Cocaine Use Modulates Neural Prediction Error During Aversive Learning

John M. Wang

Thesis submitted to the faculty of Virginia Polytechnic Institute and State University in partial fulfillment of the requirement for the degree of

Masters of Science
in
Psychology

Pearl H. Chiu, Chair
Lee D. Cooper
Brooks King-Casas
Robin K. Panneton

May 7, 2015
Blacksburg, Virginia

Keywords: reinforcement learning, prediction error, cocaine, dopamine, fMRI

Cocaine Use Modulates Neural Prediction Error During Aversive Learning

John M. Wang

Abstract

Cocaine use has contributed to 5 million individuals falling into the cycle of addiction. Prior research in cocaine dependence mainly focused on rewards. Losses also play a critical role in cocaine dependence as dependent individuals fail to avoid social, health, and economic losses even when they acknowledge them. However, dependent individuals are extremely adept at escaping negative states like withdrawal. To further understand whether cocaine use may contribute to dysfunctions in aversive learning, this paper uses fMRI and an aversive learning task to examine cocaine dependent individuals abstinent from cocaine use (C-) and using as usual (C+). Specifically of interest is the neural signal representing actual loss compared to the expected loss, better known as prediction error (δ), which individuals use to update future expectations. When abstinent (C-), dependent individuals exhibited higher positive prediction error ($\delta+$) signal in their striatum than when they were using as usual. Furthermore, their striatal $\delta+$ signal enhancements from drug abstinence were predicted by higher positive learning rate ($\alpha+$) enhancements. However, no relationships were found between drug abstinence enhancements to negative learning rates ($\alpha-$) and negative prediction error ($\delta-$) striatal signals. Abstinent (C-) individuals' striatal $\delta+$ signal was predicted by longer drug use history, signifying possible relief learning adaptations with time. Lastly, craving measures, especially the desire to use cocaine and positive effects of cocaine, also positively correlated with C- individuals' striatal $\delta+$ signal. This suggests possible relief learning adaptations in response to higher craving and withdrawal symptoms. Taken together, enhanced striatal $\delta+$ signal when abstinent and adaptations in relief learning provide evidence in supporting dependent individuals' lack of

aversive learning ability while using as usual and enhanced relief learning ability for the purpose of avoiding negative situations such as withdrawal, suggesting a neurocomputational mechanism that pushes the dependent individual to maintains dependence.

Keywords: reinforcement learning, prediction error, cocaine, dopamine, fMRI

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Cocaine Use Modulates Neural Prediction Error During Aversive Learning

Introduction

Cocaine is a central system stimulant that can create euphoric and energetic feelings in its user. These effects have led to 15 percent of Americans having experimented with cocaine (NIDA, 2010) with 5 million (UNODC, 2012) of those individuals entering into a cycle of cocaine dependence. Prior research into cocaine dependence has been focused on rewards. For example, the incentive sensitization theory of addiction specifies that the reward of the drug itself continuously increases with the development of dependence (Berridge & Robinson, 2003). However, losses also play a critical role in cocaine dependence. Dependent individuals say they prefer life as a drug user even though they realize the long-term social, health, and economic costs that comes with it (West & Hardy, 2005); they do not actively avoid these losses. Separately, dependent individuals are extremely adept at escaping negative states as they frequently choose to satisfy their immediate craving and relieve the associated effects of withdrawal (West & Hardy, 2005). It is unclear how cocaine and its long-term use may affect loss learning. To further understand whether cocaine use leads to dysfunctions in aversive learning, this paper used fMRI and an aversive learning task to examine cocaine dependent individuals abstinent from cocaine use and using as usual.

Dopamine's Role in Learning

Dopamine (DA) is the central neurotransmitter implicated in learning and is released in the striatum, nucleus accumbens (NAcc), and forebrain regions from ventral tegmental area (VTA) projections (Björklund & Dunnett, 2007; Schultz, Dayan, & Montague, 1997). Striatal DA encodes prediction errors (δ), representing the difference between expected and realized outcomes, (Montague et al., 2006). A Pavlovian-learning task in primates (Schultz et al., 1997)

using the temporal difference model of reinforcement learning (TDRL; Montague, Dayan, & Sejnowski, 1996; Montague, Hyman, & Cohen, 2004)[1] confirmed the DA δ firing patterns. Electrodes placed into the animal's ventral striatum found neuron firing when the animal encountered a reward following unconditioned stimuli. The animal had no initial expectation of a reward while receiving one, so there was a positive prediction error (δ+) associating the reward with the stimuli. When the animal encountered the stimulus again, the striatal neurons fired at the conditioned stimulus rather than at point of the reward. There was no δ, as the realized reward equaled the expectation. If the reward failed to occur, a decrease in neural firing was observed when the expected reward would have taken place, a negative prediction error (δ-). The same δ signals have been mapped onto striatal BOLD signals in humans using functional magnetic resonance imaging (fMRI), that is indirectly related to phasic striatal dopamine neuron firing (Montague, King-Casas, & Cohen, 2006). The δ reinforces the stimulus and the behaviors coupled with the outcome by updating the expectation associated with the stimulus. This expectation was also reflected in striatal BOLD signals in subsequent encounters with the coupled stimulus, which helps the agent to make choices that maximize value over time (Daw & Doya, 2006; Montague et al., 2004; Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006; Schultz et al., 1997).

[1] The TDRL framework originates from the concept of trial-and-error in the theory of "Law of Effects" (Thorndike, 1911) with key components of selection and association. Selection represents the process of the agents trying choices and comparing the consequences. Meanwhile association represents how the outcomes found through selection are associated with the selection and further used for future optimization of choices. The Rescorla-Wagner model of classical conditioning (Rescorla & Wagner, 1972) further built upon the "Law of Effects" by quantifying the values of association. More recently, learning was approached from components of mathematical psychology and machine learning by evaluating the optimal solution to a problem and fitting it to the agent's choice behavior (Sutton & Barto, 1998).

Research utilizing the TDRL framework and fMRI to examine reward learning in humans has consistently found the striatum (D'Ardenne, McClure, Nystrom, & Cohen, 2008; Pessiglione et al., 2006; Rolls, McCabe, & Redoute, 2008) reflecting δ to rewards. Furthermore, the ventral tegmental area (VTA; D'Ardenne et al., 2008) and the inferior frontal gyrus (Rolls et al., 2008) were also associated with δ, both of which are anatomically connected with the striatum (Montague et al., 2004). This research pointed to neural representations of valuation or the learning of value that are subsequently used to make future decisions.

Cocaine's Effects on Dopamine and Changes During Dependence

Acute cocaine was found to increase DA concentration in the striatum by preventing reuptake of monoamine transporters in the presynaptic cleft leading to increased amounts of dopamine and an exaggerated effect on postsynaptic neurons (Fowler, Volkow, Wang, Gatley, & Logan, 2001). With long-term drug use, cocaine dependent individuals show decreased dopaminergic receptors and lower baseline dopamine levels while abstinent (Volkow et al., 1993; 2001; Volkow, Fowler, Wang, Swanson, & Telang, 2007). The combination of changes in dopamine levels from cocaine intake and the subsequent long-term effects in the striatum may provide explanations on how aversive learning is changed in cocaine dependent individuals.

The striatal dopamine system mediates reward valuations, and the addiction literature pointed to it as a key neural system that goes awry in substance dependence (Koob & Nestler, 1997). The NAcc, central to the striatal reward system, was linked with implicit valuation from neuroimaging studies showing both NAcc responses to monetary payoff (Knutson, Adams, Fong, & Hommer, 2001; Montague et al., 2004) and the predictability changes of previously associated stimuli (Montague et al., 2004). Increased striatal activation as a result of subjective craving was suggested to lead to increased valuation of food as shown by greater consumption (Peciña,

2008). Indeed, prominent theories of addiction identified a pathological increase in implicit valuation for the drug (Robinson & Berridge, 1993; 2000). Again, the current cocaine dependence literature has mainly focused on reward learning and valuations while the central question of this paper is how cocaine use impacts dependent individuals' loss learning.

Aversive Learning

In previous studies with typical individuals and special populations, the striatal dopamine system was similarly implicated in aversive learning. Greater insensitivity was found in the striatum and VMPFC to the expected value for potential losses than potential gains (Tom, Fox, Trepel, & Poldrack, 2007). However in attempts to estimate magnitude of perceived feedback during the outcome stage of a learning task, Palminteri et al. (2012) did not report any differences between perceived rewards and punishments in their control groups, but did find dorsal striatum lesions of Huntington Disease patients decrease their ability to avoid losses. During an aversive learning task of conditioned stimulus paired with losing $2 and nothing, the caudate head (part of the dorsal striatum) and anterior cingulate correlated with prediction error (Delgado, Schotter, Ozbay, & Phelps, 2008). The striatal dopamine system is involved in aversive learning and is the hypothesized area of interest in this current study.

Competing Hypotheses for Cocaine's Effect on Aversive Learning

Based on the existing literature, two distinct and competing hypotheses were tested on how cocaine affects aversive learning in cocaine-dependent individuals while they are on and off cocaine (Fig. 1). The two proposed hypotheses are referred to as "Ceiling/Floor" and "Gain of Function".

The "Ceiling/Floor" hypothesis specifies that when the dependent individual is using as usual (C+), there is a tonic DA increase from cocaine in the striatum (Bunney, Aghajanian, &

Roth, 1973; Koob & Nestler, 1997; Schultz, 2011) and a decrease in postsynaptic DA receptors as an effect of long-term adaptation to dependence (Volkow et al., 1993; 2001; 2007). This leads to overall lower availability of postsynaptic DA receptors preventing the postsynaptic neurons from detecting prediction errors. In turn, C+ individuals would have decreased value-based learning and diminished δ. While the individual is abstinent (C-), lower tonic DA levels, as an effect of long-term adaptation to dependence (Volkow et al., 1993; 2001; 2007), enables postsynaptic DA receptor binding and the detection of δ. C- individuals in this hypothesis would display larger δ and recovery of value based learning.

In contrast, the "Gain of Function" hypothesis specifies that C+ individuals would have increased availability of DA from cocaine use (Bunney et al., 1973; Koob & Nestler, 1997; Schultz, 2011), allowing an amplification of δ because of increased postsynaptic DA receptor binding. Thus, C+ individuals would show increased value based learning and higher δ. In current literature, a theory explaining the exaggeration of cocaine valuation employed gain of function principles (Redish, 2004). Additional transient DA from cocaine will continuously reinforce the associated drug seeking behavior in the form of a positive feedback loop. Unlike typical reinforcement, simulation analyses suggest that the value of non-drug related reinforcers will asymptote as their δ nears zero, or when the value of the reward is learned (Redish, 2004). For instance, when the value of a non-drug value is learned, there will not be any δ for the typical individual as their expected value is equal to the reward received. This is not so in the case of cocaine as it adds an additional endogenous δ+ signal even when the δ from the stimulus associated with cocaine is zero. As shown using direct modulation of dopamine, the effect of additional transient dopamine via L-DOPA (a dopamine precursor and agonist) on non-drug related valuations enhanced striatal δ+ signal in comparison to a dopamine antagonist

(Pessiglione et al., 2006). For the C+ individuals in the "Gain of Function" hypothesis, the increase of DA from cocaine use may also amplify any δ. Conversely, C- individuals would not show an increase in DA availability leading to smaller δ and lower value-based learning.

To summarize the two contrasting hypotheses for C+ individuals, the "Ceiling/Floor" predicts lower δ while the "Gain of Function" predicts higher δ. For C- individuals, the "Ceiling/Floor" hypothesis predicts higher δ while the "Gain of Function" hypothesis predicts lower δ. To test the competing hypotheses, a Q-learning variant of the TDRL model was used, which specifically encodes valuations based on state-action pairs rather than the environmental state (Montague et al., 2006), to examine neural correlates involved in aversive learning. Using the Q-learning variant allows this paper to attribute value to a specific action in the process of trial and error as a function of the state-action pair (Sutton & Barto, 1998).

Valenced Prediction Error in Cocaine Dependence

While the focus of this current study was on aversive learning in dependent individuals, a key component was also on how learning differentiates between better than expected outcomes and worse than expected outcomes. In the aversive environment where all the choices are bad, a better than expected outcome means that the individual is losing less than they expected and may be relieving (positive prediction error; δ+). Meanwhile, a worse than expected outcome is one where the individual is losing more than expected, and may be accompanied by regret (negative prediction error; δ-).

Research about valenced δ in typical individuals has provided several theories that all purpose differences in δ- encoding compared to δ+ (Daw, Courville, & Touretzky, 2006; Ludvig, Sutton, & Kehoe, 2008; Niv, Duff, & Dayan, 2005). Primate studies using midbrain dopamine neuron recordings showed increased firing in relation to δ+ only and not δ- (Bayer & Glimcher,

2005; Schultz, 2010). In a neuroimaging study, ventral striatum deactivation was found to correlate with δ- of participants who did not receive juice when expected (D'Ardenne et al., 2008). Signed δ was found to correlate negatively in the insula (Pessiglione et al., 2006) and positively with the VTA and substantia nigra in the brainstem (D'Ardenne et al., 2013). These studies modeled δ+ and δ- as one linear relation, so neural signals specific to each δ could not be distinguished. Another study suggested a spatial segregation in the striatum for positive (anterior) and negative (posterior) δ (Seymour et al., 2007).

It was clear from these previous studies that δ+ and δ- needs to be distinguished and each may be differently affected in how cocaine modifies aversive learning. Specific to the "Ceiling/Floor" hypothesis, changing level of DA from cocaine intake may act as ceiling for δ+ in certain situations but not for δ- and vice versa in other situations. To distinguish possible differences between relieving (δ+) and regretful (δ-) outcomes, this paper examined an alternative model that fit separate parameters for δ+ and δ-.

Current research in cocaine dependence has mainly focused on rewards. However losses also play a critical role in cocaine dependence as dependent individuals do not actively avoid aversive social, economic, and health outcomes associated with drug use, but are extremely adept at escaping negative states like withdrawal. To further understand how cocaine and its long-term use acts upon the underlining etiology in aversive learning, this study used fMRI, an aversive learning task, and Q-learning model to examine differences in cocaine dependent individuals behavior and neural signals while abstinent from cocaine use and using as usual.

Methods

Participants

Eighteen right handed non-treatment seeking male individuals who met criteria for cocaine dependence determined by the Structured Clinical Interview for DSM-IV without comorbid Axis-I psychopathology (determined by the SCID) were recruited and included in the analysis (see Table 1 for demographic information). In a within subject design, subjects participated in two lab sessions reflecting drug state, abstinent (C-) and using as usual (C+), which were counterbalanced by the state in which they entered the study. Cocaine usage statuses were verified via cocaine metabolites urine testing for each session. Twenty-eight cocaine dependent individuals were originally recruited and participated through the entirety of the study. 10 were excluded from the analysis, 6 for only selecting one option for the entire task (preventing model fitting), 1 for pressing the same button for the entire task, and 3 due to excessive head motion (3mm or 3 degree in any direction).

Aversive Learning Task Design

The task used was a first order probabilistic instrumental learning task with losing outcomes. As illustrated in Fig. 2a, on each trial subjects chose between two abstract stimuli and subsequently observed the outcome. The trials were presented repeatedly for 36 trials or when the objectively better pattern was learned, represented by 80% selection within the past 5 trials. One of the stimuli was associated with a fixed probability (75%) of the better outcome (smaller loss) and a fixed probability (25%) of the worse outcome (larger loss). The other stimulus was associated with the opposite pattern compared with the first. The locations of the two stimuli were randomized on the screen and subjects used a button pad to select the stimuli. As implemented, this task took approximately 20 minutes and each participant underwent a thorough practice run to ensure task comprehension prior to entering the scanner. Participants

were endowed with an initial sum of $10 for participating in the study and were paid an amount proportional to their performance in the game to maximize engagement in the task.

Behavioral Analysis

Participants' behavior in the task was compared with the optimal selection for each specific trial by calculating an optimal choice percentage. The optimal selection was the choice that had a higher expected value (i.e., the optimal choice was the option that had the smaller loss for 75% of the time). Paired two sample t-tests were used to compare optimal selection differences between drug use states.

Models Fitting and Selection

Hypothesized learning models were fitted to observed behavioral data with the best fitting model used for fMRI analysis. Tested models included the standard Q-learning model (Sutton & Barto, 1998) and valence dependent model (2 Learning Rate Model) that distinguishes valenced PE (better or worse than expectations), adapted from the Risk-Sensitive TD model (Niv, Edlund, Dayan, & O'Doherty, 2012). The 2 Learning Rate model was tested due to its inclusion of separate neural mechanisms encoding δ- and δ+, and the possibility that cocaine dependent individuals would react to each in a different manner due to cocaine's short term and long effects on dopamine (see above).

In the models, the initial expected values $Q(0)$ for the possible choices a and b were set to 0. 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, which measures the difference in outcome $R_a(t)$ and expectation $Q_a(t)$, for the trial was defined as the following:

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

Q-learning model. For the standard Q-learning model algorithm, the model-based parameter estimated was learning rate α, which quantifies how much weight the prediction error $\delta(t)$ from current trials is given in updating the following trials' expected value $Q_a(t+1)$. The standard Q-learning model was the null hypothesis model. Each trial by trial expected value Q for a was calculated as follows:

$$Q_a(t+1) = Q_a(t) + \alpha * \delta(t)$$

2 Learning Rate model. In the standard Q-learning model, the learning rate does not distinguish between better or worse expected outcomes and updates the expected value for the following trial the same, regardless of the direction of outcome. Current literature on the mechanistic properties of updating expectations suggests that there are distinguishable processes. Specifically, the effects of dopamine may affect only specific directions of updating expected value. To allow this possibility, we used the 2 Learning Rate model, adapted from the Risk-Sensitive TD model (Niv et al., 2012), including separated update rules for positive and negative prediction error $\delta(t)$ in the form of positive α_+ and negative α_- learning rates, respectively. This served as the alternative hypothesized model to the standard Q-learning model (Fig. 2b):

$$Q_a(t+1) = \begin{cases} Q_a(t) + \alpha_+ * \delta(t) & if \quad \delta(t) > 0 \\ Q_a(t) + \alpha_- * \delta(t) & if \quad \delta(t) \leq 0 \end{cases}$$

All models used assumed a softmax action selection function. The probability of selecting choice a at time t was estimated as follows:

$$p_a(t) = \frac{e^{Q_a(t)\beta}}{e^{Q_a(t)\beta} + e^{Q_b(t)\beta}}$$

Inverse temperature β is an exploration parameter that quantifies the balance between the exploitation of the higher valued option and exploration of the other option for information at a cost. Lower inverse temperatures signal more exploration and are representative of more random behavior.

Model Fitting. Optimal learning rate(s) α and inverse temperature β for each model were free parameters iteratively estimated using a grid search in MATLAB using the function fminsearch that is evaluted to have the maximum log likelihood (Sutton & Barto, 1998). Learning rate(s) α for all the models were bounded between 0:1 and inverse temperature β is bounded between 0:∞. For the unchosen option b, the expected value of the subsequent trial $Q_b(t+1)$ was set to the current trial expected value $Q_b(t)$ multiplied by an additional freely estimated adjustment parameter φ included to optimize each model fit.

The model fits were evaluated using chi-square likelihood ratio tests and Bayesian information criterion (BIC), which penalizes additional parameters. Each model was fitted across all subjects' behavioral data; across subject estimates were used per previous studies (Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006; Schönberg, Daw, Joel, & O'Doherty, 2007; Schonberg et al., 2010). Using the chi-square likelihood ratio test, the Q-learning model was rejected in favor of the 2 Learning Rate Model ($p < 0.0025$, $df = 1$). Additionally, the result was further confirmed using the more conservative BIC comparison (2 Learning Rate model: 2205.3; Standard: 2207.5; Fig. 2d). The goodness of fit can be seen in Fig. 2c, which compares the calculated probabilities against percentage of time the choice was selected. The estimated parameters across all subjects were used to generate trial-by-trial Q and δ used as parametric regressors on first level imaging analysis (See Imaging Analysis Section for more detail).

Individual variances in learning rate (α+ and α-) improvement as an effect of drug abstinence were also estimated for second level fMRI analysis (See Imaging Analysis Section). A Bayesian estimation approach was used to fit within individual learning rates where individual behavioral data was conditioned on prior across subject estimates fitted for each drug condition using a bootstrap method. The difference of the best posterior fitting learning rate for both drug states (C- and C+) were taken to get the within group learning rate improvement from drug abstinence.

Imaging Analysis

Pre-processing of the imaging data was completed using statistical parametric mapping software (SPM8; Wellcome Department of Imaging Neuroscience, University College London, UK). Images were first corrected temporally for slice timing, and then for movement using least squares minimization without higher-order corrections for spin history and normalized to stereotaxic MNI (Montreal Neurological Institute) space by calculating a multiplication matrix for segmented grey and white matter and CSF separately. Images were then resampled every 3.4 mm using 4th Degree B-spline interpolation and smoothed with a 6 mm Gaussian kernel.

The general linear model (GLM) and the theory of Gaussian random fields implemented in SPM8 (Friston et al., 1995) were used to perform statistical analysis on the individual and group level. On the first level individual analysis, onset time points from stimuli, outcome events for positive prediction error outcomes, and outcome events for negative prediction error outcomes of each trial were modeled as separate events. The outcomes were categorized using the fitted estimates from the 2 Learning Rate model (see Model Fitting and Selection), where trial-by-trial δ were generated.

In addition, trial-by-trial expected values (Q) were modeled as parametric regressors onto the response events. Trial-by-trial δ and actual realized losses were modeled as parametric regressors onto the outcome events respectively for δ+ and δ-. Effects due to run number, time in scanner, and head movement parameters were modeled out as nuisance covariates for each time point.

To examine the effects of drug abstinence on neural representation of learning and valuation, using as usual (i.e., urine positive for cocaine metabolites, C+) and abstinent (urine negative for cocaine metabolites, C-) drug use states for each individual were modeled as separate first level GLMs. The effects of drug abstinence were compared using a paired between group second level contrast in SPM8 with the contrast of C- > C+. The effects of interest were δ+ and δ-. In line with the hypothesis of cocaine's effect on dopamine, the imaging analysis were masked for the striatum. Anatomical masks were constructed using WFU-pickatlas (Maldjian, Laurienti, Kraft, & Burdette, 2003) including the structures of the caudate, putamen, and globus pallidus. Also included in the mask was the nucleus accumbens, which was defined by Garrison et al., (2013) in a meta-analysis of the striatum. Results were threshold with a voxel level uncorrected $p < 0.005$ and significant clusters were defined using a family-wise-error correction on the cluster level.

To directly relate estimated learning rate (α+ and α-) enhancements from drug abstinence on the neural δ signal, separate first level and second level GLMs were created to correlate within subject α differences with neural δ differences for positive and negative δ respectively. Results were again thresholded with a voxel level uncorrected $p < 0.005$, and significant clusters were defined using family-wise-error correction on the cluster level. In addition, leave-one-out

cross-validation analyses were performed to reduce bias due to non-independence (Esterman, Tamber-rosenau, Chiu, & Yantis, 2011).

Leave-one-out cross-validation analyses involved re-estimating the same second-level analysis 18 times, leaving one subject out each time. A new set of voxels showing local maximum nearest to the group peak voxel (Left Striatum: -20, 10, -6; Right Striatum: 14, 7, 9) was calculated from these iterations. A sphere centered at each new voxel with radius of 6 mm defined the ROI and was used to extract the mean beta from the left-out subject. Correlations between the extracted betas and behavioral conformity were tested using Pearson correlation tests.

Lastly to test how neural learning signals inform the maintenance of dependence, questionnaire data characterizing individual drug use history and current cocaine craving were tested against subjects' neuroimaging data while abstaining from cocaine use based on previous results of interest involving enhanced $\delta+$ from drug abstinence. Again using leave-one-out cross-validation analysis, model-free $\delta+$ were correlated with years of drug use and subscales of the Cocaine Craving Questionnaire (Tiffany, Singleton, Haertzen, & Henningfield, 1993).

Results

Behavioral Analysis

Cocaine dependent individuals performed significantly better than chance regardless of drug use status signaling that they both understood the task and learned the better choice (C-: $p = 0.001$; C+: $p < 0.00001$.). In a paired t-test analysis on performance in the aversive learning task, the cocaine dependent individuals performed significantly better in the C+ versus the C- state (Fig. 2e).

Prediction Error Neuroimaging Analysis

To test the "Ceiling/Floor" and Gain hypotheses related to aversive learning, paired between drug states' neural learning data were compared. Calculated $\delta+$ and $\delta-$ based on estimated parameters fitted to all subjects were first regressed onto individual subjects' striatal neural signals during the outcome time point of each trial respective to the valence of the trial-by-trial δ. Then a paired two-sample t-test was used to test the difference between cocaine abstinence status for $\delta+$ and $\delta-$ respectively. While individuals were in the C- state, they showed greater signal in the striatum to positive prediction errors ($\delta+$; $p < 0.05$, small volume corrected for multiple comparisons within an anatomically defined striatum mask) than when they were in the C+ state (Fig. 3a; Table 2a). In a within status correlation analysis, C- individuals' striatum showed a positive correlation with $\delta+$ (Table 3; $p < 0.05$, small volume corrected for multiple comparisons within an anatomically defined striatum mask) in the corresponding outcomes, but C+ individuals did not show any significant correlations with $\delta+$ in the striatum in the corresponding outcomes.

In addition, there were no significant differences between C- and C+ in striatal neural signal correlated to negative prediction errors in a paired two-sample t-test analysis ($\delta-$; Fig. 3a). Neither drug states showed any significant within status correlations with $\delta-$ and striatal signal during the corresponding outcomes. These results show that while cocaine dependent individuals were abstinent from cocaine use (C-), they had enhanced $\delta+$ striatal neural signals.

Model Estimation and Correlation Analysis

Also to test the "Ceiling/Floor" and Gain hypotheses related to aversive learning, paired between drug states estimated learning parameters were compared. From a model estimation standpoint, learning rate estimates (α) paralleled the neural prediction error (δ) striatal signal as an effect of cocaine abstinence (C- > C+). Learning rate α estimates indicate how quickly the

individual updates their expected value associated with the stimuli while δ signal in the striatum represents the neural feedback that updates subsequent expected values. Cocaine use status (C- > C+) showed a significantly greater effect in $\alpha+$ than $\alpha-$ (Fig. 3a). To parse the relationship between $\alpha+$ and neural $\delta+$ enhancement as an effect of cocaine abstinence (C- > C+), we correlated individuals' α onto neural δ for positive and negative δ (Fig. 3b). $\alpha+$ enhancement predicted neural $\delta+$ ($r = 0.796$, $p = 0.00008$), while $\alpha-$ had no correlation with $\delta-$ ($r = 0.0339$, $p = 0.894$). These results show that enhanced $\alpha+$ from cocaine abstinence predicted the enhanced striatal $\delta+$ neural signal.

Expected Value Neuroimaging Analysis

In a separate analysis examining if subjects neurally encoded the values of each option and for the purpose of determining whether subjects' behavior differences were related to differences in expected value encoding, normalized expected value was regressed onto the neural activity associated with the decision phase of each trial. Subjects showed ventral medial prefrontal cortex (VMPFC) signal in both the state of C- and C+ (Fig. 4). There were no differences when comparing the neural signal of both states. These results show that the cocaine dependent individuals were able to neurally represent the valuation of the selection regardless of drug state.

Second Level Neuroimaging Analysis

To understand how neural learning signals inform the maintenance of dependence, neural learning data were correlated with drug use measures. Longer history of cocaine use predicted greater striatal model-free $\delta+$ signal in the C- state ($r = 0.643$, $p = 0.004$). In addition, anticipated positive outcome from cocaine use ($r = 0.506$, $p = 0.03$) and desire to use cocaine ($r = 0.700$, $p = 0.001$) subscales of the Cocaine-Craving Questionnaire (CCQ; Tiffany et al., 1993) positively

correlated with striatal model-free $\delta+$ signal in abstinent individuals. Both anticipated positive outcome from cocaine use ($r = 0.583$, $p = 0.01$) and desire to use cocaine ($r = 0.610$, $p = 0.007$) subscales of the CCQ were positively correlated with years of use. There were no significant correlations between striatal signals for better than expected outcomes and abstinent individuals' and CCQ subscales of anticipated withdrawal relief, intention to use, and no control of use. Leave-one-out cross-validation analyses for non-significant correlations failed to localize a maxima in the striatum, preventing the extraction of mean beta values. Lastly, using as usual individuals' striatal signals did not correlate with either history of use or CCQ measures.

Discussion

The current study examined the effects of cocaine use on aversive learning with an instrumental aversive learning task. First, this study aimed to test the proposed "Ceiling/Floor" and "Gain of Function" hypotheses (Fig. 1) related to aversive learning by comparing paired between cocaine use state behavioral learning parameters and neural learning data. Second, this study aimed to understand how neural learning signals inform the maintenance of cocaine dependence by relating neural learning data to drug use measures.

Utilizing a 2 Learning Rate computational model that distinctly estimated separate positive learning rate ($\alpha+$) for relieving outcomes and negative learning rate ($\alpha-$) for regrettable outcomes, abstinent individuals (C-) had a significantly higher $\alpha+$ in comparison to $\alpha-$ (Fig. 3a) than when they were using as usual (C+). This result indicates that C- individuals utilized behavioral $\delta+$ to update their expected values to a higher degree than when they were C+.

A parallel neural pattern was found in the striatum where cocaine abstinence (C- > C+) enhanced striatal neural $\delta+$ signals but not striatal $\delta-$ signals (Fig. 3a). Striatal neural δ signals are signals in the striatum that were significantly predicted by calculated trial-by-trial behavioral

δ using each group's estimated parameters and represent the strength in which the brain neurally updates the associated action. Thus, abstinent individuals have a stronger signal in associating actions with relieving outcomes than when they were using as usual.

To further confirm that the enhanced neural learning signal mirrors behavioral learning estimates for relieving outcomes, individual $\alpha+$ estimate enhancements from cocaine abstinence were correlated with the individual's $\delta+$ signal enhancements in the striatum (Fig. 3b). There was indeed a significant positive correlation between individual differences in estimated $\alpha+$ and their individual differences in striatal $\delta+$ signal, but not for $\alpha-$ and $\delta-$. So, improvements in cocaine abstinent individuals' abilities to learning about relieving outcomes computationally are directly related to their striatal relief learning signal. These current results are consistent with a recent study that found using as usual cocaine dependent individuals had decreased event-related potential signals compared with abstinent cocaine individuals for $\delta+$, but not $\delta-$ to unexpected gambles (Parvaz et al., 2015). Furthermore in a Pavlovian learning task, reduction in right caudate to $\delta+$ was found in cocaine dependent individuals (Rose et al., 2014).

Looking back at the original hypotheses, the current results are consistent with the "Ceiling/Floor" hypothesis for only relieving outcomes ($\delta+$). The "Ceiling/Floor" hypothesis proposed increased availability of DA while the individual is on cocaine would prevent the detection of δ and the decrease of DA while the individual is off cocaine would enable the detection of δ. The results matched the hypothesized outcomes as smaller $\delta+$ was found while dependent individuals were using as usual and larger $\delta+$ while they were abstinent. The underlying differences that lead to the results favoring the "Ceiling/Floor" hypothesis over the "Gain of Function" hypothesis may be explained by the increases in tonic DA, from cocaine use (Bunney et al., 1973; Koob & Nestler, 1997; Schultz, 2011), binding to the already lower number

of post-synaptic DA receptors (Volkow et al., 1993; 2001; 2007), which lead to a decreased availability of post-synaptic DA receptors. Thus, any prediction error signaling would not be received postsynaptically. When the individual is abstinent, the low tonic level of DA in the striatum from long-term dependence (Volkow et al., 1993; 2001; 2007) would free up post-synaptic DA receptors allowing for the detection of $\delta+$.

An alternative explanation for the lack of $\delta-$ striatal signal difference found between states and within each state may be that $\delta-$ involves a dissociable network including the habenula (Lawson et al., 2014; Matsumoto & Hikosaka, 2007; Salas, Baldwin, de Biasi, & Montague, 2010) and insula (Palminteri et al., 2012; Pessiglione et al., 2006). The habenula specifically has been found to regulate the striatum (Matsumoto & Hikosaka, 2007) and theorized to become hyperactive with continued drug use (Baldwin, Alanis, & Salas, 2011). Such regions of interest fall outside the scope of this current study and would be of interest in future studies.

The results from this current study imply that while cocaine-dependent individuals are using as usual, their neural learning signals are washed out among the noise from the endogenous effects cocaine. Without those learning signals, they may never encode the social, health, and economic losses. This possibility is consistent with Schultz's (2011) theory that specific qualities of cocaine act as inhibitory components onto other non drug-related rewards such as money and friends.

The current data did not support the "Gain of Function" hypothesis, however it is important to note that individuals in this current study are long-term cocaine dependent individuals with lower postsynaptic receptors (Volkow et al., 1993; 2001; 2007). With non-dependent individuals, a "Gain of Function" effect might be observed; there would be greater

number of post-synaptic DA receptors available to utilize the increased availability of DA to amplify the $\delta+$ signal.

Intrigued by the striatal relief learning signal ($\delta+$) in abstinent individuals (C-), this study examined whether relief learning signal is related to their drug use information in a second analysis. Specifically of interest is if these effects suggest a mechanism that drives abstinent individuals to continue cocaine use. In a correlation analysis between individuals' histories of cocaine use with their neural sensitivity of how they learned about actions that provided them relief (model free $\delta+$ striatal signal), a positive correlation was found (Fig. 5a). That is, with greater chronicity, abstinent individuals showed increased neural sensitivity to learning about actions that relieved them of aversive situations. Secondly, individuals' responses on subscales of the cocaine craving questionnaire (CCQ; Tiffany et al., 1993) were positively correlated with their striatal $\delta+$ signal (Fig. 5b). This means that abstinent individuals' neural sensitivity to learning about relieving actions was correlated with how much they craved cocaine, specifically their anticipated positive outcomes of taking cocaine and their desire to use cocaine.
A positive correlation was also found between individuals' drug use histories and their craving responses.

These results suggest that with long-term use, the brain adapts in the form of withdrawal. Craving follows withdrawal in the cycle of addiction (Koob & Volkow, 2010). The neural learning system also adapts by becoming more sensitive to learning about relieving actions. In the case of abstinent cocaine dependent individuals who might feel the negatives of withdrawal, that relieving action would be to use cocaine. This mechanism as suggested by our results may inform why dependent individuals are extremely adept at escaping negative states, such as

finding cocaine for withdrawal relief. Using cocaine to escape the negative states that comes with withdrawal leads to a continuation of the individual's dependence.

One limitation of this current study is that urine testing for cocaine byproducts does not directly map onto acute affects of cocaine as it only informed whether dependent individuals were using as usual or abstinent. Furthermore, the BOLD signal is only an indirect measure of dopamine. There was no direct measure of dopamine levels from cocaine use. Future studies may be able to provide direct application of cocaine or dopamine agonists in dependent individuals to fully map the effects of cocaine. Lastly as mentioned above, this current study focuses on effects found in the striatum. The imaging sequences were not optimized to observe differences in the habenula, an area of interest in future studies.

Conclusion

The current study extends the understanding of cocaine's effects on dependent individuals' aversive learning. First, abstinent individuals were found to show enhanced striatal $\delta+$ signal and estimated $\alpha+$ relative to when they were using as usual. This provides support for a "Ceiling/Floor" DA effect in dependent individuals. The implication of these results are that using as usual dependent individuals may not encode an aversive neural learning signal due to endogenous effects of cocaine. Without such aversive neural learning signals, dependent individuals may never encode the social, health, and economic losses that come with continued dependence. Secondly while dependent individuals were abstinent, their sensitivity to learning about actions that provide relief increased with time. Furthermore, dependent individuals' craving for cocaine also positively correlated with their sensitivity to learning about actions that provide them with relief. Taken together, these results suggest a plausible mechanism that maintains dependence. With longer duration of use and greater withdrawal, dependent

individuals' neural learning systems also adapt to be more sensitive to learning about actions that will provide the individual with relief. In withdrawal, that action includes using cocaine, and thus maintains dependence.

References

Baldwin, P. R., Alanis, R., & Salas, R. (2011). The role of the habenula in nicotine addiction. *Journal of Addiction Research & Therapy*, *S1*(2), 1–10.

Berridge, K. C. & Robinson, T. E. (2003). Parsing reward. *Trends in Neuroscience*, *26*(9), 507–513.

Björklund, A., & Dunnett, S. B. (2007). Dopamine neuron systems in the brain: an update. *Trends in Neurosciences*, *30*(5), 194–202.

Bunney, B. S., Aghajanian, G. K., & Roth, R. H. (1973). Comparison of effects of L-dopa, amphetamine and apomorphine on firing rate of rat dopaminergic neurones. *Nature: New Biology*, *245*, 123–125.

D'Ardenne, K., McClure, S. M., Nystrom, L. E., & Cohen, J. D. (2008). BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science*, *319*(5867), 1264–1267.

Daw, N. D., Courville, A. C., & Touretzky, D. S. (2006). Representation and timing in theories of the dopamine system. *Neural Computation*, *18*, 1637–1677.

Daw, N. D. & Doya, K. (2006). The computational neurobiology of learning and reward. *Current Opinion in Neurobiology*, *16*(2), 199–204.

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*(7095), 876–879.

Delgado, M. R., Schotter, A., Ozbay, E. Y., & Phelps, E. A. (2008). Understanding overbidding: using the neural circuitry of reward to design economic auctions. *Science*, *321*(5897), 1849–1852.

Esterman, M., Tamber-rosenau, B. J., Chiu, Y., & Yantis, S. (2011). Avoiding non-independence in fMRI data analysis: Leave one subject out. *NeuroImage*, *50*(2), 572–576.

Fowler, J. S., Volkow, N. D., Wang, G. J., Gatley, S. J., & Logan, J. (2001). [(11)]Cocaine: PET studies of cocaine pharmacokinetics, dopamine transporter availability and dopamine transporter occupancy. *Nuclear Medicine and Biology*, *28*(5), 561–572.

Garrison, J., Erdeniz, B., & Done, J. (2013). Prediction error in reinforcement learning: a meta-analysis of neuroimaging studies. *Neuroscience and Biobehavioral Reviews*, *37*(7), 1297–1310.

Knutson, B., Adams, C. M., Fong, G. W., & Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *The Journal of neuroscience*. *21*(159), 1-5.

Koob, G. F. & Nestler, E. J. (1997). The neurobiology of drug addiction. *The Journal of Neuropsychiatry and Clinical Neuroscience*, *9*(3), 482–497.

Koob, G. F. & Volkow, N. D. (2010). Neurocircuitry of addiction. *Neuropsychopharmacology*, *35*(1), 217–238.

Lawson, R. P., Seymour, B., Loh, E., Lutti, A., Dolan, R. J., Dayan, P., … Roiser, J. P. (2014). The habenula encodes negative motivational value associated with primary punishment in humans. *Proceedings of the National Academy of Sciences*, *111*(32), 11858–11863.

Ludvig, E. A., Sutton, R. S., & Kehoe, E. J. (2008). Stimulus representation and the timing of reward-prediction errors in models of the dopamine system. *Neural Computation*, *20*(12), 3034–3054.

Maldjian, J. A., Laurienti, P. J., Kraft, R. A., & Burdette, J. H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage*, *19*, 1233–1239.

Matsumoto, M., & Hikosaka, O. (2007). Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature*, *447*(7148), 1111–1115.

Montague, P. R., Dayan, P., & Sejnowski, T. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *The Journal of Neuroscience*, *16*(5), 1936–1947.

Montague, P. R., Hyman, S. E., & Cohen, J. D. (2004). Computational roles for dopamine in behavioural control. *Nature*, *431*(7010), 760–767.

Montague, P. R., King-Casas, B., & Cohen, J. D. (2006). Imaging valuation models in human choice. *Annual Review of Neuroscience*, *29*, 417–448.

Niv, Y., Duff, M., & Dayan, P. (2005). Dopamine, uncertainty and TD learning. *Behavioral and Brain Functions*, *1*(6), 1-9.

Niv, Y., Edlund, J. A., Dayan, P., & O'Doherty, J. P. (2012). Neural prediction errors reveal a risk-sensitive reinforcement-learning process in the human brain. *The Journal of Neuroscience*, *32*(2), 551–562.

Palminteri, S., Justo, D., Jauffret, C., Pavlicek, B., Dauta, A., Delmaire, C., … Durr, A. (2012). Critical roles for anterior insula and dorsal striatum in punishment-based avoidance learning. *Neuron*, *76*(5), 998–1009.

Parvaz, M. A., Konova, A. B., Proudfit, G. H., Dunning, J. P., Malaker, P., Moeller, S. J., … Goldstein, R. Z. (2015). Impaired neural response to negative prediction errors in cocaine addiction. *The Journal of Neuroscience*, *35*(5), 1872–1879.

Peciña, S. (2008). Opioid reward "liking" and "wanting" in the nucleus accumbens. *Physiology and Behavior*, *94*(5), 675–680.

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, *442*(7106), 1042–1045.

Redish, A. D. (2004). Addiction as a computational process gone awry. *Science*, *306*(5703), 1944–1947.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In *Classical conditioning: current research and theory*. New York, New York: Appleton-Century-Crofts.

Rolls, E. T., McCabe, C., & Redoute, J. (2008). Expected value, reward outcome, and temporal difference error representations in a probabilistic decision task. *Cerebral Cortex*, *18*(3), 652–663.

Rose, E. J., Salmeron, B. J., Ross, T. J., Waltz, J., Schweitzer, J. B., McClure, S. M., & Stein, E. a. (2014). Temporal difference error prediction signal dysregulation in cocaine dependence. *Neuropsychopharmacology*, *39*(7), 1732–1742.

Salas, R., Baldwin, P., de Biasi, M., & Montague, P. R. (2010). BOLD Responses to negative reward prediction errors in human habenula. *Frontiers in Human Neuroscience*, *4*(36), 1-7.

Schönberg, T., Daw, N. D., Joel, D., & O'Doherty, J. P. (2007). Reinforcement learning signals in the human striatum distinguish learners from nonlearners during reward-based decision making. *The Journal of Neuroscience*, *27*(47), 12860–12867.

Schonberg, T., O'Doherty, J. P., Joel, D., Inzelberg, R., Segev, Y., & Daw, N. D. (2010). Selective impairment of prediction error signaling in human dorsolateral but not ventral

striatum in Parkinson's disease patients: evidence from a model-based fMRI study. *NeuroImage*, *49*, 772–781.

Schultz, W. (2011). Potential vulnerabilities of neuronal reward, risk, and decision mechanisms to addictive drugs. *Neuron*, *69*(4), 603–617.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593–1599.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction* (Vol. 1). Cambridge, Massachusetts: A Bardford Book.

Thorndike, E. L. (1911). *Animal intelligence: Experimental studies*. Norwood, Massachusetts: Norwood Press.

Tiffany, S. T., Singleton, E., Haertzen, C. A., & Henningfield, J. E. (1993). The development of a cocaine craving questionnaire. *Drug and Alcohol Dependence*, *34*, 19–28.

Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. a. (2007). The neural basis of loss aversion in decision-making under risk. *Science*, *315*(5811), 515–518.

UNODC. (2012). World Drug Report 2011. Retrieved from http://www.unodc.org/unodc/en/data-and-analysis/WDR-2011.html

Volkow, N., Chang, L., Wang, G., Fowler, J., Ding, Y., Sedler, M., … Pappas, N. (2001). Low level of brain dopamine D2 receptors in methamphetamine abusers: association with metabolism in the orbitofrontal cortex. *The American Journal of Psychiatry*, *158*(12), 2015–2021.

Volkow, N., Fowler, J., Wang, G., Hitzemann, R., Logan, J., Schlyer, D., … Wolf, A. (1993). Decreased dopamine D2 receptor availability is associated with reduced frontal metabolism in cocaine abusers. *Synapse*, *14*(2), 169–177.

Volkow, N., Fowler, J., Wang, G., Swanson, J., & Telang, F. (2007). Dopamine in drug abuse and addiction: results from imaging studies and treatment implications. *Archives of Neurology*, *64*(11), 1575–1579.

West, R., & Hardy, A. (2005). *Theory of Addiction*. Oxford, London: Blackwell Publishing Ltd.

Table 1

Participant Characteristics.

a)

Variable	Mean (sd)
Age	45.7 (7.0)
Education - Years	12.9 (1.0)
WTAR*	98.4 (10.8)
Years of Cocaine Use	17.6 (7.9)

* Wechsler Test of Adult Reading

b)

Drug Use Information	Abstinent (se)	Using As Usual (se)
Cocaine Intake Last 48 hours	0g (0)	1.5g (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	137.9 (8.7)	145.3 (8.7)

** Cocaine Craving Questionnaire

Table 2

fMRI Results. a) Significant family wise error corrected clusters of positive prediction error ($\delta+$) neural signals enhanced with cocaine abstinence. Voxels were thresholded at T = 2.90 and results were small volume corrected for the entire striatum including nucleus accumbens defined by Garrison et al. (2013). b) Significant family wise error corrected clusters where positive learning rate ($\alpha+$) enhancement from cocaine abstinence predicted positive prediction error ($\delta+$) enhancement.

a)

Location	cluster FWE *p*	cluster size	peak T	peak MNI coordinates
Right Striatum	0.008	65	4.65	21 21 8
Left Striatum	0.016	52	4.55	-17 17 11

b)

Location	cluster FWE *p*	cluster size	peak T	peak MNI coordinates
Right Striatum	0.004	117	5.61	14 7 -9
Left Striatum	0.004	112	7.2	-20 10 -6

Table 3

First Level fMRI Results. Significant family wise error corrected clusters of positive prediction error ($\delta+$) neural signals correlated with abstinent cocaine dependent individuals' striatal neural signals during corresponding outcomes. Voxels were thresholded at T = 2.90 and results were small volume corrected for the entire striatum including nucleus accumbens defined by Garrison et al. (2013).

Location	cluster FWE *p*	cluster size	peak T	peak MNI coordinates
Right Striatum	0.005	75	5.33	24 7 -6
Left Striatum	0.08	28	4.29	-24 14 8

Figure 1

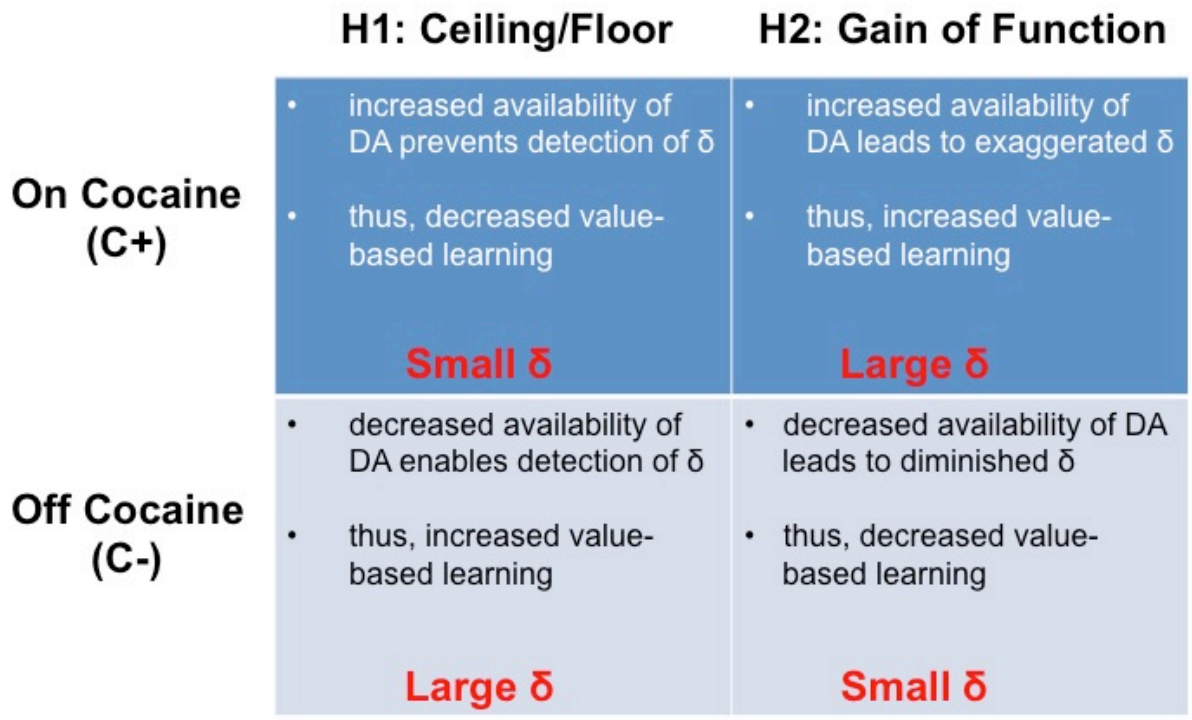

Figure 1. Contrasting prediction error (δ) hypotheses from dependence. There are two distinct and competing hypotheses of how cocaine might affect aversive learning in dependent individuals while they are using as usual (C+) or abstinent (C-). For the "Ceiling/Floor" hypothesis, it is proposed that C+ individuals have increased tonic DA from cocaine use and lower postsynaptic DA receptors from long term adaptation to dependence leading to decreased availability of postsynaptic DA and preventing the detection of δ. C- individuals have decreased tonic DA as a long term adaptation to dependence, which frees up postsynaptic DA receptors and enables the detection of δ. For the "Gain of Function" hypothesis, C+ individuals have increased tonic DA from cocaine use exaggerating δ as more DA is available to bind to postsynaptic DA receptors. C- individuals do not have increased tonic DA, so they would show diminished δ relative to C+.

Figure 2

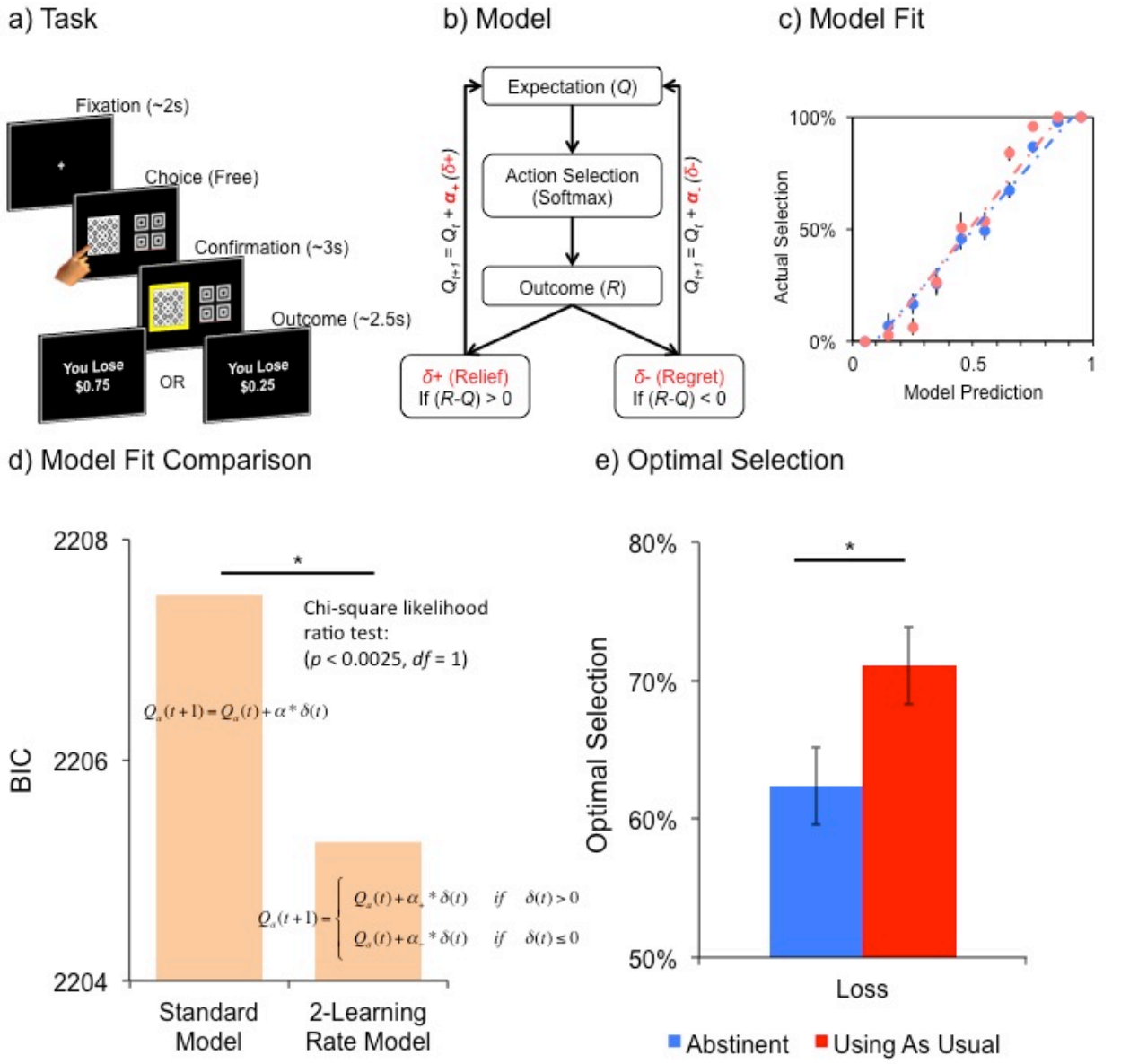

Figure 2. Experimental Design and Model Fit. a) Experimental task. Subjects selected between two distinct patterns and the subsequent outcome was shown. In this example, the selected pattern has a 75% probability of losing $0.25 and a 25% probability of losing $0.75, as it is the better of the two patterns. The participant completed up to 36 trials or until the better option was learned, defined as selecting the better option 80% of the time in the 5 most recent trials. b) Model. The best fitting Q-learning model where positive prediction error (δ+) and negative

prediction error (δ-) updates the subsequent expected value (Q) at separate estimated rates (α+ and α- respectively). c) Model Fit. Model predicted probability of selecting the better option compared to the actual ratio of the better option selected across subjects. d) The 2 learning parameter model including +/- δ significantly fit better than the standard Q-Learning model using a chi-square comparison ($p < 0.0025$, $df = 1$) and BIC comparison (2 learning parameter TD model: 2205.3; Standard: 2207.5). e) Abstinent individuals (C-) performed worse than using as usual (C+) individuals.

Figure 3

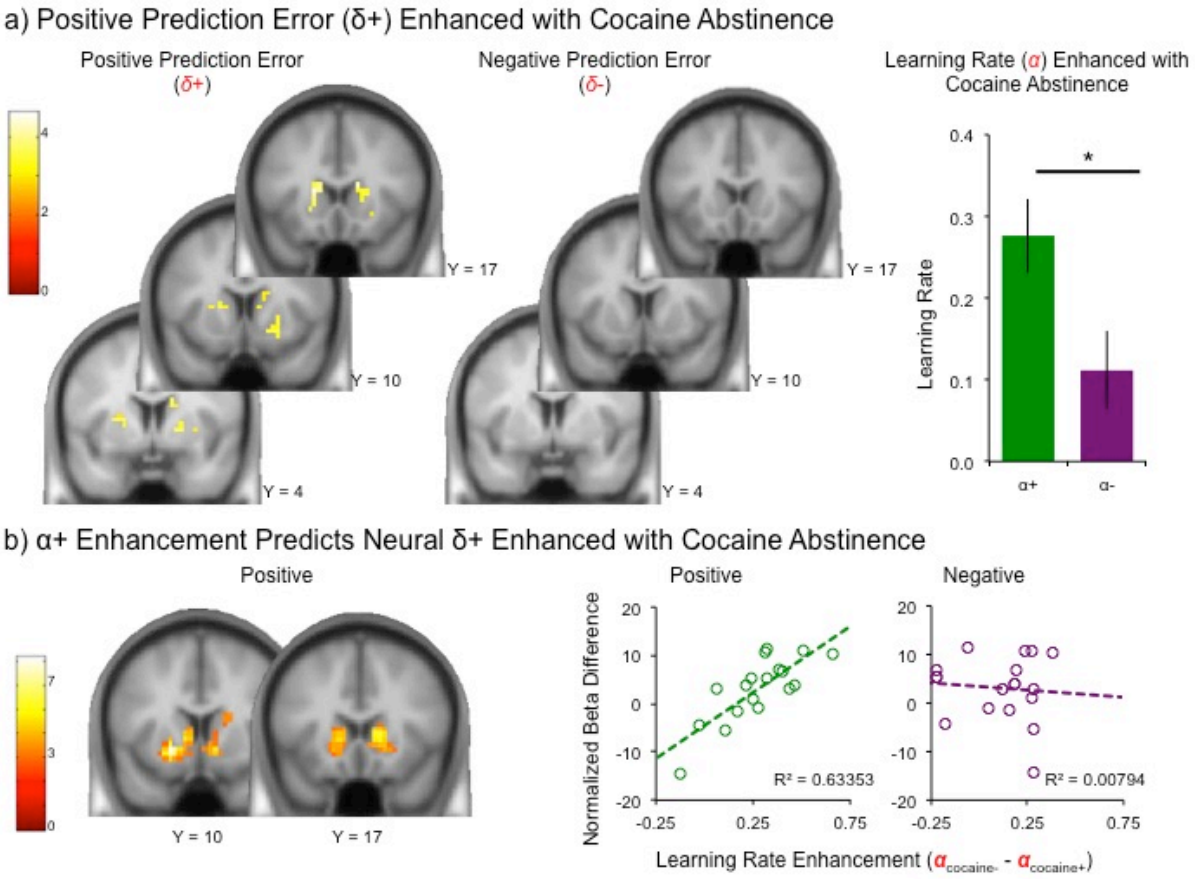

Figure 3. Imaging Results. a) When cocaine dependent individuals were abstinent from cocaine use (verified by urine testing), they showed greater striatal signals during the outcome event to positive prediction errors (δ+) only, and not to negative prediction errors (δ-). Significant clusters shown utilizes a paired two sample T test contrasting C- > C+. The neuroimaging results paralleled learning rate estimates where positive learning rate (α+) was significantly enhanced by cocaine abstinence (C- > C+) in comparison to negative learning rates (α-). b) When learning rate enhancements were correlated with striatal prediction error signals for positive and negative separately, only the increase in positive learning rate (α+) enhancement predicted striatal positive prediction error enhancement (δ-).

Figure 4

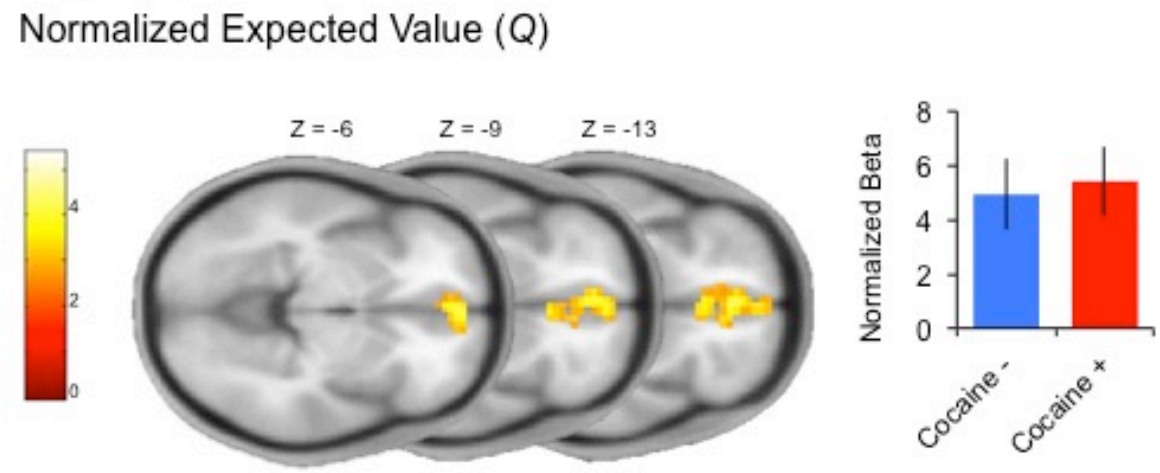

Figure 4. Expected Value. Cocaine dependent individuals showed ventral medial prefrontal cortex (VMPFC) signal during the choice event corresponding to the expected normalized expected value that was selected during that trial. There were no differences while abstinent compared to using as usual in the cocaine dependent individuals' VMPFC signals.

Figure 5

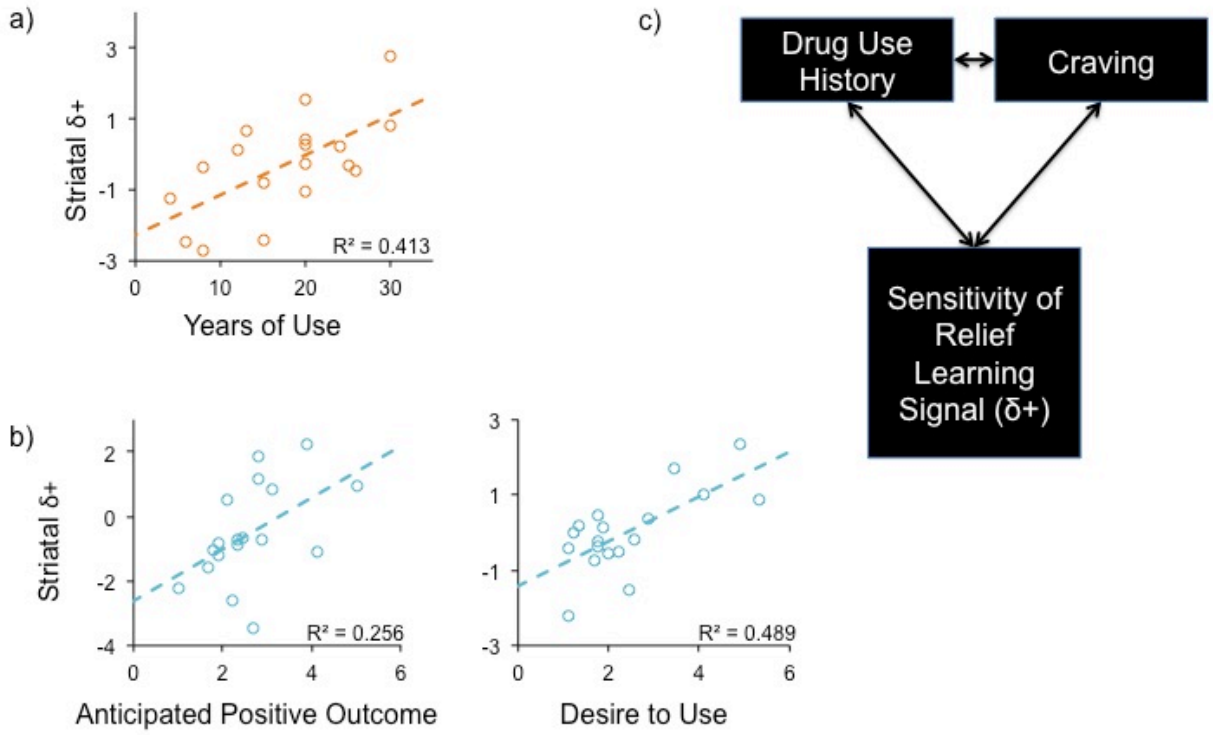

Figure 5. Cocaine Use and Craving Comparisons. a) Increased history of cocaine use predicts exacerbated neural signal to all better than expected outcomes in the striatum while abstinent. Neural signal presented are within-subject normalized betas across all voxels in the brain. b) In turn, higher better than expected striatal signal predicted higher anticipated positive outcome to cocaine use and higher desire to use cocaine in abstinent cocaine dependent individuals. c) Schematic illustration of hypothesized relationships indicating with longer duration of use and greater withdrawal, dependent individuals' neural learning system may adapt to be more sensitive to learning about actions that will provide individuals' with relief.