Midbrain D_3 Receptor Availability Predicts Escalation in Cocaine Self-administration

Supplemental Information

Experimental procedures

Subjects

Rats were pair housed in a climate-controlled room and maintained on a 12-h light/dark cycle (lights on at 7am; lights off at 7pm) with access to water *ad libitum*. Rats were given four days to acclimate to the vivarium and underwent dietary restriction to 90% of their free-feeding weight. Food was provided to rats after completing their daily behavioral testing. All experimental procedures were performed as approved by the Institutional Animal Care and Use Committee at Yale University and according to NIH and institutional guidelines and the Public Health Service Policy on Humane Care and Use of Laboratory Animals.

Assessing addiction-relevant behaviors

Motivation to obtain an infusion of cocaine or saline was assessed under a progressive ratio schedule in a single 6 h session the day following completion of the self-administration procedure. The number of active lever responses required to deliver a single infusion of cocaine (or saline) increased exponentially during the session and the final schedule achieved served as the dependent measure. Animals then underwent a 6 h extinction session in which levers were extended into the operant box and responses recorded, but had no programmed consequence. The ability of the drug-paired cues to reinstate drug-seeking behaviors in a single 1 h session was then assessed. During the

reinstatement sessions responses on the active lever resulted in the delivery of the 10 s compound cue previously associated with a cocaine (or saline) infusion and responses to the active and inactive lever recorded.

PET imaging procedure

[^{11}C]-(+)-PHNO and [^{18}F]FPEB were synthesized as previously described (1,2). Rats were transported to the Yale PET center where they were anesthetized with 2-5% isoflurane in oxygen and placed into a Focus 220 PET scanner (Siemens). Respiratory rates were monitored throughout the scan. A tail-vein catheter was secured and two rats positioned side-by-side in the bed of the scanner. A transmission scan (9 min) with ^{57}Co was acquired for attenuation correction. Rats then received a bolus injection of [^{11}C]-(+)-PHNO (injected activity 0.51 $\pm$ 0.12 mCi; injected mass: 0.00015 $\pm$ 0.00005 mg/kg) and dynamic data were acquired for 60 min. Once radioactivity had returned to baseline (30-90 min after completing the [^{11}C]-(+)-PHNO scans), rats received a bolus injection of [^{18}F]FPEB (injected activity 0.41 $\pm$ 0.16 mCi; injected mass: 0.00014 $\pm$ 0.00005 mg/kg) and dynamic data were acquired for an additional 60 min. Rats were removed from the gas anesthesia and allowed to recover overnight at the PET center before being returned to the vivarium. PET scans were conducted ~1 week before starting the cocaine or saline self-administration paradigm.

Self-administration procedures

Rats were implanted with intra-jugular catheters using surgical procedures previously described (3). Surgeries were performed under anesthesia (2-3% isoflurane).

Rats were given the analgesic Rimadyl (10 mg/kg; Henry Schein, Dublin, OH) once per day for three days and allowed to recover for 5 days before beginning the self-administration paradigm. Catheters were flushed daily with 0.1 ml of gentamicin/heparin solution (gentamicin: 0.8 mg/ml; heparin: 20 USP U/ml). Catheter patency was assessed on the last day of surgical recovery and once weekly during the self-administration procedure by infusing rats with 0.2 ml of Brevital (10 mg/ml in saline). If the catheter was defective, rats were excluded from the self-administration analyses (N=8).

On each self-administration day, rats were flushed with 0.1 ml of the gentamicin/heparin saline solution before being attached to a tether in a novel operant box. Sessions began with the extension of two levers into the operant box. Responses on the active lever resulted in a single infusion of cocaine (0.5 mg/kg/infusion; cocaine group) or saline (saline group) and presentation of a compound cue (10 s auditory cue and light cue). A 20 s timeout period followed each infusion to reduce the likelihood of a drug overdose. Responses on the inactive lever were recorded but had no programmed consequence. There was no limit on the number of infusions rats could earn in the 6 h session. Self-administration sessions were conducted daily between 700 – 1300 h for 21 consecutive days.

Data analysis

Processing of PET images

Three-dimensional sinogram files were created by binning emission data into a total of 21 frames (6 x 30 s, 3 x 60 s, 2 x 120 s, and 10 x 300 s). Emission files in list mode were reconstructed using the motion-compensation ordered subset expectation

maximization (OSEM) list-mode algorithm for resolution-recovery reconstruction, which includes corrections for normalization, dead time, scatter and attenuation (4). The resultant dynamic images had voxel dimensions of 0.949 x 0.949 x 0.796 mm and matrix dimensions of 256 x 256 x 95.

[^{11}C]-(+)-PHNO or [^{18}F]FPEB template images were developed in house using tools within the FSL suite (FMRIB Software Library, version 6) and co-registered to a rat T2-weighted structural magnetic resonance (MR) image template (5). Individual [^{11}C]-(+)-PHNO and [^{18}F]FPEB images were co-registered to the MR-aligned [^{11}C]-(+)-PHNO and [^{18}F]FPEB templates for region of interest analyses.

Comparison of reinforcement-learning models

We compared the Bayesian Information Criterion (BIC) across five different RL models: (1) standard Q-learning model, (2) forgetting Q-learning model (F Q-learning model), (3) differential-forgetting Q-learning model (DF Q-learning model), (4) forgetting RL model (F RL model), and (5) differential forgetting RL model (DF RL model). Starting action values were set to 0.33 for models 1-3, and 0 for models 4-5. Fminsearch in MATLAB (Mathworks, Inc) was used to maximize the log likelihood of the data given the parameters.

The value updating for each model are as follows:

Q-learning model: $if\ a = 1, Q_i(t+1) = (1-\alpha)Q_i(t) + \alpha(r(t) - Q_i(t))$

$if\ a \neq 1, Q_i(t+1) = Q_i(t)$

F Q-learning model: $if\ a = 1, Q_i(t+1) = (1-\alpha)Q_i(t) + \alpha(r(t) - Q_i(t))$

$$if\ a \neq 1, Q_i(t+1) = (1-\alpha)Q_i(t)$$

DF Q-learning model:

$$if\ a = 1, Q_i(t+1) = (1-\alpha)Q_i(t) + \alpha(r(t) - Q_i(t))$$

$$if\ a \neq 1, Q_i(t+1) = \gamma Q_i(t)$$

Forgetting RL model:

$$if\ a = i, Q_i(t+1) = \gamma\ Q_i(t) + \Delta(t)$$

$$if\ a \neq i, Q_i(t+1) = \gamma\ Q_i(t)$$

Differential forgetting RL model:

$$if\ a = i, Q_i(t+1) = \gamma_C\ Q_i(t) + \Delta(t)$$

$$if\ a \neq i, Q_i(t+1) = \gamma_U\ Q_i(t)$$

where $\Delta(t) = \Delta_+$ and *r(t)* = 1 for rewarded trials, and $\Delta(t) = \Delta_0$ and *r(t)* = 0 for unrewarded trials. The BIC for each model was calculated and summed across animals. The results are present in Supplemental Table 1. The BIC for the DF RL model was lower compared to all other models, indicating that this model best fit the rat choice data.

In the DF RL model the value function for the chosen option (i) was updated according to the following:

$$Q_i(t+1) = \gamma_C\ Q_i(t) + \Delta(t)$$

where the decay rate γ_C determines how quickly the chosen action value decays and Δ(t) indicates the change in the action value that depends on the outcome in trial *t*. If the outcome of the trial was rewarded, then the value function of the chosen port was updated by Δ(t) = Δ_+, the reinforcing strength of reward. If the outcome of the trial was not rewarded, then the value function of the chosen port was updated by Δ(t) = Δ_0, the

aversive strength of no reward. The value for unchosen actions was updated according to the following:

$$Q_i(t+1) = \gamma_U \, Q_i(t)$$

where the decay rate γ_U determines how quickly the unchosen action value decays. Choice probability was calculated according to a Softmax function and trial-by-trial choice data fit with these four parameters (γ_C, γ_U, Δ_+, and Δ_0) selected to maximize the likelihood of each rat's sequence of choices in each session and phase using the fminsearch function in MATLAB (2018b).

Statistical analyses

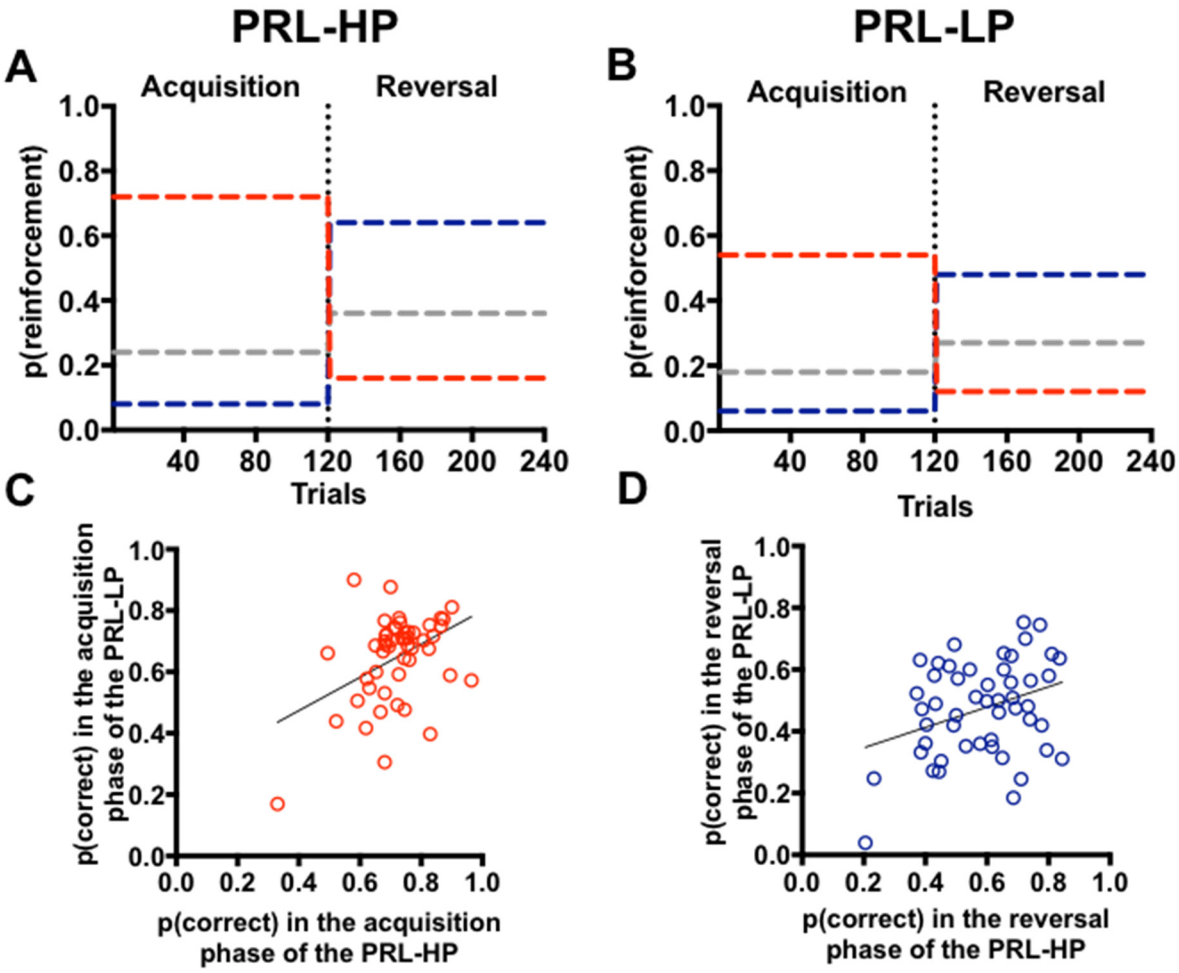
Values presented are the mean ± SEM, unless otherwise noted. Analyses were performed in SPSS (version 25; IBM Corp., Armonk, NY). Repeated measure data were analyzed using a generalized estimating equations (GEE) model using a probability distribution based on the known properties of the data. Specifically, event data (e.g., number of correct choices, number of win-stay) were analyzed using a binary logistic distribution and count data (e.g., number of infusions) were analyzed using a negative binomial distribution. The relationships between decision-making variables/self-administration data and drug-taking parameters were analyzed using linear regression models.

Supplemental references

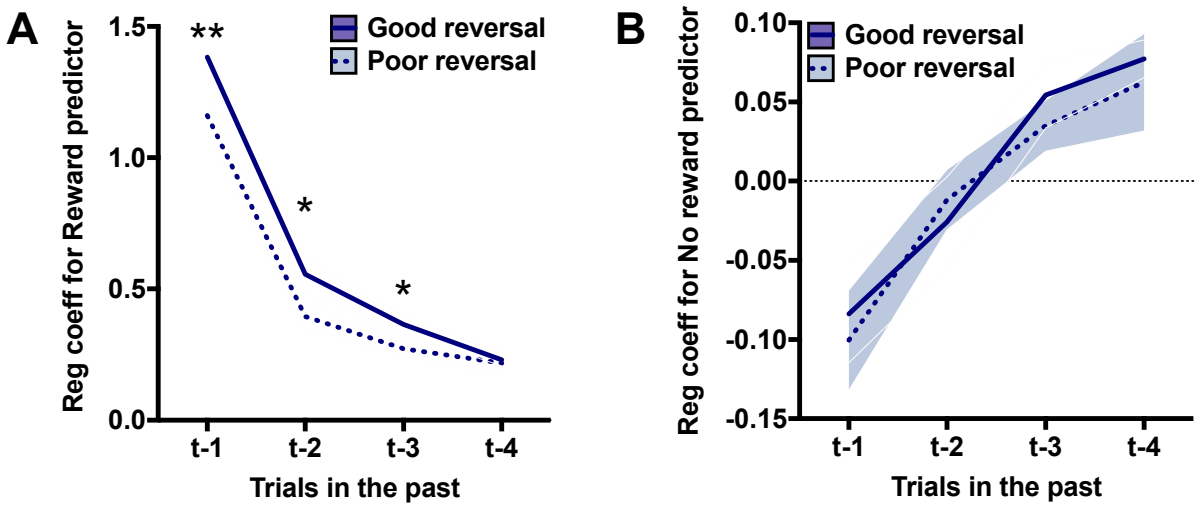
1. Gallezot J-D, Beaver JD, Gunn RN, Nabulsi N, Weinzimmer D, Singhal T, et al. Affinity and selectivity of [^{11}C]-(+)-PHNO for the D3 and D2 receptors in the rhesus monkey brain in vivo. Synapse. 2012 Jun;66(6):489–500.
2. Hamill TG, Krause S, Ryan C, Bonnefous C, Govek S, Seiders TJ, et al. Synthesis, characterization, and first successful monkey imaging studies of metabotropic glutamate receptor subtype 5 (mGluR5) PET radiotracers. Synapse. 2005 Jun 15;56(4):205–16.
3. Groman SM, Massi B, Mathias SR, Lee D, Taylor JR. Model-Free and Model-Based Influences in Addiction-Related Behaviors. Biol Psychiatry. 2019 Jun 1;85(11):936–45.
4. Carson RE, Barker WC, Johnson CA. Design of a motion-compensation OSEM list-mode algorithm for resolution-recovery reconstruction for the HRRT. In: 2003 IEEE Nuclear Science Symposium Conference Record (IEEE Cat No03CH37515). IEEE; 2003. p. 3281–5.
5. Nie B, Chen K, Zhao S, Liu J, Gu X, Yao Q, et al. A rat brain MRI template with digital stereotaxic atlas of fine anatomical delineations in paxinos space and its automated application in voxel-wise analysis. Hum Brain Mapp. 2013 Jun;34(6):1306–18.

	Model 1	Model 2	Model 3	Model 4	Model 5
BIC	127154	110192	110589	109854	**109827***

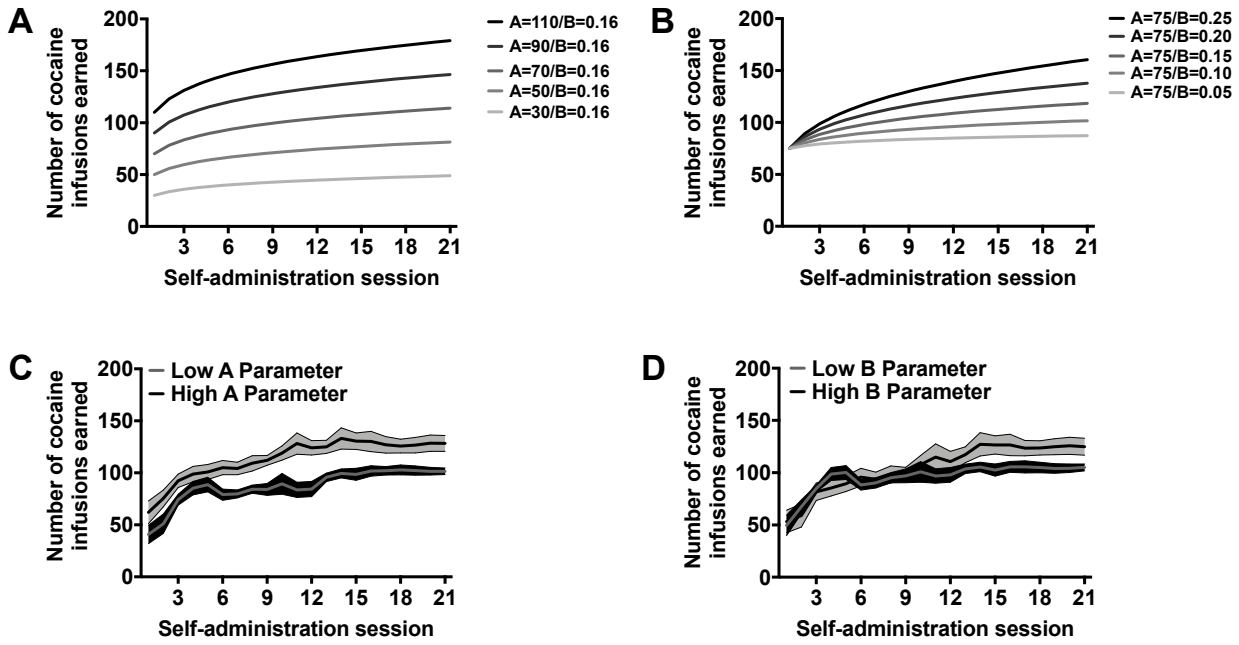
Supplemental Table S1: The sum BIC values for different reinforcement-learning algorithms. A description of each model is provided in the Supplement. The BIC for the DF RL model was lower compared to all other models, indicating that this model best fit the rat choice data.



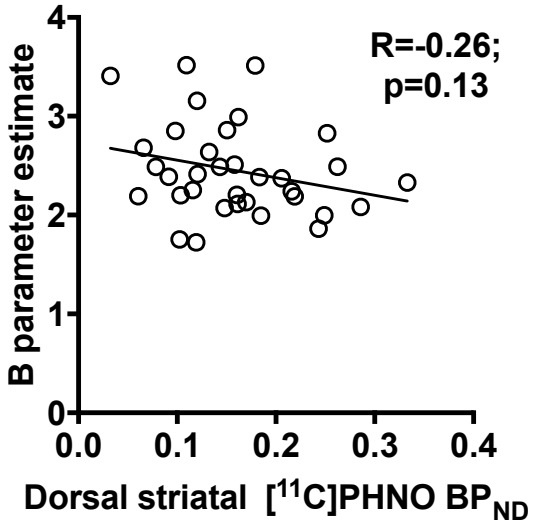
Supplemental Figure S1: The schedules of reinforcement used in the (A) PRL high probability (PRL-HP) and (B) PRL low probability (PRL-LP). The performance of rats in the (C) acquisition phase and the (D) reversal phase of the PRL-HP and PRL-LP was significantly correlated (Acquisition: R^2=0.18; p=0.002; Reversal: R^2=0.12; p=0.01), so dependent measures were collapsed across these schedules of reinforcement.



Supplemental Figure S2: Poor reversal performance is due to deficits in reward-guided decision making. (A) Post-hoc analyses of the group x trial interaction (χ^2=9.20; p=0.03) indicated that the regression coefficients for the 'Reward' predictor in the logistic regression model were significantly lower in the poor reversal group (dotted line) compared to the good reversal group at t-1, t-2 and t-3. (B) The group x trial interaction and the main effect of group were not significant for the 'No reward' predictor in the logistic regression model (group x trial: χ^2=0.65; p=0.89; group: χ^2=0.37; p=0.55). ** p<0.01; * p<0.05.



Supplemental Figure S3: Modeling cocaine self-administration data with the power function $f(x)=Ax^B$. (A)The number of cocaine infusions earned across self-administration sessions by the power function when the A parameter (e.g., initial reinforcement strength of drug) varies (A=30, 50, 70, 90, or 110) and the value of the B parameter (e.g., rate of escalation in drug use) remains constant (B=0.16). (B) The number of drug infusions earned across self-administration session by the power function when the A parameter remains constant (A=75) and the value of the B parameter varies (B=0.05, 0.10, 0.15, 0.20, or 0.25). (C) The number of drug infusions earned across the self-administration sessions in rats with with low (gray line; N = 17) or high (black line; N=17) values for the initial strength of drug reinforcement (e.g., A parameter). (D) The number of drug infusions earned across the self-administration sessions in rats with low (gray line; N = 17) or high (black line; N=17) rates of escalation in cocaine use (e.g., B parameter).



Supplemental Figure S4: The relationship between dorsal striatal [^{11}C]PHNO binding and the rate of escalation in cocaine intake (e.g., B parameter estimate).