Title

Reinforcement learning deficits exhibited by postnatal PCP-treated rats enable deep neural network classification.

Journal

Neuropsychopharmacology, 48(9)

Authors

Tranter, Michael

Aggarwal, Samarth

Young, Jared

et al.

Publication Date

2023-08-01

DOI

10.1038/s41386-022-01514-y

Peer reviewed

American College of Neuropsychopharmacology

www.nature.com/npp

ARTICLE

Check for updates

Reinforcement learning deficits exhibited by postnatal PCP-treated rats enable deep neural network classification

Michael M. Tranter[1,2,5], Samarth Aggarwal[1,5], Jared W. Young[1,2], Daniel G. Dillon[3,4] and Samuel A. Barnes[1,2]✉

The ability to appropriately update the value of a given action is a critical component of flexible decision making. Several psychiatric disorders, including schizophrenia, are associated with impairments in flexible decision making that can be evaluated using the probabilistic reversal learning (PRL) task. The PRL task has been reverse-translated for use in rodents. Disrupting glutamate neurotransmission during early postnatal neurodevelopment in rodents has induced behavioral, cognitive, and neuropathophysiological abnormalities relevant to schizophrenia. Here, we tested the hypothesis that using the NMDA receptor antagonist phencyclidine (PCP) to disrupt postnatal glutamatergic transmission in rats would lead to impaired decision making in the PRL. Consistent with this hypothesis, compared to controls the postnatal PCP-treated rats completed fewer reversals and exhibited disruptions in reward and punishment sensitivity (i.e., win-stay and lose-shift responding, respectively). Moreover, computational analysis of behavior revealed that postnatal PCP-treatment resulted in a pronounced impairment in the learning rate throughout PRL testing. Finally, a deep neural network (DNN) trained on the rodent behavior could accurately predict the treatment group of subjects. These data demonstrate that disrupting early postnatal glutamatergic neurotransmission impairs flexible decision making and provides evidence that DNNs can be trained on behavioral datasets to accurately predict the treatment group of new subjects, highlighting the potential for DNNs to aid in the diagnosis of schizophrenia.

Neuropsychopharmacology (2023) 48:1377–1385; https://doi.org/10.1038/s41386-022-01514-y

INTRODUCTION

The world is unpredictable and if a course of action is no longer profitable it is essential to be behaviorally flexible [1–4]. Flexible decision making can be evaluated using the probabilistic reversal learning (PRL) task, which has provided evidence that people with schizophrenia show decision-making deficits. For instance, people with schizophrenia exhibit abnormalities in the explore-exploit balance [5], learning rate [6], and reward prediction errors (PEs) [7, 8]. PEs serve as teaching signals to update value estimates [9, 10], ensuring that deviations in expected outcomes are detected and appropriate adaptations in subsequent behavior can be implemented. Hence, aberrant PEs may play an especially important role in impaired decision making in schizophrenia [11–13].

Deficits in reinforcement learning are evident in patients with first-episode psychosis [14, 15]. Therefore, it may be possible to use these disruptions as a tool to assist with the early detection of risk for schizophrenia. Early intervention and accurate diagnosis are crucial [16]; however, misdiagnosis or delays between symptom onset and initiation of appropriate treatment are common [17, 18]. Although the precise reasons for these diagnostic challenges are unclear, the diagnosis of schizophrenia (and mental illness in general) still relies on subjective evaluation by psychiatrists [19]. Encouragingly, recent studies have demonstrated the utility of machine learning or artificial deep neural networks (DNNs) in psychiatry [20, 21], achieving >80% accuracy in the classification of healthy participants vs. schizophrenia patients [22, 23]. Hence, this approach may complement clinical evaluations to improve the speed and accuracy of diagnosis. One limitation of this approach is that DNNs are often trained to detect differences in magnetic resonance images or electrophysiology recordings that many clinicians may not have. By contrast, the PRL task is easy to administer, and computational analysis generates a rich dataset that can be used to probe underlying mechanisms. However, it is not clear whether DNNs trained using PRL data can accurately classify controls vs. people with schizophrenia. In summary, accurate DNN-based classification based on PRL performance could be a valuable and accessible tool to assist in the initial diagnosis of schizophrenia.

As a first step toward this goal, the current study applied a DNN to data from an experimental system relevant to schizophrenia. Disrupting glutamate transmission during postnatal neurodevelopment in rodents by administering NMDA receptor antagonists (i.e., phencyclidine (PCP) or ketamine) has been used to model the neurodevelopmental origins of schizophrenia [24, 25]. Postnatal PCP-treatment induces several neuropathological abnormalities evident in schizophrenia [26], such as deficits in inhibitory GABAergic interneuron expression [27, 28]. Moreover, postnatal PCP-treatment also results in deficits in memory [29], sociability [30], and executive functioning [31, 32]. However, it remains unclear whether postnatal PCP-treatment impairs PRL performance. Thus, the goal of the current investigation was two-fold:

[1]Department of Psychiatry, University of California San Diego, La Jolla, CA 92093, USA. [2]Department of Mental Health, VA San Diego Healthcare System, La Jolla, CA 92093, USA. [3]Center for Depression, Anxiety and Stress Research, McLean Hospital, Belmont, MA 02478, USA. [4]Harvard Medical School, Boston, MA 02115, USA. [5]These authors contributed equally: Michael M. Tranter, Samarth Aggarwal. ✉email: sabarnes@ucsd.edu

Received: 30 June 2022 Revised: 21 November 2022 Accepted: 26 November 2022
Published online: 12 December 2022

(1) to test the hypothesis that postnatal PCP-treated rats exhibit impairments in PRL, and (2) to develop a DNN to identify the treatment group of rodents based on their PRL performance.

METHODS

Animals

Timed-pregnant Wistar dams (Charles River Laboratories, Raleigh, NC) were obtained at day 13 of gestation. Dams were housed individually in a climate-controlled room on a reverse 12 h light cycle (lights off at 07:00) with ad libitum access to food and water. Within 12 h of parturition, male and female pups were randomly assigned to one of six litters of four pups per sex and cross-fostered between six lactating dams. All procedures were conducted following guidelines from the National Institutes of Health and the Association for the Assessment and Accreditation of Laboratory Animal Care, and were approved by the University of California, San Diego Institutional Animal Care and Use Committee.

Drug treatment

As previously described [27, 31], rat pups were subcutaneously administered saline (0.9%) or PCP (20 mg/kg) on postnatal days (PND) 7, 9, and 11. Representatives from each treatment group were present in each litter. Pups were weaned on PND 21, split by sex, and group-housed until 4–5 weeks of age, after which they were housed in pairs of the same sex and treatment group ($n = 12$ per sex/treatment). Behavioral testing began after PND 60.

Apparatus

Behavioral training and testing were conducted in 9-hole operant boxes (Med Associates, St Albans, VT) contained within light- and sound-attenuating chambers [33]. Briefly, the rear wall was a curved array with five open response apertures. Throughout training and testing, only apertures 2 and 4 were active. A 3-W stimulus light located at the aperture rear provided visual stimuli. On the opposite wall, a receptacle delivered food rewards (45 mg sucrose pellets, Test Diet 5TUT, Richmond, IN). The apparatus was controlled by a PC running MedPC software (Med Associates, St. Albans, VT).

Probabilistic reversal learning (PRL)

Throughout training and testing, rats were food restricted to 90% of their free-feeding body weights. PRL training was conducted as previously described [34, 35]. Once the criterion during the basic training session was met, PRL testing commenced the next day. During PRL testing, rats were presented with two illuminated apertures (holes 2 and 4). One aperture was designated the target location, and the other was the non-target location. Target responses were reinforced with 80% probability, while non-target responses were reinforced with 20% probability (Fig. 1A). A 2-s inter-trial interval (ITI) separated the trials, and responses during the ITI resulted in a time-out. Once a rat made eight consecutive target responses, the contingencies switched, and the previous non-target location became the target location. Reversals continued throughout the session each time the rat made eight consecutive target responses. Sessions were terminated after 300 trials or 60 min, whichever occurred first. Rats were tested once daily for 20 days. Performance was assessed by determining the number of reversals completed per 100 trials [36]. Win-stay responding for either the target or non-target responses was calculated as the percentage of trials where the rat repeated the same choice after being rewarded on the preceding trial. Lose-shift responding for target or non-target responses was calculated as the percentage of trials where the rat switched responses after a reward was withheld during the preceding trial.

Computational modeling

To investigate mechanisms that drive behavior, we fit variants of a Rescorla–Wagner Q-learning model to the data [5, 37–39]. Briefly, on each trial (t), the decision (i.e., which action to select) is likely guided by the value (Q) assigned to each action. Action values are updated according to whether a reward was delivered or not (r; reward = 1, no reward = 0) on each trial. Q values were initialized to 0.5 (neither good nor bad as the target/non-target locations were randomly assigned for each session). Prediction errors (PEs)—the difference between the estimated value of the action taken on trial t ($Q_c(t)$) and the reward delivered on trial t ($r(t)$)—served to update the value estimate of the chosen action on each trial. The rate at which the PE signal updated value estimates was controlled by the α parameter, the learning rate. Action values were converted into choice probabilities using the softmax function, which controls the degree to which the subject engages in exploratory vs. exploitative choices. We evaluated a total of 18 model variants (see Fig. 2A for an overview of each model. A detailed description of each model variant can be found in the Supplementary Methods and Supplementary Table 1).

Deep neural network (DNN)

To predict the treatment group of subjects, we trained a DNN on the behavioral variables described in Supplementary Table 2. The model was trained using 44 of 48 subjects, with the remaining 4 subjects (2 salines, 2 PCP) set aside to evaluate the generalizability of the trained model. The model architecture is described in detail in the Supplementary Material. Briefly, our DNN architecture consisted of three one-dimensional convolutional (1D CNN) layers, two long short-term memory (LSTM) layers, and two fully connected dense layers. The output layer consisted of two nodes with a softmax activation function to predict whether the subject was saline- or PCP-treated.

Statistical analysis

PRL performance was analyzed using 3-way repeated measures ANOVAs that included treatment and sex as between-subject factors and day as the within-subject factor. The association between the two variables was determined using Pearson's correlations. Data were analyzed using R and, where appropriate, significant interactions underwent post hoc adjustments using the Bonferroni correction for multiple comparisons. P values < 0.05 were considered significant. Data were presented as the mean ± SEM and graphically displayed using GraphPad Prism or Matplotlib.

RESULTS

The number of completed reversals increased across days [$F_{(19,836)} = 65.34$, $p < 0.001$] demonstrating that performance improved with training. In addition to the main effect of treatment [$F_{(1,44)} = 12.78$, $p < 0.001$], we found that there was also a treatment × day interaction [$F_{(19,836)} = 2.01$, $p < 0.01$] showing that PCP-treated rats completed fewer reversals compared to saline-treated control rats ($p < 0.05$) (Fig. 1B). Although the main effect of sex approached significance [$F_{(1,44)} = 2.91$, $p = 0.09$] there was no interactive effect between sex and treatment [Fs < 0.52]. To gain additional insight into these performance alterations, we inspected the win-stay and lose-shift measures. Consistent with the increase in reversals, the tendency to repeat a rewarded target response increased across test days [$F_{(19,836)} = 140.97$, $p < 0.001$]. However, while we did not observe a treatment × day interaction [$F_{(19,836)} = 1.32$, $p < 0.15$], there was a main effect of treatment [$F_{(1,44)} = 17.52$, $p < 0.001$] that was mediated by an overall reduction in target win-stay responding in PCP-treated rats (Fig. 1C). A day × sex interaction was evident [$F_{(19,836)} = 3.18$, $p < 0.001$] that was driven by higher target win-stay responses in female rats during the first two days only ($p < 0.05$).

Target lose-shift (TLS), non-target lose-shift (NTLS), and non-target win-stay (NTWS) measures were also reduced in PCP-treated rats testing (Supplementary Fig. 1A–C). Although PCP-induced alterations were evident for the latency to respond (Supplementary Fig. 2A), there was no effect on the reward latency or the number of completed trials (Supplementary Fig. 2B, C), indicating that postnatal PCP-treatment did not induce a non-specific impairment in the ability to complete the task.

Performance was then analyzed using a logistic regression that used the action selected and the outcome received for the previous six trials to predict the choice on the next trial, as previously described [40, 41]. In both groups of rats, positive feedback from the previous two trials was associated with repeating the same action (positive regression coefficient) whereas negative feedback from the last trial only was associated with switching (negative regression coefficient) (Fig. 1D). Relative to saline-treated rats, PCP-treated rats were less likely to repeat an action after a reward and less likely to switch an action after no reward, corroborating the observed effects on win-stay and lose-shift responding. To determine whether repeating rewarded

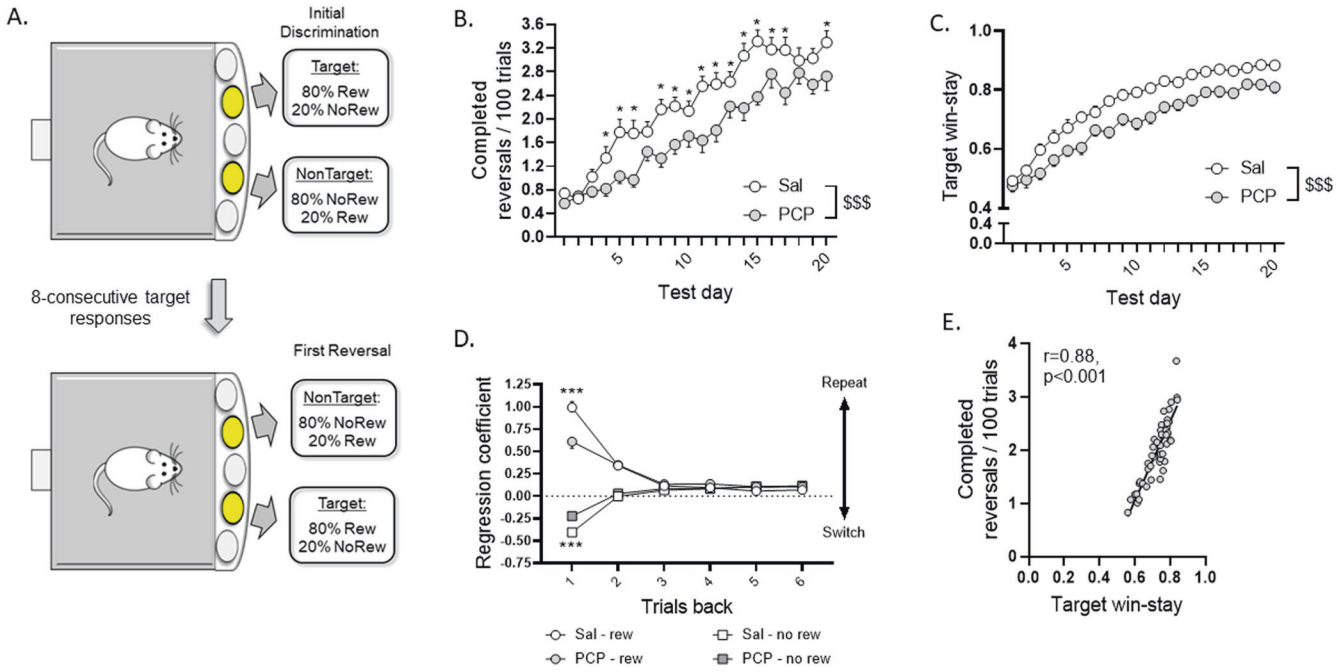


Fig. 1 Postnatal PCP-treatment impairs PRL performance. A Schematic of PRL task. The rat is presented with two illuminated nose-poke apertures. Target responses are mostly rewarded (80%), whereas non-target responses are rarely rewarded (20%). After eight consecutive target responses, the target and non-target levers reverse. **B** Across test days, the number of completed reversals in saline- and PCP-treated rats increased, but this increase was blunted in PCP-treated rats. **C** The propensity to repeat rewarded target responses also increased with training, but PCP-treated rats displayed weaker target win-stay responding than saline-treated rats. **D** Logistic regression coefficients that used the nose-poke selected and the outcome received from the previous six trials to predict the choice on the current trial. A rewarded right response was coded with +1 while a rewarded left response was coded –1 (non-rewarded responses were coded 0). As a separate predictor, non-rewarded right responses were coded +1 and non-rewarded left responses were coded –1 (rewarded responses were coded 0). Each predictor included in the model was used to determine how choices and outcomes from the previous trials influenced the choice of the current trial. A positive coefficient indicates a higher likelihood of selecting the previous nose-poke aperture whereas a negative coefficient indicates a greater likelihood of choosing the alternative nose-poke aperture. PCP-treated rats were less likely to repeat a rewarded choice and less likely to switch after an unrewarded choice, relative to saline-treated rats. **E** Scatterplot and Pearson's correlation demonstrating the relationship between target win-stay responding and the number of completed reversals. Sal, $n = 24$. PCP, $n = 24$. $^{*}p < 0.05$, $^{***}p < 0.001$. $^{\$\$\$}$ denotes main effect of treatment, $p < 0.001$.

choices or avoiding non-rewarded choices play an important role in PRL, we correlated the win-stay and lose-shift measures with the primary outcome variable. As expected, target win-stay responses were positively correlated with the number of completed reversals (Fig. 1E). By contrast, target lose-shift, non-target lose-shift, and non-target win-stay measures displayed no correlation with reversals (Supplementary Fig. 1D–F). Collectively, these findings show that a reduced tendency to repeat rewarded choices drove the overall deficit in PRL performance observed in postnatal PCP-treated rats.

To probe the behavioral mechanisms underlying these deficits, we fit several Q-learning models to behavior. Each model contained the same three essential components: choice, feedback, and learning (Fig. 2A and Supplementary Tables 1 and 3). Action values were converted into choice probabilities using the softmax function. The degree to which choices are exploratory (i.e., selecting the lower-valued action) vs. exploitative (i.e., selecting the higher-valued action) is controlled by the β parameter. A higher β parameter indicates a greater tendency to engage in exploitative choices. To determine whether the subject exhibited a bias for one side or another, a modified softmax function was used that included a bias parameter in addition to the inverse temperature parameter [42]. A bias parameter of 1 demonstrated a strong bias for the left aperture whereas a value of –1 indicated a strong bias for the right aperture. A value of 0 indicated no bias for either side was evident.

In addition to the typical delta rule that captures the difference between the expected and actual value of an action (i.e., the PE) we included two variants that would capture differences in overall outcome sensitivity or differences in reward vs. punishment sensitivity [43]. For example, if the reward sensitivity parameter equals one then a rewarded trial would generate a positive PE in a manner consistent with typical delta rule. By contrast, reducing the reward sensitivity parameter would result in a less positive PE value given the same reward. Finally, during the learning component, PEs were used to update action values. The rate at which PEs update the value of the chosen action was controlled by a learning rate, α. An alternative approach is the so-called "double update" rule, whereby the value of both the chosen and unchosen actions are updated [11]. With this variant, if a reward is received after a left response, the value of the left option will increase and the value for the right option will decrease. By contrast, if the left option was unrewarded then its value would decrease and the value of the right option would increase. The final variant included a method to decay the value of the unchosen action, designed to capture "forgetting" the value of an action if it has not been selected for several trials [42, 44]. For example, if the forget parameter is zero, then the value of the unchosen action is unaffected. By contrast, if the forget parameter equals 0.5 then the value of the unchosen action is reduced by 50% for each trial that the action is unchosen.

In total, we fit 18 models to our behavioral data, each of which used one variant from each of the three components (choice, feedback, learning) (summarized in Fig. 2A, described in detail in the Supplementary Methods). To determine the best-fitting model, we calculated the AICc, ΔAICc and relative likelihood

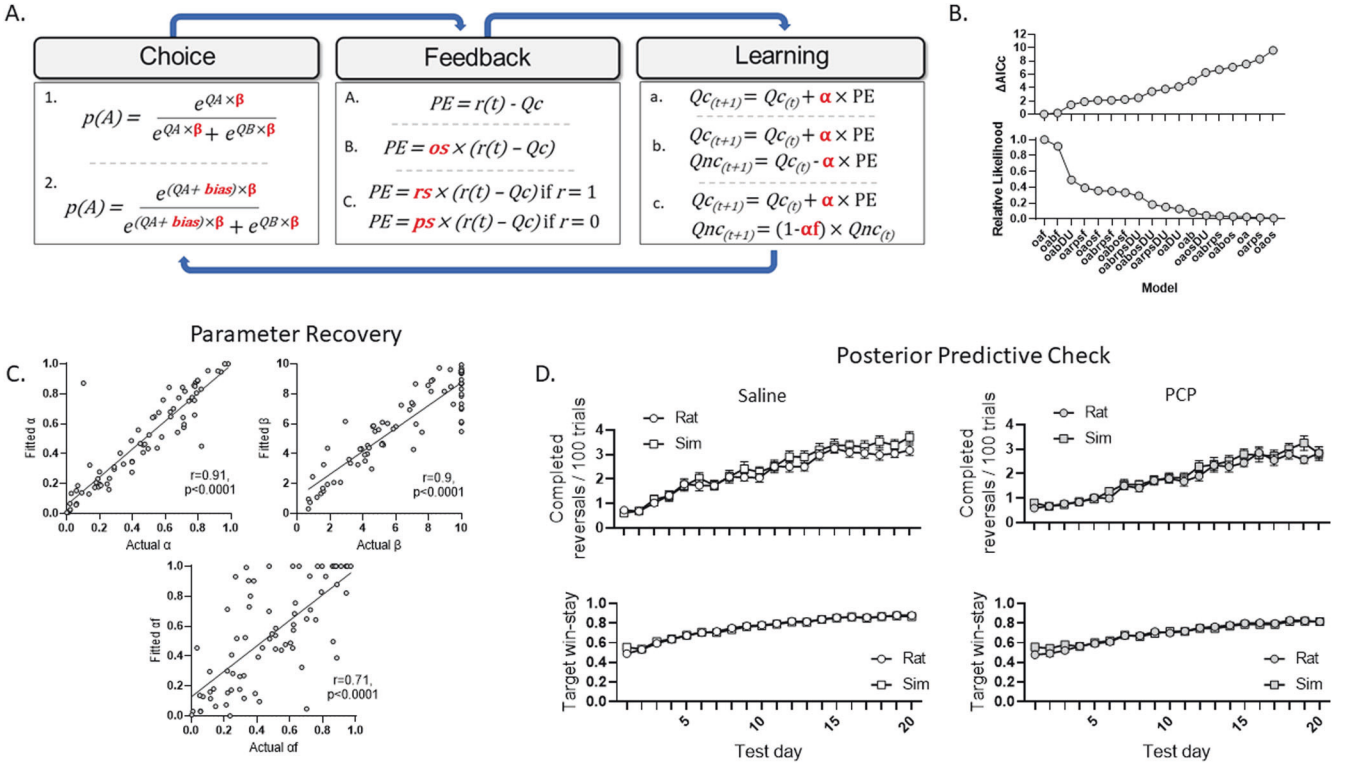


Fig. 2 Validation of Q-learning model. A Our modeling approach investigated variations within three distinct components related to decision making: choice, feedback, and learning. In total, 18 separate models were built by combing each of the options from each of the three different components. The free parameters that were estimated (panel **A**, red) include β and bias (choice component), outcome sensitivity, reward sensitivity, and punishment sensitivity (*os*, *rs*, *ps*, respectively; feedback component), and learning rate and forget rate (α and *af*, respectively; learning component). **B** We fit each model to the behavior and calculated the ΔAICc value and the relative likelihood score. The two most plausible models for our data included the one-alpha forget model (free parameters: α, β, forget) or the one-alpha bias forget model (α, β, bias, forget). **C** Using known free-parameter values, the one-alpha forget model was used to simulate PRL performance 75 times and the free parameters were then recovered. Using Pearson's correlation, we found a significant positive correlation between the actual and recovered α, β, and forget values. **D** Using the parameter estimates obtained from the rodent data, we then simulated PRL performance. This posterior predictive check revealed good correspondence between actual and simulated data for the number of completed reversals and target win-stay behavior, in both saline- and PCP-treated rats.

scores [45] (Fig. 2B and Supplementary Table 3). The best-fitting model included α, β, and forget parameters (termed the one-alpha with forget parameter model).

To determine whether we could accurately estimate the parameter values for the winning model, we performed parameter recovery by simulating PRL performance using known parameter values (Fig. 2C). Recovery of the α ($r = 0.91$, $p < 0.0001$) and β ($r = 0.9$, $p < 0.0001$) values was excellent. Recovery of the forget parameter was less robust, but the correlation between the actual and fitted *af* values was significant ($r = 0.71$, $p < 0.0001$). Finally, we performed a posterior predictive check by simulating PRL performance using the parameter values obtained from the fitted model on each day of testing. This posterior predictive check revealed a close correspondence between actual and simulated data for two key variables not explicitly included in the modeling—namely, the number of completed reversals and the extent of target win-stay responding (Fig. 2D). Thus, the one-alpha model with forget parameter was able to capture key aspects of the data, indicating that analysis of its parameters should provide insight into underlying behavioral mechanisms that drive PRL performance.

To this end, we analyzed the learning rate (α parameter) and found that it increased across the testing period [$F_{(19,836)} = 72.03$, $p < 0.001$] and was reduced in PCP-treated rats [$F_{(1,44)} = 16.97$, $p < 0.001$]. Moreover, a treatment × day interaction was evident [$F_{(19,836)} = 2.73$, $p < 0.001$] (Fig. 3A). Relative to saline-treated rats, the learning rate was reduced in PCP-treated rats from day five onwards ($p < 0.05$). By contrast, the β parameter decreased across test days [$F_{(19,836)} = 32.66$, $p < 0.001$] indicating that a greater tendency to engage in exploration was associated with improved PRL performance (Fig. 3B). In addition, a main effect of treatment revealed that the β parameter was elevated in PCP-treated rats [$F_{(1,44)} = 4.88$, $p < 0.05$], indicating an increase in exploitation. Moreover, analysis of β also revealed a treatment × day interaction [$F_{(19,836)} = 1.62$, $p < 0.05$]. Post hoc comparisons revealed that β was elevated in PCP-treated rats, relative to saline-treated rats, on days 1, 4, 5, 6, and 8. Finally, although the forget parameter increased across test days [$F_{(19,836)} = 16.84$, $p < 0.001$], there was only a trend toward a main effect of treatment [$F_{(1,44)} = 2.83$, $p = 0.09$] (Fig. 3C). No other main or interactive effect was significant [Fs < 1.03].

To illustrate how these effects affected the value attributed to each stimulus, we plotted the *Q* value for the left and right apertures from two representative animals. During the earlier test day (Fig. 3D), the value for each action changed very gradually over multiple trials, corresponding to the low learning rate. By contrast, during the later test day (Fig. 3E)—when the learning rate was increased—the value for each stimulus rapidly alternated when a reversal was triggered. Notably, this effect was blunted in the PCP-treated rat due to their relatively reduced learning rates. Furthermore, these plots highlight that the value of an action decays when it is not selected. This effect is mediated by the forget parameter decaying the value of an action in each trial that it is unselected. Hence, although the performance of PCP-treated rats improved with training (i.e., increased reversals, Fig. 1B), their

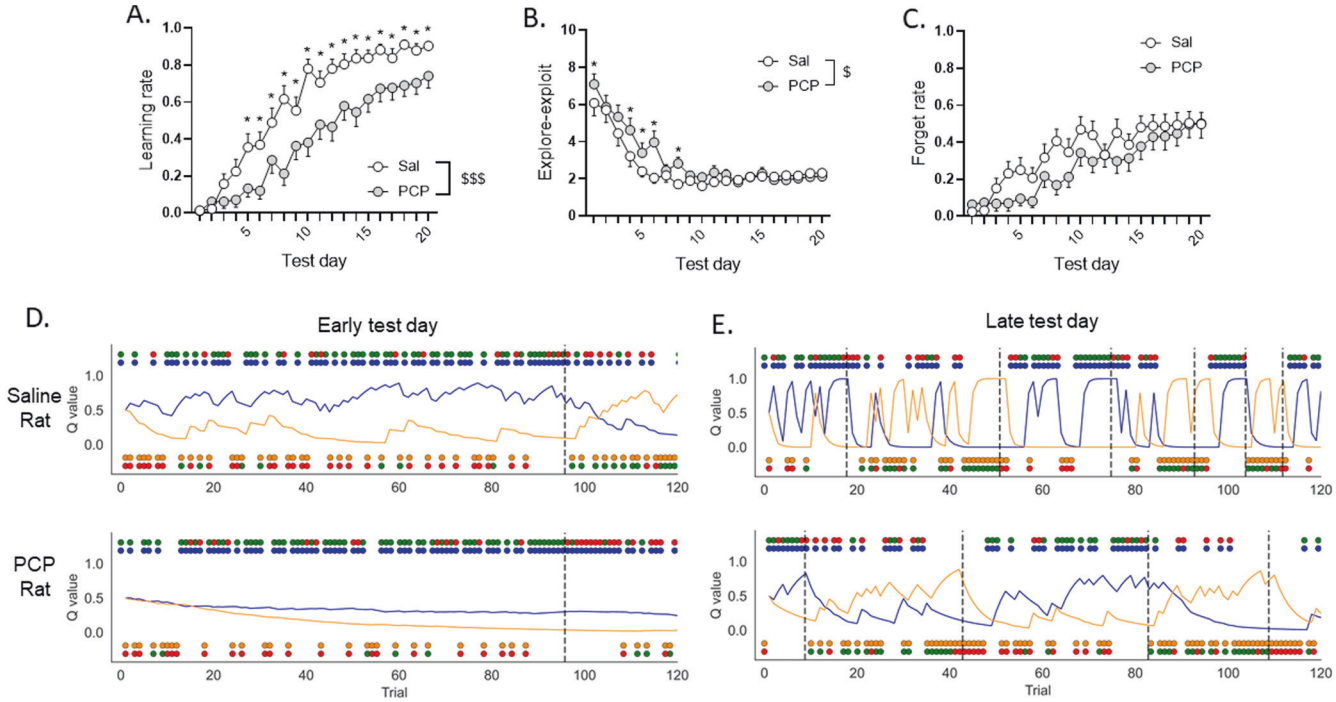


Fig. 3 Postnatal PCP-treatment disrupts Q-learning. A The learning rate increased throughout the PRL testing period, but the increase was weaker in PCP-treated vs. saline-treated rats. **B** As testing progressed, rats showed a greater tendency to explore lower-valued options (i.e., β value reduced over test days). During the first half of testing, PCP-treated rats exhibited a greater tendency to exploit higher-valued actions, but this effect was transient and only evident for several days. **C** The rate at which the value for the non-chosen action decayed increased with training. However, there was no significant difference between treatment groups. **D**, **E** To demonstrate how the changes in α, β, and the forget parameters may influence value assignment and choice behavior we plotted how the value attributed to the left and right nose-pokes (blue and orange lines, respectively) evolved throughout a session. Blue circles above the plot denote a left nose-poke response whereas orange circles below the plot denote a right nose-poke response. Green or red circles indicate whether the response was rewarded or not, respectively. **D** During an earlier test day (day 3), although both rats reached the criteria and triggered a reversal (black vertical dashed line) the low learning rate in both saline and PCP-treated rats resulted in a more gradual change in action value. This effect was also more pronounced in the PCP-treated rat. **E** During one of the later test days (day 18), the value attributed to each action was subjected to more rapid swings in response to reversals, an effect that resulted from an increase in the learning rate. Sal, $n = 24$. PCP, $n = 24$. *$p < 0.05$. $^{\$}$ denotes main effect of treatment, $p < 0.05$, $^{\$\$\$}$ denotes main effect of treatment, $p < 0.001$.

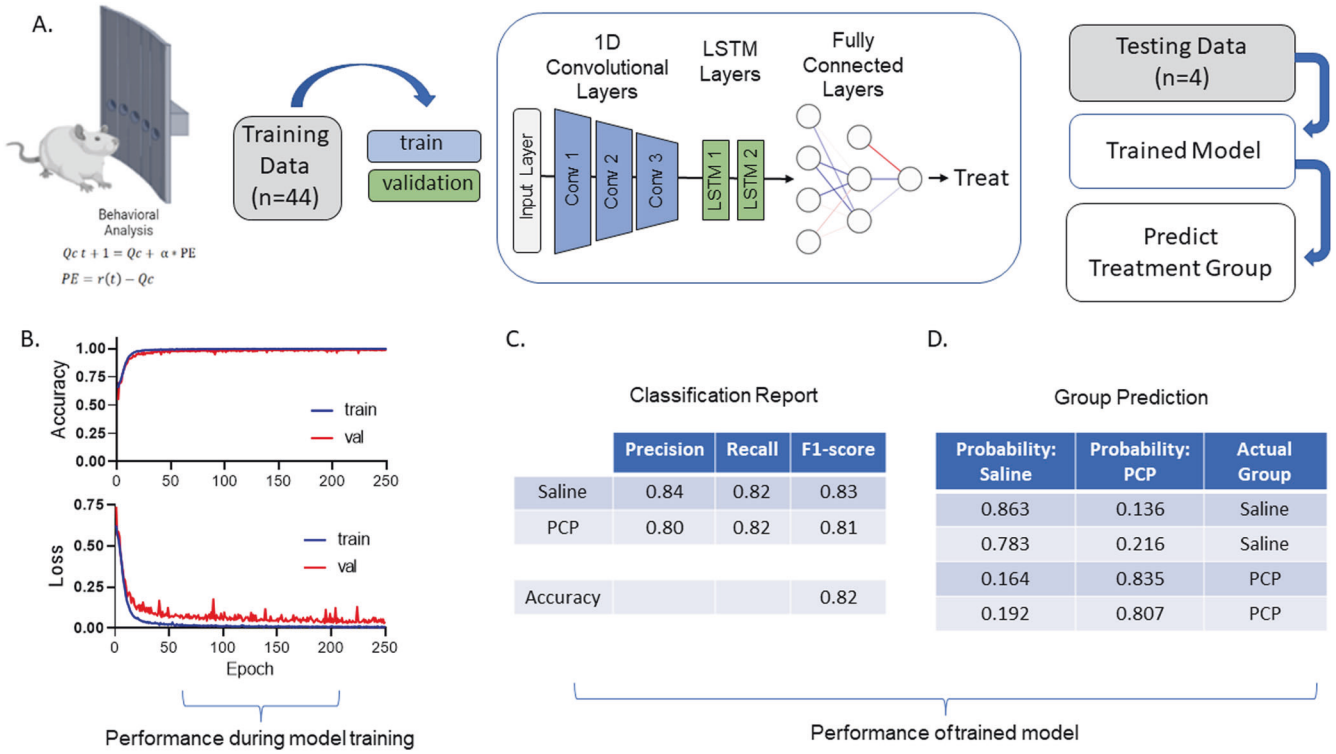
relative inability to rapidly update value estimates in response to a contingency reversal led to suboptimal performance vs. what was observed in the saline-treated controls.

Indeed, plotting the value of the selected action for trials surrounding a reversal revealed that PCP-treated rats were slower to adapt in the trials immediately after the reversal (Supplementary Fig. 3A). After the initial reduction in value (due to responding to the new non-target stimulus), saline-treated rats began selecting the stimulus with a higher value (i.e., the new target). PEs are integral to value updating [37] and were also modulated in response to the contingency reversal (Supplementary Fig. 3B). On trials immediately after the reversal, saline-treated rats showed an adaptive pattern whereby negative PEs typically predicted a response switch on the next trial, whereas positive PEs typically led to the same action being repeated (Supplementary Fig. 3C). These effects were blunted in PCP-treated rats. Moreover, the mean PEs elicited during repeat and switch trials were correlated with between-subject differences in overall target win-stay responding (Supplementary Fig. 3D). Tentatively, these data imply that PCP-induced deficits in PRL potentially result from aberrant PE signaling.

Finally, we used the trial-by-trial behavior (specific variables described in Supplementary Table 2) to train a DNN using a 44-subject training dataset to predict the treatment group of the 4-subject test dataset (Fig. 4A). The training dataset was split into training (75%) and validation (25%) sets and the model was trained for 250 epochs in batches of 500 trials. The compiled network had a total of 325,038 trainable parameters throughout all the node and layers. The model evaluates each batch of data, predicts the treatment group, and then updates the weights and biases throughout the network via backpropagation to improve its prediction of the next batch. At the end of each epoch, model performance is evaluated using the validation set. Throughout this process, the accuracy of the predictions for both the training and validation sets increased and the model loss (i.e., the degree to which the predictions were incorrect) decreased (Fig. 4B). Once trained, we then evaluated the performance using the test dataset ($n = 4$) to determine the model's ability to predict the treatment group of new subjects. The classification report revealed that the overall accuracy was 82%. Moreover, the precision, recall, and f1-score ranged between 80 and 84% (Fig. 4C). Finally, we generated the group probability for each individual subject. The predicted probability that each subject belonged to the correct treatment group ranged between 0.783 and 0.863 (Fig. 4D), demonstrating that the trained model accurately predicted the treatment group of each subject in the test dataset.

DISCUSSION

The present findings demonstrate that disrupting normal glutamatergic transmission during early postnatal neurodevelopment leads to PRL impairments in adulthood. Our computational analysis revealed that these deficits were accompanied by a reduced learning rate. Finally, these behavioral alterations were sufficient to train a deep neural network (DNN) that accurately predicted the treatment group of subjects.



	Precision	Recall	F1-score
Saline	0.84	0.82	0.83
PCP	0.80	0.82	0.81
Accuracy			0.82

Probability: Saline	Probability: PCP	Actual Group
0.863	0.136	Saline
0.783	0.216	Saline
0.164	0.835	PCP
0.192	0.807	PCP

Fig. 4 Deep neural network (DNN) predicts treatment group. A Schematic of approach and network architecture. The PRL behavioral data were analyzed and split into training ($n = 44$ rats) and testing set ($n = 4$ rats). For training of the model, the training set was split into a train and validation set (75:25 split) and passed into the DNN. The first layers consisted of one-dimensional convolutional layers. The output from the last convolutional layer was pooled, flattened, and then passed to two long short-term memory (LSTM) layers, which in turn passed its output to two fully connected dense layers. The output of the DNN consisted of two nodes and used a softmax activation function to predict the treatment group of the subject. Once the model was trained, the four subjects set aside were then used to evaluate the generalizability of the model by predicting the treatment group of new rats. **B** The model was trained in batches of 500 trials for a total of 250 epochs. The training and validation accuracy increased across epochs. In addition, the model loss for both the training and validation sets decreased across epochs. **C** Using the trained model to evaluate the performance of the four unseen rats we generated a classification report. An overall accuracy of 82% was achieved, with similarly high precision, recall, and f1-scores. **D** The predicted probability scores for individual subjects ranged from 78 to 86%.

PRL deficits in schizophrenia are characterized by impairments in reward sensitivity that result in excessive switching between stimuli [5, 11, 13, 46, 47]. We observed a similar impairment profile after postnatal PCP-treatment; rats completed fewer reversals and exhibited a pronounced reduction in win-stay responses. In addition, we also observed a reduction in lose-shift responding in PCP-treated rats, thereby reducing the tendency to switch choices that were not rewarded. Interestingly, PRL performance (i.e., completed reversals) was only correlated with correct win-stay responses. Hence, despite the observation that PCP-induced alterations were not valence-specific, the overall impairment in PRL performance evident in postnatal PCP-treated rats likely resulted from the reduced willingness or ability to repeat rewarded actions. Administration of PCP during adulthood is another commonly used approach to model aspects of schizophrenia in experimental rodents [48]. Sub-chronic PCP-treatment in adulthood has been shown to either impair [49] or have no effect [50] on reversal learning. Notably, however, these studies used a deterministic variant of the reversal learning task. The discrepancy in behavioral alterations between these studies and our findings likely results from differences in the dosing regimen (i.e., early postnatal vs. adult acute or sub-chronic) or behavioral task (i.e., deterministic vs. probabilistic reversal learning). Interestingly, social isolation rearing [35] or maternal immune activation [51] also impair PRL performance. Like postnatal PCP-treatment, these two manipulations model the neurodevelopmental origins of schizophrenia [24].

The improvement in task performance evident in both saline- and PCP-treated rats as training progressed was associated with an increased learning rate. Notably, however, the learning rate was reduced in PCP-treated vs. saline-treated rats. Moreover, during the initial test days—when performance was poor—rats tended to engage in more exploitative choices and this tendency was initially greater in PCP-treated rats. Initially, this sounds counter-intuitive as one may assume that a bias toward exploitation would be beneficial. However, the preference for exploitative choices was evident when the learning rate was comparatively low. Importantly, this combination (low α, high β) would contribute to rigid, inflexible behavior [52]. The values of the target and non-target stimuli need to update rapidly after a reversal, but with a low α this updating occurs too gradually, which causes the value for the former target stimulus to remain elevated for too long post-reversal. When paired with a high β (i.e., exploit the high-valued action), the subject would persistently choose the former target stimulus for too many trials since its value would remain elevated. As training progressed, the β parameter was reduced indicating that action selection had shifted toward exploration and the difference between treatment groups diminished. However, the reduction in the learning rate evident in PCP-treated rats persisted throughout testing.

It has long been established that PEs highlight whether an outcome was better or worse than expected [53]. Moreover, PE signaling is disrupted in schizophrenia [8] and in human subjects administered NMDA receptor antagonists [54, 55]. Normalizing PE

signaling in schizophrenia patients, via transcranial stimulation, attenuated learning deficits [56]. For the trials immediately after a reversal, we found that responses eliciting a positive PE were likely to be repeated whereas responses that elicited a negative PE were likely to be avoided on the next trial. The PEs for these trials were correlated with the propensity to repeat rewarded target responses, consistent with the notion that PEs serve as a teaching signal [9] and guide the subject to make the appropriate decision when reward contingencies changed. These considerations point to a fundamental role for abnormal PEs in the performance deficits observed in the PCP group. However, two important caveats must be noted. First, our study is strictly behavioral and thus we cannot make strong claims about underlying neural mechanisms. It is well known that changes in dopamine cell firing represent the neurophysiological marker of reward PEs [9, 53], and although several studies have evaluated PEs using behavior alone [57–60], many studies have successfully related PEs to activity within the dopaminergic mesolimbic pathway [12, 41, 61–63]. Consequently, we suggest that the performance reductions we observed in the PCP group may reflect aberrant activity of dopaminergic networks; the fact that perturbation of early postnatal NMDA receptor signaling also disrupts dopaminergic neurotransmission [64–66] is consistent with this suggestion. However, this interpretation is speculative given the absence of neurophysiology in the current study.

Second, although we found that postnatal PCP-treatment-induced alterations in PEs may have influenced subsequent choices, dissociating these alterations from changes in the learning rate is challenging, and indeed—as emphasized above—we found that learning rates were reduced in the PCP group. On the one hand, learning rates and PEs are dissociable in the sense that learning rates —as fit in the models we used—are stable within a session whereas PEs change on a trial-by-trial basis. Moreover, as just discussed, there is extensive evidence linking trial-level fluctuations in mesolimbic dopamine signaling to PEs, whereas the neural systems that control individual differences in learning rates are less well-understood. Critically, however, PEs reflect the difference between outcomes and the expected values of actions, and because learning rates control the rate for which value estimates are updated, they also affect PE sign and magnitude—thus, these two measures are not fully dissociable in the current study. Therefore, although alterations in PEs may influence subsequent choices and win-stay performance, these changes may result from a downstream consequence of the PCP-induced reduction in the learning rate; additional work is needed to disentangle these two mechanisms. Encouragingly, EEG measures of PEs during PRL tasks have been obtained in humans and rodents [67] and are similarly pharmacological sensitive across species [68], thus the PRL task may continue to provide a useful method for pursuing these more complex questions.

Deep learning approaches are becoming more common in psychiatry [20, 21, 69] and have been applied to solve a variety of complex problems from predicting treatment efficacy [70, 71] to assisting with diagnostics [72, 73]. Deep learning algorithms can identify unique features within large, complex datasets [21], and have been successful in classifying schizophrenia patients using neuroimaging data [74]. However, the rising costs [75] and the requirement for specialized technical staff [76] associated with neuroimaging preclude their routine use during a psychiatric evaluation. Using our behavioral dataset, we were able to train a DNN to classify the treatment group of new subjects with approximately 82% accuracy, and we correctly predicted each treatment group with a high probability. Importantly, this classification accuracy is consistent with previous studies that used neuroimaging data [77–81]. Hence, our findings demonstrate the potential utility for deep learning approaches to assist in diagnostics using behavioral data easily obtained using a simple computer-based task.

This work is not without limitations that warrant discussion. First, although the one-alpha model with a forget parameter used here was able to capture the key behavioral processes underlying PRL performance, it is possible that another model would provide an even better fit and potentially lead to different conclusions. There are many model variants that could be used to evaluate PRL performance, such as Q-learning models with separate learning rates for reward vs. non-rewards [82], Q-learning models with a dynamic learning rate [83–85], actor-critic models [86], or hybrid models incorporating elements of Q-learning and actor-critic models [60]. Hence, in addition to exploring a variety of Q-learning variants (as we have done, here), future investigations that evaluate conceptually distinct reinforcement learning models may provide additional insight into behavior. For example, the one-alpha model with the forget parameter had a fixed learning rate that precluded the dissociation of the α parameter from trial-by-trial PEs. Using a Q-learning model with a dynamic learning rate [83, 85] while also measuring dopaminergic activity may provide a method to disentangle changes in PEs from the learning rate and provide confirmation of the hypothesis that PCP-induced alterations in PE signaling play a central role in dysfunctional PRL.

Second, while our DNN accurately classified saline- vs. PCP-treated rats, these remain rodent outcomes. Thus, future work is required to confirm whether this approach can achieve a similar accuracy using schizophrenia patients. Another avenue for future work involves mental disorders other than schizophrenia. PRL performance is disrupted in several psychiatric conditions, including major depression [87], obsessive-compulsive disorder [88], and substance use disorder [89]. Since the pattern of PRL deficits appears to vary somewhat across disorders, these differences open the possibility that deep learning approaches could identify these differences and assist with diagnostic classification. Moreover, in addition to the PRL task, other tasks that evaluate distinct constructs such as risky or effortful decision making, executive functioning, or cognitive control are impaired in several mental illnesses [90–95]. Training a DNN on behavior obtained from a battery of tasks would undoubtedly improve the robustness of diagnostic classifications based on behavioral profiles. Third, due to their nature, when using deep learning algorithms it is unclear exactly what combination of parameters helps to discriminate between groups [69]. The initial layers of our model were one-dimensional convolutional layers that would have likely learned how the relationship between each of the variables differed between groups for each trial [96], whereas the subsequent LSTM layers would have likely captured how these relationships varied across trials [97]. Future research involving feature importance could help delineate which of the behavioral parameters are essential for detecting the differences that exist between groups.

In conclusion, our findings show that disrupting early postnatal glutamate transmission in rats leads to impaired PRL performance. Although pinpointing the exact neurocognitive mechanisms that most strongly contribute to these impairments remains a key goal, the differences in behavior were sufficient to train a DNN to accurately predict the treatment group of rodents in a test dataset. Consequently, combining PRL with a DNN has the potential to ultimately assist with the initial diagnosis of psychiatric conditions in a manner more accessible than previous DNN approaches.

REFERENCES

1. 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 Ser B Biol Sci. 2007;362:933–42.
2. Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ. Cortical substrates for exploratory decisions in humans. Nature. 2006;441:876–9.
3. Ebitz RB, Albarran E, Moore T. Exploration disrupts choice-predictive signals and alters dynamics in prefrontal cortex. Neuron. 2018;97:450–61.e9
4. Murray EA, Rudebeck PH. Specializations for reward-guided decision-making in the primate ventral prefrontal cortex. Nat Rev Neurosci. 2018;19:404–17.
5. Culbreth AJ, Westbrook A, Xu Z, Barch DM, Waltz JA. Intact ventral striatal prediction error signaling in medicated schizophrenia patients. Biol Psychiatry Cogn Neurosci Neuroimaging. 2016;1:474–83.

6. Hernaus D, Xu Z, Brown EC, Ruiz R, Frank MJ, Gold JM, et al. Motivational deficits in schizophrenia relate to abnormalities in cortical learning rate signals. Cogn Affect Behav Neurosci. 2018;18:1338–51.
7. Ermakova AO, Knolle F, Justicia A, Bullmore ET, Jones PB, Robbins TW, et al. Abnormal reward prediction-error signalling in antipsychotic naive individuals with first-episode psychosis or clinical risk for psychosis. Neuropsychopharmacology. 2018;43:1691–99.
8. Yaple ZA, Tolomeo S, Yu R. Abnormal prediction error processing in schizophrenia and depression. Hum Brain Mapp. 2021;42:3547–60.
9. Nasser HM, Calu DJ, Schoenbaum G, Sharpe MJ. The dopamine prediction error: contributions to associative models of reward learning. Front Psychol. 2017;8:244.
10. Schultz W. Dopamine reward prediction-error signalling: a two-component response. Nat Rev Neurosci. 2016;17:183–95.
11. Schlagenhauf F, Huys QJM, Deserno L, Rapp MA, Beck A, Heinze H-J, et al. Striatal dysfunction during reversal learning in unmedicated schizophrenia patients. NeuroImage. 2014;89:171–80.
12. Katthagen T, Kaminski J, Heinz A, Buchert R, Schlagenhauf F. Striatal dopamine and reward prediction error signaling in unmedicated schizophrenia patients. Schizophr Bull. 2020;46:1535–46.
13. Reddy LF, Waltz JA, Green MF, Wynn JK, Horan WP. Probabilistic reversal learning in schizophrenia: stability of deficits and potential causal mechanisms. Schizophr Bull. 2016;42:942–51.
14. Strauss GP, Datta R, Armstrong W, Raugh IM, Kraguljac NV, Lahti AC. Reinforcement learning abnormalities in the attenuated psychosis syndrome and first episode psychosis. Eur Neuropsychopharmacol. 2021;47:11–19.
15. Chang WC, Waltz JA, Gold JM, Chan TCW, Chen EYH. Mild reinforcement learning deficits in patients with first-episode psychosis. Schizophr Bull. 2016;42:1476–85.
16. McGorry PD, Killackey E, Yung A. Early intervention in psychosis: concepts, evidence and future directions. World Psychiatry. 2008;7:148–56.
17. Norman RM, Malla AK, Verdi MB, Hassall LD, Fazekas C. Understanding delay in treatment for first-episode psychosis. Psychological Med. 2004;34:255–66.
18. Tzur Bitan D, Grossman Giron A, Alon G, Mendlovic S, Bloch Y, Segev A. Attitudes of mental health clinicians toward perceived inaccuracy of a schizophrenia diagnosis in routine clinical practice. BMC Psychiatry. 2018;18:317.
19. Frances A. The past, present and future of psychiatric diagnosis. World Psychiatry. 2013;12:111.
20. Quaak M, van de Mortel L, Thomas RM, van Wingen G. Deep learning applications for the classification of psychiatric disorders using neuroimaging data: systematic review and meta-analysis. NeuroImage: Clin. 2021;30:102584.
21. Chen L, Xia C, Sun H. Recent advances of deep learning in psychiatric disorders. Precis Clin Med. 2020;3:202–13.
22. Kalmady SV, Greiner R, Agrawal R, Shivakumar V, Narayanaswamy JC, Brown MRG, et al. Towards artificial intelligence in mental health by improving schizophrenia prediction with multiple brain parcellation ensemble-learning. npj Schizophr. 2019;5:2.
23. Chou P-H, Yao Y-H, Zheng R-X, Liou Y-L, Liu T-T, Lane H-Y, et al. Deep neural network to differentiate brain activity between patients with first-episode schizophrenia and healthy individuals: a multi-channel near infrared spectroscopy study. Front Psychiatry. 2021;12:655292.
24. Fatemi SH, Folsom TD. The neurodevelopmental hypothesis of schizophrenia, revisited. Schizophr Bull. 2009;35:528–48.
25. Grayson B, Barnes SA, Markou A, Piercy C, Podda G, Neill JC. Postnatal phencyclidine (PCP) as a neurodevelopmental animal model of schizophrenia pathophysiology and symptomatology: a review. Curr Top Behav Neurosci. 2015;29:403–28.
26. Lewis DA, Curley AA, Glausier JR, Volk DW. Cortical parvalbumin interneurons and cognitive dysfunction in schizophrenia. Trends Neurosci. 2012;35:57–67.
27. Kaalund SS, Riise J, Broberg BV, Fabricius K, Karlsen AS, Secher T, et al. Differential expression of parvalbumin in neonatal phencyclidine-treated rats and socially isolated rats. J Neurochem. 2013;124:548–57.
28. Kjaerby C, Broberg BV, Kristiansen U, Dalby NO. Impaired GABAergic inhibition in the prefrontal cortex of early postnatal phencyclidine (PCP)-treated rats. Cereb Cortex. 2014;24:2522–32.
29. Wang C, McInnis J, Ross-Sanchez M, Shinnick-Gallagher P, Wiley JL, Johnson KM. Long-term behavioral and neurodegenerative effects of perinatal phencyclidine administration: implications for schizophrenia. Neuroscience. 2001;107:535–50.
30. Terranova JP, Chabot C, Barnouin MC, Perrault G, Depoortere R, Griebel G, et al. SSR181507, a dopamine D(2) receptor antagonist and 5-HT(1A) receptor agonist, alleviates disturbances of novelty discrimination in a social context in rats, a putative model of selective attention deficit. Psychopharmacology (Berl). 2005;181:134–44.
31. Broberg BV, Dias R, Glenthøj BY, Olsen CK. Evaluation of a neurodevelopmental model of schizophrenia—early postnatal PCP treatment in attentional set-shifting. Behav Brain Res. 2008;190:160–63.
32. Broberg BV, Glenthoj BY, Dias R, Larsen DB, Olsen CK. Reversal of cognitive deficits by an ampakine (CX516) and sertindole in two animal models of schizophrenia—sub-chronic and early postnatal PCP treatment in attentional set-shifting. Psychopharmacology (Berl). 2009;206:631–40.
33. Amitai N, Markou A. Disruption of performance in the five-choice serial reaction time task induced by administration of N-Methyl-D-aspartate receptor antagonists: relevance to cognitive dysfunction in schizophrenia. Biol Psychiatry. 2010;68:5–16.
34. Bari A, Theobald DE, Caprioli D, Mar AC, Aidoo-Micah A, Dalley JW, et al. Serotonin modulates sensitivity to reward and negative feedback in a probabilistic reversal learning task in rats. Neuropsychopharmacology. 2010;35:1290–301.
35. Amitai N, Young JW, Higa K, Sharp RF, Geyer MA, Powell SB. Isolation rearing effects on probabilistic learning and cognitive flexibility in rats. Cogn Affect Behav Neurosci. 2014;14:388–406.
36. Dalton GL, Wang NY, Phillips AG, Floresco SB. Multifaceted contributions by different regions of the orbitofrontal and medial prefrontal cortex to probabilistic reversal learning. J Neurosci. 2016;36:1996–2006.
37. Sutton RS, Barto AG. Reinforcement learning: an introduction. MIT Press; 2018.
38. Correa CMC, Noorman S, Jiang J, Palminteri S, Cohen MX, Lebreton M, et al. How the level of reward awareness changes the computational and electrophysiological signatures of reinforcement learning. J Neurosci. 2018;38:10338–48.
39. 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 USA. 2007;104:16311–16.
40. Parker NF, Baidya A, Cox J, Haetzel LM, Zhukovskaya A, Murugan M, et al. Choice-selective sequences dominate in cortical relative to thalamic inputs to NAc to support reinforcement learning. Cell Rep. 2022;39:110756.
41. Parker NF, Cameron CM, Taliaferro JP, Lee J, Choi JY, Davidson TJ, et al. Reward and choice encoding in terminals of midbrain dopamine neurons depends on striatal target. Nat Neurosci. 2016;19:845–54.
42. Hattori R, Danskin B, Babic Z, Mlynaryk N, Komiyama T. Area-specificity and plasticity of history-dependent value coding during learning. Cell. 2019;177:1858–72.e15
43. Huys QJ, Pizzagalli DA, Bogdan R, Dayan P. Mapping anhedonia onto reinforcement learning: a behavioural meta-analysis. Biol Mood Anxiety Disord. 2013;3:1–16.
44. Katahira K. The relation between reinforcement learning parameters and the influence of reinforcement history on choice behavior. J Math Psychol. 2015;66:59–69.
45. Burnham KP, Anderson DR, Huyvaert KP. AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. Behav Ecol Sociobiol. 2011;65:23–35.
46. Waltz JA, Gold JM. Probabilistic reversal learning impairments in schizophrenia: further evidence of orbitofrontal dysfunction. Schizophr Res. 2007;93:296–303.
47. Waltz JA, Kasanova Z, Ross TJ, Salmeron BJ, McMahon RP, Gold JM, et al. The roles of reward, default, and executive control networks in set-shifting impairments in schizophrenia. PLoS One. 2013;8:e57257.
48. Neill JC, Barnes S, Cook S, Grayson B, Idris NF, McLean SL, et al. Animal models of cognitive dysfunction and negative symptoms of schizophrenia: focus on NMDA receptor antagonism. Pharm Ther. 2010;128:419–32.
49. Abdul-Monim Z, Neill JC, Reynolds GP. Sub-chronic psychotomimetic phencyclidine induces deficits in reversal learning and alterations in parvalbumin-immunoreactive expression in the rat. J Psychopharmacol (Oxf, Engl). 2007;21:198–205.
50. Brigman JL, Ihne J, Saksida LM, Bussey TJ, Holmes A. Effects of Subchronic Phencyclidine (PCP) Treatment on Social Behaviors, and Operant Discrimination and Reversal Learning in C57BL/6J Mice. Front Behav Neurosci. 2009;3:2
51. Amodeo DA, Lai CY, Hassan O, Mukamel EA, Behrens MM, Powell SB. Maternal immune activation impairs cognitive flexibility and alters transcription in frontal cortex. Neurobiol Dis. 2019;125:211–18.
52. Bari BA, Moerke MJ, Jedema HP, Effinger DP, Cohen JY, Bradberry CW. Reinforcement learning modeling reveals a reward-history-dependent strategy underlying reversal learning in squirrel monkeys. Behav Neurosci. 2022;136:46–60
53. Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. Science. 1997;275:1593–99.
54. Corlett PR, Honey GD, Fletcher PC. Prediction error, ketamine and psychosis: an updated model. J Psychopharmacol. 2016;30:1145–55.
55. Corlett PR, Honey GD, Aitken MRF, Dickinson A, Shanks DR, Absalom AR, et al. Frontal responses during learning predict vulnerability to the psychotogenic effects of ketamine: linking cognition, brain activity, and psychosis. Arch Gen Psychiatry. 2006;63:611–21.
56. Reinhart RMG, Zhu J, Park S, Woodman GF. Medial–frontal stimulation enhances learning in schizophrenia by restoring prediction error signaling. J Neurosci. 2015;35:12232–40.
57. Jang AI, Nassar MR, Dillon DG, Frank MJ. Positive reward prediction errors during decision-making strengthen memory encoding. Nat Hum Behav. 2019;3:719–32.
58. Bai Y, Katahira K, Ohira H. Dual learning processes underlying human decision-making in reversal learning tasks: functional significance and evidence from the model fit to human behavior. Front Psychol. 2014;5:871.

59. Cinotti F, Fresno V, Aklil N, Coutureau E, Girard B, Marchand AR, et al. Dopamine blockade impairs the exploration-exploitation trade-off in rats. Sci Rep. 2019;9:6770.
60. Strauss GP, Thaler NS, Matveeva TM, Vogel SJ, Sutton GP, Lee BG, et al. Predicting psychosis across diagnostic boundaries: behavioral and computational modeling evidence for impaired reinforcement learning in schizophrenia and bipolar disorder with a history of psychosis. J Abnorm Psychol. 2015;124:697–708.
61. Oyama K, Hernádi I, Iijima T, Tsutsui K-I. Reward prediction error coding in dorsal striatal neurons. J Neurosci. 2010;30:11447–57.
62. Cooper JC, Dunne S, Furey T, O'Doherty JP. Human dorsal striatum encodes prediction errors during observational learning of instrumental actions. J Cogn Neurosci. 2012;24:106–18.
63. D'Ardenne K, McClure SM, Nystrom LE, Cohen JD. BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. Science. 2008;319:1264–67.
64. Semba J, Tanaka N, Wakuta M, Suhara T. Neonatal phencyclidine treatment selectively attenuates mesolimbic dopamine function in adult rats as revealed by methamphetamine-induced behavior and c-fos mRNA expression in the brain. Synapse. 2001;40:11–8.
65. Nakao K, Jeevakumar V, Jiang SZ, Fujita Y, Diaz NB, Pretell Annan CA, et al. Schizophrenia-like dopamine release abnormalities in a mouse model of NMDA receptor hypofunction. Schizophr Bull. 2019;45:138–47.
66. Uehara T, Sumiyoshi T, Seo T, Matsuoka T, Itoh H, Suzuki M, et al. Neonatal exposure to MK-801, an N-methyl-d-aspartate receptor antagonist, enhances methamphetamine-induced locomotion and disrupts sensorimotor gating in pre- and postpubertal rats. Brain Res. 2010;1352:223–30.
67. Cavanagh JF, Gregg D, Light GA, Olguin SL, Sharp RF, Bismark AW, et al. Electrophysiological biomarkers of behavioral dimensions from cross-species paradigms. Transl Psychiatry. 2021;11:482.
68. Cavanagh JF, Olguin SL, Talledo JA, Kotz JE, Roberts BZ, Nungaray JA, et al. Amphetamine alters an EEG marker of reward processing in humans and mice. Psychopharmacology (Berl). 2022;239:923–33.
69. Koppe G, Meyer-Lindenberg A, Durstewitz D. Deep learning for small and big data in psychiatry. Neuropsychopharmacology. 2021;46:176–90.
70. Lin E, Kuo PH, Liu YL, Yu YW, Yang AC, Tsai SJ. A deep learning approach for predicting antidepressant response in major depression using clinical and genetic biomarkers. Front Psychiatry. 2018;9:290.
71. Jaworska N, de la Salle S, Ibrahim MH, Blier P, Knott V. Leveraging machine learning approaches for predicting antidepressant treatment response using electroencephalography (EEG) and clinical data. Front Psychiatry. 2018;9:768.
72. Oh J, Oh B-L, Lee K-U, Chae J-H, Yun K. Identifying schizophrenia using structural MRI with a deep learning algorithm. Fron Psychiatry. 2020;11:16.
73. Sun J, Cao R, Zhou M, Hussain W, Wang B, Xue J, et al. A hybrid deep neural network for classification of schizophrenia using EEG Data. Sci Rep. 2021;11:4706.
74. Cortes-Briones JA, Tapia-Rivas NI, D'Souza DC, Estevez PA. Going deep into schizophrenia with artificial intelligence. Schizophr Res. 2022;245:122–40.
75. Hill CE, Reynolds EL, Burke JF, Banerjee M, Kerber KA, Magliocco B, et al. Increasing out-of-pocket costs for neurologic care for privately insured patients. Neurology. 2021;96:e322–32.
76. van Beek EJR, Kuhl C, Anzai Y, Desmond P, Ehman RL, Gong Q, et al. Value of MRI in medicine: More than just another test? J Magn Reson Imaging. 2019;49:e14–25.
77. Yan W, Calhoun V, Song M, Cui Y, Yan H, Liu S, et al. Discriminating schizophrenia using recurrent neural network applied on time courses of multi-site FMRI data. EBioMedicine. 2019;47:543–52.
78. Salvador R, Canales-Rodríguez E, Guerrero-Pedraza A, Sarró S, Tordesillas-Gutiérrez D, Maristany T, et al. Multimodal integration of brain images for MRI-based diagnosis in schizophrenia. Front Neurosci. 2019;13:1203.
79. Kim J, Calhoun VD, Shim E, Lee JH. Deep neural network with weight sparsity control and pre-training extracts hierarchical features and enhances classification performance: evidence from whole-brain resting-state functional connectivity patterns of schizophrenia. Neuroimage. 2016;124:127–46. Pt A
80. Niu YW, Lin QH, Qiu Y, Kuang LD, Calhoun VD. Sample Augmentation for Classification of Schizophrenia Patients and Healthy Controls Using ICA of fMRI Data and Convolutional Neural Networks. Tenth International Conference on Intelligent Control and Information Processing. 2019:297–302. https://doi.org/10.1109/ICICIP47338.2019.9012169
81. Zhao J, Huang J, Zhi D, Yan W, Ma X, Yang X, et al. Functional network connectivity (FNC)-based generative adversarial network (GAN) and its applications in classification of mental disorders. J Neurosci Methods. 2020;341:108756.
82. Metha JA, Brian ML, Oberrauch S, Barnes SA, Featherby TJ, Bossaerts P, et al. Separating probability and reversal learning in a novel probabilistic reversal learning task for mice. Front Behav Neurosci. 2020;13:270–70.
83. Ito M, Doya K. Validation of decision-making models and analysis of decision variables in the rat basal ganglia. J Neurosci. 2009;29:9861–74.
84. Pearce JM, Hall G. A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. Psychological Rev. 1980;87:532–52.
85. Bartolo R, Averbeck BB. Prefrontal cortex predicts state switches during reversal learning. Neuron. 2020;106:1044–54.e4
86. Gold JM, Waltz JA, Matveeva TM, Kasanova Z, Strauss GP, Herbener ES, et al. Negative symptoms and the failure to represent the expected reward value of actions: behavioral and computational modeling evidence. Arch Gen Psychiatry. 2012;69:129–38.
87. Taylor Tavares JV, Clark L, Furey ML, Williams GB, Sahakian BJ, Drevets WC. Neural basis of abnormal response to negative feedback in unmedicated mood disorders. NeuroImage. 2008;42:1118–26.
88. Smith P, Benzina N, Vorspan F, Mallet L, N'Diaye K. Compulsivity and probabilistic reversal learning in OCD and cocaine addiction. Eur Psychiatry. 2015;30(Supplement):S110–11.
89. Ersche KD, Roiser JP, Robbins TW, Sahakian BJ. Chronic cocaine but not chronic amphetamine use is associated with perseverative responding in humans. Psychopharmacology (Berl). 2008;197:421–31.
90. Barnes SA, Der-Avakian A, Young JW. Preclinical models to investigate mechanisms of negative symptoms in schizophrenia. Schizophr Bull. 2017;43:706–11.
91. Kovács I, Richman MJ, Janka Z, Maraz A, Andó B. Decision making measured by the Iowa Gambling Task in alcohol use disorder and gambling disorder: a systematic review and meta-analysis. Drug Alcohol Depend. 2017;181:152–61.
92. Modirrousta M, Meek BP, Sareen J, Enns MW. Impaired trial-by-trial adjustment of cognitive control in obsessive compulsive disorder improves after deep repetitive transcranial magnetic stimulation. BMC Neurosci. 2015;16:63.
93. Dillon DG, Wiecki T, Pechtel P, Webb C, Goer F, Murray L, et al. A computational analysis of flanker interference in depression. Psychological Med. 2015;45:2333–44.
94. Beard C, Donahue RJ, Dillon DG, Van't Veer A, Webber C, Lee J, et al. Abnormal error processing in depressive states: a translational examination in humans and rats. Transl Psychiatry. 2015;5:e564.
95. Strauss GP, Whearty KM, Morra LF, Sullivan SK, Ossenfort KL, Frost KH. Avolition in schizophrenia is associated with reduced willingness to expend effort for reward on a Progressive Ratio task. Schizophr Res. 2016;170:198–204.
96. Han H, Li Y, Zhu X. Convolutional neural network learning for generic data classification. Inf Sci. 2019;477:448–65.
97. Livieris IE, Pintelas E, Pintelas P. A CNN–LSTM model for gold price time-series forecasting. Neural Comput Appl. 2020;32:17351–60.

AUTHOR CONTRIBUTIONS

SAB designed and ran the experiments, analyzed the data, and primarily wrote the manuscript. MMT assisted in writing the manuscript. SA assisted with implementing the deep neural network and with writing the manuscript. JWY assisted with writing the manuscript. DGD assisted with implementing the Q-learning analysis and writing the manuscript.

FUNDING

This work was funded by NIMH grants R01MH108653 (SAB) and R01MH111676 (DGD).

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41386-022-01514-y.

Correspondence and requests for materials should be addressed to Samuel A. Barnes.