Published in final edited form as:

Psychopharmacology (Berl). 2013 October ; 229(3): 493–501. doi:10.1007/s00213-013-3222-6.

Disruption of model-based behavior and learning by cocaine self-administration in rats

Heather M. Wied,

Cellular Neurobiology Research Branch, Behavioral Neurophysiology Research Section, National Institute on Drug Abuse Intramural Research Program, 251 Bayview Blvd., Baltimore, MD 21224, USA, Department of Anatomy and Neurobiology, University of Maryland School of Medicine, 20 Penn St., Baltimore, MD 21201, USA

Joshua L. Jones,

Cellular Neurobiology Research Branch, Behavioral Neurophysiology Research Section, National Institute on Drug Abuse Intramural Research Program, 251 Bayview Blvd., Baltimore, MD 21224, USA, Department of Anatomy and Neurobiology, University of Maryland School of Medicine, 20 Penn St., Baltimore, MD 21201, USA

Nisha K. Cooch,

Cellular Neurobiology Research Branch, Behavioral Neurophysiology Research Section, National Institute on Drug Abuse Intramural Research Program, 251 Bayview Blvd., Baltimore, MD 21224, USA, Department of Anatomy and Neurobiology, University of Maryland School of Medicine, 20 Penn St., Baltimore, MD 21201, USA

Benjamin A. Berg, and

Cellular Neurobiology Research Branch, Behavioral Neurophysiology Research Section, National Institute on Drug Abuse Intramural Research Program, 251 Bayview Blvd., Baltimore, MD 21224, USA

Geoffrey Schoenbaum

Cellular Neurobiology Research Branch, Behavioral Neurophysiology Research Section, National Institute on Drug Abuse Intramural Research Program, 251 Bayview Blvd., Baltimore, MD 21224, USA, Department of Anatomy and Neurobiology, University of Maryland School of Medicine, 20 Penn St., Baltimore, MD 21201, USA, Solomon H. Snyder Department of Neuroscience, The Johns Hopkins University, Baltimore, MD, USA

Geoffrey Schoenbaum: geoffrey.schoenbaum@nih.gov

Abstract

Rationale—Addiction is characterized by maladaptive decision-making, in which individuals seem unable to use adverse outcomes to modify their behavior. Adverse outcomes are often infrequent, delayed, and even rare events, especially when compared to the reliable rewarding drug-associated outcomes. As a result, recognizing and using information about their occurrence put a premium on the operation of so-called model-based systems of behavioral control, which allow one to mentally simulate outcomes of different courses of action based on knowledge of the underlying associative structure of the environment. This suggests that addiction may reflect, in part, drug-induced dysfunction in these systems. Here, we tested this hypothesis.

Objectives—This study aimed to test whether cocaine causes deficits in model-based behavior and learning independent of requirements for response inhibition or perception of costs or punishment.

Methods—We trained rats to self-administer sucrose or cocaine for 2 weeks. Four weeks later, the rats began training on a sensory preconditioning and inferred value blocking task. Like devaluation, normal performance on this task requires representations of the underlying task structure; however, unlike devaluation, it does not require either response inhibition or adapting behavior to reflect aversive outcomes.

Results—Rats trained to self-administer cocaine failed to show conditioned responding or blocking to the preconditioned cue. These deficits were not observed in sucrose-trained rats nor did they reflect any changes in responding to cues paired directly with reward.

Conclusions—These results imply that cocaine disrupts the operation of neural circuits that mediate model-based behavioral control.

Keywords

Addiction; Cocaine; Orbitofrontal; Sensory preconditioning; Blocking

Introduction

Addiction is characterized by a pattern of maladaptive decision-making in which subjects continue to seek drugs despite serious adverse consequences (Gawin 1991; Mendelson and Mello 1996). This pattern of behavior is captured in the diagnostic criteria for substance dependence in the DSM-IV-TR (APA 2000), which *require* drug-seeking to be persistent in the face of attempts to stop or cut back in order to avoid various negative outcomes or costs. While there are a number of ways to describe such apparent deficits—inability to inhibit established behaviors (Jentsch and Taylor 1999; Winstanley et al. 2010), increased desire or wanting of the drug (Robinson and Berridge 2000, 2008), loss of sensitivity to punishment or costs (Deroche-Gamonet et al. 2004; Vanderschuren and Everitt 2004), or disruption of the normal homeostatic set point of reward and anti-reward systems (Koob and Le Moal 2008)—none of these sufficiently capture the complex array of the behavioral deficits. Addicted individuals clearly can act affirmatively to obtain desired outcomes and can even learn relatively complex strategies or policies in order to obtain drugs. They can also redirect behavior when confronted with immediate consequences or costs. Yet, despite these apparently intact abilities, they are unable to modify drug seeking in response to real-world events that most people would respond to appropriately.

One distinguishing feature of these real-world events—one which is typically not captured by experimental manipulations—is that the certain consequences of drug use tend to be highly variable and unpredictable, delayed, and even rare events. Often, these consequences are adverse, such as financial and family problems, physical injury, or incarceration. All of these are all highly likely over the long term; however, they tend to be unlikely in the very short term, at least as compared to the immediate goal and certain experience associated with the drug-associated cues and drug high or alleviation of withdrawal symptoms. Further, these long-term events are unlikely to happen over and over again, so it is difficult to learn about them in the same way that one can learn about the repetitive, immediate rewarding aspects of drug use. As a result, the successful application of information about these consequences likely requires so-called model-based systems of behavioral control, which allow one to utilize information about the causal structure of the environment, which may be only anecdotally or incidentally related to the behavior at hand, in order to mentally simulate the possible outcomes of a behavior (Daw et al. 2005; Niv et al. 2006). Failure of these

systems would result in behavior—and learning—that is heavily dependent on cached or precomputed policies developed through prior direct experience. To that end, learning and behavior after cocaine exposure appear normal or even enhanced under conditions in which model-free systems are sufficient (Mendez et al. 2010; Roesch et al. 2007; Simon et al. 2007; Taylor and Horger 1999; Taylor and Jentsch 2001; Wyvell and Berridge 2001).

Thus, the core behavioral features of addiction may reflect, in part, either a preexisting or drug-induced dysfunction in the model-based control of behavior (Lucantonio et al. 2012). Consistent with this idea, studies using reinforcer devaluation have shown that exposure to psychostimulants, either passively or by active self-administration, causes long-lasting deficits in the ability of rats to adjust learned behaviors after changes in the value of the associated reward (Nelson and Killcross 2006; Schoenbaum and Setlow 2005). These deficits are observed spontaneously after devaluation (i.e., without new learning) and appear in the face of apparently normal acquisition of both the behavior (CS→US or R→US) and of the new reward value (US→illness).

While these studies are consistent with the idea that some drugs affect model-based processing, they are similar and thus provide only a single data point. Further, they require the animal to appreciate an adverse outcome in order to redirect or withhold responding. This leaves open, at least somewhat, questions over whether the deficits reflect problems with specific functions, such as sensitivity to punishment or the ability to inhibit responding, rather than a more general impairment in the integrity of model-based processing, which would be orthogonal to these idiosyncratic task requirements. The case that drugs can disrupt model-based processing would be much strengthened by the provision of convergent evidence from a task that shares the requirement for inferential reasoning that defines reinforcer devaluation, but does not depend on either of these more specific capabilities. Here, we provide such evidence, using a sensory preconditioning and inferred value blocking task in which we have previously demonstrated the role of the orbitofrontal cortex in model-based behavior and learning (Jones et al. 2012).

Materials and methods

Subjects

Forty-five adult male Long-Evans rats (Charles River Laboratories, Wilmington, MA) ($n=30$ controls including 13 naïve, nonsurgical controls, and 17 sucrose-trained, surgical controls; $n=15$ cocaine trained) weighing 275–325 g on arrival were individually housed and given ad libitum access to food and water, except during behavioral training and testing. During these procedures, rats were food deprived to 85–90 % of their baseline body weight. Rats were maintained on a 12-h light/dark cycle and trained and tested during the light cycle. Experiments were performed at the University of Maryland School of Medicine and National Institute on Drug Abuse, Intramural Research Program, in accordance with the University of Maryland and NIH guidelines.

Apparatus

Self-administration and behavioral training and testing were conducted in standard behavior boxes measuring 30.48×25.40×30.48 cm (Coulbourn Instruments) enclosed within sound-attenuating boxes. During self-administration, each chamber was equipped with two 4-cm-wide levers on opposite walls located 8 cm above the grid floor. The lever located on the right side of the chamber was active and retractable and led to activation of the infusion pump, while the lever on the left side of the box was inactive and stationary. For cocaine self-administration, each rat had silastic tubing that was protected by a metal spring that extended from the intravenous catheter to a liquid swivel (Instech Laboratories, Plymouth

Meeting, PA) mounted to an arm fixed outside the behavior chamber. Tygon tubing extended from the liquid swivel to an infusion pump located outside the chamber. For sucrose self-administration, 0.04 mL of a 10 % sucrose solution was placed in a recessed dipper in the center of the right wall approximately 6 cm above the floor grid.

Sensory preconditioning and blocking were conducted in a different room, but using the same type of standard apparatus in sound-attenuating boxes. A recessed food cup was placed in the center of the right wall approximately 2 cm above the floor. The feeder was mounted outside the behavior chamber and delivered 45 mg sucrose pellets into the food cup. Auditory cues were used during the behavioral training and testing. The tone speaker was located on the wall above the food cup and calibrated to ~70 dB. The tone cues were produced by an audio generator that produced three tones, one as a pure tone frequency, and a siren tone that alternated between two frequencies every 100 ms. In addition, a clicker (2 Hz) was mounted on the wall above the food cup, and a white noise speaker calibrated to ~65 dB was placed external to the behavior box behind the right wall. Visual cues that were used during inferred value blocking included a house light and cue light that were placed on the wall to the left or the right of the food cup approximately 10 cm above the floor.

Surgical procedures

Thirty-two rats underwent surgery for implantation of chronic intravenous jugular catheters. Rats were anesthetized with ketamine (100 mg/kg, i.p., Sigma) and xylazine (10 mg/kg, i.p., Sigma). A silastic catheter was inserted into the right jugular vein and passed subcutaneously to either the top of the skull where it was attached to a modified 22-gauge cannula (Plastics One, Roanoke, VA) and mounted to the rat's skull with dental cement and skull screws or fixed to cannula on nylon mesh which was sutured subcutaneously between the scapulae. Carprofen (5 mg/kg, s.c., Pfizer) was given during surgery and for 2 days postsurgery to relieve postoperative pain. After surgery, and throughout self-administration training, catheters were flushed daily with gentamicin/saline solution (0.16 % gentamicin, BioSource International) to prevent infection.

Self-administration

Rats recovered for 14 days following surgery before starting self-administration training. These rats were trained to lever press for cocaine-HCl (0.75 mg/kg/infusion; $n=15$) or sucrose (10 % w/v; $n=17$) for 14 days (Fig. 1). Rats were trained to press the active lever to self-administer cocaine under a fixed ratio (FR1) schedule of reinforcement with each press delivering a 4-s infusion of cocaine. After each infusion, the active lever retracted and was followed by a 40-s timeout period. All self-administration sessions lasted 3 h, with 15-min timeout periods after each hour. To prevent overdose, the number of cocaine infusions was limited to 20/h such that when the rat had received 20 cocaine infusions, the active lever was retracted for the remainder of the hour. The same procedures were followed for animals that self-administered sucrose, except the rats were shaped to press the active lever to receive 0.04 mL of a 10 % sucrose solution delivered in a recessed dipper.

Sensory preconditioning

Catheterized rats ($n=32$) began sensory preconditioning 4 weeks after the end of self-administration. At the same time, we also trained naïve rats ($n=13$) who had received neither surgery nor any prior training. The sensory preconditioning procedure consisted of three phases (Fig. 2), using procedures identical to those in our prior study (Jones et al. 2012):

1. *Preconditioning*: Rats were shaped to retrieve sucrose pellets from the food cup in a single session that included ten deliveries of two 45-mg sucrose pellets ($n=45$ total subjects; $n=30$ controls; $n=15$ cocaine). After shaping, all rats received 2 days

of sensory preconditioning involving auditory cues. Each day, rats were exposed to two blocks of six trials. A block of trials consisted of presentations of paired auditory cues (A→B and C→D). Specifically, a 10-s presentation of cue A would be followed immediately by a 10-s presentation of cue B, in a six-trial block. The second block would consist of six presentations of the pairing a 10-s presentation of cue C followed immediately by a 10-s presentation of cue D. The inter-trial intervals varied from 3 to 6 min. Auditory cues A and C were either white noise or clicker counterbalanced, and auditory cues B and D were either siren or tone counterbalanced. The order of cue blocks was counterbalanced across each of the two sensory preconditioning days.

2. *Conditioning:* After 2 days of preconditioning, rats underwent Pavlovian conditioning for six consecutive days. Each day, rats received a single training session, in which a total of 12 trials were presented: six trials of cue B paired with a food reward and six trials of D paired with no food reward. Auditory cue B was paired with three sucrose pellets that were delivered into the food cup during the 10-s presentation (one pellet was delivered after 3, 3, and 4 s time elapsed). Alternatively, auditory cue D was presented alone for 10 s with no sucrose delivered. Trials were counterbalanced including three trials of alternating blocks for a total of six reinforced auditory cue B presentations and six unreinforced auditory cue D presentations. The inter-trial intervals varied between 3 and 6 min.
3. *Probe test:* After conditioning, a single probe test day included reminder training of three trials of B paired with sucrose and three trials of D unpaired. After the presentations of B and D, rats were exposed to two blocks of counterbalanced auditory cues A and C presented in extinction (six presentations each). The inter-trial intervals varied between 3 and 6 min.

Inferred value blocking

This procedure consisted of two training phases (Fig. 3), using procedures identical to those in our prior study (Jones et al. 2012):

1. *Blocking:* A subset of rats from the prior experiment ($n=35$; $n=22$ controls; $n=13$ cocaine) underwent additional compound conditioning to test whether cocaine would also disrupt the normal ability of the inferred value to block new learning. These rats were trained for 2 days with preconditioned auditory cues A and C paired with novel light cues X and Y (house light; flashing cue light; counterbalanced), with all presentations reinforced with a food reward. Auditory cues A and C were played together in compound with visual cues X and Y (i.e., AX and CY) for 10 s each. Each compound cue was presented in two blocks of three trials, alternating cue blocks, for a total of 12 trials each day (six AX and six CY, counterbalanced). Each 10-s compound cue was reinforced by the delivery of a sucrose pellet at 3, 6, and 10 s after cue onset. Inter-trial intervals were between 3 and 6 min.
2. *Probe test:* After compound conditioning, rats were given a single probe test consisting of reminder training of three trials each of AX and CY paired with normal sucrose delivery, followed by two blocks of visual cues X and Y presented alone (four presentations each) for 10 s each without sucrose delivery. The inter-trial intervals varied between 3 and 6 min. The first three trial presentations of X and Y were analyzed.

Response measures

During self-administration, active and inactive lever pressing was recorded. During sensory preconditioning, the percentage of time rats spent with their head in the food cup during cue presentation was measured by an infrared photo beam positioned at the front of the food cup. The analysis for auditory and visual cues presented alone was restricted to the last 5 s of cue presentation. Baseline response measures in the corresponding pre-cue periods were subtracted from the cue-evoked response to account for individual variation in baseline activity. Rearing behavior to visual stimuli during inferred value blocking in both the compound conditioning sessions and blocking probe test was scored from video recordings with a handheld timer. In order to correct for time spent rearing (Takahashi et al. 2009; Jones et al. 2012), the percentage of time responding during the visual stimuli, and corresponding baseline periods, was calculated as follows:

$$\% \text{ of responding} = 100 \times ([\% \text{ of time in food cup}] / [100 - \% \text{ of time rearing}])$$

Statistical analyses

Data were acquired using Coulbourn Graphic State 2 software (Coulbourn Instruments). Raw data were processed in Matlab to extract lever presses in self-administration and percentage time spent in the food cup for sensory preconditioning and inferred value blocking. For all statistical tests, an alpha level of 0.05 was used. The naïve and surgical control groups did not differ at the end of conditioning or during probe testing; thus, their data were collapsed to form a single control group. All data were analyzed in Statistica, using repeated measures analysis of variance with Bonferroni post hoc testing when appropriate ($p < 0.05$).

Results

Self-administration

Chronic intravenous jugular catheters were implanted in rats (17 sucrose controls; 15 cocaine). After recovering from surgery, these rats self-administered either cocaine or sucrose for 14 days (Fig. 1). Rats acquired cocaine and sucrose self-administration as demonstrated by a significant interaction between lever and day in the analysis of active and inactive lever presses for cocaine ($F_{(13, 338)} = 2.01, p < 0.05$) and sucrose ($F_{(13, 403)} = 5.99, p < 0.01$). In addition, there was a significant effect of training day for the number of infusions of cocaine ($F_{(13, 182)} = 15.96, p < 0.01$) and number of sucrose deliveries ($F_{(13, 208)} = 5.24, p < 0.01$).

Sensory preconditioning

1. *Preconditioning*: In preconditioning, rats were taught to associate two pairs of neutral auditory cues ($A \rightarrow B$ and $C \rightarrow D$). Food cup responding during the presentation of each cue was taken as an index of conditioning. Rats in both groups responded at or near baseline levels to all cues (Fig. 2a, b). A two-factor ANOVA (cue $\times$ treatment) comparing the percent of time the rats spent in the food cup demonstrated no significant main effects or any interactions ($F_s < 2.08; p_s > 0.16$) (Fig. 2a, b).
2. *Conditioning*: During the 6 days of conditioning, the rats learned to discriminate between the rewarded auditory cue B and the non-rewarded auditory cue D (Fig. 2c, d). Rats in both groups increased responding to B, but not to D during these six conditioning sessions. A three-factor ANOVA (cue $\times$ treatment $\times$ session) revealed significant main effects of cue ($F_{(1, 86)} = 221.99, p < 0.01$) and session ($F_{(5, 430)} = 27.62, p < 0.01$) and significant interactions between treatment and cue ($F_{(1, 86)} = 9.88, p < 0.01$), treatment and session ($F_{(5, 430)} = 2.34, p < 0.05$), cue and

session ($F_{(5, 430)}=13.01, p<0.01$), and cue, treatment, and session ($F_{(5, 430)}=2.38, p<0.05$). Effects involving treatment were primarily due to differences between the two groups in the initial rate of conditioning; controls acquired the discrimination between B and D faster than cocaine animals ($p<0.05$) (session 1 versus session 3). Thus, while control and cocaine-experienced rats responded similarly to cue D in all sessions (Bonferroni post hoc correction; $ps>0.05$), responding to cue B differed in the first session ($p<0.05$) but not thereafter ($ps>0.05$). This likely reflected the prior experience with reward retrieval for controls. Accordingly, there were no significant differences between groups by the end of conditioning ($ps>0.05$).

3. *Probe test:* In the probe test, both control and cocaine-experienced rats responded more to B, the cue previously paired with sucrose, than to D, the non-reinforced cue (Fig. 2e, 2f). Controls also exhibited significantly greater responding to the preconditioned cue, A, which had been previously paired with the rewarded cue, B, than to the preconditioned control cue, C (Fig. 2e). Cocaine-experienced rats showed no difference in responding to the preconditioned cues A and C (Fig. 2f). A two-factor ANOVA (cue $\times$ treatment) indicated a significant main effect of cue ($F_{(3, 129)}=85.76, p<0.01$) and a significant interaction between cue and treatment ($F_{(3, 129)}=2.90, p<0.05$). Bonferroni post hoc testing showed that both groups responded significantly more to B than D ($ps<0.05$), and controls responded significantly more to A than to either C or D ($ps<0.05$), while cocaine-experienced rats responded similarly to A, C, and D ($ps>0.05$).

Inferred value blocking

1. *Blocking:* During the 2 days of compound conditioning, the preconditioned cues, A and C, were presented in compound with new visual cues, X and Y. AX and CY were both reinforced with the same sucrose reward previously paired with B. A three-factor ANOVA (treatment $\times$ cue $\times$ session) demonstrated significant main effects of session ($F_{(1, 66)}=22.69, p<0.01$) and treatment ($F_{(1, 66)}=9.62, p<0.01$) and a significant interaction between cue and session ($F_{(1, 66)}=6.35, p<0.05$). While cocaine-experienced rats exhibited slightly lower responding overall, there were no significant interactions with treatment and other factors ($F_s<0.91, ps>0.34$) (Fig. 3a, b).
2. *Probe test:* In the probe test, rats were presented with reinforced AX and CY reminder trials, followed by probe trials, in which X and Y were presented alone. A two-factor ANOVA (treatment $\times$ cue) on conditioned responding during the probe trials revealed a significant interaction between cue and treatment ($F_{(1, 33)}=8.83, p<0.05$). Bonferroni post hoc testing showed that controls responded significantly more to Y than to X ($p<0.05$) (Fig. 3c), whereas rats exposed to cocaine responded similarly to these two cues ($p>0.05$) (Fig. 3d). Subsequent comparisons of cue-evoked responding to responding in the baseline period showed that while controls exhibited both a significant main effect of period ($F_{(1, 42)}=19.7, p<0.01$) and a significant period by cue interaction ($F_{(1, 42)}=7.00, p<0.05$), cocaine-experienced rats exhibited only a main effect of period ($F_{(1, 24)}=7.68, p<0.05$; other $ps>0.66$).

Discussion

Here, we have shown that cocaine self-administration disrupts behavior and learning that depends upon inferred value. In a sensory preconditioning task, cocaine-experienced rats were not able to use the relationship between the preconditioned cue and the subsequently conditioned cue to infer value in order to guide behavior, nor were they able to mobilize this information in order to modulate learning during the blocking task. These effects were

enduring, appearing during probe tests conducted 5–6 weeks after the last contact with cocaine. In addition, they were specific inasmuch as the cocaine-experienced rats responded like controls to the cues that were directly paired with reward.

These data add to prior reports that psychostimulants affect reinforcer devaluation (Nelson and Killcross 2006; Schoenbaum and Setlow 2005). As in the sensory preconditioning task used here, normal changes in behavior after reinforcer devaluation require access to a model-based representation of the causal structure of the task, so that never before experienced consequences can be mentally simulated. As noted earlier, rats that have been exposed to psychostimulants show long-lasting deficits in their ability to inhibit learned Pavlovian and instrumental behaviors after reinforcer devaluation. Our current results show a similar deficit, but in a setting that requires neither response inhibition nor the ability to perceive and respond appropriately to aversive information. In addition, the current data extend these findings and show that these operations are impacted both for the purpose of guiding behavior appropriately as well as for new learning derived from this information, since cocaine-experienced rats were also unable to use the inferred value of the preconditioned cue to block learning. Overall, these results substantially strengthen the idea that exposure to some addictive drugs disrupts the neural circuits that mediate model-based processing.

Currently, candidates for mediating this model-based processing include prefrontal regions, such as the medial and orbital prefrontal cortices. Changes in instrumental responding after reinforcer devaluation as well as extinction learning, particularly recall, depend critically on input from medial prefrontal cortex to dorsal striatal regions (Corbit and Balleine 2003; Killcross and Coutureau 2003; Milad and Quirk 2002; Quirk et al. 2000; Rhodes and Killcross 2004, 2007; Zapata et al. 2010). Similarly, flexible responding to Pavlovian cues after devaluation and in other settings requires orbitofrontal output that impacts areas in dorsal and ventral striatum (Izquierdo et al. 2004; Machado and Bachevalier 2007; Pickens et al. 2003; West et al. 2011). Altered prefrontal processing would therefore be expected to lead to behavioral deficits in these areas. Indeed, orbitofrontal inactivation causes precisely the deficits reported here after cocaine use in the sensory preconditioning and inferred value blocking tasks (Jones et al. 2012).

Despite the similarity of our cocaine-induced deficits to those seen in OFC-inactivated rats (Jones et al. 2012), the current data cannot eliminate the possibility that the deficit occurs because cocaine-experienced rats are unable to acquire the stimulus–stimulus association. That is, because cocaine exposure occurred prior to any conditioning, it is possible that the probe test deficits are a result of an inability to learn the stimulus–stimulus contingency in the first phase of the experiment. Such an effect might reflect an impact of cocaine on hippocampal regions known to be important for model-based behaviors and stimulus–stimulus learning (Wimmer and Shohamy 2011, 2012). However, we are not aware of evidence that stimulus–stimulus learning is impaired by cocaine exposure, and given the strong evidence implicating the prefrontal regions in addiction and in the type of processing isolated in later phases of the task, we would view this as a more parsimonious explanation of the current results.

It is also worth noting that although deficits in some model-based behaviors have been conceived of as reflecting strengthened processing of the countervailing “habits” in striatal regions (Belin and Everitt 2008; Robbins and Everitt 1999), the current results cannot be easily explained by such a mechanism. This is because stronger encoding of the cached value or policy associated with the rewarded cue, in concert with an intact model-based system, should have led to stronger, not weaker, responding to the preconditioned cue and better, not worse, blocking. Thus, the only reasonable explanation that accounts for the

current results is alterations in model-based processing. While this does not preclude a contribution of stronger habits to other aspects of drug seeking, it is in fact consistent with what we and others have reported, which is that psychostimulants cause rather modest changes in information processing in dorsal striatum (Takahashi et al. 2007) while substantially altering that in orbitofrontal cortex (Homayoun and Moghaddam 2006; Stalnaker et al. 2006).

So, what are the implications of these findings for understanding and, more importantly, addressing behavioral issues in addiction? One implication is that individuals with a history of psychostimulant use will be less responsive than other patients to so-called talking therapies, which require insight and mental simulation to have their effects. This would be consistent with the high rates of relapse after even apparently successful psychotherapeutic interventions (Dutra et al. 2008). A second implication is that psychostimulant users are likely to be less adept than other patients at learning when outcomes are not as rewarding as expected, especially to the extent that these differences are defined by observational and inference-based evidence. As noted in our introduction, we believe that this type of deficit is apparent in the diagnostic criteria for addiction.

Identifying a functional deficit and associated circuit can allow for better insight into these behaviors and perhaps targeting of affected individuals with pharmacologic agents known to enhance prefrontal function. For example, imaging research has long identified frontostriatal, and particularly orbitofrontal, circuits as abnormally active (or inactive) in psychostimulant users (Goldstein and Volkow 2011). Abnormal activity patterns have been associated with reactivity to drug-associated cues, craving, and also deficits in reward perception. One particularly interesting result is that blood oxygen level-dependent (BOLD) signal in psychostimulant users fails to reflect monetary reward gradients normally (Goldstein et al. 2007b). That is, while BOLD signal increases generally to a monetary reward in cocaine users, it does not scale with the amount or magnitude. This result has been replicated several times and is associated with altered P300 and striatal signaling, chronicity of drug use, impairments in self-control, and even lack of insight (Goldstein et al. 2007a; Konova et al. 2012; Moeller et al. 2012; Parvaz et al. 2012). The latter speculation is particularly intriguing in light of the current results, since insight arguably depends upon the same process of mental simulation that characterizes model-based processing. Accordingly, we would suggest that the failure of BOLD signal in frontal areas to scale with reward likely reflects dysfunctional processing and failure of operation of these model-based systems. This is because the primary difference between a \$10 bill and a \$100 bill lies not in its sensory qualities, but rather in the quantity and quality of goods that one can obtain with each. Fully appreciating this difference requires inference and mental simulation. Lacking this ability, patients and orbitofrontal BOLD signal fail to identify any difference between these two green pieces of paper—both are similarly significant.

If this is correct, then these imaging findings provide a marker for altered prefrontal function in specific individuals. Such a diagnostic tool might then be combined with pharmacologic agents known to enhance prefrontal-dependent processing. This effort has been the focus of recent endeavors with methylphenidate (Goldstein et al. 2010; Moeller et al. 2013), but additional candidates include nicotinic and alpha-2 adrenergic receptor agents. These have recently been shown to have highly specific effects on the maintenance of information in working or scratchpad memory in prefrontal cortex (Castner et al. 2011; Wang et al. 2011). Similarly, serotonergic tone has been shown to be necessary for reversal learning and devaluation (Clarke et al. 2004; Nonkes et al. 2010; Schilman et al. 2010). Our results suggest that approaches that combine these agents with typical psychotherapeutic interventions that appropriately target individuals who show abnormal processing in orbital or other prefrontal areas might prove especially beneficial.

Acknowledgments

This work was supported by the Intramural Research Program (G.S.) and grants from the National Institute on Drug Abuse (H.M.W. F30-DA033100, J.L.J. F32-DA031517). The opinions expressed in this article are the authors' own and do not reflect the view of the NIH/DHHS.

References

- APA. Diagnostic and statistical manual of mental disorders. 4. American Psychiatric Association; Washington DC: 2000. (text revision)
- Belin D, Everitt BJ. Cocaine seeking habits depend upon dopamine-dependent serial connectivity linking the ventral with the dorsal striatum. *Neuron*. 2008; 57:432–441. [PubMed: 18255035]
- Castner SA, Smagin GN, Piser TM, Wang Y, Smith JS, Christian EP, Mrzljak L, Williams GV. Immediate and sustained improvements in working memory after selective stimulation of alpha7 nicotinic acetylcholine receptors. *Biol Psychiatry*. 2011; 69:12–18. [PubMed: 20965497]
- Clarke HF, Dalley JW, Crofts HS, Robbins TW, Roberts AC. Cognitive inflexibility after prefrontal serotonin depletion. *Science*. 2004; 304:878–880. [PubMed: 15131308]
- Corbit LH, Balleine BW. The role of prelimbic cortex in instrumental conditioning. *Behav Brain Res*. 2003; 146:145–157. [PubMed: 14643467]
- Daw ND, Niv Y, Dayan P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci*. 2005; 8:1704–1711. [PubMed: 16286932]
- Deroche-Gamonet V, Belin D, Piazza PV. Evidence for addiction-like behavior in the rat. *Science*. 2004; 305:951–953. [PubMed: 15310881]
- Dutra L, Stathopoulou G, Basden SL, Leyro TM, Powers MB, Otto MW. A meta-analytic review of psychosocial interventions for substance use disorders. *Am J Psychiatry*. 2008; 165:179–187. [PubMed: 18198270]
- Gawin FH. Cocaine addiction: psychology and neurophysiology. *Science*. 1991; 251:1580–1586. [PubMed: 2011738]
- Goldstein RZ, Volkow ND. Dysfunction of the prefrontal cortex in addiction: neuroimaging findings and clinical implications. *Nat Rev Neurosci*. 2011; 12:652–669. [PubMed: 22011681]
- Goldstein RZ, Alia-Klein N, Tomasi D, Zhang L, Cottone LA, Maloney T, Telang F, Caparelli EC, Chang L, Ernst T, Samaras D, Squires NK, Volkow ND. Is decreased prefrontal cortical sensitivity to monetary reward associated with impaired motivation and self-control in cocaine addiction? *Am J Psychiatry*. 2007a; 164:43–51. [PubMed: 17202543]
- Goldstein RZ, Tomasi D, Alia-Klein N, Cottone LA, Zhang L, Telang F, Volkow ND. Subjective sensitivity to monetary gradients is associated with frontolimbic activation to reward in cocaine abusers. *Drug Alcohol Depend*. 2007b; 87:233–240. [PubMed: 16997508]
- Goldstein RZ, Woicik PA, Maloney T, Tomasi D, Alia-Klein N, Shan J, Honorio J, Samaras D, Wang R, Telang F, Wang GJ, Volkow ND. Oral methylphenidate normalizes cingulate activity in cocaine addiction during a salient cognitive task. *Proc Natl Acad Sci U S A*. 2010; 107:16667–16672. [PubMed: 20823246]
- Homayoun H, Moghaddam B. Progression of cellular adaptations in medial prefrontal and orbitofrontal cortex in response to repeated amphetamine. *J Neurosci*. 2006; 26:8025–8039. [PubMed: 16885216]
- Izquierdo AD, Suda RK, Murray EA. Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. *J Neurosci*. 2004; 24:7540–7548. [PubMed: 15329401]
- Jentsch JD, Taylor JR. Impulsivity resulting from frontostriatal dysfunction in drug abuse: implications for the control of behavior by reward-related stimuli. *Psychopharmacology*. 1999; 146:373–390. [PubMed: 10550488]
- Jones JL, Esber GR, McDannald MA, Gruber AJ, Hernandez G, Mirenzi A, Schoenbaum G. Orbitofrontal cortex supports behavior and learning using inferred but not cached values. *Science*. 2012; 338:953–956. [PubMed: 23162000]

- Killcross S, Coutureau E. Coordination of actions and habits in the medial prefrontal cortex of rats. *Cerebral Cortex*. 2003; 13:400–408. [PubMed: 12631569]
- Konova AB, Moeller SJ, Tomasi D, Parvaz MA, Alia-Klein N, Volkow ND, Goldstein RZ. Structural and behavioral correlates of abnormal encoding of money value in the sensorimotor striatum in cocaine addiction. *Eur J Neurosci*. 2012; 36:2979–2988. [PubMed: 22775285]
- Koob GF, Le Moal M. Review. Neurobiological mechanisms for opponent motivational processes in addiction. *Phil Trans R Soc Lond B*. 2008; 363:3113–3123. [PubMed: 18653439]
- Lucantonio F, Stalnaker TA, Shaham Y, Niv Y, Schoenbaum G. The impact of orbitofrontal dysfunction on cocaine addiction. *Nat Neurosci*. 2012; 15:358–366. [PubMed: 22267164]
- Machado CJ, Bachevalier J. The effects of selective amygdala, orbital frontal cortex or hippocampal formation lesions on reward assessment in nonhuman primates. *Eur J Neurosci*. 2007; 25:2885–2904. [PubMed: 17561849]
- Mendelson JH, Mello NK. Management of cocaine abuse and dependence. *N Engl J Med*. 1996; 334:965–972. [PubMed: 8596599]
- Mendez IA, Simon NW, Hart N, Mitchell MR, Nation JR, Wellman PJ, Setlow B. Self-administered cocaine causes long-lasting increases in impulsive choice in a delay discounting task. *Behav Neurosci*. 2010; 124:470–477. [PubMed: 20695646]
- Milad MR, Quirk GJ. Neurons in medial prefrontal cortex signal memory for fear extinction. *Nature*. 2002; 420:70–74. [PubMed: 12422216]
- Moeller SJ, Hajcak G, Parvaz MA, Dunning JP, Volkow ND, Goldstein RZ. Psychophysiological prediction of choice: relevance to insight and drug addiction. *Brain*. 2012; 135:3481–3494. [PubMed: 23148349]
- Moeller SJ, Honorio J, Tomasi D, Parvaz MA, Woicik PA, Volkow ND, Goldstein RZ. Methylphenidate enhances executive function and optimizes prefrontal function in both health and cocaine addiction. *Cereb Cortex*. 2013; 1093/cercor/bhs345
- Nelson A, Killcross S. Amphetamine exposure enhances habit formation. *J Neurosci*. 2006; 26:3805–3812. [PubMed: 16597734]
- Niv Y, Joel D, Dayan P. A normative perspective on motivation. *Trends Cogn Sci*. 2006; 10:375–381. [PubMed: 16843041]
- Nonkes LJ, Tomson K, Maertins A, Dederen J, Maes JH, Homberg J. Orbitofrontal cortex and amygdalar over-activity is associated with an inability to use the value of expected outcomes to guide behavior in serotonin transporter knockout rats. *Neurobiol Learn Mem*. 2010; 94:65–72. [PubMed: 20388545]
- Parvaz MA, Maloney T, Moeller SJ, Woicik PA, Alia-Klein N, Telang F, Wang G-J, Squires NK, Volkow ND, Goldstein RZ. Sensitivity to monetary reward is most severely compromised in recently abstaining cocaine addicted individuals: a cross-sectional ERP study. *Psychiatry Res Neuroimaging*. 2012; 203:75–82.
- Pickens CL, Setlow B, Saddoris MP, Gallagher M, Holland PC, Schoenbaum G. Different roles for orbitofrontal cortex and basolateral amygdala in a reinforcer devaluation task. *J Neurosci*. 2003; 23:11078–11084. [PubMed: 14657165]
- Quirk GJ, Russo GK, Barron JL, Lebron K. The role of ventro-medial prefrontal cortex in the recovery of extinguished fear. *J Neurosci*. 2000; 20:6225–6231. [PubMed: 10934272]
- Rhodes SEV, Killcross SA. Lesions of rat infralimbic cortex enhance recovery and reinstatement of an appetitive Pavlovian response. *Learn Mem*. 2004; 11:611–616. [PubMed: 15466316]
- Rhodes SEV, Killcross SA. Lesions of rat infralimbic cortex enhance renewal of extinguished Pavlovian responding. *Eur J Neurosci*. 2007; 25:2498–2503. [PubMed: 17445245]
- Robbins TW, Everitt BJ. Drug addiction: bad habits add up. *Nature*. 1999; 398:567–570. [PubMed: 10217139]
- Robinson TE, Berridge KC. The psychology and neurobiology of addiction: an incentive-sensitization view. *Addiction*. 2000; 95:S91–S117. [PubMed: 11002906]
- Robinson TE, Berridge KC. Review. The incentive sensitization theory of addiction: some current issues. *Phil Trans R Soc Lond B*. 2008; 363:3137–3146. [PubMed: 18640920]

- Roesch MR, Takahashi Y, Gugs N, Bissonette GB, Schoenbaum G. Previous cocaine exposure makes rats hypersensitive to both delay and reward magnitude. *J Neurosci.* 2007; 27:245–250. [PubMed: 17202492]
- Schilman EA, Klavir O, Winter C, Sohr R, Joel D. The role of the striatum in compulsive behavior in intact and orbitofrontal-cortex-lesioned rats: possible involvement of the serotonergic system. *Neuropsychopharmacology.* 2010; 35:1026–1039. [PubMed: 20072118]
- Schoenbaum G, Setlow B. Cocaine makes actions insensitive to outcomes but not extinction: implications for altered orbitofrontal-amygdalar function. *Cerebral Cortex.* 2005; 15:1162–1169. [PubMed: 15563719]
- Simon NW, Mendez IA, Setlow B. Cocaine exposure causes long-term increases in impulsive choice. *Behav Neurosci.* 2007; 121:543–549. [PubMed: 17592945]
- Stalnaker TA, Roesch MR, Franz TM, Burke KA, Schoenbaum G. Abnormal associative encoding in orbitofrontal neurons in cocaine-experienced rats during decision-making. *Eur J Neurosci.* 2006; 24:2643–2653. [PubMed: 17100852]
- Takahashi Y, Roesch MR, Stalnaker TA, Haney RZ, Calu DJ, Taylor AR, Burke KA, Schoenbaum G. The orbitofrontal cortex and ventral tegmental area are necessary for learning from unexpected outcomes. *Neuron.* 2009; 62:269–280. [PubMed: 19409271]
- Takahashi Y, Roesch MR, Stalnaker TA, Schoenbaum G. Cocaine exposure shifts the balance of associative encoding from ventral to dorsolateral striatum. *Front Integr Neurosci.* 2007; 1:11.10.3389/neuro.07/011.2007 [PubMed: 18958239]
- Taylor JR, Horger BA. Enhanced responding for conditioned reward produced by intra-accumbens amphetamine is potentiated after cocaine sensitization. *Psychopharmacology.* 1999; 142:31–40. [PubMed: 10102780]
- Taylor JR, Jentsch JD. Repeated intermittent administration of psychomotor stimulant drugs alters the acquisition of Pavlovian approach behavior in rats: differential effects of cocaine, d-amphetamine and 3,4-methylenedioxymethamphetamine (“ecstasy”). *Biol Psychiatry.* 2001; 50:137–143. [PubMed: 11526995]
- Vanderschuren LJMJ, Everitt BJ. Drug seeking becomes compulsive after prolonged cocaine self-administration. *Science.* 2004; 305:1017–1019. [PubMed: 15310907]
- Wang M, Gamo NJ, Yang Y, Jin LE, Wang XJ, Laubach M, Mazer JA, Lee D, Arnsten AF. Neuronal basis of age-related working memory decline. *Nature.* 2011; 476:210–213. [PubMed: 21796118]
- West EA, DesJardin JT, Gale K, Malkova L. Transient inactivation of orbitofrontal cortex blocks reinforcer devaluation in macaques. *J Neurosci.* 2011; 31:15128–15135. [PubMed: 22016546]
- Wimmer GE, Shohamy D. The striatum and beyond: hippocampal contributions to decision making. In: Delgado, M.; Phelps, EA.; Robbins, TW., editors. *Attention and performance XXII.* Oxford University Press; Oxford: 2011.
- Wimmer GE, Shohamy D. Preference by association: how memory mechanisms in the hippocampus bias decisions. *Science.* 2012; 338:270–273. [PubMed: 23066083]
- Winstanley CA, Olausson P, Taylor JR, Jentsch JD. Insight into the relationship between impulsivity and substance abuse from studies using animal models. *Alcohol Clin Exp Res.* 2010; 34:1306–1318. [PubMed: 20491734]
- Wyvell CL, Berridge KC. Incentive sensitization by previous amphetamine exposure: increased cue-triggered “wanting” for sucrose reward. *J Neurosci.* 2001; 21:7831–7840. [PubMed: 11567074]
- Zapata A, Minney VL, Shippenberg TS. Shift from goal-directed to habitual cocaine seeking after prolonged experience in rats. *J Neurosci.* 2010; 30:15457–15463. [PubMed: 21084602]

Fig. 1.

Rats were trained to press the active lever to self-administer a 10 % liquid sucrose solution or cocaine-HCl (0.75 mg/kg/infusion). Figures show the mean number of lever presses on the active and passive levers for sucrose controls (**a**) and cocaine rats (**b**) over 14 days. The mean number of sucrose or cocaine reward deliveries during each session is also shown. Error bars = SEM

Fig. 2.

Cocaine self-administration disrupts the expression of behavior that depends upon inferred values. The experimental timeline is schematized at the *bottom of the figure*. Figures show the percentage of time spent in food cup during presentation of the cues during each of the three phases of training: preconditioning (**a, b**), conditioning (**c, d**), and the probe test (**e, f**). Control rats are presented in the *top panels* (**a, c, e**) and the cocaine-experienced rats are shown in the *bottom panels* (**b, d, f**). * $p < 0.05$, ** $p < 0.01$ or better. Error bars = SEM

Fig. 3. Cocaine self-administration disrupts learning that depends upon inferred values. The experimental timeline is schematized at the bottom of the figure. Figures show the percentage of time spent in food cup during presentation of the cues during blocking conditioning (**a**, **b**) and the subsequent probe test (**c**, **d**). Control rats are presented in the *top panels* (**a**, **c**) and the cocaine-experienced rats are shown in the *bottom panels* (**b**, **d**). ** $p < 0.01$ or better. Error bars = SEM