient reward-related BOLD activity in basal ganglia structures. These changes in basal ganglia (BG) BOLD signal do not themselves necessarily indicate that the BG are the source of the deficits. For example, reduced top-down input from OFC to BG during reward would result in reduced BG activations. Thus, the neural underpinnings of reinforcement learning deficits and negative symptoms are at present unclear.

In the current study, we examined reinforcement learning in schizophrenia using a task that required subjects to modulate the response time of a single motor response, where both reinforcement probabilities and magnitudes varied as a function of RT (27). Behavioral task conditions were designed to assess Go learning, NoGo learning, and relative sensitivity to reward frequency vs. magnitude.

Computational modeling was performed in an effort to obtain a richer understanding of the behavioral findings. The modeling approach allows us to assess the degree to which participants adjust response times as a function of positive and negative prediction errors across all four of the task conditions (not solely in the conditions where it is most advantageous to do so), while distinguishing these measures from several other components to RT adjustment in the task (captured by other model parameters).

Based upon theories suggesting that SZ is associated with abnormal dopaminergic signaling (high tonic DA but low phasic DA) and impaired D1 function in particular (28–30), we hypothesized that SZ patients would fail to show relative speeding when rewards are most available for fast responses, and thus earlier responding results in better-than-expected outcomes (i.e., positive prediction errors). Conversely, based on our previous findings (19) and the fact that patients were medicated with D2 antagonists, we hypothesized that SZ patients would show intact slowing as a function of negative prediction errors when early responses produced outcomes that were lower than expected on average. The latter interpretation is supported by computational simulations showing that D2 blockade enhances NoGo learning and RT slowing (see (31)) and recent demonstration that D2 blockade enhances NoGo learning in Tourette's syndrome patients (16).

Evidence for such a pattern of spared sensitivity to negative outcomes, coupled with reduced ability to learn to approach responses leading to positive reinforcement could be considered a perfect neurobehavioral recipe for avolition and anhedonia. In light of previous studies indicating that reinforcement learning impairments are most severe in high negative symptom patients (32,33), we therefore also examined the role of negative symptoms in Go and NoGo learning, with the prediction that high negative symptom patients would show the greatest Go learning impairment and comparatively intact NoGo performance. Alternatively, rather than resulting from reinforcement *learning* deficits *per se*, some aspects of negative symptoms may be characterized by a reduced tendency to appropriately *explore* alternative actions in the hope that they might produce better outcomes. Notably, the computations of outcome uncertainty used to guide exploration are thought to depend on neuromodulation within the prefrontal cortex (PFC) (34–36). Accordingly, a recent genetic study with the same task used here showed that, consistent with striatal DA genetic effects on Go/NoGo learning, individual differences in uncertainty-driven exploration were predicted by COMT val/met genotype (36). We thus applied the same computational analyses of trial-by-trial responses in the current study to investigate the prediction that patients would show not only differences in speeding and slowing as a function of prediction errors, but also whether they would exhibit uncertainty-driven exploration.

Methods and Materials

Participants

Participants included 51 patients meeting DSM-IV-TR criteria for schizophrenia and 39 healthy controls (CN). The patients were recruited from the outpatient clinics at the Maryland Psychiatric Research Center and were studied during a period of clinical stability. All patients met DSM-IV diagnostic criteria for schizophrenia or schizoaffective disorder. Consensus diagnosis was established with a best-estimate approach based on medical

records and confirmed using the Structured Clinical Interview for DSM-IV (SCID) (37). All patients were receiving antipsychotic medications.

Control subjects were recruited through random digit dialing and word of mouth among individuals recruited through random digit dialing. All controls underwent a screening interview and denied lifetime and family history of psychosis and any active Axis I disorder on the SCID. All participants denied lifetime history of significant neurological conditions and recent substance abuse as determined by the SCID (none within 6 months). Upon entry to our subject pool, we routinely screen for substance use via urine toxicology testing. In the current study, targeted urine toxicology testing was performed in instances where there were suspicions of substance use. Patient and control groups did not significantly differ in age, parental education, gender, or ethnicity. Patients had fewer years of total education and lower WASI estimated full-scale IQ's than controls (see Table 1).

Schizophrenia patients were also divided into High (HI-NEG) and Low Negative (LOW-NEG) symptom groups based upon a median split on the Scale for the Assessment of Negative Symptoms (SANS: (38,39)) total score. The 22 item version of the SANS developed in the CONSIST clinical trial was used (39), which has fewer items than the original 30- or 25-item version, with total scores ranging from 0–110. The three groups did not significantly differ on age, parental education, gender, or ethnicity; however, they did differ on IQ, such that CN had significantly higher IQ than both schizophrenia groups. There were no differences in IQ between the HI-NEG and LOW-NEG patients. HI-NEG and LOW-NEG patients significantly differed on the BPRS negative symptom factor score, but not on positive symptoms, disorganization, or total scale score. HI-NEG and LOW-NEG patients were also prescribed a similar regimen of antipsychotic medications at the time of testing and did not differ on chlorpromazine (CPZ) equivalent dosage (40) (see Table 1).

General Procedures

The current tests were administered as part of a larger battery of reward-learning, symptom interview, and neuropsychological measures. For each subject, demographic, diagnostic, and symptom ratings were completed prior to administration of the neurocognitive evaluations. Symptom interviews included the SANS and Brief Psychiatric Rating Scale (BPRS: (41)). Patient and control participants recruited from the community received monetary compensation for participation. Study personnel administering the neurocognitive tasks included B.A. and M.A. level research assistants.

Temporal Utility Integration Task

Participants completed the “temporal utility integration task” designed by Moustafa et al. (27). In this task, subjects were presented a clock face, which had a single arm that made a full turn over the course of 5 seconds. Participants were asked to press a button on a response pad at some point before the arm made a full turn. Following each response, participants were informed whether they had won points, and if so, how many. The trial ended once the subject responded using the game pad or if the 5 s duration elapsed and the subject did not respond. The inter-trial-interval (ITI) was set at 1 s. Participants completed four separate conditions, each consisting of 50 trials, in which reward probability and magnitude varied as a function of time elapsed on the clock until response. In the three primary conditions (DEV, CEV, and IEV), the number of points (reward magnitude) increased, whereas the probability of receiving the reward decreased, over time within each trial. Feedback was presented on the screen in the format of “You win XX points!” Functions within each condition were designed such that the expected value (probability*magnitude) either decreased (DEV), increased (IEV), or remained constant (CEV), across the 5 s trial duration (Figure 1). Thus in the IEV condition, early responses

yielded a small number of points (lower than expected on average), and the associated negative prediction errors should lead to NoGo learning and slowed responses. In contrast, early responses in the DEV condition yielded a higher number of points than expected, and should therefore lead to Go learning/speeding. Slower responses in the IEV condition yielded more points on average, whereas in the DEV condition faster responses yielded more points.

In addition to these primary conditions, we also included a condition where expected value remains constant (like CEV), but reward probability increases and magnitude decreases as time elapses on the clock (i.e., the opposite to CEV). Since both CEV and CEVR have equal expected values across the entire clock face, any difference in response time in these two conditions can be attributed to a participants' potential bias to learn more about reward probability than about magnitude or vice-versa. Specifically, if a subject waits longer to respond in CEVR than in CEV, it can be inferred that the participant is risk averse as they value higher probabilities of reward more than higher magnitudes of reward.

Order of condition (CEV, DEV, IEV, CEVR) was counterbalanced across participants, and a rest break was given between each of the conditions (i.e., after every 50 trials). At the beginning of each condition, subjects were instructed to respond at different times in order to find the interval on the clock that would allow them to win the most points; however, they were not told about the different rules for each condition (e.g., IEV, DEV). Each condition also had a different color clock face to highlight the uniqueness of each context, and the assignment of color was counterbalanced across conditions. The task was presented using E-Prime software.

Computational modeling was used as a tool to more specifically probe aspects of behavior in this task (36). The model allows us to estimate the degree to which individuals adjust their response times as a function of accumulated reward prediction errors, and uncertainty-driven exploration, distinctly from other components (see Table 2 for description of model parameters and Supplement 1 and ref. ³⁶ for mathematical details). The major parameters of interest for the current study are α_G , α_N , and ϵ . The α_G and α_N parameters were used to test whether patients have deficits in learning from gains vs. losses more fully than what can be surmised from the behavioral data as the model estimates on average the degree to which subjects speed up or slow down and use positive and negative prediction errors across all conditions. The ϵ parameter was used to test the possibility that individuals with SZ have a reduced tendency to appropriately *explore* alternative actions in the hope that they might produce better outcomes.

Data Analysis

Behavioral analyses examined RT for each condition, either for the entire block or the difference score between the second and first half of trials in each condition as indicated in the text. Repeated measures ANOVAs, one-way ANOVAs, t-tests, and chi-square analyses were calculated to determine group differences. Spearman correlations were calculated to examine relationships between test data and symptoms. The Greenhouse-Geisser correction was applied in instances when the assumptions of sphericity or covariance were violated. Scheffe contrasts were additionally performed as post hoc tests. Wilcoxon Mann-Whitney tests were used to examine group differences on modeling parameters. Initial analyses examined between-group differences in patients and controls. However, given that SANS scores are typically bimodally distributed, we examined the role of negative symptoms using between-group analyses (i.e., comparing high negative symptom, low negative symptom, and control groups), but also reported correlations for completeness. Data were analyzed using SPSS version 12 software.

Results

Go vs. No Go Learning and Uncertainty-Driven Exploration

Analysis of behavioral data indicated that in both SZ patients and healthy controls, RTs in the IEV condition (No Go learning) were significantly slower than the DEV condition (Go Learning) (CN: $t = -4.48, p < 0.001$; SZ: $t = -4.99, p < .001$), suggesting that both groups learn to adapt RTs in the expected direction (see Table S1 in Supplement 1). However, these overall means calculated across the entire block of trials mask differences in learning from the beginning to the end of the condition. As such, difference scores were computed separately for each condition to estimate RT adaptation from the first half of trials to the second half of trials (2nd half of trials – 1st half of trials). Consistent with hypotheses, SZ patients fail to learn to speed up by the end of the block in the DEV (Go learning) condition as much as CN, but perform similarly to CN in the IEV (No Go) and CEV (Control) conditions (see Figure 2). This was confirmed statistically by separate repeated measures ANOVAs, which indicated that groups significantly differed on the DEV condition, $F(1, 88) = 9.49, p = 0.003$ ($\eta^2 = .10$), but not the IEV, $F(1, 88) = 0.01, p = 0.913$ ($\eta^2 = .00$), or CEV conditions, $F(1, 88) = 1.86, p = 0.176$ ($\eta^2 = .02$). These behavioral findings are consistent with intact No Go, but impaired Go learning in schizophrenia¹.

Computational modeling was used to obtain a richer understanding of these behavioral findings on Go and No Go learning. Overall, the computational model encompassing the specified combination of parameters (see Supplemental Results in Supplement 1) and which was the best fit to the data in our previous study, also provided a reasonable fit to the behavioral data here (Figure 3).

Significant parameter differences between SZ patients and controls were observed for ϵ , the degree to which exploration occurs in proportion to relative uncertainty about reward outcomes, $F(1, 88) = 9.1, p = .003$. These differences remain significant after bonferroni correction (Figure 4). Additional analyses also confirmed that the exploration effect in schizophrenia was specific to uncertainty, as groups did not differ in measures of overall RT variability or RT swings (see Figure S4 in Supplement 1). There was also a trend for αG to be smaller in patients (Wilcoxon Mann-Whitney test, two-tailed $p = 0.07$), consistent with the behavioral results, whereby patients exhibited deficits in learning to speed responses in the DEV condition. A follow-up logistic regression with both parameters entered as predictors confirmed that both the explore ($p < .02$) and αG parameters ($p = .028$) were independently predictive of SZ. There were no other significant differences between patients and controls in any of the other parameters (see Table 3).

A regression analysis shed further light onto this interpretation, revealing that individual differences in the tendency to speed up to maximize rewards in the DEV condition is predicted by αG ($p = .018$), such that higher parameters were associated with increased speeding. No such difference was seen in terms of the model parameter estimating the degree to which participants slow down as a function of negative prediction errors.

Negative Symptoms

We also conducted analyses examining behavioral and modeling parameters in patients with High Negative Symptoms (HI-NEG), Low Negative Symptoms (LOW-NEG), and Controls (CN). As can be seen in Figure 5, the HI-NEG group showed significantly reduced speeding

¹To further examine the specificity of Go and No Go learning performance in SZ and CN, we conducted a 3 condition (CEV, DEV, IEV) \times 2 Time (Block 1, Block 2) \times 2 Group (SZ, CN) repeated measures ANOVA, and found a nonsignificant interaction ($p = 0.16$). Nonetheless, the analyses were consistent in direction with the results of the ANOVAs performed on the learning measures. We suspect that the additional variance introduced into these more complex ANOVA models resulted in less observed power.

from the first to second half of trials in the DEV condition, consistent with a Go learning deficit in high negative symptom patients. One-way ANOVAs, conducted with difference scores (2nd half of trials – 1st half of trials) as the dependent variable supports this interpretation, indicating that the 3 groups significantly differed on DEV, $F(2, 85) = 4.78$, $p = 0.01$ ($\eta^2 = .11$), but not IEV, $F(2, 85) = 0.23$, $p = 0.37$ ($\eta^2 = .01$), or CEV, $F(2, 85) = 0.99$, $p = 0.73$ ($\eta^2 = .02$). Post hoc Scheffe contrasts conducted for the DEV change condition were significant between the HI-NEG and CN ($p = .02$), but not the LOW-NEG and CN ($p = .14$) or HI-NEG and LOW-NEG ($p = .68$) groups. The correlation between the SANS total score and DEV and IEV conditions was nonsignificant. The discrepancy between the significant between-subjects analysis on the DEV condition and nonsignificant correlation between negative symptoms and DEV learning is likely explained by the fact that Go learning deficits were most pronounced in HI-NEG patients, but still present in LOW-NEG patients as well, thereby attenuating the strength of the correlation.

In a separate analysis of behavioral data, HI-NEG patients also failed to show either a probability or magnitude bias, whereas CN and LOW-NEG both showed a bias to learn more about probability than magnitude (see Figure S1 in Supplement 1 for discussion).

One-way ANOVAs indicated that the 3 groups failed to significantly differ on any parameter other than exploration, $F(2, 85) = 5.02$, $p = 0.009$ (HI-NEG: $M = 1187$, $SD = 1561$; LOW-NEG: $M = 1323$, $SD = 1678$). Post hoc Scheffe contrasts indicated significant differences between HI-NEG and CN ($p = 0.03$) subjects; however, LOW-NEG and CN ($p = 0.06$) and HI-NEG and LOW-NEG ($p = 0.97$) did not significantly differ. Interestingly, correlational analyses indicated that the dramatic reduction in exploration was most severe in patients with high avolition-anhedonia SANS summary scores ($r = -0.28$, $p < 0.05$). There were no significant correlations between ϵ and the restricted affect (SANS alolia + blunted affect items) summary score ($r = 0.05$) or the SANS total, suggesting that the relationship may be specific to the avolition-anhedonia domain. Follow-up correlational analyses using the avolition and anhedonia global scores indicated that the relationship with exploration was specific to anhedonia (Anhedonia $r = -0.44$, $p < 0.01$; Avolition $r = -0.15$, $p > 0.3$) (see Figure 6). The test for significant differences between these correlations approached significance ($z = -1.54$, $p = 0.06$).

Given the unique association with anhedonia, we further investigated whether the association between anhedonia and exploration was specific to uncertainty, and determined that anhedonia was only associated with uncertainty-driven exploration, and not overall RT variability or consecutive RT variance. Furthermore, control model simulations revealed that other models of RT swings, including parameters for lose-switch, or regression to the mean, did not correlate with anhedonia (see Supplemental Results and Figure S2 in Supplement 1).

Antipsychotic Medication

Correlational analyses indicated that CPZ dosage was not significantly correlated with behavioral performance in any of the conditions (all p 's > 0.16) or modeling parameters (all p 's > 0.14). Analyses examining between-group differences in patients categorized as a function of low and high potency D2 blockade antipsychotics indicated no differences between medication groups in behavioral task conditions (see Supplemental Results in Supplement 1).

Discussion

Two main findings emerged from the current study. First, behavioral data indicated that patients were less able to learn to speed up to maximize rewards, which is consistent with a Go learning deficit. The model simulations suggest this deficit may at least in part be due to

lower αG parameter, as a regression analysis revealed that individual differences in the tendency to speed up to maximize rewards in the DEV condition is predicted by αG , such that higher parameters were associated with increased speeding. Given that SZ showed a deficit in both αG and the DEV, but not αN or IEV, we feel that the results of the computational model provide further confidence that the deficits specific to Go learning in schizophrenia are reliable. Furthermore, symptom sub-group analyses revealed that, in terms of DEV performance, Go learning deficits are most severe in patients exhibiting greater severity of negative symptoms.

These findings are consistent with our previous probabilistic selection study indicating that SZ is associated with impaired Go learning and intact No Go learning (19). When viewed in conjunction with neurocomputational models of corticostriatal circuitry in reinforcement learning (7,8), the current behavioral and modeling findings are suggestive of potential dysfunction in the direct D1 driven BG pathway leading to abnormalities in using positive feedback to guide behavior, with relatively intact function in the D2 driven, indirect pathway leading to normal use of probabilistic negative feedback in decision making. This BG-based account is supported by other evidence indicating that BG dopamine acts to speed responding toward rewarding cues (42,43), as well as pharmacological and animal studies showing that this process likely relies on D1-driven activation and Go learning (44–46). However, this interpretation is of course speculative, and cannot be confirmed without conducting a study on unmedicated 1st episode patients to see if No Go learning improves when patients are treated with D2 blocking antipsychotics.

A second major finding was that SZ patients exhibited a large and reliable reduction in the tendency to make exploratory behavioral adjustments toward responses that could potentially yield larger expected values than those obtained by staying with the status quo. Additionally, given that there was no association between anhedonia and overall RT variability or consecutive variance, anhedonia appears to be selectively associated with the failure to initiate the *proactive* strategy of adjusting responses to gather more information in order to reduce uncertainty about potential benefits of alternative behaviors. These findings demonstrate the usefulness of computational modeling approaches to psychiatry (25,47–51).

We posit that these effects are related to degradations in prefrontal cortical dopaminergic function, often attributed as a source of negative symptoms (28,30,52,53). This interpretation is supported by our recently reported gene-dose effect of the val/met polymorphism of the COMT gene in healthy individuals performing this same task (36), which indicated that the val/val genotype was characterized by the lowest degree of uncertainty-driven exploration and the met/met genotype with the greatest degree of exploration. Variations in COMT affect prefrontal, and particularly orbitofrontal, dopamine levels (22), and a recent study reported a COMT gene dose effect on orbitofrontal activity during reward receipt (54). Thus together, these studies support the assertion that the val/val genotype shares features of cognitive dysfunction observed in SZ (55). Finally, ongoing imaging work in healthy individuals (56), together with other related studies (35,57), suggest that relative uncertainty computations associated with exploration are represented in prefrontal cortical activation patterns. Finally, even if the computations of expected reward values are relatively intact in SZ, it is possible that patients with anhedonia explicitly assign a negative expected value to uncertain outcomes, due to their prior expectations (see (51) for a related model of depression). Regardless of the neural mechanism, our findings suggest that anhedonia may result from an inability to determine when to explore actions that might improve one's ability to obtain rewards.

Of particular interest was that reduced uncertainty-driven exploration correlated with the Avolition-Anhedonia domain on the SANS, but not the Restricted Affect factor.

Additionally, the effect was more highly related to anhedonia than avolition. This result is potentially informative about differences in the pathology of these symptom domains. As rated by the SANS, anhedonia reflects a behavioral component of reward seeking (e.g., initiating social activities, sexual interest/and or activity, pursuing recreational activities, number of close relationships), rather than the capacity to experience pleasure, which is often inferred from behavior. Avolition items on the SANS are less related to reward seeking behavior, and more broadly related to the frequency with which patients initiate and persist in many kinds of tasks which is likely to be influenced by a number of factors, such as disorganization, generalized cognitive impairment, and sedation. The significant correlation with anhedonia, but not avolition may therefore reflect that reduced reward seeking behavior in schizophrenia is critically related to the extent to which patients make exploratory choices when they are uncertain about the value of alternative actions and whether they might produce better outcomes than the status quo.

Results should be viewed with certain limitations in mind. First, analyses regarding the role of medication on task performance should be viewed with caution, as CPZ equivalents for atypical medications may not be appropriate and D2 potency classifications provide only a gross estimate of the effects of different antipsychotics. A more definitive test of antipsychotic effects should be conducted in first episode patients tested on and off medications. Second, we did not collect DNA in this study and it is unclear whether the COMT genetic effect observed in healthy individuals on exploration may partially contribute to the effects of anhedonia and SZ reported here. Finally, although the SANS is still the gold standard negative symptom assessment in the field, it has recently been suggested that newer measures being developed in response to the NIMH MATRICS (e.g., (60)) initiative may provide a more comprehensive and current assessment of negative symptom dimensions. As such, it is unclear whether the relationship reported between SANS anhedonia and exploration may actually reflect some other component of negative symptoms on these newer scales.

In summary, the current findings have important implications for understanding the etiology of schizophrenia. Results from the computational model and behavioral data indicate that patients have deficits in Go learning, which appear to be due to reduced sensitivity to positive prediction errors. Thus patients show a reduced sensitivity to the impact of rewarding outcomes on future behavioral choices. Furthermore, patients display reduced uncertainty-driven exploration, which was specifically associated with greater severity of anhedonia. Thus, patients are less likely to explore, and therefore less likely to discover, that an alternative response might yield more rewarding outcomes. While these deficits are independent of one another in the model, at a clinical level it is easy to imagine how these impairments might amplify one another and result in a narrow behavioral repertoire and a lack of goal-directed, reward-seeking behavior.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by US National Institutes of Mental Health grant R01 MH080066-01. We would also like to thank the subjects who participated in the study and staff at the Maryland Psychiatric Research Center who made the completion of the study possible. We are especially thankful to members of Dr. Gold's lab, Jackie Kiwanuka, Sharon August, Lindsay Phebus, Leeka Hubzin, and Tatyana Matveeva, who conducted subject recruitment and testing. We also thank Thomas Wiecki for helpful comments on an earlier draft of this manuscript.

References

1. Elliott R, Newman JL, Longe OA, Deakin JF. Differential response patterns in the striatum and orbitofrontal cortex to financial reward in humans: a parametric functional magnetic resonance imaging study. *J Neurosci* 2003;23(1):303–307. [PubMed: 12514228]
2. Frank MJ, Claus ED. Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. *Psychol Rev* 2006;113(2):300–326. [PubMed: 16637763]
3. O'Doherty J, Kringelbach ML, Rolls ET, Hornak J, Andrews C. Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat Neurosci* 2001;4(1):95–102. [PubMed: 11135651]
4. Roesch MR, Olson CR. Neuronal activity related to reward value and motivation in primate frontal cortex. *Science* 2004;304(5668):307–310. [PubMed: 15073380]
5. Wallis JD. Orbitofrontal cortex and its contribution to decision-making. *Annu Rev Neurosci* 2007;30:31–56. [PubMed: 17417936]
6. Wallis JD, Miller EK. Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *Eur J Neurosci* 2003;18(7):2069–2081. [PubMed: 14622240]
7. Frank MJ. Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. *J Cogn Neurosci* 2005;17(1):51–72. [PubMed: 15701239]
8. Frank MJ, Seeberger LC, O'Reilly RC. By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* 2004;306(5703):1940–1943. [PubMed: 15528409]
9. Reynolds JN, Hyland BI, Wickens JR. A cellular mechanism of reward-related learning. *Nature* 2001;413(6851):67–70. [PubMed: 11544526]
10. Wickens JR, Begg AJ, Arbuthnott GW. Dopamine reverses the depression of rat corticostriatal synapses which normally follows high-frequency stimulation of cortex in vitro. *Neuroscience* 1996;70(1):1–5. [PubMed: 8848115]
11. Cohen MX, Frank MJ. Neurocomputational models of basal ganglia function in learning, memory and choice. *Behav Brain Res* 2009;199(1):141–156. [PubMed: 18950662]
12. Pavese N, Evans AH, Tai YF, Hotton G, Brooks DJ, Lees AJ, et al. Clinical correlates of levodopa-induced dopamine release in Parkinson disease: a PET study. *Neurology* 2006;67(9):1612–1617. [PubMed: 17101892]
13. Tedroff J, Pedersen M, Aquilonius SM, Hartvig P, Jacobsson G, Langstrom B. Levodopa-induced changes in synaptic dopamine in patients with Parkinson's disease as measured by [¹¹C]raclopride displacement and PET. *Neurology* 1996;46(5):1430–1436. [PubMed: 8628494]
14. Bodi N, Keri S, Nagy H, Moustafa A, Myers CE, Daw N, et al. Reward-learning and the novelty-seeking personality: a between- and within-subjects study of the effects of dopamine agonists on young Parkinson's patients. *Brain* 2009;132(Pt 9):2385–2395. [PubMed: 19416950]
15. Frank MJ, Moustafa AA, Haughey HM, Curran T, Hutchison KE. Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proc Natl Acad Sci U S A* 2007;104(41):16311–16316. [PubMed: 17913879]
16. Palminteri S, Lebreton M, Worbe Y, Grabli D, Hartmann A, Pessiglione M. Pharmacological modulation of subliminal learning in Parkinson's and Tourette's syndromes. *Proc Natl Acad Sci U S A* 2009;106(45):19179–19184. [PubMed: 19850878]
17. Voon V, Pessiglione M, Brezing C, Gallea C, Fernandez HH, Dolan RJ, et al. Mechanisms underlying dopamine-mediated reward bias in compulsive behaviors. *Neuron* 2010;65(1):135–142. [PubMed: 20152119]
18. Strauss GP, Robinson BM, Waltz JA, Frank MJ, Kasanova Z, Herbener ES, et al. Patients With Schizophrenia Demonstrate Inconsistent Preference Judgments for Affective and Nonaffective Stimuli. *Schizophr Bull.* 2010
19. Waltz JA, Frank MJ, Robinson BM, Gold JM. Selective reinforcement learning deficits in schizophrenia support predictions from computational models of striatal-cortical dysfunction. *Biol Psychiatry* 2007;62(7):756–764. [PubMed: 17300757]

20. Waltz JAFM, Wiecki TV, Gold JM. Altered probabilistic learning and response biases in schizophrenia: Behavioral evidence and neurocomputational modeling. *Neuropsychology*. In Press.
21. Chase HW, Clark L, Myers CE, Gluck MA, Sahakian BJ, Bullmore ET, et al. The role of the orbitofrontal cortex in human discrimination learning. *Neuropsychologia* 2008;46(5):1326–1337. [PubMed: 18242647]
22. Slifstein M, Kolachana B, Simpson EH, Tabares P, Cheng B, Duvall M, et al. COMT genotype predicts cortical-limbic D1 receptor availability measured with [¹¹C]NNC112 and PET. *Mol Psychiatry* 2008;13(8):821–827. [PubMed: 18317466]
23. Juckel G, Schlagenhauf F, Koslowski M, Wustenberg T, Villringer A, Knutson B, et al. Dysfunction of ventral striatal reward prediction in schizophrenia. *Neuroimage* 2006;29(2):409–416. [PubMed: 16139525]
24. Simon JJ, Biller A, Walther S, Roesch-Ely D, Stippich C, Weisbrod M, et al. Neural correlates of reward processing in schizophrenia--relationship to apathy and depression. *Schizophr Res* 2010;118(1–3):154–161. [PubMed: 20005675]
25. Waltz JA, Schweitzer JB, Gold JM, Kurup PK, Ross TJ, Salmeron BJ, et al. Patients with schizophrenia have a reduced neural response to both unpredictable and predictable primary reinforcers. *Neuropsychopharmacology* 2009;34(6):1567–1577. [PubMed: 19052540]
26. Moran PM, Owen L, Crookes AE, Al-Uzri MM, Revely MA. Abnormal prediction error is associated with negative and depressive symptoms in schizophrenia. *Prog Neuropsychopharmacol Biol Psychiatry* 2008;32(1):116–123. [PubMed: 17764799]
27. Moustafa AA, Cohen MX, Sherman SJ, Frank MJ. A role for dopamine in temporal decision making and reward maximization in parkinsonism. *J Neurosci* 2008;28(47):12294–12304. [PubMed: 19020023]
28. Abi-Dargham A, Mawlawi O, Lombardo I, Gil R, Martinez D, Huang Y, et al. Prefrontal dopamine D1 receptors and working memory in schizophrenia. *J Neurosci* 2002;22(9):3708–3719. [PubMed: 11978847]
29. Abi-Dargham A, Moore H. Prefrontal DA transmission at D1 receptors and the pathology of schizophrenia. *Neuroscientist* 2003;9(5):404–416. [PubMed: 14580124]
30. Weinberger DR. Implications of normal brain development for the pathogenesis of schizophrenia. *Arch Gen Psychiatry* 1987;44(7):660–669. [PubMed: 3606332]
31. Wiecki TV, Riedinger K, von Ameln-Mayerhofer A, Schmidt WJ, Frank MJ. A neurocomputational account of catalepsy sensitization induced by D2 receptor blockade in rats: context dependency, extinction, and renewal. *Psychopharmacology (Berl)* 2009;204(2):265–277. [PubMed: 19169674]
32. Farkas M, Polgar P, Kelemen O, Rethelyi J, Bitter I, Myers CE. Associative learning in deficit and nondeficit schizophrenia. *Neuroreport* 2008;19(1):55–58. [PubMed: 18281892]
33. Polgar P, Farkas M, Nagy O, Kelemen O, Rethelyi J, Bitter I, et al. 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* 2008;99(1–3):200–207. [PubMed: 17693060]
34. Cohen JD, McClure SM, Yu AJ. Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philos Trans R Soc Lond B Biol Sci* 2007;362(1481):933–942. [PubMed: 17395573]
35. Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ. Cortical substrates for exploratory decisions in humans. *Nature* 2006;441(7095):876–879. [PubMed: 16778890]
36. Frank MJ, Doll BB, Oas-Terpstra J, Moreno F. Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. *Nat Neurosci* 2009;12(8):1062–1068. [PubMed: 19620978]
37. First, MB.; Spitzer, RL.; Gibbon, M.; Williams, JBW. Structured Clinical Interview for DSM-IV-TR Axis I Disorders- Patient Edition (SCID-I/P 2/2001 Revision). New York: Biometrics Research Department, New York State Psychiatric Institute; 2001.
38. Andreasen NC. The Scale for the Assessment of Negative Symptoms (SANS): conceptual and theoretical foundations. *Br J Psychiatry Suppl* 1989;(7):49–58. [PubMed: 2695141]

39. Buchanan RW, Javitt DC, Marder SR, Schooler NR, Gold JM, McMahon RP, et al. The Cognitive and Negative Symptoms in Schizophrenia Trial (CONSIST): the efficacy of glutamatergic agents for negative symptoms and cognitive impairments. *Am J Psychiatry* 2007;164(10):1593–1602. [PubMed: 17898352]
40. Woods SW. Chlorpromazine equivalent doses for the newer atypical antipsychotics. *J Clin Psychiatry* 2003;64(6):663–667. [PubMed: 12823080]
41. Overall JEGD. The Brief Psychiatric Rating Scale. *Psychol Rep* 1962;10:799–812.
42. Berridge KC. The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology (Berl)* 2007;191(3):391–431. [PubMed: 17072591]
43. Niv Y, Rivlin-Etzion M. Parkinson's disease: fighting the will? *J Neurosci* 2007;27(44):11777–11779. [PubMed: 17978012]
44. Dalley JW, Laane K, Theobald DE, Armstrong HC, Corlett PR, Chudasama Y, et al. Time-limited modulation of appetitive Pavlovian memory by D1 and NMDA receptors in the nucleus accumbens. *Proc Natl Acad Sci U S A* 2005;102(17):6189–6194. [PubMed: 15833811]
45. Everitt BJ, Robbins TW. Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat Neurosci* 2005;8(11):1481–1489. [PubMed: 16251991]
46. Nakamura K, Hikosaka O. Role of dopamine in the primate caudate nucleus in reward modulation of saccades. *J Neurosci* 2006;26(20):5360–5369. [PubMed: 16707788]
47. Braver TS, Barch DM, Cohen JD. Cognition and control in schizophrenia: a computational model of dopamine and prefrontal function. *Biol Psychiatry* 1999;46(3):312–328. [PubMed: 10435197]
48. Smith AJ, Li M, Becker S, Kapur S. Linking animal models of psychosis to computational models of dopamine function. *Neuropsychopharmacology* 2007;32(1):54–66. [PubMed: 16710321]
49. Frank MJ, Santamaria A, O'Reilly RC, Willcutt E. Testing computational models of dopamine and noradrenaline dysfunction in attention deficit/hyperactivity disorder. *Neuropsychopharmacology* 2007;32(7):1583–1599. [PubMed: 17164816]
50. Murray GK, Corlett PR, Clark L, Pessiglione M, Blackwell AD, Honey G, et al. Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. *Mol Psychiatry* 2008;13(3):239–67–76. [PubMed: 17684497]
51. Huys QJ, Dayan P. A Bayesian formulation of behavioral control. *Cognition* 2009;113(3):314–328. [PubMed: 19285311]
52. Kasanova Z, Waltz JA, Strauss GP, Frank MJ, Gold JM. Optimizing vs. Matching: Response Strategy in a Probabilistic Learning Task is associated with Negative Symptoms of Schizophrenia. under review.
53. Rolls ET, Loh M, Deco G, Winterer G. Computational models of schizophrenia and dopamine modulation in the prefrontal cortex. *Nat Rev Neurosci* 2008;9(9):696–709. [PubMed: 18714326]
54. Dreher JC, Kohn P, Kolachana B, Weinberger DR, Berman KF. Variation in dopamine genes influences responsivity of the human reward system. *Proc Natl Acad Sci U S A* 2009;106(2):617–622. [PubMed: 19104049]
55. Egan MF, Goldberg TE, Kolachana BS, Callicott JH, Mazzanti CM, Straub RE, et al. Effect of COMT Val108/158 Met genotype on frontal lobe function and risk for schizophrenia. *Proc Natl Acad Sci U S A* 2001;98(12):6917–6922. [PubMed: 11381111]
56. Long. in preparation.
57. Boorman ED, Behrens TE, Woolrich MW, Rushworth MF. How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. *Neuron* 2009;62(5):733–743. [PubMed: 19524531]
58. Cohen AS, Minor KS. Emotional experience in patients with schizophrenia revisited: meta-analysis of laboratory studies. *Schizophr Bull* 2010;36(1):143–150. [PubMed: 18562345]
59. Kring AM, Moran EK. Emotional response deficits in schizophrenia: insights from affective science. *Schizophr Bull* 2008;34(5):819–834. [PubMed: 18579556]
60. Kirkpatrick B, Strauss GP, Nguyen L, Fischer BA, Daniel DG, Cienfuegos A, et al. The Brief Negative Symptom Scale: Psychometric Properties. *Schizophr Bull*. 2010

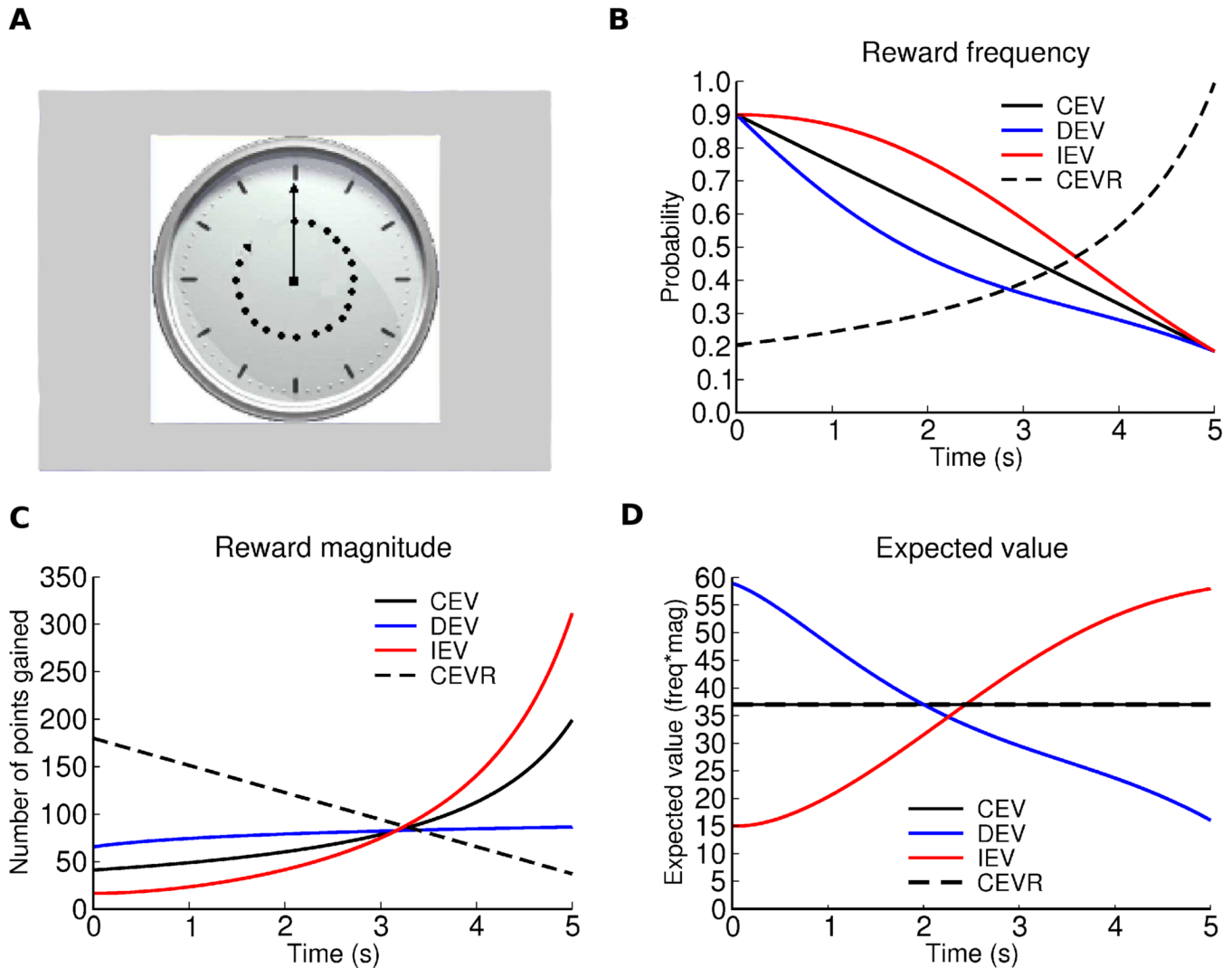


Figure 1. Depiction of Task Conditions

Task conditions: DEV, CEV, IEV, and CEVR. The x axis in all plots corresponds to the time after onset of the clock stimulus at which the response is made. The equations are designed such that the expected value in the beginning in DEV is approximately equal to that at the end in IEV so that if optimal, subjects should obtain the same average reward in both IEV and DEV. **a)** Example clock-face stimulus. **b)** Probability of reward occurring as a function of response time; **c)** Reward magnitude (contingent on a)); **d)** Expected value across trials for each time point. Note that CEV and CEVR have the same EV so the black line represents EV for both conditions.

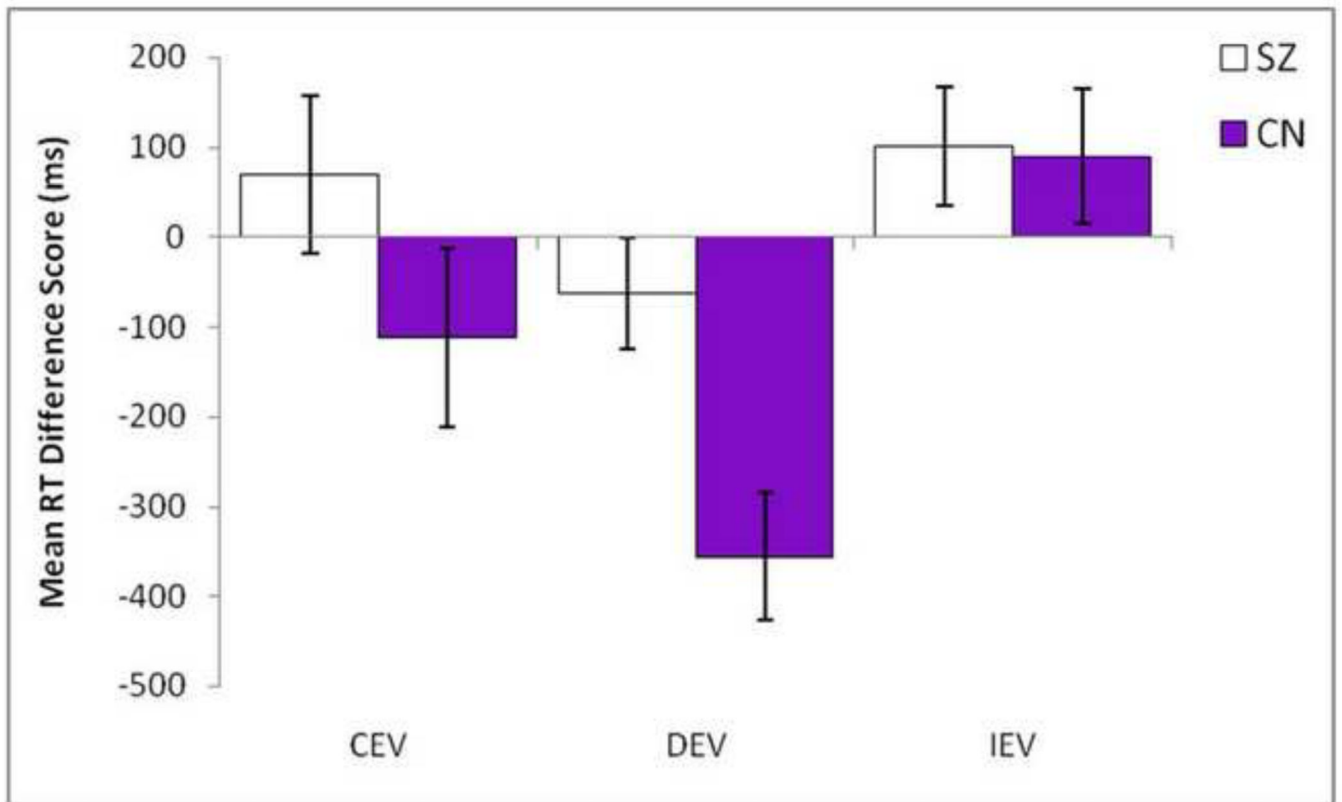
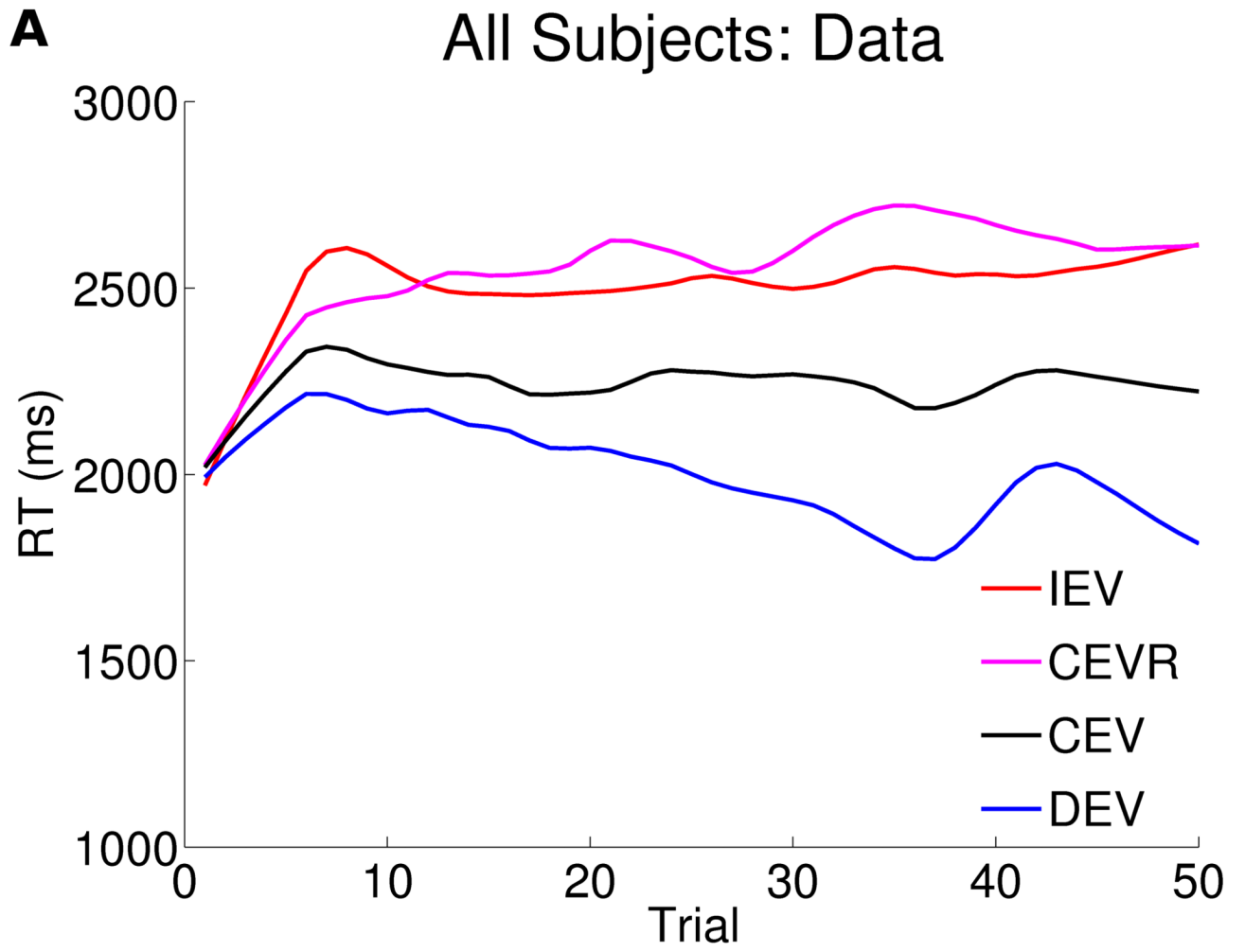


Figure 2. Mean RT Difference Score from First Half of Trials to Second Half of Trials for CEV, DEV, and IEV conditions in Schizophrenia Patients and Controls

SZ = White bars; CN = Purple bars; CEV = Constant expected value; DEV = Decreasing expected value (i.e., Go Learning); IEV = Increasing expected value (i.e., No Go Learning). Values reflect the mean RT change from beginning to end of the block for each condition (i.e., the relative learning within that condition). More negative values reflect that subjects learn to speed responses (Go), and more positive values reflect learning to slow down to obtain rewards within that condition (NoGo).



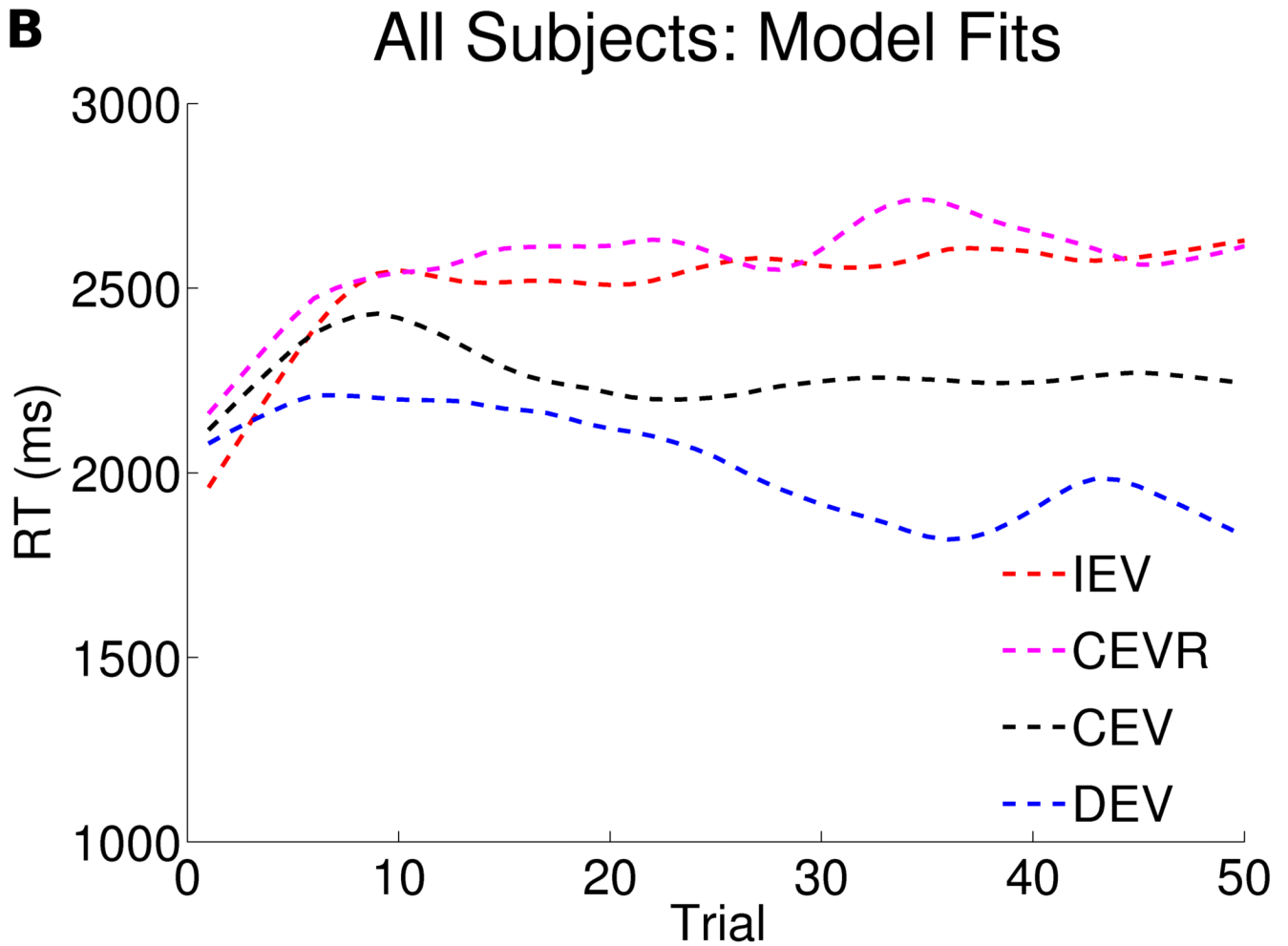


Figure 3. Response times as a function of trial in all 90 subjects (panel a) and computational model fits (panel b)

The figures depict response times as a function of trial number, smoothed (with weighted linear least squares fit) over a 10 trial window in **a)** all 90 participants, **b)** computational model fits. Overall, relative to baseline CEV response times, participants speed up in DEV and slow down in IEV. CEVR response times are also slowed in both data and model due to a high frequency of negative prediction errors for early responses (see Moustafa et al, 2008 and Frank et al, 2009).

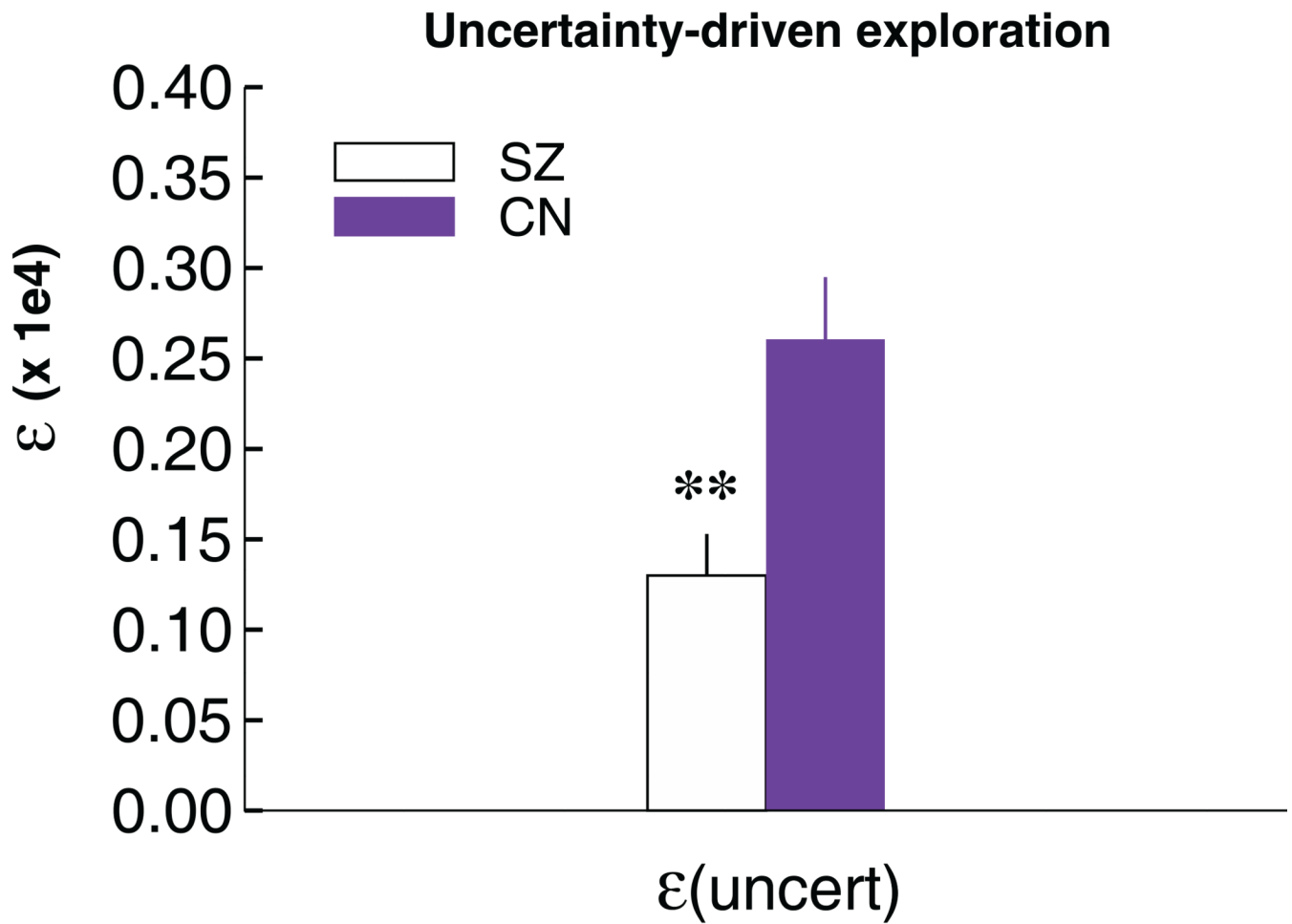


Figure 4. Uncertainty-driven exploration in individuals with schizophrenia and controls
The explore parameter estimated from the model is reduced in schizophrenia ($p < .01$).

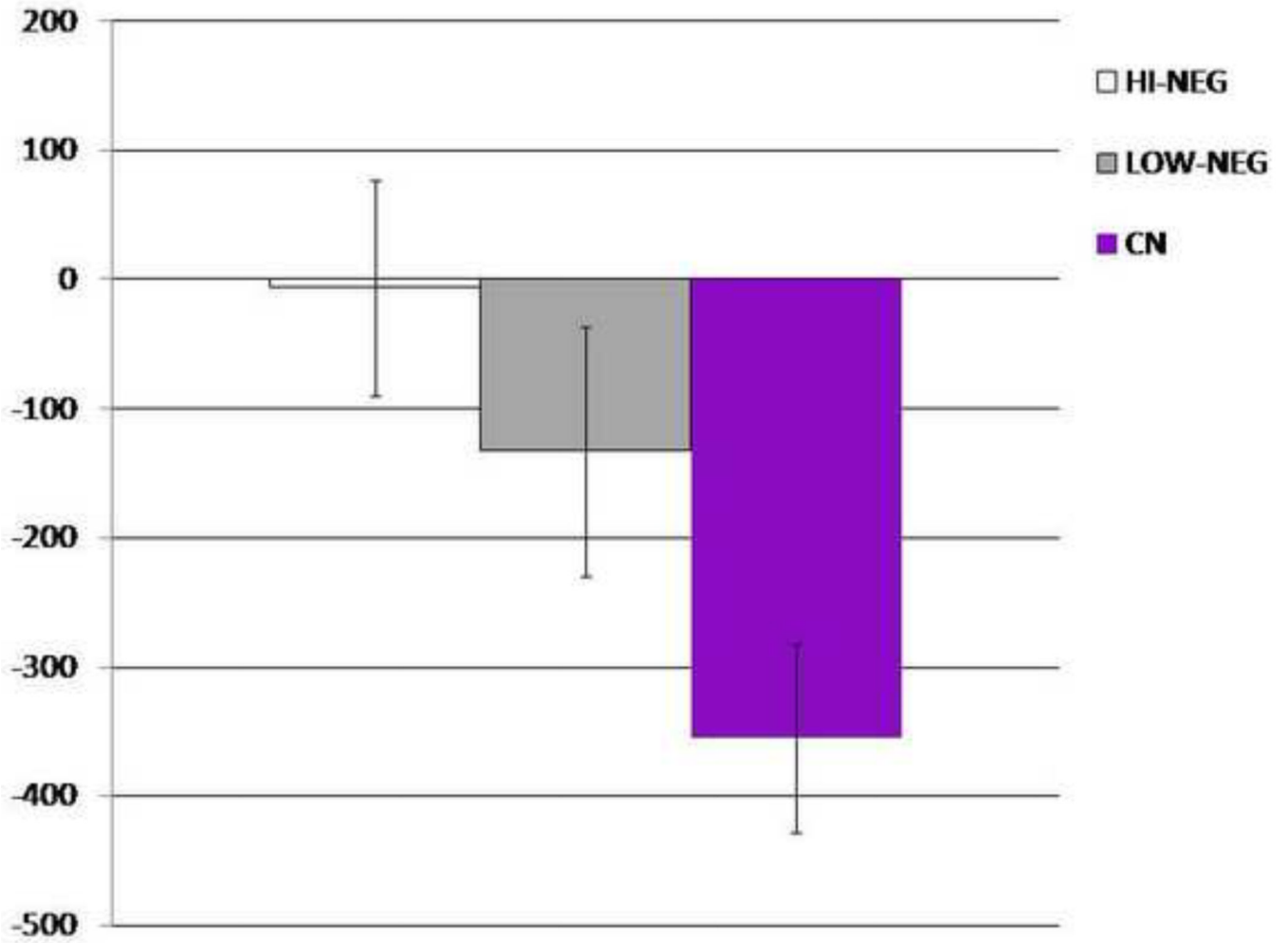


Figure 5. Go Learning HI-NEG, LOW-NEG, and CN Subjects
Mean RT change from beginning to end of block for the DEV condition (Go Learning). More negative values reflect better Go Learning (i.e., learning to speed up from then first half of trials to the second half of trials).

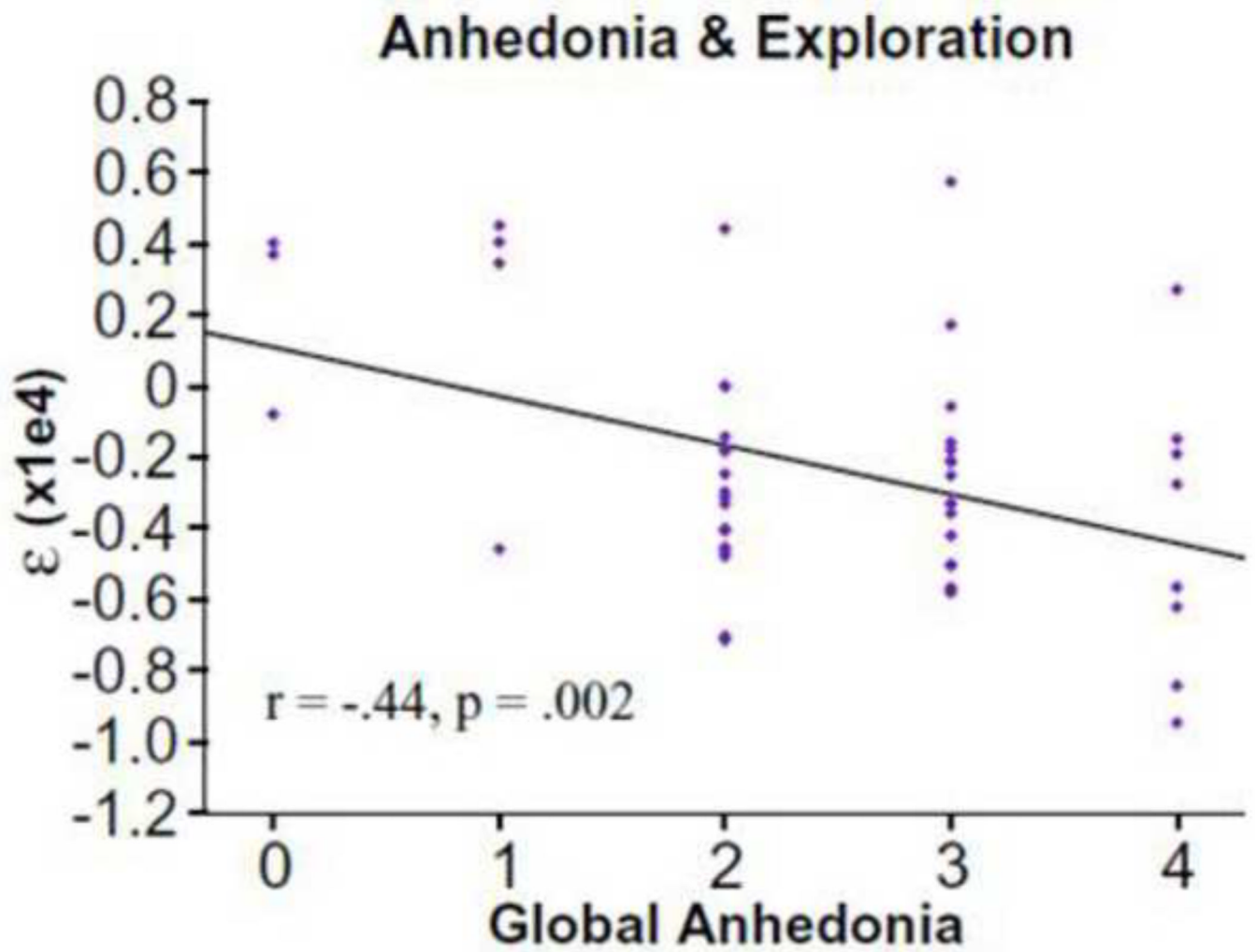


Figure 6. Uncertainty-driven exploration (ϵ parameter) as a function of anhedonia. Left: scatter plot across all patients. Right: means for each level of anhedonia. Error bars reflect s.e.m.

Table 1

Demographic and Clinical Characteristics of Patients and Controls.

	SZ (n = 51)	CN (n = 39)	<i>p</i> -value
Age	42.77 (10.08)	43.49 (10.68)	<i>p</i> = 0.74
Education	12.80 (2.27)	14.89 (2.06)	<i>p</i> < 0.001
Parental Education	13.56 (3.07)	13.26 (2.26)	<i>p</i> = 0.67
WASI Estimated Full-scale IQ	96.86 (13.71)	113.13 (12.01)	<i>p</i> < 0.001
% Male	72.5%	66.7%	<i>p</i> = 0.35
Ethnicity			<i>p</i> = 0.35
Caucasian	56.9%	64.1%	
African-American	35.3%	35.9%	
Asian	02.0%	00.0%	
Other	00.0%	05.9%	
	HI-NEG (n = 25)	LOW-NEG (n = 24)	
Age*	41.29 (10.76)	45.47 (8.27)	<i>p</i> = 0.36
Education*	13.04 (2.55)	12.62 (2.06)	<i>p</i> < 0.001
Parental Education*	13.29 (10.76)	13.68 (8.27)	<i>p</i> = 0.82
WASI Estimated Full-scale IQ*	96.92 (15.39)	96.08 (11.93)	<i>p</i> < 0.001
% Male*	75%	64%	<i>p</i> = 0.71
Ethnicity*			<i>p</i> = 0.08
Caucasian	60.0%	54.2%	
African-American	28.0%	41.7%	
Asian	04.0%	0.00%	
Other	08.0%	04.2%	
Antipsychotic Medications			
% Conventional	04%	04%	
% Atypical	100%	96%	
Clozapine	54%	56%	
CPZ Equivalent Dosage	578 (394)	462 (371)	<i>p</i> = 0.24
BPRS Symptoms			
Negative	2.28 (0.84)	1.38 (0.48)	<i>p</i> < 0.001
Positive	2.50 (1.22)	2.48 (0.99)	<i>p</i> = 0.77
Disorganized	1.45 (0.71)	1.29 (0.47)	<i>p</i> = 0.35
Total	38.37 (9.30)	34.4 (6.70)	<i>p</i> = 0.09

Note.

* Analyses conducted on HI-NEG, LOW-NEG, and CN Groups. The most frequently used medication was clozapine, used alone (n = 21), in conjunction with risperidone (n = 6), or in conjunction with aripiprazole (n = 1). Risperidone prescribed alone (n = 6), or in conjunction with olanzapine (n = 2) was the second most frequently prescribed antipsychotic. Patients were also prescribed olanzapine (n = 8), fluphenazine (n = 1), ziprasidone (n = 2), quetiapine (n = 3). One patient was prescribed haloperidol in conjunction with quetiapine. Mean BPRS scores indicated that

patients experienced a moderate level of symptom severity at the time of testing: Total ($M = 36.34, SD = 8.26$); Positive ($M = 2.42, SD = 1.10$); Negative ($M = 1.73, SD = 0.71$); Disorganized ($M = 1.41, SD = 0.51$).

Table 2

Reinforcement Learning Domains Assessed by Computational Modeling Parameters

Modeling Parameter	Description
K	Estimates baseline motor response tendency independent of other factors.
λ	A response recency parameter scaling the impact of the previous response's RT on the current choice, independent of any change in value.
ϵ	Predicts trial-by-trial RT swings to occur when one is relatively more uncertain about the reward probabilities for fast or slow responses. Thus, with sufficiently high values, RT swings are predicted to occur in the direction of greater uncertainty about the likelihood that outcomes might be better than the status quo.
αG	The degree to which individuals speed RTs as a function of positive prediction errors.
αN	The degree to which individuals slow RTs as a function of negative prediction errors.
ρ	Predicts the extent to which individuals adjust RTs in the direction of greater probability of obtaining a positive outcome based on the observed reward statistics.
v	A "going for gold" parameter, which predicts that participants will adjust RT's toward that which has produced the single largest reward experienced thus far.

Table 3

Best-fitting model parameters for patients and controls.

K	λ	αG	αN	ρ	ν	ε	
SZ	1532 (67)	0.31 (.02)	0.17 (.03) *	0.31 (.05)	583 (73)	0.12 (.02)	1306 (228) **
NC	1558 (80)	0.34 (.03)	0.26 (.04) *	0.29 (.06)	580 (72)	0.12 (.01)	2593 (351) **

Note. Values reflect mean (SE).

** $p < 0.01$;

* $p = 0.07$