Title

Using Computational Modeling to Capture Schizophrenia-Specific Reinforcement Learning Differences and Their Implications on Patient Classification

Journal

Biological Psychiatry Cognitive Neuroscience and Neuroimaging, 7(10)

ISSN

2451-9022

Authors

Geana, Andra

Barch, Deanna M

Gold, James M

et al.

Publication Date

2022-10-01

DOI

10.1016/j.bpsc.2021.03.017

Peer reviewed

HHS Public Access

Author manuscript

Biol Psychiatry Cogn Neurosci Neuroimaging. Author manuscript; available in PMC 2023 October 01.

Published in final edited form as:

Biol Psychiatry Cogn Neurosci Neuroimaging. 2022 October ; 7(10): 1035–1046. doi:10.1016/j.bpsc.2021.03.017.

Using Computational Modeling to Capture Schizophrenia-Specific Reinforcement Learning Differences and Their Implications on Patient Classification

Andra Geana,

Deanna M. Barch,

James M. Gold,

Cameron S. Carter,

Angus W. MacDonald III,

J. Daniel Ragland,

Steven M. Silverstein,

Michael J. Frank

Department of Cognitive, Linguistic, and Psychological Sciences (AG, MJF), Brown University, Providence, Rhode Island; Department of Psychological & Brain Sciences (DMB), Washington University in St. Louis; Department of Radiology (DMB), Mallinckrodt Institute of Radiology, and Department of Psychiatry (DMB), Washington University School of Medicine, St. Louis, Missouri; Department of Psychiatry (JMG), Maryland Psychiatric Research Center, Baltimore, Maryland; Department of Psychiatry (CSC, JDR), University of California, Davis, California; Department of Psychology (AWM), University of Minnesota, Minneapolis, Minnesota; and Department of Psychiatry (SMS), University of Rochester, Rochester, New York.

Abstract

BACKGROUND.—Psychiatric diagnosis and treatment have historically taken a symptom-based approach, with less attention on identifying underlying symptom-producing mechanisms. Recent efforts have illuminated the extent to which different underlying circuitry can produce phenotypically similar symptomatology (e.g., psychosis in bipolar disorder vs. schizophrenia). Computational modeling makes it possible to identify and mathematically differentiate behaviorally unobservable, specific reinforcement learning differences in patients with schizophrenia versus other disorders, likely owing to a higher reliance on prediction error–driven learning associated with basal ganglia and underreliance on explicit value representations associated with orbitofrontal cortex.

METHODS.—We used a well-established probabilistic reinforcement learning task to replicate those findings in individuals with schizophrenia both on (n = 120) and off (n = 44) antipsychotic medications and included a patient comparison group of bipolar patients with psychosis (n = 60) and healthy control subjects (n = 72).

Address correspondence to Andra Geana, Ph.D., at andra_geana@brown.edu.

Supplementary material cited in this article is available online at https://doi.org/10.1016/j.bpsc.2021.03.017.

RESULTS.—Using accuracy, there was a main effect of group ($F_{3,279} = 7.87$, $p < .001$), such that all patient groups were less accurate than control subjects. Using computationally derived parameters, both medicated and unmediated individuals with schizophrenia, but not patients with bipolar disorder, demonstrated a reduced mixing parameter ($F_{3,295} = 13.91$, $p < .001$), indicating less dependence on learning explicit value representations as well as greater learning decay between training and test ($F_{1,289} = 12.81$, $p < .001$). Unmedicated patients with schizophrenia also showed greater decision noise ($F_{3,295} = 2.67$, $p = .04$).

CONCLUSIONS.—Both medicated and unmedicated patients showed overreliance on prediction error–driven learning as well as significantly higher noise and value-related memory decay, compared with the healthy control subjects and the patients with bipolar disorder. Additionally, the computational model parameters capturing these processes can significantly improve patient/control classification, potentially providing useful diagnosis insight.

Schizophrenia (SZ) is a complex disorder comprising different classes of symptoms (positive: delusions, hallucinations; negative: lack of motivation, anhedonia) and associated cognitive deficits in learning, memory, and decision making (1–3). The heterogeneity in clinical presentation is so marked that the disease has been proposed to be not a singular disorder but a constellation of disorders sharing phenotypic features (4). The specific mechanisms underlying different symptom classes are as yet unclear, and the complexity of the disease leaves etiology, diagnosis, and outcome measurements challenging—particularly given the symptom overlap with other disorders, such as major depressive disorder [also characterized by motivation deficits (5)] or bipolar disorder (which can express with psychosis). Further, it is often difficult to differentiate behavioral deficits across disorders on a variety of cognitive tasks (6,7).

Recently, computational psychiatry has emerged as a bridge between clinical and theory-driven approaches, using computational methods to classify and uncover patterns not observable from descriptive statistics and other conventional analysis approaches (8). For example, machine-learning algorithms have been tested on their ability to classify patients with SZ from healthy control subjects (HCs) (9). Computational modeling can identify distinct underlying mechanisms that produce similar observable behaviors [for instance, distinguishing among sources of decision noise—motor noise, random exploration, or directed exploration—that produce similar error responses but rely on different processes (10)], making it a useful tool for distinguishing subtle differences in the causes of or the way symptoms and deficits interact in SZ compared with other disorders.

Ample empirical and computational work on schizophrenia has focused dopamine circuitry abnormalities as a key mechanism underlying the disorder. Dopamine has been linked to both positive and negative SZ symptoms (11), with irregular dopamine release hypothesized to ascribe undue salience to irrelevant stimuli (leading, for instance, to the formation of delusions). Disrupted dopamine function has also been hypothesized to prevent appropriate learning from reward feedback, potentially causing negative symptoms (12). Although findings on specific domains (gains/losses) and impairment severity have been mixed (13), convincing evidence shows decreased ability for rapid behavioral adjustments [for instance, in extradimensional/intradimensional set-shifting tasks (14)] as well as impaired explicit

value representations (explicit knowledge about specific benefits, such as monetary reward, of different choice options) to guide decision making (15,16). Performance impairments in gradual, trial-by-trial learning have been more subtle, with evidence for spared prediction error (PE)–based learning (17) but differences in the ability to use PEs in positive-feedback (gain) and negative-feedback (loss) domains [an asymmetry specifically linked to high negative symptoms (18)].

Recent work explains these deficits as dysfunctions in frontostriatal circuitry, specifically in the interaction of two systems involved in value representation and reinforcement learning (RL): the orbitofrontal cortex (OFC), involved in mapping expected values and updating representations in response to feedback, and the basal ganglia (BG), involved in PE computations (13,16). This interaction dictates the amount to which decisions rely on explicit value representations (e.g., context-independent representations of the estimated reward associated with a stimulus) versus PE-based representations (e.g., context-dependent estimations of which stimulus provides more reward).

Using a probabilistic reward task and a computational model that allowed differentiating (OFC-driven) stimulus value representation from (BG-driven) PE-based action selection, Gold *et al.* (3) showed that SZ patients with high negative symptoms relied more heavily than HCs and patients with low negative symptoms on learning which actions avoided punishment but had difficulty representing positive expected values to guide novel choices. However, all patients in the study were on antipsychotic medication, leaving open the question of how/whether treatment [known to impact trial-by-trial learning via D2 blockade (19)] might shift the learning strategy.

The present study aimed to replicate these findings and test whether they hold in unmedicated patients and to examine to what extent medication impacts not only performance but also the underlying learning strategies. In the previous work above, RL-related effects were linked largely to negative symptoms [potentially because of motivational/reward sensitivity effects, as well as the fact that the standard antipsychotic medication asymmetrically impacts positive more than negative symptoms (20)]. We therefore also focused analyses on the negative symptom spectrum. We expected to find similar effects to Gold *et al.* (3) on deficits in explicit value representation, which we believe is as SZ-specific impairment driven by abnormalities in the interaction of the OFC-based and BG-based computations. To test this, we recruited a sample of patients with bipolar disorder (BI), who share some symptoms and cognitive deficits of SZ patients [such as impaired behavioral adjustments following changes in outcome contingencies (21) believed to also rely on reward processing abnormalities] but do not show similar RL impairments (22).

We employed computational modeling to gain further insight into underlying RL strategies and impairments and uncovered individual-specific parameters coding for learning, value representation, reward sensitivity, and decision noise. To test whether model-based differences reliably inform SZ-specific deficits, we compared how model parameters affected classification (compared with model-free measures) between patients and HCs and between different patients groups.

METHODS AND MATERIALS

Participants and Clinical Assessment

Participants for the study were recruited as part of the Cognitive Neuroscience Test Reliability And Clinical Applications for Schizophrenia (CNTRACS) Consortium, which included five research sites (see the Supplement for details). Recruiting and informed consent procedures for each site were approved by each site's institutional review board.

A masters-level clinician conducted or supervised diagnostic assessments that included the Structured Clinical Interview for DSM-IV-TR (23), 24-item Brief Psychiatric Rating Scale (24–27), Young Mania Rating Scale (28), Bipolar Depression Rating Scale (29), and Clinical Assessment Interview for Negative Symptoms (30). Analyses examining relationships to negative symptom severity used the motivation and pleasure subscale of Clinical Assessment Interview for Negative Symptoms. Analyses examining relationships to community function used the participant and informant versions of the Specific Levels of Functioning Scale (31). The groups were recruited to be as similar as possible on sex, age, race, and parental socioeconomic status, measured using the Hollingshead Index (32) as updated using occupational prestige ratings based on the 1989 general social survey (33). See the Supplement for details on exclusion criteria.

Procedural Task

We used the same probabilistic learning task as Gold *et al.* (3). Participants learned associations between eight different stimuli, *a*–*h* (natural landscapes) (Figure 1), and their associated monetary value.

For the learning phase (160 trials), participants were always presented the same pairings (*ab*, *cd*, *ef*, *gh*), in randomized order. For the transfer phase (72 trials), in addition to learned pairings, participants were shown novel pairings (e.g., *ae*). This allowed us to test to what extent individuals learned the specific value associated with a stimulus (e.g., *a*: $0.045, *e*: −$0.045), as opposed to simply learning the context-dependent action strategy (e.g., never choose *b* because *b* was always paired with *a*, and *a* was better).

Computational Model

We fit several RL models to participants' data, varying model structure (Q-learning vs. actor-critic vs. mixed strategy) (see the Supplement for details). The best-fitting model was a variant of the hybrid model of Gold *et al.* (3), which assumed a mixed strategy in which "pure" basal ganglia–dependent learning (in the form of an actor-critic algorithm) is complemented by top-down representation of expected reward value for each choice (via Q-learning architecture). The present model included further updates to enhance fit and dissociate potential sources of variance in patient populations.

Actor-Critic Architecture.—The actor-critic component of the model computes expected values for each state and updates them at each time step based on observed PEs.

In this architecture, the state refers to the subset of stimuli observed on a specific trial. In the learning phase, there are four possible states, one for each stimulus pair. *ab* (state 1), *cd* (state 2), *ef* (state 3), and *gh* (state 4). Each state is assigned the same initial value.

That value is updated by the critic component each time the participant observes that state, based on the reward and subsequent PE, as

$$V_{s,t+1} = V_{s,t} + \alpha_C * PE_t \quad (1)$$

where α_C is the critic learning rate determining how much the most recent observed outcome contributes to update the current estimate of state value (higher α_C translates to larger updates), and PE_t is the PE computed as $PE_t = outcome_t - V_{s,t}$.

This model chooses based on actor weights $w_{s,a,t}$, which quantify the propensity toward an action under a specific state. There are two available actions in this scenario—choose stimulus 1 or choose stimulus 2—but their weights change depending on state (analogous, for instance, to the tendency to turn left or right depending on whether one is at one street corner or another).

When a given action leads to a positive outcome, the corresponding weight is augmented in proportion to the positive PE it generates, scaled by an actor learning rate α_A:

$$W_{s,a,t+1} = W_{s,a,t} + \alpha_A * PE_t \quad (2)$$

Thus, the actor selects actions based only on their relative winner status in the choice context (e.g., if stimulus *a* loses 10 points and *b* loses 2 points, *b* is the relative winner; similarly, if *c* rewards 2 points and *d* rewards 4 points, *d* is the winner. *b* and *d* have similar status in this type of strategy). Winner status is updated via reward prediction errors and not directly based on value. As per Gold *et al.* (3), actor weights are normalized on each trial, so their relative scale matches that of Q values (otherwise, actor weights grow without bound).

Q-learning Architecture.—The Q-learning component learns the expected reward value of each state-action pair and chooses based on these predicted rewards. Critically, as highlighted in (3,34), this strategy allows for differentiation between rewards that were generated from truly positive outcomes (in the gain conditions) versus those that merely resulted from avoidance of negative outcomes (in the loss conditions).

Q values update with each trial, based on a prediction error scaled by a learning rate:

$$Q_{a,t+1} = Q_{a,t} + \alpha_Q \times (outcome_t - Q_{a,t}) \quad (3)$$

Q values were initialized at zero (see the Supplement for details on model with nonzero initial values).

Mixed Strategy.—Gold *et al.* (3) and Hernaus *et al.* (34) reported that performance was best described by a mixture of actor-critic and Q-learning strategies. The actor weights and the Q-learning values are mixed into a hybrid value, scaled by a mixing factor c:

$$H_{s,a,t} = w_{s,a,t} \times (1 - c) + Q_{a,t} \times c \quad (4)$$

This mixing factor is a quantitative indicator of the extent to which a participant relies on explicit value representations or simpler, PE-based computations. Higher values for c indicate a strategy more consistent with explicit representations (such as Q-learning); lower values indicate a strategy closer to action-value representations (such as actor-critic); critically, these strategies make different predictions under certain choice contexts.

In the present work, we included two additional parameters that accounted for variance independently from the relative contributions of actor-critic versus Q-learning. First, we included an irreducible-noise/epsilon parameter, ε, that accounts for overall response variability thought to be due to attentional lapses. This ε, generally part of the ε-greedy choice function in RL, implements the degree to which choice is reward-maximizing versus random between all available options and can be considered a random exploration parameter (e.g., choosing among the options that are not highest-reward, with equal probability). It can also be used to capture a proportion of trials that are not well explained by the model (for instance, due to lapses in attention). That is how we use it in the present model, in which the softmax parameter captures exploration; previous work (35) has shown an ε-softmax mix allows for better estimates of other model parameters when there are outlier choices.

Choice probabilities were computed via ε-softmax.

$$P_{a,t} = (1 - \varepsilon)(Softmax_{a,t}) + \varepsilon \times u_t \quad (5)$$

where $Softmax_{a,t} = \frac{e^{B * H_{s,a,t}}}{\sum_a e^{B * H_{s,a,t}}}$, where temperature B determines how much value impacts choice probability, noise parameter ε determines value-based versus random choice, and u_t is a uniform distribution across all available actions.

Second, we included a decay parameter quantifying forgetting of learned values between the training and test phases. Decayed values were adjusted as follows:

$$Q_{fin}(all) = Q_{fin}(all) \times (1 - d) + Q_{init}(all) \times d \quad (6)$$

where Q_{fin} represents the final learned values, Q_{init} the original (uniform) value priors. Thus, between training and testing, values could decay back toward the original prior, allowing more recently seen stimuli a more accurate (updated) value representation, whereas stimuli seen more trials ago get noisier representations. This has been found to better account for memory and attentional effects (36).

Fitting Procedures and Model Comparison.—All models were fit using standard maximum likelihood procedures (see the Supplement for details on fitting procedure implementation and model comparison details).

Data Analysis

We analyzed accuracy data (defined as choosing the more rewarding stimulus) from training and test phases. Participants were split into groups based on clinical diagnosis and medication status (HCs, BI, medicated SZ [SZON], and unmedicated SZ [SZOFF]), and task conditions were defined as gains (all stimuli associated with positive monetary rewards) or losses (all stimuli associated with monetary loss). Analyses of variance (ANOVAs) were conducted using SPSS Statistics (version 25; IBM Corp., Armonk, NY). Post hoc tests were conducted using Tukey's honestly significant difference test.

RESULTS

There were 72 HCs, 60 individuals with BI with psychosis, 120 with SZ/schizoaffective disorder who were taking antipsychotic medications (SZON), and 44 with SZ/schizoaffective disorder who had not taken antipsychotic medications for at least 1 month (SZOFF). Demographics and clinical characteristics for each group are presented in Table S1.

Learning

All participants learned to perform above chance (Figure 2), with a significant main effect of time on accuracy (moving time window of 20 trials [8] × group [4] × condition [4] mixed ANOVA; Hyunh-Feldt correction to account for sphericity violation; $F_{5.1,1995} = 117.86$, $p < .001$).

Although it is possible that some participants in all groups were still learning (see the Supplement for details on late-stage learning), the HCs reached higher accuracies on average by the end of training than all patients groups ($F_{3,279} = 7.87$, $p < .001$); no significant interaction between time step and group was observed ($F_{33,3069} = 1.16$, $p = .238$).

Test Phase

The test phase presented participants with both old pairings from the learning phase and novel stimulus pairings, requiring representation of absolute (not just relative, state-dependent) stimulus value to make a correct choice. As previously found (3), the novel pairing, which required value comparison between the two best stimuli in different gains/losses conditions (i.e., frequent winner [FW] associated with high chance of reward, vs. frequent loss avoider [FLA] associated with high chance of loss avoidance), proved more difficult than within-domain pairings (FW vs. infrequent winner [IW]) or the cross-domain FW versus frequent loser (FL) pairing.

We compared the FW-FLA condition with both FW-FL and FW-IW. Mixed group (4) by condition (2) ANOVA on average test phase accuracy revealed a significant main effect of condition in both cases (FW-FLA vs. FW-FL, $F_{1,293} = 119.89$, $p < .001$; FW-FLA vs. FW-IW, $F_{1,293} = 69.58$, $p < .001$), with average accuracy higher in FW-FL and FW-IW conditions than in FW-FLA ($M_{FWvFLA} = 0.62$, $SD_{FWvFLA} = 0.254$, $M_{FWvIW} = 0.831$, $SD_{FWvIW} = 0.264$, $M_{FWvFL} = 0.845$, $SD_{FWvFL} = 0.249$) (Figure 3A). There was no main effect of condition between FW-FL and the FW-IW ($F_{1,293} = 1.96$, $p = .163$).

The FW-FLA condition also showed a significant main effect of group, with controls showing significantly higher accuracy than the SZOFF group (M_{HC} = 0.645, SD_{HC} = 0.30, M_{SZOFF} = 0.545, SD_{SZOFF} = 0.29, $F_{3,296}$ = 2.85, p = .037).

We also observed differences on the learned (old) pairings, with a significant main effect of group (condition [2] by group [4] repeated measures ANOVA, $F_{1,287}$ = 6084.43, p < .001), with Tukey's test revealing that HCs had significantly higher accuracy than all patient groups (M_{HC-BI} = 0.0701, p = .021; $M_{HCSZOFF}$ = 0.0798, p = .018; $M_{HC-SZON}$ = 0.1241, p < .01). No significant main effect of condition ($F_{1,287}$ = 0.451, p = .502) and no interaction was observed ($F_{3,287}$ = 1.745, p = .158) (Figure 4).

Model Fits

Figure 5A shows parameter fits. As previously found by Gold *et al.* (3), the mixing coefficient differs significantly by group (one-way ANOVA, $F_{3,295}$ = 13.91, p < .001), with post hoc tests showing that HCs have significantly higher coefficients (M_{HC} = 0.689, SD_{HC} = 0.22) than SZON (M_{SZON} = 0.478, SD_{SZON} 0.21, p < .001) and SZOFF (M_{SZOFF} 0.55, SD_{HC} = 0.23, p = .013) groups. The bipolar group did not differ significantly from either SZ or HC.

The irreducible-noise parameter ε also showed a significant main effect of group (Figure 5A) ($F_{3,295}$ = 2.67, p = .04), with HCs significantly less noisy than the SZOFF group (M_{HC} = 0.043, SD_{HC} = 0.16, M_{SZOFF} = 0.159, SD_{SZOFF} = 0.29, p = .03). The decay parameter coding forgetting between training and test differed significantly between the two SZ groups and the two others groups ($F_{1,289}$ = 12.81, p < .001), with higher decay rates in the SZ groups.

The softmax temperature parameter showed no significant effect of group; however, the recovery for this was less reliable than for the others (Supplement) (Figure 3), potentially because of collinearity between this parameter and others involved in the value function scaled by the softmax temperature (37) (see the Supplement for details of how temperature covaries with other parameters).

Test Phase

Model simulations showed that our model accurately captured behavior (Figure 3B), including the main effect of condition on accuracy between FW-FLA and the easier FW-IW and FW-FL conditions. This difference in accuracy between conditions was correlated with the mixing parameter (Figure 3C) (r^2 = 0.45, p = .039) but not with the random noise parameter (Figure 3D) (r^2 = 0.1, p = .088). The difference was not correlated with any other parameters.

Using Computational Modeling to Improve Classification—The utility of a model in computational psychiatry can be quantified by the degree to which key model parameters are diagnostic of clinical status with greater reliability than could be achieved from model-free performance measures. Highly accurate model-based classification can improve diagnosis, and knowing which parameters aid classification can shed light on underlying mechanisms in a way that raw-data measures cannot (9).

We set up classifiers to predict group identity (HC, BI, SZON/OFF) based on either accuracy (a model-agnostic quantity) or fit model parameters. Classifiers were trained on four quantiles of data and tested on the remaining quantile (leave-one-out cross-validation); this process was iterated until all data had been used for testing, and the classifier accuracy measures were averaged.

We used precision-recall curves to quantify classifier performance. Precision (positive predictive value) refers to a classifier's ability to correctly predict positive cases and avoid false positives. Recall (sensitivity) refers to the ability to predict true positives and not miss negatives. Precision-recall evaluates predictive success similarly to receiver operating characteristic curves but is more suitable for imbalanced class sets (17).

We trained model-based classifiers using all parameters as well as the most predictive parameters for group differences (the mix and noise described above). Model-agnostic classifiers used measures of accuracy across training and testing, including the most predictive ones for group differences (overall accuracy in training and testing, accuracy in the FW-FLA condition) as well as overall training/testing accuracies in early learning (first quarter) or late learning (fourth quarter).

Figure 6 shows classifiers predicting SZON versus HC, SZON versus BI, and SZON versus SZOFF. As expected, all methods (model-agnostic and model-based) performed best when classifying HCs from patients with SZ and were less effective classifying between patient groups. The model-agnostic, accuracy-based classifiers were most effective when including only late-stage learning, but they generally underperformed compared with the model-based classifiers (Figure 6B, D, F).

In the HC versus SZON case (Figure 6A, B), the best prediction was obtained using the hybrid mix and noise parameters to train the classifier; however, all model-based classifiers outperformed accuracy-based classification. Classification for BI versus SZON groups was not as accurate as HC versus SZON (Figure 6C, D) ($M_{AUCHCvSZON} = 0.776$, $SD_{HCvSZON} = 0.148$, $M_{AUCBIvSZON} = 0.642$, $SD_{AUCBIvSZON} = 0.12$); however, both model-based and accuracy-based classifiers performed significantly better than chance, and using model parameters still improved classification over using only model-free measures ($t_{17} = 4.23$, $p < .001$).

The model-agnostic classifier mixing fourth-quarter learning across phases with FW-FLA accuracy performed close to the best model-based classifier in the HC versus SZON case (Figure 6B) ($M_{Mix + noise} = 0.776$, $SD_{Mix + noise} = 0.148$, $M_{Acc4qmix} = 0.728$, $SD_{AAcc4qmix} = 0.119$). However, when classifying between patient groups, the best-performing model-based classifier was significantly better than the best-performing model-agnostic classifier ($t_{BIvSZ} = 3.0157$, $p = .012$, $t_{SZON\text{-}OFF} = 3.242$, $p = .011$).

Lastly, we examined whether parameters related to deficits in anhedonia/amotivation in patients. Higher decay was correlated with greater anhedonia/amotivation across all patients ($r = 0.20$, $p = .003$), with a similar association within the patients with SZ ($r = 0.18$, $p = .027$), though not within patients with BI ($r = 0.10$, $p = .44$). A lower mixing parameter was trend-level associated with greater anhedonia/amotivation among all patients ($r = -0.13$, $p =$

.06), but not within the SZ or BI groups individually (r_{SZ} = −0.10, p = .21, r_{BI} = −0.003, p = .98). See the Supplement for positive symptom analyses.

DISCUSSION

Our study replicates and extends previous work showing specific differences in RL in patients with SZ compared with controls and patients with BI. In line with Gold *et al.* (3), we found that a mixed strategy of BG-driven PE updating and OFC-driven explicit value representation best fitted choice data in the probabilistic selection task. SZON and SZOFF groups showed significantly lower mix parameters compared with HC and BI groups (Figure 5A), indicating a maladaptive overreliance on PE-updating to the detriment of using value representations. This was confirmed by significantly lower accuracies in the test phase, when choosing stimuli associated with common positive outcomes (FW) over those associated with rare negative outcomes (FLA).

These findings provide an important replication of evidence for impairments in explicit value representation in SZ and significantly extend previous work to show that this holds true for both medicated and unmedicated patients with SZ but not for those with bipolar disorder. This suggests this pattern is not secondary to medication effects and provides initial evidence for specificity to nonaffective psychosis.

Two more parameters in our mixed model showed a significant group difference. The noise parameter, coding for non–exploration-directed response variability and thought to correspond to either internal neural variability (38) or perhaps suboptimal inference by the brain (39), was higher in both medicated and unmedicated patients with SZ than in controls. Random decision noise has been proposed to correspond to tonic levels of norepinephrine (40); although there is no clear link yet established, our findings are consistent with recent theories suggesting a link between overactive noradrenergic pathways and cortical dopamine dysregulation, leading to SZ-like behaviors (especially in the positive symptom domain—e.g., psychosis) (41–43).

Together with the correlations with anhedonia/amotivation, these mix parameter and noise results imply a maladaptive overreliance on BG-driven PE-based updating, with insufficient frontal contributions coding explicit value representation. This is consistent with accounts that link frontostriatal abnormalities and motivational deficits in SZ (44), suggesting altered involvement of ventral striatum and frontal areas (including ventromedial prefrontal cortex and OFC) in signaling the expected value of observed stimuli in gain and loss domains. We did not, however, find consistent gain/loss differences across SZ groups (potentially because of medication effects. Figure 4 shows a reversed trend in accuracy in gains vs. losses, depending on medication status), and the learning performance (including model-based learning rates) did not offer clear evidence for impaired striatal contributions. Rather, gradual PE-based learning appeared relatively spared, although asymptotic learning, as has been reported previously (45), was more likely impaired; however, as discussed in Results and the Supplement, it is possible that some subjects did not reach asymptote, making it difficult to interpret this finding. This fits the reported mixed evidence regarding striatal learning in SZ and aligns our current results with the theory that striatally mediated RL

mechanisms may be spared, while cortically mediated, more rapid, and explicit learning systems are impaired (46).

The other parameter found to vary between participants without SZ (including BI) and those with SZ was the between-phase decay on learned values. Both SZON and SZOFF groups showed significantly higher decay than the BI and control groups, indicating perhaps a deficit in working memory necessary for carrying over learned values from training to test phase. Previous work has found impairments in working memory contributions to learning in SZ (47,48), but results are mixed on the retention of information, with some studies finding intact retention from learning to test phase but impaired generalization to novel contexts (49). Our participants with SZ were less accurate than controls on both old and novel pairings in the test phase (Figure 4A), suggesting that the deficit does not lie solely in the generalization to novel contingencies that require explicit value representation but is likely compounded by other effects, such as memory decay (Figure 4B).

Model parameter fits suggest that this effect is likely a mix of impaired between-phase generalization and within-phase memory decay, as accuracy on the old pairings in the test phase correlated with both the decay and the mix parameter. However, our task was not designed to test working memory, and from the current structure, it is difficult to discern to what extent the difference in accuracy on old pairings is due to memory effects or deficits in value representation; furthermore, the decay was implemented in the winning model as occurring between the training and testing phase, and thus it precludes any conclusions on within-phase memory decay.

Although learning rate differences have been found in the literature (50,51), we saw no group differences here in model-based learning rates. There was also no significant difference in the softmax temperature quantifying sensitivity to value differences (although a trend existed, showing marginally lower softmax temperatures for SZ groups than for HC and BI), strengthening our hypothesis that learning deficits in SZ are more likely to stem from differences in value representation (and possibly generalization) than from different reward sensitivity.

The behavioral and computational results in the SZON and SZOFF groups hint at potentially distinct effects of medication on different aspects of learning. Although the overall accuracies did not differ [there is a trend toward higher accuracies in the unmedicated patients, consistent with previous findings on the effects of antipsychotics on gradual learning (21), but it is not significant], we did observe lower learning in the loss domain in medicated patients (Figure 4B). This was only significant in the test phase (although the same pattern is visible in the learning phase; in Figure 3B, the contrast is not significant). This effect could be due to a differential impact of medication on the underlying RL mechanism [for instance, antipsychotic medication has been found to increase sensitivity to losses but not gains, promoting maladaptively high negative learning rates and lose-shift strategies (52)] or on different encoding of gains and losses contingencies into long-term memory.

Our present study does not allow us to differentiate the source of this finding, as pairwise contrasts in model parameters were not different between the two groups (Figure 5A). However, examining trending differences suggests a higher reliance on Q-learning in the unmedicated group (as evidenced by higher mixing parameter and higher Q-learning rate) along with higher decision noise (via the noise parameter and lower temperature parameters). However, because of the imbalanced group sizes and other potential comorbidities in the unmedicated group, further work would be required to accurately specify the effect of different types of antipsychotic medication on Q-learning rate and value-based RL strategies.

Finding that model-based classification outperformed model-free (in this case, accuracy-based) classification (Figure 6) is a promising step toward improved diagnosis and treatment. Often, the similarity in behavior performance measures—encountered in other cognitive tasks as well (6,7)— and the overlap in clinical symptoms between patients with BI and those with SZ (6,51) make diagnosis difficult and may play a role in the cases in which treatment is ineffective. The leading assessment measures (21) indicate that timeline and frequency of positive and negative symptoms also be taken into account (e.g., in the case of a patient presenting with symptoms of psychosis and mania, the initial classification may be schizoaffective disorder; if the psychosis symptoms disappear with time, the diagnosis may be reclassified as BI, while conversely, if mania symptoms disappear and psychosis becomes chronic, it is reclassified as SZ), leading to potential delay in accurate diagnosis.

In the present work, we classified using both model-agnostic, accuracy-based behavioral measures and model-based parameters and found that using parameters improved classification both between HCs and patients, and more importantly, between patients with BI and those with SZ. This result strengthens the existing evidence that computational methods may provide crucial insight into clinical practice (8,9,48).

The difference in performance between the best-performing accuracy-based classifier and the best-performing model-based classifier varied among the three group comparisons, with the lowest difference in the HC versus SZON comparison.This, along with the fact that the mixed, late-stage learning accuracy classifier performed similarly to the model-based classifier for HCs versus SZON (Figure 6B), suggests that the advantage of model-based classifiers is highest when behavioral or neuropsychological measures might not be sufficient for diagnosis, such as in the case of psychiatric disorders with overlapping symptoms (such as BI and SZ). This is consistent with previous findings in the literature on the utility of summary measures of overall intelligence and cognition in classifying patients from HCs (26) but the limited utility of these measures in the case of mixed patient groups (27).

The fact that combining three different measures of accuracy was required to come close to the performance of model-based classifiers is consistent with our rationale for using modeling—which is to access and quantify underlying cognitive processes that are difficult to extract from behavioral data alone. In the present case, using one single model parameter (the mix parameter) matched or outperformed a classifier using three carefully selected model-free measures. Thus, theoretically, model-agnostic classifiers could come even closer

to model-based ones, if we extracted and processed the data further and found other quantities that approximate underlying learning strategies (for instance, by further narrowing the trial types used to compute accuracy, or adding in further predictors such as time since that trial type was encountered). The difficulty lies in determining precisely which types of measures would be best fit for this—a difficulty resolved with computational modeling, which extracts quantities of interest in a straightforward, normative way

Finally, although these results stand as proof of concept for the utility of employing computational modeling, under the current framework, our model cannot prospectively predict diagnosis or dictate treatment; further work is required to build on this, toward more compelling applications. We hope that, particularly given the breadth of relatively easy-to-administer cognitive tasks available, using computational modeling to better pin down task strategies and their underlying mechanisms could be an accessible tool to assist psychiatric diagnosis and treatment.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS AND DISCLOSURES

Funding for this study was provided by the National Institute of Mental Health ROI1s (Grant No. MH084840 [to DMB], Grant No. MH084826 [to CSC], Grant No. MH084821 [to JMG], Grant No. MH084861 [to AWM], and Grant No. MH084828 [to SMS]).

AG had full access to all study data and takes responsibility for the integrity of the data and accuracy of the data analysis. AG performed the data analysis. All authors developed the study concept and design and aided in interpretation and provided critical revisions. All authors approved the final version of the paper for submission.

We wish to thank the participants in this study, who gave generously of their time, and the many staff who helped complete this project.

MJF reports consulting fees from Hoffman LaRoche Pharmaceuticals. JMG has consulted with Acadia Pharmaceuticals and reports royalty payments from the Brief Assessment of Cognition in Schizophrenia. All other authors report no biomedical financial interests or potential conflicts of interest.

REFERENCES

1. Barch DM, Ceaser A (2012). Cognition in schizophrenia. Core psychological and neural mechanisms. Trends Cogn Sci 16.27–34. [PubMed: 22169777]
2. Gold JM, Waltz JA, Prentice KJ, Morris SE, Heerey EA (2008). Reward processing in schizophrenia. A deficit in the representation of value. Schizophr Bull 34.835–847. [PubMed: 18591195]
3. Gold JM, Waltz JA, Matveeva TM, Kasanova Z, Strauss GP, Herbener ES, et al. (2012). Negative symptoms and the failure to represent the expected reward value of actions. Behavioral and computational modeling evidence. Arch Gen Psychiatry 69.129–138. [PubMed: 22310503]
4. Strauss GP, Waltz JA, Gold JM (2014). A review of reward processing and motivational impairment in schizophrenia. Schizophr Bull 40(suppl 2).S107–S116. [PubMed: 24375459]
5. Barch DM, Pagliaccio D, Luking K (2015). Mechanisms underlying motivational deficits in psychopathology. Similarities and differences in depression and schizophrenia. In. Simpson E, Balsam P, editors. Behavioral Neuroscience of Motivation. Cham. Springer, 411–449.

6. Whitton AE, Treadway MT, Pizzagalli DA (2015). Reward processing dysfunction in major depression, bipolar disorder and schizophrenia. Curr Opin Psychiatry 28.7–12. [PubMed: 25415499]
7. Zhao Q, Ma YT, Lui SS, Liu WH, Xu T, Yu X, et al. (2013). Neurological soft signs discriminate schizophrenia from major depression but not bipolar disorder. Prog Neuropsychopharmacol Biol Psychiatry 43.72–78. [PubMed: 23266480]
8. Huys QJ, Maia TV, Frank MJ (2016). Computational psychiatry as a bridge from neuroscience to clinical applications. Nat Neurosci 19.404–413. [PubMed: 26906507]
9. Wiecki TV, Poland J, Frank MJ (2015). Model-based cognitive neuroscience approaches to computational psychiatry. Clustering and classification. Clin Psychol Sci 3.378–399.
10. Wilson RC, Collins AG (2019). Ten simple rules for the computational modeling of behavioral data. Elife 8.e49547.
11. Howes OD, Kapur S (2009). The dopamine hypothesis of schizophrenia. Version III–the final common pathway. Schizophr Bull 35.549–562. [PubMed: 19325164]
12. Howes OD, Kambeitz J, Kim E, Stahl D, Slifstein M, Abi-Dargham A, Kapur S (2012). The nature of dopamine dysfunction in schizophrenia and what this means for treatment. Arch Gen Psychiatry 69.776–786. [PubMed: 22474070]
13. Deserno L, Boehme R, Heinz A, Schlagenhauf F (2013). Reinforcement learning and dopamine in schizophrenia. Dimensions of symptoms or specific features of a disease group? Front Psychiatry 4.172. [PubMed: 24391603]
14. Ceaser AE, Goldberg TE, Egan MF, McMahon RP, Weinberger DR, Gold JM (2008). Set-shifting ability and schizophrenia. A marker of clinical illness or an intermediate phenotype? Biol Psychiatry 64.782–788. [PubMed: 18597738]
15. Dowd EC, Barch DM (2012). Pavlovian reward prediction and receipt in schizophrenia. Relationship to anhedonia. PLoS One 7.e35622.
16. Kim H, Lee D, Shin YM, Chey J (2007). Impaired strategic decision making in schizophrenia. Brain Res 1180.90–100. [PubMed: 17905200]
17. Saito T, Rehmsmeier M (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLoS One 10.e0118432.
18. Strauss GP, Frank MJ, Waltz JA, Kasanova Z, Herbener ES, Gold JM (2011). Deficits in positive reinforcement learning and uncertainty-driven exploration are associated with distinct aspects of negative symptoms in schizophrenia. Biol Psychiatry 69.424–431. [PubMed: 21168124]
19. Morris SE, Holroyd CB, Mann-Wrobel MC, Gold JM (2011). Dissociation of response and feedback negativity in schizophrenia. Electrophysiological and computational evidence for a deficit in the representation of value. Front Hum Neurosci 5.123. [PubMed: 22065618]
20. Fusar-Poli P, Papanastasiou E, Stahl D, Rocchetti M, Carpenter W, Shergill S, McGuire P (2015). Treatments of negative symptoms in schizophrenia. Meta-analysis of 168 randomized placebo-controlled trials. Schizophr Bull 41.892–899. [PubMed: 25528757]
21. Insel C, Reinen J, Weber J, Wager TD, Jarskog LF, Shohamy D, Smith EE (2014). Antipsychotic dose modulates behavioral and neural responses to feedback during reinforcement learning in schizophrenia. Cogn Affect Behav Neurosci 14.189–201. [PubMed: 24557585]
22. McKirdy J, Sussmann JED, Hall J, Lawrie SM, Johnstone EC, McIntosh AM (2009). Set shifting and reversal learning in patients with bipolar disorder or schizophrenia. Psychol Med 39.1289–1293. [PubMed: 19105856]
23. First MB, Gibbon M (2004). The structured clinical interview for DSM-IV axis I disorders (SCID-I) and the structured clinical interview for DSM-IV axis II disorders (SCID-II). In. Hilsenroth MJ, Segal DL, editors. (2004), Comprehensive Handbook of Psychological Assessment, 2. New Jersey. John Wiley & Sons, Inc., 134–143
24. Overall JE, Gorham DR (1962). The brief psychiatric rating scale. Psychol Rep 10.799–812.
25. Andersen J, Larsen JK, Schultz V, Nielsen BM, Kørner A, Behnke K, et al. (1989). The Brief Psychiatric Rating Scale. Dimension of schizophrenia–reliability and construct validity. Psychopathology 22.168–176. [PubMed: 2762476]

26. Ventura J, Lukoff D, Nuechterlein KH, Liberman RP, Green M, Shaner A (1993). Appendix 1. Brief Psychiatric Rating Scale (BPRS) expanded version (4.0) scales, anchor points and administration manual. Int J Methods Psychiatr Res 3.43.
27. Dingemans PM, Linszen DH, Lenior ME, Smeets RM (1995). Component structure of the expanded Brief Psychiatric Rating Scale (BPRS-E). Psychopharmacology (Berl) 122.263–267. [PubMed: 8748395]
28. Young RC, Biggs JT, Ziegler VE, Meyer DA (2000). Young Mania Rating Scale. In. Handbook of Psychiatric Measures. Washington, DC. American Psychiatric Association, 540–542.
29. Berk M, Malhi GS, Cahill C, Carman AC, Hadzi-Pavlovic D, Hawkins MT, et al. (2007). The Bipolar Depression Rating Scale (BDRS). Its development, validation and utility. Bipolar Disord 9.571–579. [PubMed: 17845271]
30. Kring AM, Gur RE, Blanchard JJ, Horan WP, Reise SP (2013). The Clinical Assessment Interview for Negative Symptoms (CAINS). Final development and validation. Am J Psychiatry 170.165–172. [PubMed: 23377637]
31. Schneider LC, Struening EL (1983). SLOF. A behavioral rating scale for assessing the mentally ill. Soc Work Res Abstr 19.9–21. [PubMed: 10264257]
32. Hollingshead AB (1975). Four Factor Index of Social Status. New Haven. Yale University.
33. Nakao K, Treas J (1992). The 1989 Socioeconomic Index of Occupations. Construction from the 1989 Occupational Prestige Scores. Chicago. National Opinion Research Center.
34. Hernaus D, Gold JM, Waltz JA, Frank MJ (2018). Impaired expected value computations coupled with overreliance on stimulus-response learning in schizophrenia. Biol Psychiatry Cogn Neurosci Neuroimaging 3.916–926. [PubMed: 29735404]
35. Nassar MR, Frank MJ (2016). Taming the beast. Extracting generalizable knowledge from computational models of cognition. Curr Opin Behav Sci 11 49–54. [PubMed: 27574699]
36. Radulescu A, Daniel R, Niv Y (2016). The effects of aging on the interaction between reinforcement learning and attention. Psychol Aging 31.747–757. [PubMed: 27599017]
37. Daw ND (2011). Trial-by-trial data analysis using computational models. In. Delgado MR, Phelps EA, Robbins TW, editors. Decision Making, Affect, and Learning. Attention and Performance XXIII. Oxford. Oxford University Press, 3–38.
38. Wilson RC, Geana A, White JM, Ludvig EA, Cohen JD (2014). Humans use directed and random exploration to solve the explore-exploit dilemma. J Exp Psychol Gen 143.2074–2081. [PubMed: 25347535]
39. Gershman SJ (2019). Uncertainty and exploration. Decision (Wash D C) 6.277–286. [PubMed: 33768122]
40. Aston-Jones G, Cohen JD (2005). An integrative theory of locus coeruleus-norepinephrine function. Adaptive gain and optimal performance. Annu Rev Neurosci 28.403–450. [PubMed: 16022602]
41. Warren CM, Wilson RC, Van Der Wee NJ, Giltay EJ, Van Noorden MS, Cohen JD, Nieuwenhuis S (2017). The effect of atomoxetine on random and directed exploration in humans. PLoS One 12.e0176034.
42. Seeman P (2011). All roads to schizophrenia lead to dopamine supersensitivity and elevated dopamine D2(high) receptors. CNS Neurosci Ther 17.118–132. [PubMed: 20560996]
43. Grace AA (2010). Dopamine system dysregulation by the ventral subiculum as the common pathophysiological basis for schizophrenia psychosis, psychostimulant abuse, and stress. Neurotox Res 18.367–376. [PubMed: 20143199]
44. Waltz JA, Xu Z, Brown EC, Ruiz RR, Frank MJ, Gold JM (2018). Motivational deficits in schizophrenia are associated with reduced differentiation between gain and loss-avoidance feedback in the striatum. Biol Psychiatry Cogn Neurosci Neuroimaging 3.239–247. [PubMed: 29486865]
45. Weickert TW, Terrazas A, Bigelow LB, Malley JD, Hyde T, Egan MF, et al. (2002). Habit and skill learning in schizophrenia. Evidence of normal striatal processing with abnormal cortical input. Learn Mem 9.430–442. [PubMed: 12464703]

46. Dowd EC, Frank MJ, Collins A, Gold JM, Barch DM (2016). Probabilistic reinforcement learning in patients with schizophrenia. Relationships to anhedonia and avolition. Biol Psychiatry Cogn Neurosci Neuroimaging 1.460–473. [PubMed: 27833939]

47. Johnson MK, McMahon RP, Robinson BM, Harvey AN, Hahn B, Leonard CJ, et al. (2013). The relationship between working memory capacity and broad measures of cognitive ability in healthy adults and people with schizophrenia. Neuropsychology 27.220–229. [PubMed: 23527650]

48. Collins AG, Brown JK, Gold JM, Waltz JA, Frank MJ (2014). Working memory contributions to reinforcement learning impairments in schizophrenia. J Neurosci 34.13747–13756. [PubMed: 25297101]

49. Shohamy D, Mihalakos P, Chin R, Thomas B, Wagner AD, Tamminga C (2010). Learning and generalization in schizophrenia. Effects of disease and antipsychotic drug treatment. Biol Psychiatry 67.926–932. [PubMed: 20034612]

50. Waltz JA, Frank MJ, Wiecki TV, Gold JM (2011). Altered probabilistic learning and response biases in schizophrenia. Behavioral evidence and neurocomputational modeling. Neuropsychology 25.86–97. [PubMed: 21090899]

51. Weickert TW, Goldberg TE, Egan MF, Apud JA, Meeter M, Myers CE, et al. (2010). Relative risk of probabilistic category learning deficits in patients with schizophrenia and their siblings. Biol Psychiatry 67.948–955. [PubMed: 20172502]

52. Reddy LF, Waltz JA, Green MF, Wynn JK, Horan WP (2016). Probabilistic reversal learning in schizophrenia. Stability of deficits and potential causal mechanisms. Schizophr Bull 42.942–951. [PubMed: 26884546]

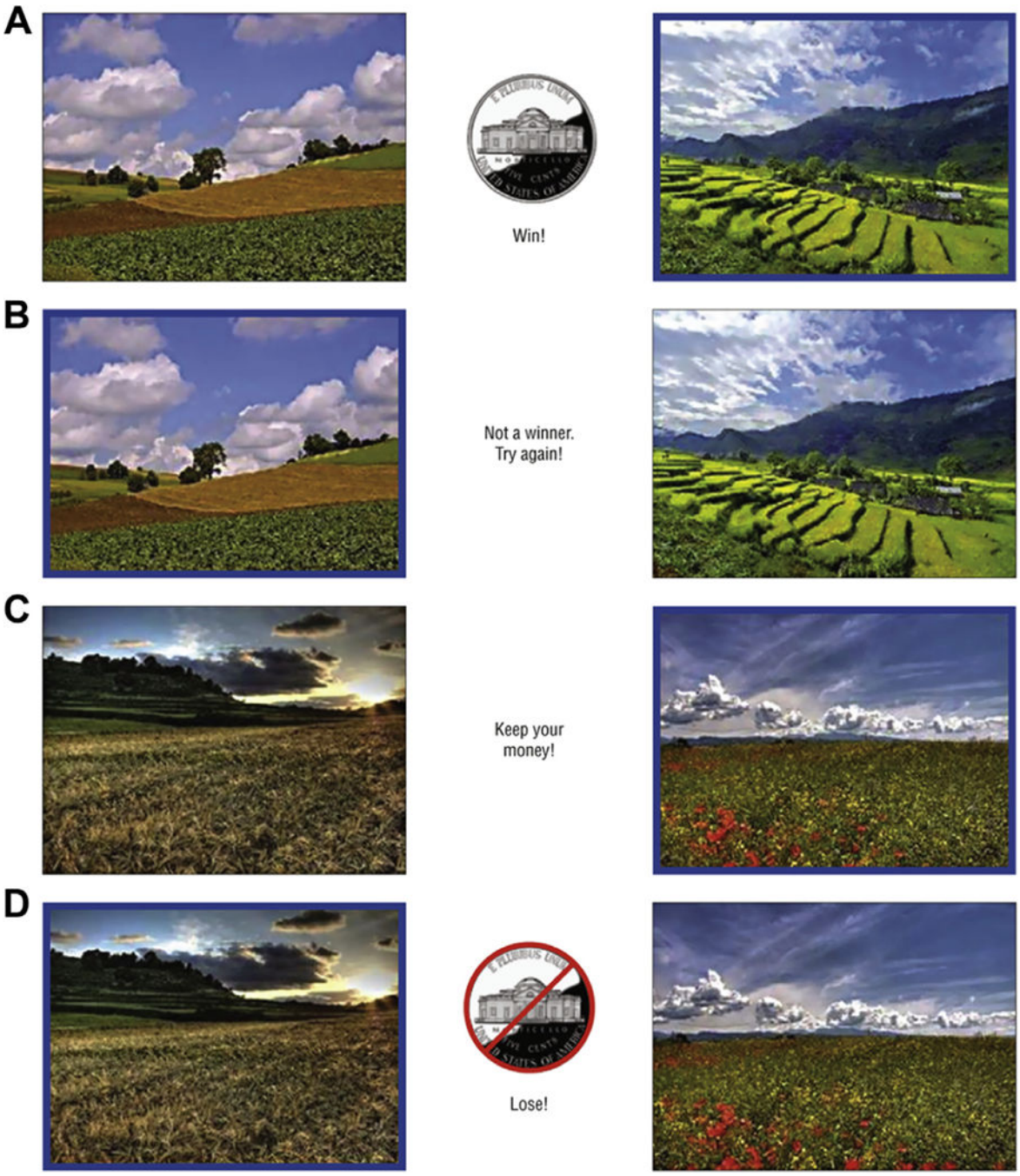

Figure 1.

Reinforcement learning task [from Gold *et al.* (3)]. On every trial, participants were shown two stimuli and required to choose one so as to maximize reward (or, depending on the available stimuli, minimize loss). Each condition (gain or loss avoidance) had two possible pairs of stimuli. The figure shows one example pair from a gain trial [with a win **(A)** or not-win **(B)** outcome] and one loss-avoidance trial [with an avoid-loss **(C)** or loss **(D)** outcome]. These stimuli could be associated with positive (e.g., winning $0.05) or negative (losing $0.05) value and had different probabilities of reward. The gain condition stimuli (*a, b, c,*

d) were always associated with positive reward, either high probability (frequent winner) or low probability (infrequent winner). The other four stimuli always yielded negative reward with high probability (frequent loser) or low probability (frequent loss avoider).

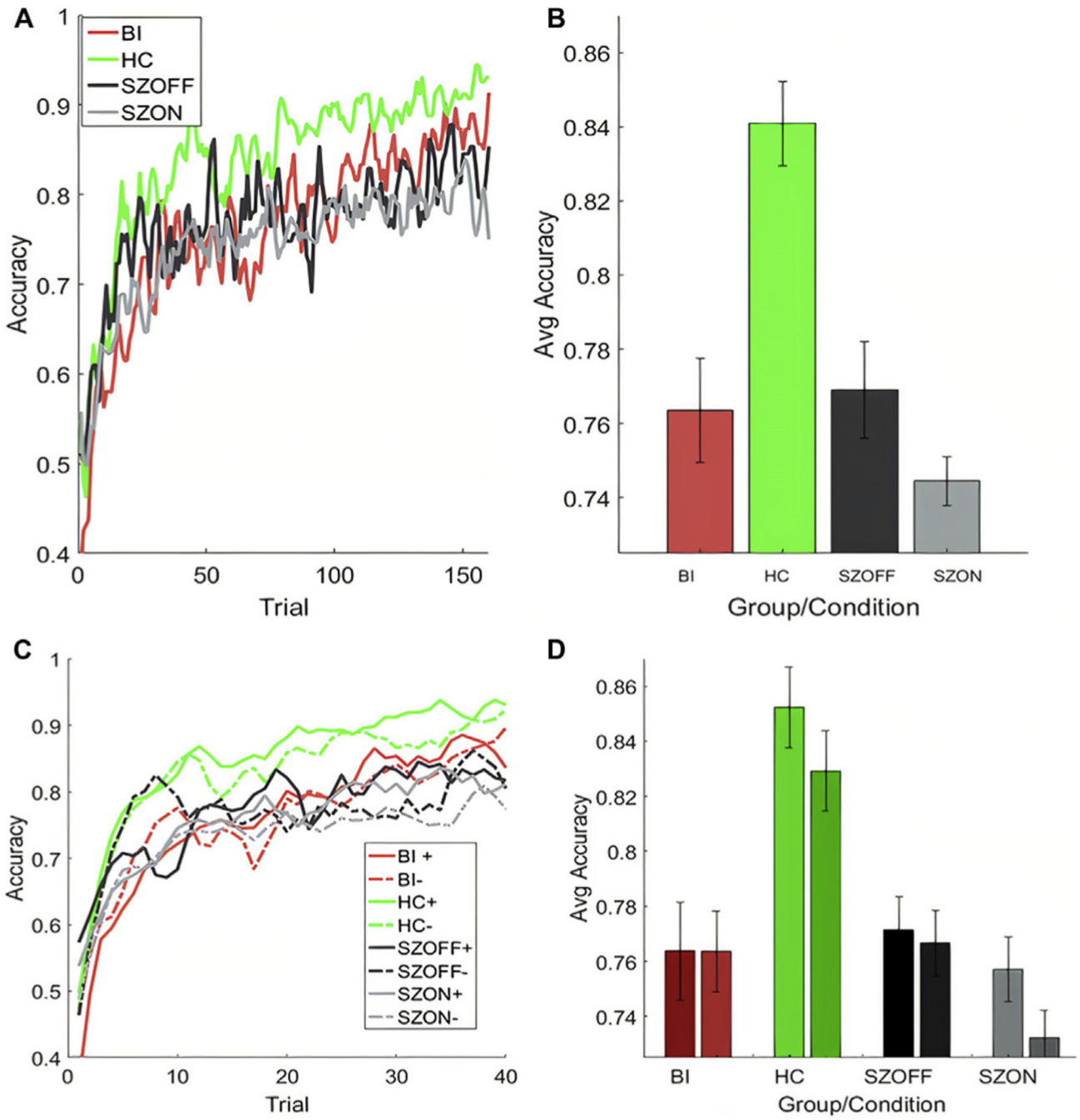

Figure 2.

Accuracy in the training phase. **(A)** Overall learning curves and (inline) average accuracy across the training phase. Green. healthy control (HC), red. bipolar disorder (BI), dark gray. unmedicated schizophrenia (SZOFF), light gray. medicated schizophrenia (SZON). **(B)** Average accuracy across learning phase trials. **(C)** Learning curves split by condition (gains [+] or losses [−]). **(D)** Average accuracy split by condition. Post hoc tests revealed that HCs had an average of 12.3% higher accuracy ($p < .001$) than the SZON group. None of the other contrasts were significant.

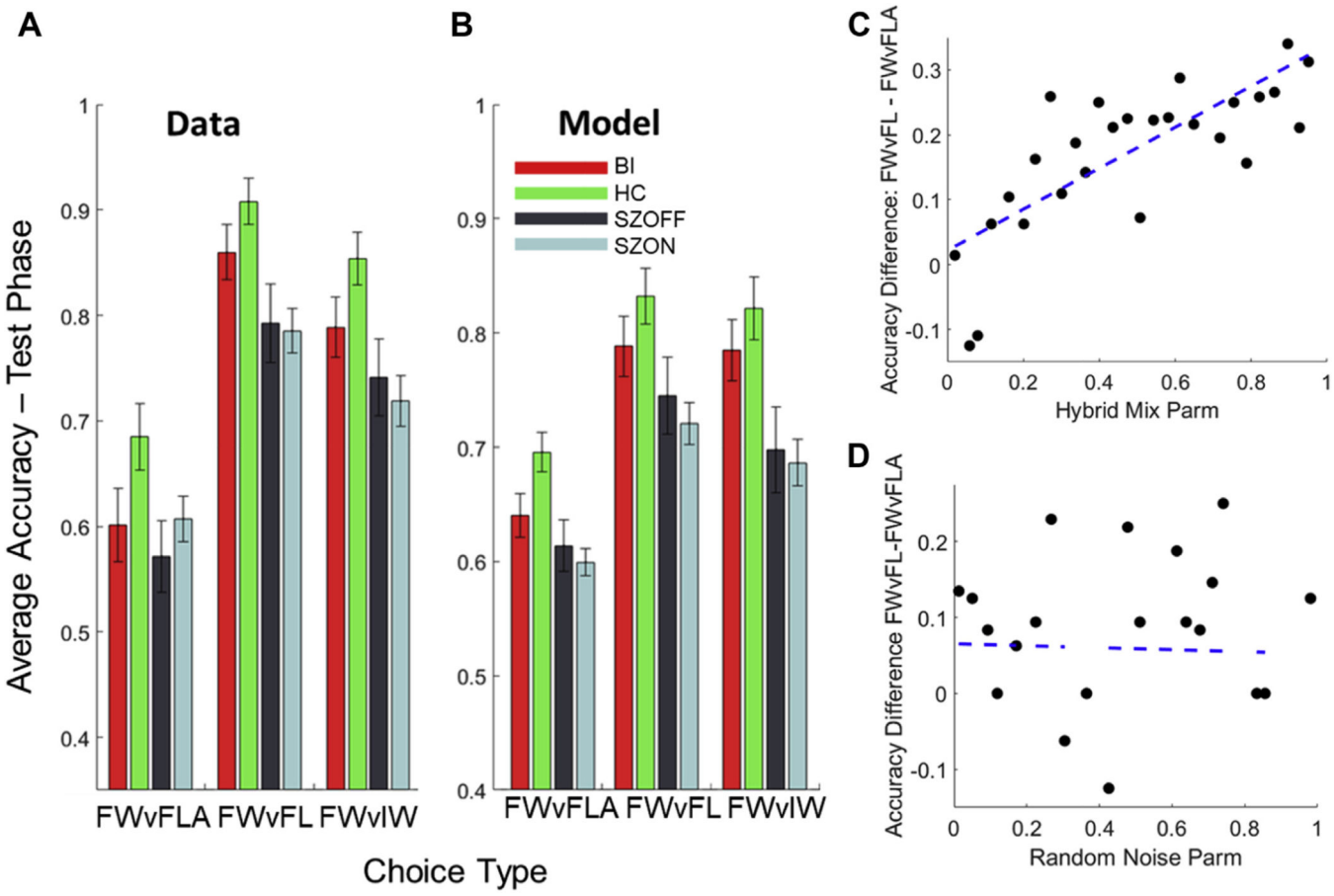

Figure 3.

Test phase performance for the participants **(A, C, D)** and the model simulations **(B)**. **(A)** Average accuracy on the 72 trials of test phase, on three new stimulus pairing conditions (see text). **(B)** Model simulation results on same three new stimulus pairing conditions. **(C)** Significant positive correlation between hybrid mix parameter and performance differences on easy new condition (frequent winner [FW]-frequent loser [FL]) vs. hard new condition (FW-frequent loss avoider [FLA]). Each point represents the average value in a bin of size 0.05. Blue line represents linear regression. **(D)** No correlation between noise parameter and performance. BI, bipolar disorder; HC, healthy control; IW, infrequent winner; Parm, parameter; SZOFF, unmedicated schizophrenia; SZON, medicated schizophrenia.

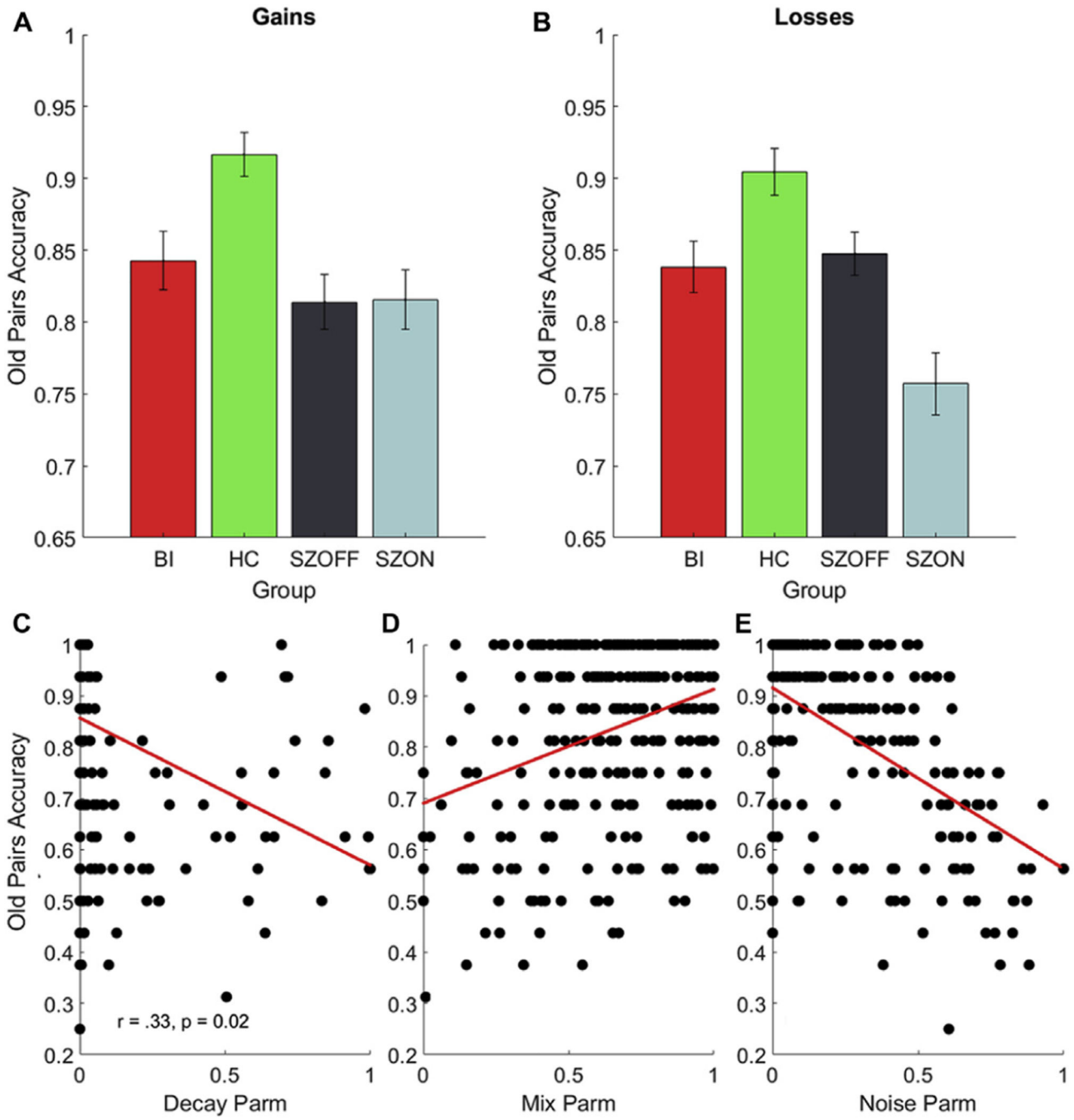

Figure 4.

Performance on learned (old) pairings in the testing phase. **(A)** Average accuracy on trials in which stimuli had been previously associated with rewards (*ab, cd*). **(B)** Average accuracy on trials in which stimuli had been previously associated with losses (*ef, gh*). **(C–E)** Average accuracy on the learned pairings correlates with decay, mix, and noise parameters. BI, bipolar disorder; HC, healthy control; Parm, parameter; SZOFF, unmedicated schizophrenia; SZON, medicated schizophrenia.

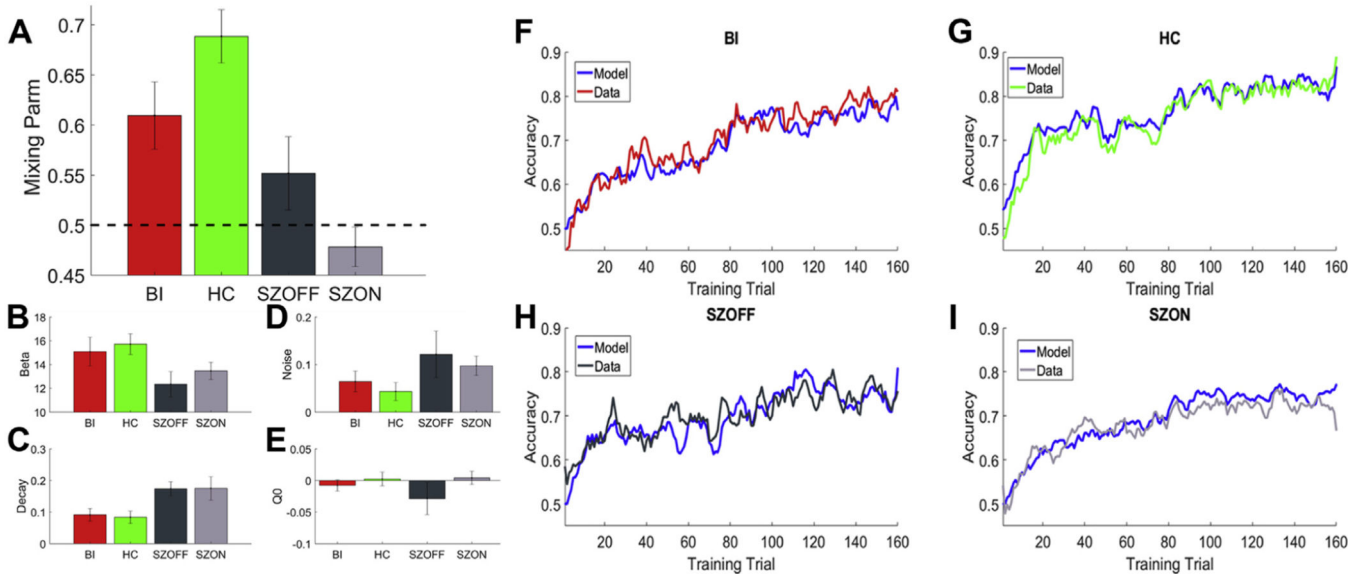

Figure 5.

Model parameters across groups **(A–E)** and model checks **(F–I)**. **(A)** Hybrid mix parameter (determining the reliance on Q-learning strategy vs. actor-critic strategy) is highest in healthy control (HC) group and lowest in medicated schizophrenia (SZON) group. A value higher than 0.5 means higher reliance on Q-learning. Also shown are softmax temperature beta **(B)**, noise parameter **(D)**, memory decay **(C)**, and initial Q values **(E)**. **(F–I)** Model (blue lines) captures learning behavior in all four participant groups; shown here is training phase; parameters used to simulate model behavior are fit group parameters. BI, bipolar disorder; Parm, parameter; SZOFF, unmedicated schizophrenia.

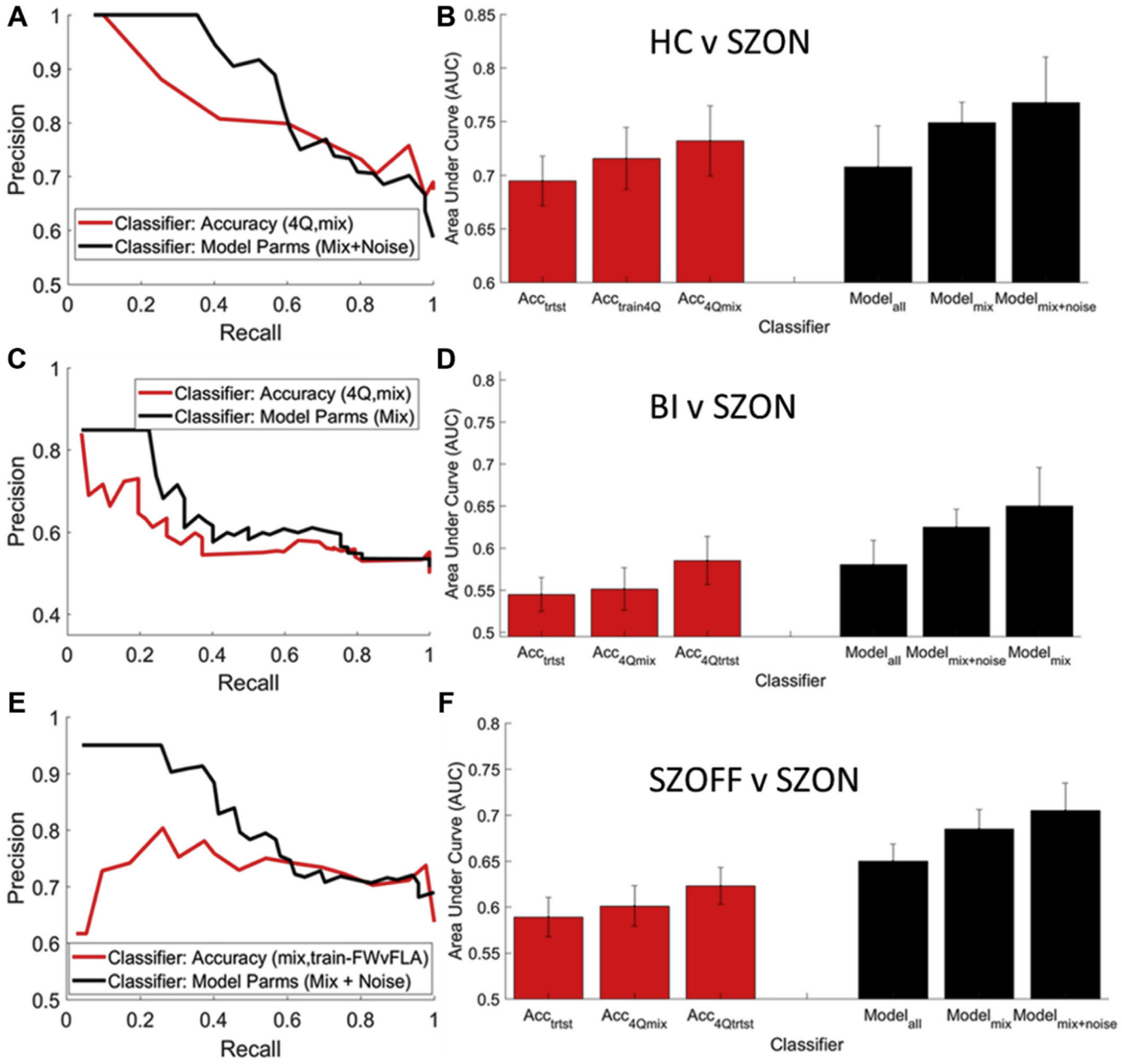

Figure 6.

Classification results for model-based vs. model-agnostic (accuracy-based) classifiers. **(A, B)** Classifying healthy control subjects (HC) from medicated patients with schizophrenia (SZON). **(C, D)** Classifying patients with bipolar disorder (BI) vs. SZON. **(E, F)** Classifying unmedicated patients with schizophrenia (SZOFF) and SZON. (Left panels) Precision-recall curves for the best classifier in each group. (Right panels) Areas under the curve for the top three performing classifiers in each group. Red bars represent model-agnostic classifiers and black bars represent model-based classifiers. Parms, parameters.