

In Cocaine Dependence, Neural Prediction Errors During Loss Avoidance Are Increased With Cocaine Deprivation and Predict Drug Use

John M. Wang, Lusha Zhu, Vanessa M. Brown, Richard De La Garza II, Thomas Newton, Brooks King-Casas, and Pearl H. Chiu

ABSTRACT

BACKGROUND: In substance-dependent individuals, drug deprivation and drug use trigger divergent behavioral responses to environmental cues. These divergent responses are consonant with data showing that short- and long-term adaptations in dopamine signaling are similarly sensitive to state of drug use. The literature suggests a drug state-dependent role of learning in maintaining substance use; evidence linking dopamine to both reinforcement learning and addiction provides a framework to test this possibility.

METHODS: In a randomized crossover design, 22 participants with current cocaine use disorder completed a probabilistic loss-learning task during functional magnetic resonance imaging while on and off cocaine (44 sessions). Another 54 participants without Axis I psychopathology served as a secondary reference group. Within-drug state and paired-subjects' learning effects were assessed with computational model-derived individual learning parameters. Model-based neuroimaging analyses evaluated effects of drug use state on neural learning signals. Relationships among model-derived behavioral learning rates ($\alpha+$, $\alpha-$), neural prediction error signals ($\delta+$, $\delta-$), cocaine use, and desire to use were assessed.

RESULTS: During cocaine deprivation, cocaine-dependent individuals exhibited heightened positive learning rates ($\alpha+$), heightened neural positive prediction error ($\delta+$) responses, and heightened association of $\alpha+$ with neural $\delta+$ responses. The deprivation-enhanced neural learning signals were specific to successful loss avoidance, comparable to participants without psychiatric conditions, and mediated a relationship between chronicity of drug use and desire to use cocaine.

CONCLUSIONS: Neurocomputational learning signals are sensitive to drug use status and suggest that heightened reinforcement by successful avoidance of negative outcomes may contribute to drug seeking during deprivation. More generally, attention to drug use state is important for delineating substrates of addiction.

Keywords: Addiction, Cocaine, Computational psychiatry, Dopamine, fMRI, Prediction error, Reinforcement learning

<https://doi.org/10.1016/j.bpsc.2018.07.009>

In substance-dependent individuals, responses to negative environmental cues appear to both vary with state of drug use and contribute to continued drug seeking. Specifically, when drug deprived, substance-dependent individuals are adept at avoiding negative states, such as withdrawal and isolation, through drug use (1,2). At the same time, when drug using, dependent individuals ignore negative outcomes, including serious social, health, and economic costs (1,3,4). The divergence between these behavioral responses to negative consequences suggests a drug state-dependent role for loss learning (i.e., learning about negative outcomes) in maintaining substance use. In particular, these clinical and behavioral data suggest the hypothesis that heightened reinforcement from avoiding negative states may facilitate drug seeking during deprivation relative to during substance use. A largely parallel but related literature linking neural dopamine (DA) systems in

reinforcement learning (5–7) and cocaine addiction (8–10) provides a framework within which to examine this possibility.

Extant data show that while acute cocaine use increases striatal DA (9), chronic cocaine use decreases postsynaptic DA receptor availability (9). These short- and long-term neurophysiological adaptations together contribute to changes in DA signaling and detection that are sensitive to state of drug use (8–11) [for reviews, see (12,13)]. In addition, related evidence indicates that during healthy reinforcement learning, DA release encodes prediction errors (signaling better or worse than expected) that have detectable correlates in the human striatum (14,15). Together, this literature suggests that in the case of substance dependence, DA-related learning signals are likely to be enhanced with drug deprivation because DA receptors are relatively freed, allowing the detection of prediction errors. A few previous studies that examined neural

SEE COMMENTARY ON PAGE 218

substrates of contingency learning in cocaine dependence primarily focused on reward learning and found decreased prediction error signaling in cocaine-dependent individuals compared with control individuals [(16–18); see also Stewart *et al.* (19), who reported greater baseline neural win responses in successful future abstainers].

To evaluate cocaine state modulation of learning signals and assess the potential drug state–dependent role of loss-learning mechanisms in maintaining substance dependence, we tested cocaine-dependent individuals in a loss-learning task both during cocaine deprivation and when using cocaine as usual. Using a computational psychiatry approach, we assessed behavioral and neural learning substrates, in the form of model-derived learning rate parameters and striatal encoding of prediction errors, respectively, and tested the relationship of neural learning signals to measures of drug use and dependence.

METHODS AND MATERIALS

Participants and Experimental Design

A total of 22 right-handed, non-treatment-seeking male individuals who met criteria only for current cocaine use disorder without other substance dependencies or comorbid Axis I psychopathology were enrolled from a larger study on biomarkers of substance use (see Tables 1 and 2 for demographic information and self-reported craving, and see Supplemental Methods for inclusion/exclusion criteria). Following an initial lab visit to assess cocaine use and entrance criteria, eligible individuals participated in two subsequent scanning sessions in a within-subject design. In one session participants were instructed to use cocaine as usual (C+), and in a second session participants were instructed to abstain from cocaine use for at least 72 hours (C–). Cocaine use state was verified at each lab visit with urine testing for cocaine metabolites (National Institute on Drug Abuse 5-panel drug test; Alere Toxicology, Waltham, MA); C+ and C– sessions were counterbalanced for order. All participants provided informed consent, and all procedures were approved by the institutional review boards of Baylor College of Medicine and Virginia Tech.

Participants completed a probabilistic loss-learning task (Figure 1A) during functional magnetic resonance imaging (fMRI) scanning in two separate lab sessions ($N = 22$, with each participant scanned in both states of cocaine use; see Supplemental Methods for scanning parameters and preprocessing procedures). The task entailed learning from repeated choices between two losing options [two-arm bandit

Table 1. Participant Characteristics

Variable	Cocaine-Dependent Individuals ($N = 22$)
Age	45.7 (7.0)
Education, Years	12.9 (1.0)
WTAR	98.4 (10.8)
Cocaine Use, Years	17.6 (7.9)

Values are mean (SD).

WTAR, Wechsler Test of Adult Reading, a standardized score representing verbal IQ.

Table 2. Participant Self-reported Cocaine Use and Craving

Drug Use Information	Deprived (C–)	Using as Usual (C+)
Estimated Cocaine Intake Last 48 Hours, g	0 (0)	1.5 (0.5)
CCQ Grand Total	137.9 (8.7)	145.3 (8.7)
CCQ Anticipated Positive Outcome	2.4 (0.3)	2.7 (0.3)
CCQ Desire to Use	3.2 (0.2)	3.4 (0.2)
CCQ Intention to Use	3.0 (0.3)	3.0 (0.3)
CCQ Anticipated Withdrawal Relief	3.9 (0.4)	3.5 (0.4)
CCQ No Control	3.8 (0.4)	3.5 (0.4)

Values are mean (SD). Grand total is composed of total raw score from individual items; subscale scores are averaged across items within subscale.

CCQ, Cocaine Craving Questionnaire.

in the loss domain; see details in Supplemental Methods; adapted from (7,20)], with one having a higher probability of producing the better outcome (i.e., smaller loss). On each trial, subjects chose between two abstract stimuli and subsequently observed the outcome (Figure 1A). Participants were instructed that one option was better than the other and that payment was related to their choices, but they were not explicitly informed of the outcome probabilities or loss framework. Trials were presented for a maximum of 36 trials per block or when sufficient learning had occurred (see Supplemental Methods for learning criteria). Each block consisted of novel stimuli that required participants to learn the contingencies between stimuli and outcomes within each block.

An additional group of 54 male individuals (see Supplemental Table S1A for demographics) with no history of Axis I psychopathology was used as an independent nonpsychiatric control sample to identify learning-related striatum activation that could be used as a reference against which to interpret any neural effects observed in the individuals with cocaine use disorder and to compute nonpsychiatric individual parameter estimates for model validation and parameter recovery.

Behavioral Analyses

Model-Agnostic Behavioral Analyses. To verify that participants learned during the task, the behavioral choices of each individual in each drug use state were examined over time and quantified as the percentage of trials that the objectively better choice was selected. Within-sample and paired t tests on performance were implemented in MATLAB (The MathWorks, Inc., Natick, MA).

Computational Model–Based Behavioral Analyses. To assess model-derived behavioral learning effects for each participant in each drug use state, participants' behavioral choices were fit to a basic reinforcement Q-learning model that included two learning rates (α) that provided separate update rules for positive ($\delta+$) and negative ($\delta-$) prediction errors (better or worse than expectations, mapping onto successful and unsuccessful loss avoidance [$\delta+$ and $\delta-$, adapted from previous studies (21–24)]) in the form of positive ($\alpha+$) and negative ($\alpha-$) learning rates (Figure 1B). The

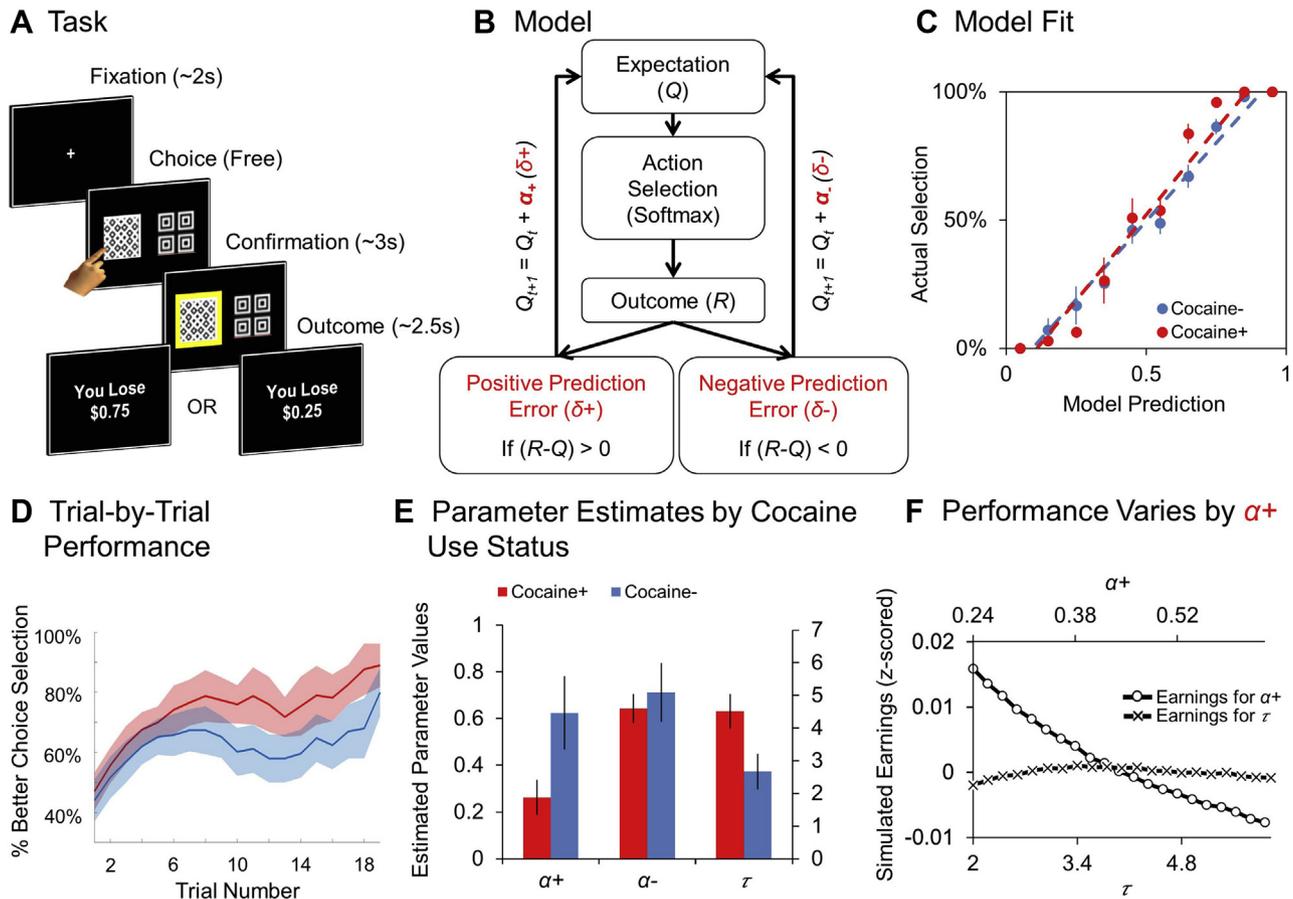


Figure 1. Experimental design and cocaine modulation of learning. **(A)** Participants performed a probabilistic loss-learning task where they made a series of choices between two options and were shown the outcome of each choice. In this example, the selected option has a 75% chance of losing \$0.25 and a 25% chance of losing \$0.75 and is the better of the two options (less loss). Participants completed trials until learning occurred or up to 36 trials per block. **(B)** The reinforcement learning model incorporated positive and negative prediction errors ($\delta+$ and $\delta-$, respectively), which updated the subsequent expected value (Q) with separately estimated learning rates ($\alpha+$ and $\alpha-$, respectively). Trial-by-trial prediction errors are computed as the difference between the outcome and the expected value ($R - Q$). **(C)** The model-predicted probability of selecting the better option was a good fit with both C- and C+ participants' actual behavior (percentage better option selected across subjects). Model prediction and actual selection for participants in the C- (blue) and C+ (red) drug use states were similar (no difference of average log likelihood per trial in a paired comparison across states; $t_{17} = 1.47, p = .14$). **(D)** In both states of drug use, individuals showed learning, improving in choosing the better option as trials progressed. C- individuals relative to C+ individuals showed diminished overall accuracy ($t_{17} = 2.62, p = .01$). **(E)** Bootstrapped group estimates for positive learning rate ($\alpha+$), negative learning rate ($\alpha-$), and inverse temperature (τ) suggested higher $\alpha+$ and lower τ during cocaine deprivation. **(F)** To clarify whether the drug state modulation of $\alpha+$ and/or τ was associated with the cocaine state differences in model-free behavioral performance (panel D), we simulated behavioral choices iterating through the observed ranges of $\alpha+$ and τ parameter values and show that increasing $\alpha+$ was associated with diminished total earnings, while performance did not vary with changes in τ .

separate learning rates allowed us to evaluate asymmetries in learning from $\delta+$ and $\delta-$ [e.g., (16–20)] and assess the possibility that cocaine's effects on DA systems differentially affect these components. The model was a good fit of participants' choices during the loss-learning task (Figure 1C), it was a better fit than a single learning rate model, and model and parameter recovery using simulated data further verified the fit of the model to the observed behavior (see Behavioral Results in Results section). A third estimated parameter, inverse temperature (τ), provided a measure of exploration and indicated the sensitivity of choice probabilities to differences in values. See Model Fitting and Selection in Supplemental Methods for additional descriptions of model selection, model validation, and parameter recovery procedures.

For the two-learning-rate model, the initial expected values $Q(0)$ for the possible choices a and b were set to 0 because participants were not instructed a priori about the range of possible outcomes. For trial number t , the outcome for the chosen option a was represented by $R_a(t)$, with the expected value represented by $Q_a(t)$. The prediction error $\delta(t)$, which measures the difference in outcome $R_a(t)$ and expectation $Q_a(t)$ for a trial, was defined as follows:

$$\delta(t) = R_a(t) - Q_a(t)$$

The parameter estimation procedures included separate update rules for positive and negative prediction errors $\delta(t)$ in the form of positive ($\alpha+$) and negative ($\alpha-$) learning rates, respectively (Figure 1B). The learning rate parameters quantified how much

weight the prediction error $\delta(t)$ from current trials was given in updating the following trials' expected value $Q_a(t + 1)$:

$$Q_a(t+1) = \begin{cases} Q_a(t) + \alpha_+ * \delta(t) & \text{if } \delta(t) < 0 \\ Q_a(t) + \alpha_- * \delta(t) & \text{if } \delta(t) \geq 0 \end{cases}$$

A standard softmax action selection function was used to calculate the probability of selecting choice a at time t and was implemented as follows:

$$P_a(t) = \frac{e^{Q_a(t)*\tau}}{e^{Q_a(t)*\tau} + e^{Q_b(t)*\tau}}$$

Positive and negative learning rates (α_+ and α_-) and inverse temperature (τ) were free parameters, iteratively estimated in MATLAB using the function `fminsearch`, that were evaluated to have the maximum log likelihood (25). Learning rates were bounded between 0 and 1, and inverse temperature was bounded between 0 and ∞ . For the unchosen option b , the expected value of the subsequent trial $Q_b(t + 1)$ was set to the current trial's expected value $Q_b(t)$ multiplied by an additional decay parameter (φ , bounded between 0 and ∞), similar to previous studies (26–29).

Individual variances in learning rates (α_- and α_+) as an effect of drug state (C– or C+) were estimated for second-level fMRI analysis (see Imaging Analysis section below). First, the prior mean and distribution of learning rates for participants in each drug state (C– or C+) were estimated using bootstrapped maximum likelihood created via sampling, approximating integration, around a bootstrapped maximum likelihood estimation across subjects (30,31). Individual learning rates were subsequently estimated by conditioning individuals' behavioral data on the respective drug state (C– or C+) group's prior distribution to account for drug use status differences. For each participant, individual C+ learning rates were then subtracted from the individual C– learning rates to compute a cocaine deprivation-enhanced learning rate for each participant. Group-specific bootstrapped estimates were used for the inverse temperature and decay parameters during individual estimation of learning rates.

Imaging Analyses

To examine neural substrates of loss learning associated with cocaine use state in dependent individuals, model-derived learning variables fit across all participants (as described above) were first correlated with fMRI data collected during the loss-learning task. Next, the model-based neural prediction error signals were related to participants' self-reported cocaine use measures.

First-Level fMRI Processing. The general linear model implemented in SPM8 (32) was used to perform neuroimaging analyses at the individual and group levels. For the first-level analyses, onset times for stimuli, outcome events for δ_+ outcomes, and outcome events for δ_- outcomes for each trial were modeled as separate punctate events. The outcomes were categorized based on the sign of the prediction error ($\delta > 0$ or $\delta < 0$, indicating δ_+ or δ_- , respectively), using the fitted estimates of the two-learning-rate model, in which trial-by-trial δ s were generated (see procedures in Supplemental Methods). To examine the first-level effects of drug use

status on neural representation of learning and valuation, cocaine-positive (C+, urine positive for cocaine metabolites) and cocaine-negative (C–, urine negative for cocaine metabolites) drug use states for each individual were modeled as separate first-level general linear models. Trial-by-trial expected values (Q) were modeled as parametric regressors onto the response events. Trial-by-trial δ_+ and δ_- and the actual outcomes were modeled as parametric regressors onto separate δ_+ and δ_- outcome events, respectively. Effects due to run number, time in scanner, and head movement parameters were modeled as nuisance covariates for each time point.

Within and Paired Drug State Analyses. The within-drug state (C– or C+) and paired-subjects (C– > C+) effects of cocaine use were compared using one-sample and paired-subjects' second-level contrasts in SPM8. The effects of interest were neural responses to δ_+ and δ_- . In line with previous data demonstrating the role of the striatum and DA in learning (7,14,15), the imaging analyses were masked for the striatum. Anatomical masks were constructed using WFU PickAtlas (33) including the structures of the caudate, putamen, and globus pallidus. Also included in the striatum mask was the nucleus accumbens [per Garrison *et al.* (34)]. Results were thresholded with a voxel-level uncorrected $p < .001$ unless otherwise noted, and significant clusters were defined using familywise error correction.

Correlation Analyses Between Cocaine State-Modulated Learning Rate and Neural Prediction Error Signals.

To relate drug state effects on behavioral learning rates (α_+ and α_-) with the corresponding neural δ signals, separate first-level and second-level general linear models were created to correlate within-subject drug-modulated α_+ and α_- differences (C– > C+ for α_+ and C– > C+ for α_-) with the corresponding neural differences for positive and negative δ (C– > C+ for δ_+ and C– > C+ for δ_-). Results were again thresholded with a voxel-level uncorrected $p < .001$, and significant clusters were defined using familywise error correction. In addition, leave-one-out cross-validation analyses were performed in regions of interest (see Supplemental Methods) to reduce bias due to nonindependence (35).

Relationships Between Neural Prediction Error Responses and Behavioral Cocaine Use Measures.

To test relationships between the observed neural learning signals and cocaine use measures, questionnaire data characterizing individual drug use history and current cocaine craving were tested against subjects' C– neural prediction error responses (given the primary results of interest involving enhanced δ_+ from drug deprivation). Again, using leave-one-out cross-validation analysis, neural signals from trials with δ_+ were correlated with years of drug use and subscales of the Cocaine Craving Questionnaire (36). The analyses identified relationships among years of drug use, δ_+ neural signal, and the desire to use cocaine. Based on the results of the correlation analysis, a mediation analysis was performed testing whether neural learning signals mediated the relationship between duration of drug use and individuals' desire to use cocaine or expected positive outcome from cocaine use (C– measures).

A bootstrap approach to mediation (37) was implemented in R to calculate a 95% confidence interval with 10,000 bootstrapped resamples.

RESULTS

Behavioral Results

Model-Agnostic Behavioral Results. In both drug use states, participants demonstrated learning and performed significantly above chance (percentage of trials on which the “better” option was chosen; C−: 62.76%, SE = 3.22, $t_{17} = 3.91$, $d = 0.96$, $p < .01$; C+: 73.61%, SE = 2.69, $t_{17} = 8.70$, $d = 2.12$, $p < .01$; chance: 50%) (Figure 1D). In addition, participants in the C− state showed diminished accuracy relative to C+ participants ($t_{17} = 2.62$, $d = 0.30$, $p = .01$).

Computational Model-Derived Behavioral Results.

Computational model-based analyses, using bootstrapped group parameters [per (38); 200 estimation iterations within each drug use state with subjects drawn with replacement for each iteration] for positive and negative learning rate ($\alpha+$ and $\alpha-$, respectively) and inverse temperature (τ), suggested increased positive learning rates ($\alpha+$) and decreased τ in C− participants relative to C+ participants; $\alpha-$ did not differ between participants in the C− and C+ states (Figure 1E). To clarify whether cocaine state modulation of $\alpha+$ or τ was associated with the diminished behavioral accuracy in C− participants, we simulated behavioral choices holding $\alpha+$ constant (iterating through the ranges of the observed parameter values) while allowing τ to vary and similarly holding τ constant and allowing $\alpha+$ to vary. As shown in Figure 1F, these simulations revealed increased $\alpha+$ to be associated with decreased performance and no relationship between τ and performance (see simulation details in Supplemental Methods). Together, these data provide initial evidence of drug state modulation of learning, where cocaine deprivation-related increases in positive learning rates are associated with diminished behavioral performance.

Imaging Results

Effects of Cocaine Use State on Neural Prediction Error Signals. Significant neural correlates of positive prediction error were observed in the striatum for C− participants ($\delta+$ for C−) (Figure 2A and Supplemental Table S2A) but not for C+ participants. In addition, no significant neural correlates of negative prediction error ($\delta-$) were found during either drug use state. Positive prediction error ($\delta+$) responses were verified in the nonpsychiatric participants (Supplemental Figure S1 and Supplemental Table S2B), and post hoc analyses using an independent striatum region of interest indicated that $\delta+$ in C− participants was comparable to this control cohort, whereas C+ participants showed significantly diminished $\delta+$ responses (C− vs. nonpsychiatric control participants, $t_{70} = 0.15$, $d = 0.004$, $p = .87$; C+ vs. nonpsychiatric control participants, $t_{70} = 3.22$, $d = 0.09$, $p = .001$) (Supplemental Figure S1; see analytic details in Supplemental Methods).

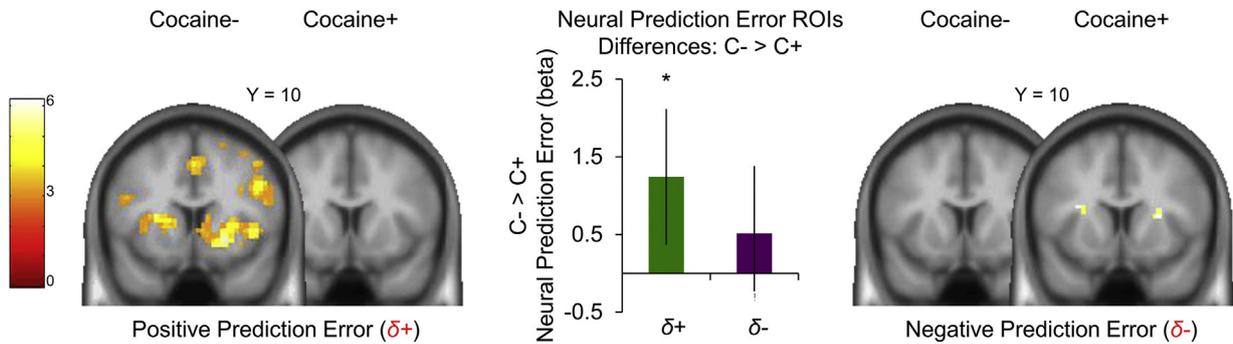
The specificity of the neural encoding of positive prediction errors in the C− participants (Figure 2A) was striking in its

parallel with the increased positive learning rate in these participants. Thus, to test for a neural instantiation of the deprivation-increased positive learning rate, we first computed individual behavioral learning rate estimates for each participant in the C− and C+ states [see Methods and Materials and (30)] and generated for each participant deprivation-enhanced positive and negative learning rate metrics (i.e., C− > C+ for $\alpha-$ and $\alpha+$, respectively, for each individual). For positive learning rate, the degree of participants' deprivation enhancement was significantly associated with the degree of deprivation enhancement of positive prediction error responses in the striatum (C− > C+ for $\alpha+$ and neural C− > C+ for $\delta+$) (Figure 2B and Supplemental Table S2C) ($r = .79$, $p < .01$; using leave-one-out cross-validation to avoid potential bias due to nonindependence) (35). No relationship between drug state modulation of negative learning rates and their associated neural prediction error signals was observed (C− > C+ for $\alpha-$ and neural C− > C+ for $\delta-$; $r = -.08$, $p = .72$) (Figure 2C). For C− > C+, contrasts further show deprivation enhancement (see Supplemental Figure S2A and Supplemental Table S2D). Supplemental Figure S3 shows similar imaging results when using group estimates from within-status behavioral estimates. In addition, no effects of cocaine deprivation on neural expected value signals were detected (Supplemental Figure S2B), indicating generally intact outcome valuation unaffected by drug use status.

Results Relating Neural Prediction Error Signals and Behavioral Cocaine Use Measures.

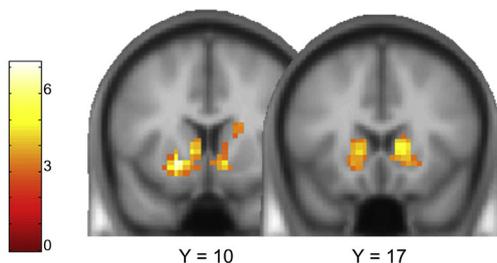
As described above, the specificity of drug state modulation and deprivation enhancement to positive (i.e., successful loss avoidance) prediction errors ($\delta+$) was consistent with the hypothesis that reinforcement from successfully avoiding negative states contributes to continued drug seeking in addiction. In this case, successful loss avoidance in cocaine-deprived participants should be further related to aspects of real-world cocaine use. To test this possibility, we regressed C− individuals' neural $\delta+$ responses (β values from outcomes with $\delta+$) against self-reported drug craving [subscales of Cocaine Craving Questionnaire (36)] (Supplemental Figure S4A and Supplemental Table S3) and observed that neural $\delta+$ responses were related specifically to the desire to use cocaine (Figure 3A and Supplemental Table S3) [$r = .70$, $p < .01$; correlations again performed using neural signals obtained from leave-one-out cross-validation analyses and Bonferroni corrected for multiple comparisons as described in Methods and Materials (35)]. These relationships were also present using the deprivation-enhanced neural $\delta+$ signal (i.e., C− > C+, extracted from outcomes with $\delta+$) ($r = .67$, $p < .01$) and not observed in the C+ state (i.e., signal from outcomes with $\delta+$ while C+) (Supplemental Figure S4B) ($r = -.20$, $p = .42$). Greater neural $\delta+$ responses during cocaine deprivation were also associated with greater years of cocaine use (C−) (Figure 3A) ($r = .64$, $p < .01$); no relationship between neural $\delta+$ and chronicity of use was observed for participants in the C+ state (Supplemental Figure S4C) ($r = -.09$, $p = .71$). No other subscales of the Cocaine Craving Questionnaire were correlated with striatal $\delta+$ signals (Supplemental Table S3). Lastly, desire to use cocaine (Figure 3A) ($r = .61$, $p < .01$) was also positively

A Neural Prediction Error (δ)



B

C- > C+: Positive Learning Rate ($\alpha+$)
Correlated with Positive Prediction Error ($\delta+$)



C

C- > C+: Correlations between α and neural δ (peak from leave-one-out analysis)

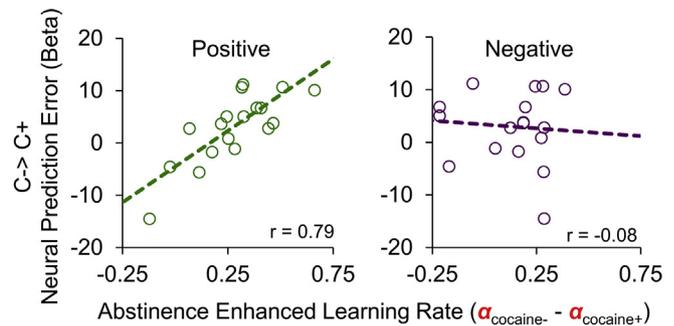


Figure 2. Cocaine use status modulates neural learning signals and reveals increased positive prediction error signaling during deprivation. **(A)** Positive prediction error signal ($\delta+$) was found in the right striatum for cocaine-deprived participants (C-; peak voxel at $t = 5.33$; cluster familywise error corrected $p = .005$; thresholded at $t = 2.9$ for visualization) (see Supplemental Table S2A). No significant neural $\delta+$ signals were found for participants using cocaine as usual (C+). Neither C- nor C+ participants showed significant neural responses to negative prediction errors ($\delta-$). See also Supplemental Figure S2A and Supplemental Table S2D for C- > C+ contrasts that further show deprivation enhancement of learning signals. **(B)** Deprivation enhancement of positive learning rate ($\alpha+$) was correlated with deprivation enhancement of positive prediction error ($\delta+$) in the striatum (C- > C+ for $\alpha+$ and neural C- > C+ for $\delta+$; left striatum peak at $t = 7.20$, right striatum peak at $t = 5.61$; both p s < .0001; thresholded at $t = 2.9$ with striatum mask for visualization) (see Supplemental Table S2C). **(C)** Drug state-dependent (i.e., C- > C+) neural β values were extracted for both $\delta+$ and $\delta-$ from the bilateral striatum and correlated with their corresponding deprivation-enhanced learning rates. For positive learning rate, the degree of participants' deprivation enhancement was significantly associated with the degree of deprivation enhancement of positive prediction error responses in the striatum (C- > C+ for $\alpha+$ and neural C- > C+ for $\delta+$) (Supplemental Table S2C) ($r = .79, p < .01$) (27). No relationship between drug state modulation of negative learning rates and their associated neural prediction error signals was observed (C- > C+ for $\alpha-$ and neural C- > C+ for $\delta-$; $r = -.08, p = .72$). β value differences (C- > C+; whole brain normalized) are plotted. * $p < .05$, relative to zero. ROI, region of interest.

correlated with participants' years of cocaine use. Following these observed relationships, a mediation analysis [(37); see Methods and Materials] revealed that the deprivation-enhanced neural $\delta+$ signal fully mediated the relationship between years of cocaine use and desire to use cocaine while deprived (Figure 3B) (path c: $\beta = .09, p < .01$; path a: $\beta = .06, p = .04$; path b: $\beta = .66, p = .01$; path c': $\beta = .05, p = .07$; mediation effect $a \times b$: 95% confidence interval = 0.0008–0.0953).

DISCUSSION

Using a computational psychiatry approach (39–41), we show drug state modulation of learning signals in cocaine-dependent participants, such that successful loss avoidance signals are greater during deprivation and the neural responses

are associated with both longer history of drug use and greater desire for cocaine. The specificity of the deprivation enhancement to positive neural prediction error signals during loss avoidance appears to parallel clinical descriptions of addiction as a cycle maintained by negative reinforcement where drug-deprived dependent individuals seek drugs and thus successfully avoid negative states (e.g., withdrawal, isolation); such successful loss avoidance has been posited to reinforce continued drug seeking [for relevant discussions, see (1,2)].

These data are consonant with prior studies showing that with greater chronicity of cocaine use, physiological adaptations occur in DA systems (9,11). In particular, the enhanced neural positive prediction error ($\delta+$) encoding in C- participants relative to C+ participants is consistent with studies

Drug State-Dependent Learning Predicts Cocaine Use

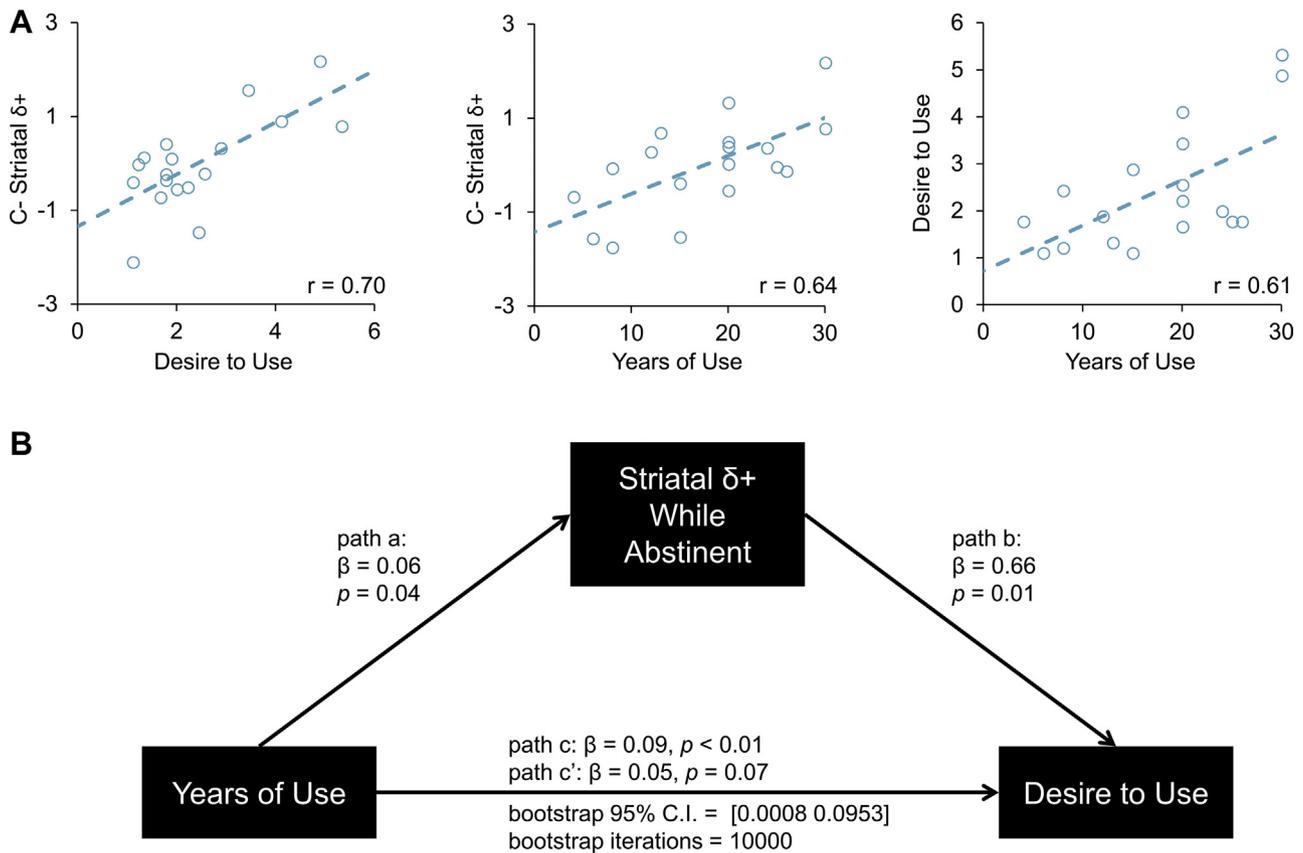


Figure 3. During cocaine deprivation, neural positive prediction error signals mediate relationship between chronicity of drug use and desire to use cocaine. In deprived cocaine-dependent individuals (C-): **(A)** Positive prediction error signal (normalized $\delta+$) in the striatum is associated with greater desire to use cocaine ($r = .70$, $p < .01$), longer history of cocaine use predicted increased neural prediction error signal (normalized $\delta+$; $r = .64$, $p < .01$), and longer history of cocaine use predicted higher desire to use cocaine ($r = .61$, $p < .01$). **(B)** Following these correlational results, a mediation analysis found that the relationship between individuals' years of cocaine use and their desire to use while deprived (path c: $\beta = .09$, $p < .01$) was fully mediated by individuals' deprivation-enhanced neural $\delta+$ signal (path a: $\beta = .06$, $p = .04$; path b: $\beta = .66$, $p = .01$; path c': $\beta = .05$, $p = .07$; mediation effect $a \times b$: 95% confidence interval [C.I.] = 0.0008–0.0953).

showing that humans with long-term cocaine dependency have decreased density of striatal DA receptors and lower tonic DA levels (9) and that acute cocaine intake in chronically cocaine-treated mice reduces DA signaling (11). Following from these studies, $\delta+$ signals ought to be more evident during drug deprivation (as observed here) than during drug use because DA receptors, although diminished in density, are free in the deprived state to detect $\delta+$ fluctuations. We note that in the current cocaine-dependent participants, neural $\delta+$ responses in the drug-deprived state are comparable to the $\delta+$ observed in nonpsychiatric control participants, whereas $\delta+$ signaling in the drug-using state was diminished relative to the control participants. Together, these data suggest that although learning signal impairments appear to be restored by cocaine deprivation in dependent participants, such intact learning can have increasingly detrimental consequences in the context of unhealthy reinforcers, negative environmental states, and adverse outcomes (e.g., when dependent individuals are faced with withdrawal avoidance, drug-available environments, and drug use).

The current data are also relevant for closely related reports of significant increase in prediction error correlates following DA

agonist administration (7) and computational model-based theories that drug use exacerbates prediction errors (42) or triggers false phasic activation of DA neurons (43). A key difference between these previous reports and the current findings is the incorporation of the consequences of long-term drug dependence (i.e., diminished DA functioning) into an understanding of learning in addiction (42,44). In addition, the current diminished $\delta+$ signaling in C+ individuals and enhancement in C- individuals is consistent with related work showing DA drug state modulation of learning signals in participants with Parkinson's disease (who are known to have impaired DA function); these participants similarly show reduced prediction error-related blood oxygen level-dependent responses when on DA-enhancing medication (levodopa) and greater prediction error responses while off medication, specifically to positive prediction errors [$\delta+$; (45,46)].

Finally, we show that greater neural loss-learning $\delta+$ (signaling successful loss avoidance) during deprivation mediates a relationship between chronicity of drug use and desire for cocaine. This relationship supports the hypothesis that drug state-dependent learning signals play a role in maintaining drug use. The current data thus emphasize that

both drug use chronicity and the context in which learning is assessed (e.g., loss, gain) may be critical for identifying neurobehavioral mechanisms that maintain drug use [for related data indicating differences in neural substrates of loss and gain learning, see (47–50)].

The limitations of the current work provide avenues for further study. First, a relatively small number of male participants were included in this study ($N = 22$). While the within-subjects design and advantages of sample homogeneity partially mitigate the sample size, replication in a larger, more diverse sample would address questions regarding generalizability. In addition, the current study identified drug state modulation of responses to negative outcomes but did not evaluate the degree to which the physical consequences per se (i.e., small or large monetary loss), emotions associated with the consequences, or other aspects of the outcomes contribute to the reinforcement provided by successful loss avoidance. Clarifying the role of components of negative outcomes in maintaining substance use in dependence ought to be a focus of future studies. Finally, we focused our neural analyses primarily on regions of the striatum, given previous work linking learning mechanisms and cocaine pharmacodynamics to these regions (5–7,9). Supplemental analyses found no effects of cocaine state on encoding of expected value (see Supplemental Figure S2B), indicating generally intact outcome valuation unaffected by drug use status; nonetheless, other neural regions implicated in learning may be of interest in future investigations.

In summary, in cocaine-dependent participants, we show that drug deprivation enhances neural signaling of successful loss avoidance, which in turn predicts increased desire to use cocaine. The deprivation-enhanced neural prediction error is in line with prior reports of DA adaptations associated with chronic substance use (2,9,11) and also points to a potential mechanism by which drug seeking is maintained. That is, when dependent individuals are at their most vulnerable (i.e., during drug deprivation), reward signals associated with successful avoidance of negative states are at their greatest and may contribute to a pernicious cycle of drug seeking in the face of quit attempts. Of note, DA dysregulation has been associated with poor response to behavioral treatments in addiction (10), and innovative behavioral training protocols have identified learning systems as potential new mechanistic treatment targets for cocaine dependence (51). More generally, the current results support targeting learning-based therapies to identify goal-driven behaviors that provide relief from the negative outcomes of drug deprivation and indicate that attention to drug state may be critical for understanding neural mechanisms of addiction and refining learning-based therapies.

ACKNOWLEDGMENTS AND DISCLOSURES

This work was supported in part by the National Institutes of Health (Grant Nos. R01MH091872 and R21DA042274 [to PHC], Grant No. R01DA036017 to [BK-C], and Grant Nos. RC1DA028387 and R01DA023624 [to RDLG]).

PHC, BK-C, RDLG, and TN designed the experiments. JMW analyzed the data with input from LZ, VMB, PHC, and BK-C. PHC, BK-C, RDLG, and TN supervised this work. JMW and PHC drafted the manuscript with input from all authors. All authors edited and approved the final version.

We acknowledge the technical assistance of George Christopoulos, Dongil Chung, Jacob Lee, James Mahoney, Dharol Tankersley, Katherine

McCurry, Nina Lauharatanahirun, and members of the Chiu, De La Garza, King-Casas, and Newton Labs.

The authors report no biomedical financial interests or potential conflicts of interest.

ARTICLE INFORMATION

From the Virginia Tech Carilion Research Institute (JMW, LZ, VMB, BK-C, PHC), Roanoke, Department of Psychology (JMW, VMB, BK-C, PHC), Virginia Tech, and Virginia Tech–Wake Forest University School of Biomedical Engineering and Science (BK-C), Blacksburg, Virginia; and Baylor College of Medicine (RDLG, TN), Houston, Texas.

RDLG and TN contributed equally to this work.

Address correspondence to Pearl H. Chiu, Ph.D., or Brooks King-Casas, Ph.D., Virginia Tech Carilion Research Institute, 2 Riverside Circle, Roanoke, VA 24016; E-mail: chiup@vtc.vt.edu or bkcasas@vtc.vt.edu.

Received Mar 13, 2018; revised Jul 2, 2018; accepted Jul 11, 2018.

Supplementary material cited in this article is available online at <https://doi.org/10.1016/j.bpsc.2018.07.009>.

REFERENCES

- West R, Hardy A (2005): *Theory of Addiction*. Malden, MA: Blackwell.
- Potenza MN, Sofuoglu M, Carroll KM, Rounsaville BJ (2011): Neuroscience of behavioral and pharmacological treatments for addictions. *Neuron* 69:695–712.
- Lucantonio F, Stalnaker TA, Shaham Y, Niv Y, Schoenbaum G (2012): The impact of orbitofrontal dysfunction on cocaine addiction. *Nat Neurosci* 15:358–366.
- Hyman SE, Malenka RC, Nestler EJ (2006): Neural mechanisms of addiction: The role of reward-related learning and memory. *Annu Rev Neurosci* 29:565–598.
- Montague PR, Hyman SE, Cohen JD (2004): Computational roles for dopamine in behavioural control. *Nature* 431:760–767.
- Schultz W, Dayan P, Montague PR (1997): A neural substrate of prediction and reward. *Science* 275:1593–1599.
- Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD (2006): Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* 442:1042–1045.
- Acevedo-Rodriguez A, Zhang L, Zhou F, Gong S, Gu H, De Biasi M, et al. (2014): Cocaine inhibition of nicotinic acetylcholine receptors influences dopamine release. *Front Synaptic Neurosci* 6:19.
- Volkow ND, Fowler JS, Wang G, Swanson JM (2004): Dopamine in drug abuse and addiction: Results from imaging studies and treatment implications. *Mol Psychiatry* 9:557–569.
- Martinez D, Carpenter KM, Liu F, Slifstein M, Broft A, Friedman AC, et al. (2011): Imaging dopamine transmission in cocaine dependence: Link between neurochemistry and response to treatment. *Am J Psychiatry* 168:634–641.
- Park K, Volkow ND, Pan Y, Du C (2013): Chronic cocaine dampens dopamine signaling during cocaine intoxication and unbalances D1 over D2 receptor signaling. *J Neurosci* 33:15827–15836.
- Keramati M, Durand A, Girardeau P, Gutkin B, Ahmed SH (2017): Cocaine addiction as a homeostatic reinforcement learning disorder. *Psychol Rev* 124:130–153.
- Willuhn I, Wanat MJ, Clark JJ, Phillips PEM (2010): Dopamine signaling in the nucleus accumbens of animals self-administering drugs of abuse. *Curr Top Behav Neurosci* 11:29–71.
- Montague PR, King-Casas B, Cohen JD (2006): Imaging valuation models in human choice. *Annu Rev Neurosci* 29:417–448.
- Jocham G, Klein TA, Ullsperger M (2014): Differential modulation of reinforcement learning by D2 dopamine and NMDA glutamate receptor antagonism. *J Neurosci* 34:13151–13162.
- Parvaz MA, Konova AB, Proudfit GH, Dunning JP, Malaker P, Moeller SJ, et al. (2015): Impaired neural response to negative prediction errors in cocaine addiction. *J Neurosci* 35:1872–1879.
- Tanabe J, Reynolds J, Krmpotich T, Claus E, Thompson LL, Du YP, Banich MT (2013): Reduced neural tracking of prediction error in substance-dependent individuals. *Am J Psychiatry* 170:1356–1363.

Drug State-Dependent Learning Predicts Cocaine Use

18. Rose EJ, Salmeron BJ, Ross TJ, Waltz J, Schweitzer JB, McClure SM, Stein EA (2014): Temporal difference error prediction signal dysregulation in cocaine dependence. *Neuropsychopharmacology* 39: 1732–1742.
19. Stewart JL, Connolly CG, May AC, Tapert SF, Wittmann M, Paulus MP (2014): Cocaine dependent individuals with attenuated striatal activation during reinforcement learning are more susceptible to relapse. *Psychiatry Res Neuroimaging* 223:129–139.
20. Brown VM, Zhu L, Wang JM, Frueh BC, King-Casas B, Chiu PH (2018): Associability-modulated loss learning is increased in posttraumatic stress disorder. *eLife* 7:1–27.
21. Niv Y, Edlund JA, Dayan P, O'Doherty JP (2012): Neural prediction errors reveal a risk-sensitive reinforcement-learning process in the human brain. *J Neurosci* 32:551–562.
22. Christakou A, Gershman SJ, Niv Y, Simmons A, Brammer M, Rubia K (2013): Neural and psychological maturation of decision-making in adolescence and young adulthood. *J Cogn Neurosci* 25:1807–1823.
23. Li J, Delgado MR, Phelps EA (2011): How instructed knowledge modulates the neural systems of reward learning. *Proc Natl Acad Sci U S A* 108:55–60.
24. Palminteri S, Lefebvre G, Kilford EJ, Blakemore S (2017): Confirmation bias in human reinforcement learning. *PLoS Comput Biol* Aug 11; 13(8):e1005684.
25. Sutton RS, Barto AG (1998): *Reinforcement Learning: An Introduction*, Vol. 1. Cambridge, MA: MIT Press.
26. Niv Y, Daniel R, Geana A, Gershman SJ, Leong YC, Radulescu A, Wilson RC (2015): Reinforcement learning in multidimensional environments relies on attention mechanisms. *J Neurosci* 35:8145–8157.
27. Boorman ED, Behrens TEJ, Woolrich MW, Rushworth MFS (2009): How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. *Neuron* 62:733–743.
28. Cavanagh JF (2015): Cortical delta activity reflects reward prediction error and related behavioral adjustments, but at different times. *NeuroImage* 110:205–216.
29. Collins AGE, Frank MJ (2016): Neural signature of hierarchically structured expectations predicts clustering and transfer of rule sets in reinforcement learning. *Cognition* 152:160–169.
30. 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 & Performance XXIII*. Oxford, UK: Oxford University Press, 3–38.
31. Ahn W-YW, Krawitz A, Kim W, Busemeyer JR, Brown JW (2011): A model-based fMRI analysis with hierarchical Bayesian parameter estimation. *J Neurosci Psychol Econ* 4:95–110.
32. Friston KJ, Holmes AP, Worsley KJ, Poline J-P, Frith CD, Frackowiak RSJ (1994): Statistical parametric maps in functional imaging: A general linear approach. *Hum Brain Mapp* 2:189–210.
33. Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH (2003): An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage* 19:1233–1239.
34. Garrison J, Erdeniz B, Done J (2013): Prediction error in reinforcement learning: A meta-analysis of neuroimaging studies. *Neurosci Biobehav Rev* 37:1297–1310.
35. Esterman M, Tamber-Rosenau BJ, Chiu Y, Yantis S (2010): Avoiding non-independence in fMRI data analysis: Leave one subject out. *NeuroImage* 50:572–576.
36. Tiffany ST, Singleton E, Haertzen CA, Henningfield JE (1993): The development of a Cocaine Craving Questionnaire. *Drug Alcohol Depend* 34:19–28.
37. Efron B, Tibshirani R (1986): Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Stat Sci* 1:54–75.
38. Martinez WL, Martinez AR (2008): *Computational Statistics Handbook with MATLAB*, 2nd ed. London: Chapman and Hall/CRC.
39. Maia TV, Frank MJ (2011): From reinforcement learning models to psychiatric and neurological disorders. *Nat Neurosci* 14:154–162.
40. Friston KJ, Stephan KE, Montague R, Dolan RJ (2014): Computational psychiatry: The brain as a phantastic organ. *Lancet Psychiatry* 1: 148–158.
41. Friston KJ, Redish AD, Gordon JA (2017): Computational nosology and precision psychiatry. *Comput Psychiatry* 1:2–23.
42. Redish AD (2004): Addiction as a computational process gone awry. *Science* 306:1944–1947.
43. Schultz W (2011): Potential vulnerabilities of neuronal reward, risk, and decision mechanisms to addictive drugs. *Neuron* 69:603–617.
44. Keiflin R, Janak PH (2015): Dopamine prediction errors in reward learning and addiction: From theory to neural circuitry. *Neuron* 88: 247–263.
45. Voon V, Pessiglione M, Brezing C, Gallea C, Fernandez HH, Dolan RJ, Hallett M (2010): Mechanisms underlying dopamine-mediated reward bias in compulsive behaviors. *Neuron* 65:135–142.
46. Schmidt L, Braun EK, Wager TD, Shohamy D (2014): Mind matters: Placebo enhances reward learning in Parkinson's disease. *Nat Neurosci* 17:1793–1797.
47. Palminteri S, Justo D, Jauffret C, Pavlicek B, Dauta A, Delmaire C, *et al.* (2012): Critical roles for anterior insula and dorsal striatum in punishment-based avoidance learning. *Neuron* 76:998–1009.
48. Seymour B, Daw N, Dayan P, Singer T, Dolan R (2007): Differential encoding of losses and gains in the human striatum. *J Neurosci* 27:4826–4831.
49. Cox SML, Frank MJ, Larcher K, Fellows LK, Clark CA, Leyton M, Dagher A (2015): Striatal D1 and D2 signaling differentially predict learning from positive and negative outcomes. *NeuroImage* 109: 95–101.
50. Pessiglione M, Delgado MR (2015): The good, the bad and the brain: Neural correlates of appetitive and aversive values underlying decision making. *Curr Opin Behav Sci* 5:78–84.
51. Ersche KD, Gillan CM, Jones PS, Williams GB, Ward LHE, Luijten M, *et al.* (2016): Carrots and sticks fail to change behavior in cocaine addiction. *Science* 352:1468–1471.