LOST IN STATES: EFFECTS OF COCAINE EXPERIENCE ON REPRESENTATIONS SUPPORTING DECISION MAKING

By

Lauren Mueller

A DISSERTATION

Presented to the Department of Behavioral Neuroscience at Oregon Health & Science University School of Medicine

In partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

Table of Contents

List of Figures

List of Abbreviations

AMY, amygdala

ANCOVA, analysis of covariance

ANOVA, analysis of variance

ASST, attentional set-shift task

BA, badass

BBB, blood brain barrier

BOLD, blood oxygenation level-dependent

CANTAB, Cambridge Neuropsychological Test Automated Battery

CIN, cholinergic interneuron

CPP, conditioned place preference

CS, conditioned stimulus

dmPFC, dorsomedial prefrontal cortex

DMS, dorsomedial striatum

ED, extra-dimensional

FDG, fluorodeoxyglucose

Fig, figure

fMRI, functional magnetic resonance imaging

FSI, fast-spiking interneuron

GABA, γ-aminobutyric acid

GLM, generalized linear model

HPC, hippocampus

ID, intra-dimensional

I.M., intramuscular

I.P., intraperitoneal

ITI, intertrial interval

I.V., intravenous

LDTg, laterodorsal tegmentum

LH, lateral habenula

LiCl, lithium chloride

MSN, medium spiny neuron

mPFC, medial prefrontal cortex

NAc, nucleus accumbens

NS, narrow spike

OFC, orbitofrontal cortex

PET, positron-emitting tracer

PETH, peri-event time histogram

PL, prelimbic cortex

RSA, representational similarity analysis

SEM, standard error of the mean

SUD, substance use disorder

UID, unidentified

UPGMA, unweighted pair group method with arithmetic mean

US, unconditioned stimulus

VP, ventral pallidum

VTA, ventral tegmental area

WS, wide spike

Acknowledgements

Foremost, I would like to thank my two mentors, Geoffrey Schoenbaum and Bita Moghaddam. From the moment I stepped into Geoff's lab as a summer student I felt like I had arrived. His ability to attract and select kind, thoughtful, and diverse individuals who are equally clever as they are motivated speaks volumes about the good human and scientist that he is. He truly has established a fantastically unique and nurturing scientific environment where anything is possible – within the confines of governmental red tape. My atypical graduate school journey has certainly been trying at times but Geoff has supported me through it all. It is because of Geoff that I felt strong enough to use my voice, and through his mentorship and interactions with Schoenbaum Lab members that I have been able to thrive both as a person and a scientist. I would like to thank Bita Moghaddam for being a badass (BA) scientist and mentor. 2020 was a hard year for countless reasons and she made it easier; I have so much gratitude for her and the other BA people at OHSU who made my program transition possible. It was through her weekly virtual lab meetings with other Moghaddam Lab members that I felt socially connected during the pandemic, helping me stay grounded through hard times. She has truly been an excellent person to talk to both about life and science. Both Geoff and Bita are extraordinary people and mentors.

I would like to thank Andrew Wikenheiser for taking me as a summer student, not realizing it would be a decade-long commitment. It is because of his style of teaching and patience that I have become an electrophysiologist and a semi-capable programmer. I look forward to his continued mentorship during the next phase of my scientific career.

I would like to thank Melissa Sharpe for epitomizing what it means to be great female scientist. Her sense of direction, confidence and level of productivity serve as an example for what I hope to achieve in the coming phases of my career.

I am grateful for my DAC committee: Matt Lattal, Marina Wolf, and John Williams for being a BA group of scientists, helping me transition schools smoothly and provide a critical eye for my work. I have greatly appreciated their time and input throughout this process.

Finally, I would like to thank my family and friends for their love and support. Craig Werner and Jennifer Martin Schieber for their support and friendship throughout my journey at Buffalo – they have taught me what it truly means to be supportive lab mates. My parents, who listened when I asked for help, footed my medical expenses for various rehab facilities, remained patient when I relapsed, and learned to trust me again when I got back on my feet. Without their support I would not have been able to start my journey into recovery over 13 years ago and realize my passion for research. Lastly, I would like to thank my husband, Justin Siemian, who has been a great teammate. His voice of reason, mellow demeanor, and ability to join me in running behavioral paradigms on weekends have been critical to my success. His great human nature and intelligence have helped me grow as a person and a scientist.

The work included in this dissertation was supported by the Intramural Research Program at the National Institute on Drug Abuse (ZIA-DA000587).

Chapter 1: General Introduction

Overview

Understanding why and how decisions are made has been a central focus of neuroscience for several decades. The decision-making process helps humans and animals to decide the optimal course of action amongst other potential options, requiring the constant updating and integration of previous and current information to forecast future outcomes. To assist with decision making, organisms form states, each of which contains the associative information necessary for guiding behaviors under a certain circumstance. Impairments to state formation and selection can cause suboptimal decision making, leading to ineffective actions and poor outcomes.

Many psychiatric disorders are associated with decision making deficits. For substance use disorders (SUDs), which involve the seeking and taking of a substance despite known negative consequences, a central component of the disease etiology seems to be altered decision making. While SUD criteria are complex, central components of the disorder include drug use at the expense of other life activities and losses of control over the amount of and time over which the drug is used. These criteria suggest that decision making deficits may prevent negative health, social, and financial consequences from deterring future drug use. This is also highlighted in findings that drug users exhibit deficits across many decision-making metrics, even those not explicitly pertaining to drug. Thus, although traditional drug abuse research has focused on the role of the brain's reward system as the central treatment target for SUDs, it may be that targeting non-mesolimbic regions to strengthen decision-making functions offers an advantageous strategy for managing SUDs or other psychiatric disorders.

Among the brain's many important regions for decision making, the orbitofrontal cortex (OFC), prelimbic cortex (PL), and dorsomedial striatum (DMS) are a few that are particularly important. As described in detail below, the OFC and PL appear to govern somewhat distinct aspects of decision-making strategies, yet both feed information to the DMS, which functions as a

mediator between different courses of action. While many studies have investigated the effects of impaired regional activity on behavior, or how drugs of abuse affect each region, there is scant literature describing how previous drug use affects the function of these brain areas in a subsequent decision-making task.

Here, we build on several decades of human and animal research to investigate how previous cocaine self-administration manifests in persistent changes in decision-making behavior as well as the associated neural activity in the OFC, PL, and DMS. We propose that extended drug experience induces long-term neural changes that prevent one or more of these areas from tracking state and determining appropriate courses of action to achieve optimal outcomes. To test this prediction, we recorded single unit activity in the DMS, or in the OFC and PL simultaneously, during a complex decision-making task in rats that had previously self-administered cocaine for 14 days. Our investigations link drug-induced behavioral and neurophysiological changes, allowing us to examine how deficits in neural state representations within each brain area were accompanied by real-time behavioral deficits. Importantly, these changes were often not evident in gross neurophysiological readouts such as firing rate, and instead required a very detailed analysis of how each unit responds during different portions of the behavioral task. Thus, our approach goes beyond a simple description of coarse changes in both behavior and brain function, and instead attempts to link the two as inseparable components of the same overall process.

Decision making

Decision making is the selection of a particular course of action among alternative possibilities. This cognitive process could be simplified into three predominant facets (Izquierdo and Belcher, 2012). The first is attention, or the capacity to attend to pertinent environmental information, known as cues, and disregard those that are irrelevant. The second is valuation, or the

appraisal of the cost or effort required to make a choice. The third is experience, or knowledge about past actions and the outcomes to which they were associated which can be utilized to anticipate the outcomes of prospective choices. Since the world is constantly changing, good decision making requires that the attention, valuation, and experience that go into the cognitive process get frequently updated and integrated so that animals can adapt behavior in a changing environment.

When the information that governs a decision changes abruptly, subjects must adjust their responding to achieve the best outcome possible, a process that requires both the exploitation of previously successful decisions and the exploration of alternate solutions. To do this, the attention, valuation, and experience associated with each context can be assigned to compartmentalized "states". The information contained in each state defines the rules appropriate for guiding decision making under those conditions; however, the usefulness of this ability is predicated on appropriate control of state formation and its use. If such control is weak, breakthrough of incompletely-segregated contradictory rules might lead to situationally-inappropriate behavior, whereas if control is too strong, over-protected rules might they fail to change in the face of new information indicating they should be permanently altered, leading to inflexible responding.

Animal models of decision making

The study of laboratory animals on decision-making tasks has been influential in discovering the neural underpinnings of adaptive decision making, particularly as they pertain to attention, valuation, and experience. Here, we broadly describe a few behavioral assays that are particularly relevant to these attributes of the decision-making process, which will provide a background for the experiments performed for this thesis.

Reversal learning

Reversal learning tasks are one of the most common assays used to examine behavioral flexibility (Izquierdo and Jentsch, 2012). While there are several different varieties of the task (e.g., odor discrimination reversals, visual discrimination reversals, spatial reversals, response reversals, etc.) they all function to measure the ability of subjects to adjust responding following changes in reward contingency. For the example of odor discrimination, the task generally consists of recurrent pairings of an action (e.g., responding to a particular fluid delivery well) with an outcome (e.g., delivery of reward), allowing subjects to learn about reward contingencies through the sensory features of stimuli that predict reward availability and the necessary actions that must be taken to earn reward. To accomplish this, subjects must first learn that discriminative odor cues carry information regarding whether a specific response procures a reward (e.g., presentation of odor cue 1 means respond to left, but not right, fluid delivery well to receive reward). Over time, subjects become well-trained on the task and demonstrate discriminative behavior, suggesting that they have learned the necessary associative rule. Once this happens and subjects reach a given discriminative accuracy threshold, antecedents are reversed. When such associations are switched, although the initially trained action no longer procures the reward, subjects will continue to use the expired contingencies until they learn the new associative rules. Unlike the first acquisition phase of the task when rats learn the first action-outcome association, the reversal requires subjects to suppress the initially acquired response and instead express responses that were previously not reinforced. Because of this, reversal learning assays are often used to quantify flexible or adaptive decision making.

Proper OFC function is important for reversal learning. In one study, rats with OFC lesions were conditioned on a series of go, no-go odor discrimination problems, during which one odor cue (positive) indicated the accessibility of a water reinforcer following responding at the fluid

delivery, and the second odor cue (negative) indicated no reinforcer availability. The task required that rats learn to suppress responding at the fluid delivery well following presentation of the negative odor cue to forgo the longer duration intertrial intervals (ITI) associated with these inappropriate responses. Once rats acquired initial odor discriminations, they underwent sequential switches of the last problem to examine how well they could withhold responding following changes in the associative contingencies between the odor cues and the outcomes. Experimenters observed that rats with OFC lesions properly responded for positive odors and withheld responding for negative odors, suggesting that they were able to learn new discrimination problems and that their general ability to inhibit responding was not disrupted. However, further examination of behavior during the serial reversals showed that rats with OFC lesions demonstrated impaired acquisition of switching on previously learned discriminations, requiring a greater amount of trials than controls to relearn discriminations after the cue-outcome associations were switched. These results show that OFC function is important for adapting behavior during reversal learning (Schoenbaum et al., 2002).

Pavlovian reinforcer devaluation

Another procedure used to examine the neural machinery involved in adaptive decision making is Pavlovian reinforcer devaluation. While devaluation can be accomplished using different methods (e.g., malaise, satiety-induced, etc.) their purpose is largely the same, in that they demonstrate that subjects learn to predict the current value of the expected outcome and that changing such representations alters the value of the outcome. A classic example of this assay (Rescorla, 1990) involves first training subjects over several sessions that one response (i.e., lever press) leads to a specific outcome (i.e., water containing quinine), and a second response (i.e., chain pull) leads to a separate outcome (i.e., water containing weak hydrochloric acid). Following

training, the devaluation procedure is implemented in which only one of the reinforcers (water containing quinine or water containing weak hydrochloric acid) is paired with lithium chloride (LiCl)-induced malaise. After several exposures to the outcome-LiCl pairing, subjects are tested for their preference between the two responses (level press vs. chain pull). Selective depression of the response paired with the devalued outcome suggests that behavior is guided using knowledge of associative contingencies, the ability of which is critical for adaptive decision making.

Proper OFC function has also been shown to be important for Pavlovian reinforcer devaluation. One study in rats examined the impact of OFC lesions on actions that are influenced by associations between a conditioned stimulus (CS) and unconditioned stimulus (US) and found that improper OFC function impairs the ability of a stimulus to access associative information about the value of the associated reinforcement. Briefly, during the first phase, rats were conditioned to associate a CS (e.g., cue light) with delivery of a US (e.g., food pellet reward) in a cup. As rats learned the association, they began approaching the cup during the presentation of the CS expecting the delivery of the US. During the second phase, rats were given access to and allowed to consume the US in their home cage, after which experimenters administered LiCl to induce malaise, thereby reducing the incentive value of the US. Finally, rats were placed back into the initial experimental chamber and approaches to the food cup in response to the CS without the US was probed. Researchers found that while lesions of the OFC did not alter the acquisition of conditioned responding to the CS during the first phase or LiCl- induced taste version, unlike their control counterparts they were not able to properly devalue the outcome, showing no change in responding to the CS. Impaired devaluation shows a deficit in the ability of OFC-lesioned subjects to predict the current value of the anticipated outcome (Gallagher et al., 1999).

Contingency degradation

Contingency degradation is another paradigm used to examine associative processes that contribute to decision making (Yin et al., 2005), specifically functioning to measure subjects' sensitivity to alterations in the contingency between performing the action and procuring the outcome. Operationalized, subjects are usually first trained over several sessions that one response (i.e., left lever press) leads to one reward (i.e., food pellet) and a second response (i.e., right lever press) leads to a separate distinct reward (i.e., sucrose solution) until an experimenter-set threshold of responses for reward is achieved. Once criteria are met, degradation training begins during which one outcome (i.e., food pellet) is delivered non-contingently such that the probability of outcome delivery is equally as likely if subjects respond appropriately (i.e., left lever press) or not, while delivery of the other outcome (i.e., sucrose solution) is still contingent upon a response (i.e., right lever press). After degradation, subjects undergo a choice extinction test during which responding on both levers without reward delivery (i.e., left vs. right lever) is quantified. If action-outcome associations are properly learned and intact, then the expectation is that contingency degradation will result in suppression of responding for the degraded response-reward (left lever press-food pellet) association, making it a useful tool for the assessment of adaptive behavior.

The DMS is one region important for performance of contingency degradation. One study examining the role of the DMS in the representation of associations in instrumental conditioning found that lesion or inactivation of the region abolished contingency degradation. Specifically, rats were trained to respond on two levers, each associated with a different outcome, and then tested using different procedures to measure action-outcome learning. To examine the role of the DMS in the acquisition and expression of action-outcome associations, experimenters lesioned the DMS prior to lever training and found that lesions of the posterior DMS (pDMS) at this time point impaired sensitivity of instrumental responding to both the devalued outcome and the degraded

contingency during extinction. Similar findings were observed when the pDMS was lesioned following lever training. Further, transient inactivation of the pDMS with an infusion of muscimol showed selective impairment of sensitivity to devaluation and contingency degradation, without disrupting the ability of subjects to discern instrumental responses or the identifying features of the procured outcomes (Yin et al., 2005).

Attentional set-shift task

Attentional set-shift tasks (ASST) are commonly employed to examine the ability of animals to pay attention to relevant features of the environment. While variations of such set-shifting assays exist (e.g., Wisconsin Card Sort Task, intra-dimensional (ID) /extra-dimensional (ED) component of the Cambridge Neuropsychological Test Automated Battery (CANTAB)) they are all predicated on the idea that in order to make adaptive decisions, subjects must alter their behavioral policies in response to altered environmental contingencies. First, this task requires that attentional sets must be formed; sets are created when subjects learn that a set of rules can be applied to complex stimuli to discriminate relevant from extraneous cues. For example, by pairing certain texture cues with food reward in a task that presents texture, digging medium, and scent cues, subjects become conditioned to attend and respond to a reward-predictive cue of the relevant set (e.g., textures) and disregard the cues of the irrelevant set (i.e., digging media and scents). The contingency is then reinforced in subsequent blocks where the rewarded member of the relevant set is changed, but the relevant nature of the texture and irrelevant nature of the digging media and scents are held constant. Such reinforced rules create a cognitive set. Three phases of ASST can be utilized to examine adaptive behavior: (a) reversals, which according to the above example are defined as a switch between which relevant cue of a particular dimension (texture) is rewarded, (b) intra-dimensional shifts, which refers to a novel discrimination rule being based on the same perceptual

dimension as the one to which subjects had been attending (i.e., a shift from texture to digging media since both are based on tactile information), and (c) extra-dimensional shifts, in which the discrimination rule shifts to the other perceptual dimension (i.e., scent in this example, which is based on olfactory instead of tactile information). Reversals and intra-dimensional shifts test adaptive decision making, forcing subjects to maintain the attentional set or perceptual dimension while changing the stimulus and reward paired rule. On the other hand, extra-dimensional shifts challenge behavioral flexibility since the subject is forced to switch its focus to the previously irrelevant set of stimuli. Continued choices according to the previously relevant set but now irrelevant set indicates perseveration.

PL function has been shown to be important for the performance of ED shifts during the ASST. Briefly, one study conditioned rats on a variation of an ASST in which subjects had to dig in bowls for food reinforcers. During the task, bowls were accessible in pairs, only one of which contained the reinforcer. Subjects chose the dish to dig based on its defining sensory features including the texture that covered its surface, the medium contained within the dish, or its scent. Within one session, subjects responded during a series of discriminations, including reversals, an ID shift, and an ED shift. While mPFC lesions did not disrupt the initial acquisition or reversal learning, significant behavioral deficits were observed following the ED shift, such that mPFC-lesioned rats required twice as many trials to acquire new discriminations as compared to controls. These results show that lesions to the mPFC selectively impair the ability to shift attentional sets (Birrell and Brown, 2000), an effect similar to what had previously been observed in marmosets (Dias et al., 1996).

Brain areas involved in decision making

While many brain regions have been implicated in decision making, here we discuss a few which have been shown to be involved in the aforementioned decision-making tasks (Izquierdo and Belcher, 2012) and which will be the focus of the experiments of this thesis:

Decision making and the OFC

The OFC has been shown to be critical for decision making and behavioral flexibility. OFC lesions result in the inability to quickly adjust behavioral responding following changes in response-reward contingencies (Teitelbaum, 1964; Schoenbaum et al., 2003; Izquierdo et al., 2004). Further, reversal learning deficits are linked to neural correlates in the OFC (Thorpe et al., 1983; Schoenbaum et al., 1999; Wallis and Miller, 2003). Importantly, OFC lesion-induced reversal deficits are not simply due to impaired response inhibition (Chudasama et al., 2007), suggesting a critical role for OFC in the guidance of behavior when outcomes change. Consistent with this, reinforcer devaluation studies, in which learned responses are monitored following devaluation of the anticipated outcome, show that OFC-lesioned subjects exhibit behavioral inflexibility. While sham animals are able to integrate preexisting knowledge of value with new outcome value and appropriately respond less for the devalued outcome, OFC-lesioned animals are unable to integrate and update value to alter their learned responding (Gallagher et al., 1999; Izquierdo et al., 2004). Additionally, changes in OFC single-units in response to predictive cues and food reward reflect modifications in the current value of the expected outcomes (Tremblay and Schultz, 1999; Hikosaka and Watanabe, 2000).

Although often taken as evidence of an exclusive role in value-guided behavior, devaluation and similar deficits after OFC lesions have more recently been linked to a larger function, namely tracking the current location within the cognitive map of the current task space (Wilson et al.,

2014; Sharpe et al., 2019). In rats, following the acquisition of spatial-specific object-reward associations, OFC neural ensembles encode value-based schemas that represented objects in the context and location where they are associated with reward or non-reward. Similarly, OFC neural ensembles signal distinct spatial contexts that delineate the mapping of specific stimuli to actions and outcomes (Farovik et al., 2015). A human neuroimaging study utilizing pattern classification on data collected during a 16-state decision-making task found that OFC activity decoded unobservable task states and that the accuracy of decoding was related to decision-making performance. Additionally, behavioral accuracy following state transitions correlated with the similarity between the neural representations of those states (Schuck et al., 2016). Data such as these support the idea that the OFC provides an essential representation of the current state of the environment which guides outcome-oriented actions.

Decision making and the PL

The PL is also important for adaptive decision making but in a manner that is distinct from OFC. PL is required during changes in the attentional demands of a task, for rats to express an instrumental response reflective of the current optimal goal value, and for using environmental stimuli to resolve response conflict (Balleine and Dickinson, 1998; Birrell and Brown, 2000; Marquis et al., 2007). Additionally, PL is required for switching between response and strategy sets (Ragozzino et al., 2003; Stefani et al., 2003; Stefani and Moghaddam, 2005; Floresco et al., 2008). Consistent with this, activity in PL neurons reflect changes in task strategy. For instance, PL neurons recorded from rats alternating between place and response strategies in a plus maze task increased activity when a new strategy was adopted, but remained stable when the same strategy was utilized by different behaviors (Rich and Shapiro, 2009). Taken together, these

findings suggest that PL may recognize changes in the environment to facilitate the transition from one state to another (Sharpe et al., 2019).

Decision making and the DMS

Receiving inputs from both the OFC and the PL, the DMS is one region known to encode the action-outcome associations that are required for decision making, functioning as a mediator of choice between specific courses of action (Yin et al., 2005; Balleine et al., 2009; Balleine and O'Doherty, 2010). DMS cholinergic interneurons (CINs), comprising only ~1% of the total neuron population within this region (Bradfield et al., 2013), have been implicated in the specific organization of associative information and have been shown to track information about outcomes that could be used to guide appropriate decision making. In particular, one study showed that when rats had lesions or pharmacological manipulations that reduced DMS CIN activity, they could initially acquire goal-directed outcome associations, but were unable to adapt their behavior to subsequent changes to those associations. These changes included contingency degradation and contingency reversals (Bradfield et al., 2013). Although the exact mechanism underlying this effect is as yet unknown, these findings suggest that DMS CINs function to allow striatum to maintain information from different states simultaneously, and thereby prevent interference between learning that has occurred in different states, a fundamental basis of adaptive decision making. Furthermore, individual DMS CINS have been shown to encode real-time state representations during a decision-making task, the ability of which is dependent on OFC input (Stalnaker et al., 2016). Taken together, prior work suggests that both the OFC and PL could provide unique information to the DMS relevant for deploying situationally-appropriate associative information required for optimal decision making.

SUD

In its most severe form, SUD, is a chronic, relapsing disorder characterized by compulsive drug seeking and taking despite adverse consequences (NIDA, 2020a). In the United States alone, approximately 21.5 million people age 12 and older suffer from SUDs, of which 913,000 involve stimulants such as cocaine (SAMHSA, 2015, 2016). Substance abuse – including tobacco, alcohol, and illicit drugs – results in increased healthcare costs, crime rates and lost work productivity, costing the United States more than $740 billion annually (NIDA, 2020b). Despite the adverse consequences it has on our society, mechanisms leading to and perpetuating SUDs, remain a key and unresolved question, making it difficult to develop proper therapeutics to treat such disorders.

Neural circuitry of SUDs

Clinical and preclinical studies demonstrate that subjects who suffer from SUDs experience major disruptions within neurological reward circuitry, which represents a functional disturbance of their attention towards and intake of rewarding drugs (Nestler and Carlezon, 2006). As one of the most recognized reward pathways, the mesolimbic dopamine system is greatly affected by use of addictive drugs. This reward circuit features two central areas: the ventral tegmental area (VTA), which plays a role in the recognition of the rewarding effects of drugs, and the nucleus accumbens (NAc), which acts as mediator of the rewarding effects of drugs. Although the mesolimbic dopamine system is considered by many to be the most critical reward pathway, there are several circuits running in parallel that also play important roles in the development of SUDs. These parallel circuits and the psychobiological integration of their signals are each implicated in the development and progression of the disease, making it difficult to parse the independent contribution of each circuit.

As one of the main players, the VTA is primarily composed of dopaminergic neurons and γ-aminobutyric acidergic (GABAergic) interneurons (Russo and Nestler, 2013). To name a few,

the VTA receives glutamatergic inputs from the lateral habenula (LH) and laterodorsal tegmentum (LDTg) and sends dopaminergic outputs to the hippocampus (HPC), OFC, medial prefrontal cortex (mPFC), and the amygdala (AMY; Wise, 2004; Russo and Nestler, 2013). Efferent VTA dopaminergic neurons have been repeatedly demonstrated to play a central role in reward. Dopaminergic cell bodies in the VTA send their projections via the medial forebrain bundle to the NAc (Pierce and Kumaresan, 2006). The release of dopamine from the VTA into the NAc allows these two areas to communicate, which facilitates the detection of rewarding stimuli controlled by GABAergic interneurons. These interneurons of the VTA release GABA which holds the dopaminergic neurons projecting to the NAc under tonic inhibition. In this way, some substances that interrupt GABA release from interneurons (Heilig et al., 2011) lead to the disinhibition of dopaminergic neurons, allowing for rewarding dopamine release in the NAc. However, different drug classes affect dopamine neurons through different mechanisms, as reviewed in depth previously (Lüscher and Ungless, 2006).

As another major contributor, the NAc is the site of integration of communication between limbic and cortical structures, establishing it as a major regulator of goal-directed behavior and critical for reward. The NAc is comprised of approximately 95% GABAergic medium spiny neurons, which express either dopamine D1 receptors, called D1-type MSNs, or dopamine D2 receptors, called D2-type MSNs. In addition to afferent VTA dopaminergic projections, the MSNs of the NAc receive glutamatergic input from several regions including, but not limited to, the HPC, mPFC, and the AMY, all of which are thought to play critical roles in the processing of learned information contributing to the progression of addiction. These excitatory glutamatergic inputs have been shown to mediate context conditioning, sensitization, reinforced learning, and relapse behaviors (Everitt et al., 1991; Everitt et al., 1999; Kalivas et al., 2003; Kalivas, 2004; Lüscher,

2013; Wolf, 2016; Bobadilla et al., 2017). Additionally, the NAc sends GABAergic outputs to the VTA either directly by D1-type MSNs, or by either D1- or D2-type MSNs via the ventral pallidum (VP) (Russo and Nestler, 2013; Pardo-Garcia et al., 2019), which are believed to act as a negative feedback mechanism to regulate VTA activity (Rahman and McBride, 2000).

Pharmacological mechanism of cocaine

Cocaine is a monoamine transporter inhibitor with four typical routes of administration including intranasal, intravenous, inhalation, and oral. Following administration, cocaine passes the blood brain barrier (BBB) by way of a proton-coupled organic cation antiporter, as well as passive diffusion. Once past the BBB, cocaine blocks the dopamine transporter, preventing the reuptake of dopamine from the synaptic cleft into the pre-synaptic axon terminal. Increased dopamine in the synaptic cleft leads to greater dopamine receptor activation in the post-synaptic neuron causing the rewarding euphoric effects of the drug. Additionally, cocaine blocks both norepinephrine and serotonin transporters preventing their reuptake; increased amounts of these neurotransmitters in the post-synaptic cleft result in increased stimulation of their respective receptors leading to the modulation of other psychological processes such as emotions and arousal.

Animal models of SUD

Studies in psychopharmacology examine the acute effects of drug administration as well as the extended consequences of drug exposure. In acute drug effect studies, researchers study the acute behavioral, physiological, or neurochemical changes that occur during the period when the drug is biologically active in the organism. While this is of clear importance to understanding SUD (we need to know the acute effects of a drug to understand why a human might repeatedly take it), these acute examinations are generally too narrow in scope to draw conclusions related to abuse potential. For example, smoking one cigarette and experiencing nicotine exposure does not make

an individual a smoker, which is a label typically reserved for more continual users over a period of weeks or months. As such, animal models that assess changes in either baseline animal behavior or responses to a drug or drug-related stimuli following repeated drug exposure are better suited to studying mechanisms of SUD. Importantly, while no single behavioral paradigm fully captures all the intricacies of SUD in the human condition, existing procedures can model particular aspects of it, uncovering crucial insights into the behavioral, anatomical, cellular, and genetic mechanisms of SUDs. Broadly, there are two categories of assays used to examine the effects of repeated drug exposure, (1) experimenter administration and (2) volitional (self-) administration.

Experimenter-administered drug

Experimenter-administered models of SUD have become popular due to their ease of implementation and tight control of drug exposure since the experimenter determines the drug dose and gives a forced injection to which the animal has no choice. Two procedures that have become particularly popular and offer distinct insight into understanding SUD are (1) conditioned place preference and (2) locomotor sensitization.

Conditioned place preference

Conditioned place preference (CPP) is a form of Pavlovian conditioning used to indirectly assess the rewarding properties of a drug by measuring the amount of a time a subject spends in a repeatedly drug-paired context. For CPP, two discriminable contexts typically defined by wall patterns, odors, and floor textures are created which serve as the conditioned stimuli. The two compartments are separable by a closable door or hallway which can allow free access to both chambers or restrict movement to one chamber at a time (Mueller et al., 2015). Although there are variations to the task, the first day of the paradigm usually entails a "pretest" or baseline measurement of side preference, during which the subject is placed into the apparatus and allowed

to freely explore it with the door open. In unbiased CPP designs, subjects will on average spend a roughly equal amount of time in both chambers. In a biased CPP design, subjects will on average spend more time in one chamber than the other. Subsequent conditioning sessions involve closing the door or hallway to divide the contexts and placing subjects in one chamber following experimenter-administered drug injection as the unconditioned stimulus (e.g., cocaine, morphine, nicotine) on one conditioning-session type, and in the opposite chamber following a saline/vehicle injection as the control on the opposing conditioning-session type. Over multiple conditioning sessions, generally alternating between drug-paired and saline-paired training sessions across days, subjects learn the association between the context (conditioned stimulus) and the drug (unconditioned stimulus), such that the context alone becomes preferred by the subject. Following conditioning, subjects undergo a "test" session in which the door is opened, allowing subjects to freely move between both chambers, and time in each chamber is again quantified. Although many resulting metrics can be drawn from CPP, pretest and test values allow for the calculation of a preference score (test versus pretest time spent in the drug-paired chamber), the expectation being that subjects demonstrate a greater preference for the drug-paired context following conditioning. Such a result gives an indication of how rewarding the drug stimulus was for the subject throughout conditioning.

Locomotor sensitization

Locomotor sensitization refers to an increased locomotor response to a drug following its repeated exposure via experimenter administration. The neuroadaptations underlying increased drug-induced locomotion are thought to run in parallel, or at least in part reflect, those occurring in reward pathways. That is, such drug-induced behavioral changes are indicative of hypersensitivity of the reward circuitry, thereby allowing experimenters to examine the rewarding

properties of drugs and how repeated exposure effects such pathways. For this procedure, subjects are typically habituated to locomotor chambers to reduce novelty-induced locomotion. Over the next several days, animals are given a forced drug injection before being placed in the chamber and allowed to move freely around it, while locomotor activity is recorded. Depending on the experiment details, a withdrawal period may be imposed, in which the animal is left drug-untreated for one or more days, prior to a final "challenge test", where a final drug injection is given and locomotor activity is recorded. Often, locomotion on the challenge test is higher than on the final day of continuous drug treatment, indicating some mechanism by which sensitivity is further enhanced by drug cessation.

Both CPP and locomotor sensitization are amenable to intervention studies, since potential therapeutic compounds can be given concurrently with the drug in the training or continuous exposure "development" phase, or instead during the test or challenge "expression" day. However, the downside of these procedures is they do not incorporate an aspect of will or action on the subjects' part. Just as we would not force-feed a person and suggest they had overeaten, it would be dubious to refer to subjects given experimenter-administered drug injections as being in a willingly drug-associated state. To capture the interaction between *choosing* to take drug and the exposure of the drug itself, models of volitional drug intake are needed.

Volitional drug intake

As the current "gold standard" of drug abuse models, self-administration requires the subject to take action in order to receive drug delivery. Commonly performed via intravenous self-administration, an animal undergoes surgical implantation of a chronic indwelling catheter (e.g., jugular vein of a rat), which is linked to an external access port. During experimental sessions, the external port is connected to a drug supply line, such than some defined action expressed by the

animal (e.g., lever press) triggers the dispensation of a volume of drug into the bloodstream and delivery of a drug-paired cue (e.g., cue light). Three decades of research have revealed that neuroadaptations as a consequence of drug intake differ between active self-administration and passive administration, either experimenter delivered or yoked. This was most elegantly demonstrated in studies that utilized a yoking procedure, where one animal actively self-administers the drug while another "yoked" subject receives passive identical infusions. One study found changes in dopamine turnover rates in several brain regions in animals that had self-administered cocaine, which were independent from changes as a result of cocaine exposure alone (Smith et al., 2003). These findings suggest that there is a unique interaction between the choice to take a drug and its effect, which are absent when the volitional aspect is removed as in the experimenter-administered models described above. Thus, the strong face validity has led many to believe that self-administration offers the best model to study drug abuse liability and potential intervention treatments.

Moreover, self-administration studies can be divided into different phases, depending on experimental timeline, and types, defined by length of drug access per day. First, the animal must learn the action that produces drug intake. Once this initial association is discovered, the animal should increase its performance of that action if the drug effect is reinforcing. This phase is referred to acquisition. Depending on the experimental details, acquisition is generally followed by maintenance, in which the amount of drug taken stabilizes over several sessions, or escalation, in which drug intake continually increases. While these phenomena both indicate that the drug effect is reinforcing, escalation may indicate a more pathological hallmark of abuse liability. Until this point, self-administration studies can be classified as "short-access" (e.g., 1 h/day) as opposed to "long-access" (e.g., 6 h/day); different neurobiological adaptations have been found between these

setups. Next, animals may either be completely removed from the operant chambers (forced abstinence) or placed in the operant chambers without access to drug, and trained to press the lever without reinforcer delivery ("extinction"). The goal of extinction is to degrade the animals' associations between the action (e.g., lever press) and drug delivery to produce low responding in the future. The goal of abstinence experiments is to observe plasticity that is not contaminated by the degradation associated with extinction training.

After various timepoints following removal of drug access, animals can be placed back into the chamber for a "seeking" test, which are interpreted as models of relapse. Such tests can be performed by presenting subjects with the previously drug-paired cue ("cue-primed") during which responses on the active lever are recorded but no drug is delivered. The number of responses performed can be interpreted as a measure of how much the animal "wanted" the drug. Importantly, since drug is either not given or is administered in a fixed dose by an experimenter, it is possible to design experiments that incorporate more than one reinstatement session.

Many interesting findings have come from these procedures. Notably, the phenomenon of incubation has been described, where the drug-seeking responses of rats markedly increased longer into drug abstinence (e.g., greater on day 30 than on day 1), which are also associated with various molecular changes in the brain. These findings indicate long-term consequences of drug exposure that continue to occur despite the absence of ongoing drug exposure. Since the largest hurdle to remaining drug-free is avoiding relapse after periods of sobriety, it is possible that these long-term changes offer a starting point for investigating reinstatement/relapse propensity and targets for therapeutics.

Effects of repeated drug experience on behavior

There is a vast literature base on the effects of drugs on behavior. As described previously, these can be divided into examinations of acute drug effects (i.e., when the drug is biologically present in the organism) and chronic drug effects (i.e., how a single or repeated drug exposure alters the biological system over time). In the latter category, many animal studies within the field of substance abuse research have examined how drug experience changes behavior related to the drug itself or drug-related stimuli. The phenomenon of tolerance to pharmacological compounds is one example. Tolerance refers to the decreasing response to a drug stimulus over its repeated exposure, thereby necessitating higher doses over time to achieve the same biological response. Tolerance is found in both therapeutic situations, such as opioid management of pain, as well as in cases of SUD, where a subject will require larger amounts of substance to achieve the same euphoric effects.

As a parallel but contrasting phenomenon to tolerance, sensitization is another way that behavior associated directly with the drug is affected over drug exposure; sensitization refers to the increasing response to a drug stimulus over its repeated exposure. Sensitization likely reflects a subject's system becoming more attuned to detecting drug availability and seems to occur through different pathways than tolerance. For instance, when given repeated injections of an opioid, rats will become tolerant to the antinociceptive effects but sensitized to the locomotor-stimulating effects. As described previously, sensitization can be revealed most simply by increased locomotor responses to drug administration. However, other features of drug-related behavior have also been interpreted as sensitization. CPP, for example, where an animal learns to prefer a drug-paired environment, can be viewed as the animal becoming more "sensitive" to the drug context, causing it to spend more time there. Likewise, an animal learning to perform an action to receive drug administration or associating a cue stimulus with the delivery of a drug, are

each evidence of the animal becoming sensitized to those events as a function of drug exposure. This sensitization wherein the experience of drug exposure is transferred to other external stimuli can explain why a subject attributes importance to an associated context, cue, or operant device even after extinction.

In humans, sensitization might be related to how drug users experience cycles of abstinence followed by relapse (Hunt et al., 1971; Leshner, 1997; O'Brien, 2005). Although individuals are able to maintain periods of abstinence while in treatment, the abstinence is often forced, and upon release from treatment and return to their initial environments containing familiar drug-associated cues such as paraphernalia or drug-associated social connections or contexts, they relapse (O'Brien et al., 1992). In the preclinical lab setting, the relapse phenomenon can be modeled using context- or cue-induced reinstatement, described as the recommencement of drug seeking after extinction following exposure to drug-related contexts or stimuli (Shaham et al., 2003). To examine cue-induced seeking test, subjects first typically undergo a self-administration paradigm during which a response (e.g., lever press, nose poke, etc.) leads to delivery of a drug concurrently with the delivery of a discrete stimulus (e.g., light, tone, etc.). Contingent presentation of the discrete cue with the delivery of reward enables subjects to learn the association between the reinforcing effects of drug and the stimulus. Following self-administration, subjects undergo extinction, in which responses are no longer reinforced with either drug or cue and responding decreases. Subsequently, subjects undergo a reinstatement test during which only the drug-paired cue is presented, which reinvigorates responding. Typically, it is observed that presentation of the cue alone elicits responding, termed reinstatement of drug seeking, and quantification of responses for the drug-paired cue provides a readout of the magnitude of drug-seeking behavior. Thus, operationally the cue serves as a stand-in for the potential real-world drug-associated cues that a person might

encounter to trigger relapse. In one seminal study, subjects underwent several days of self-administration during which responses on a lever led to the contingent delivery of a buzzer tone with an infusion of morphine. After responding on the previously drug-paired lever was extinguished, subjects underwent the reinstatement test, during which experimenters found that subjects still responded for the cue alone. The authors asserted that learning to make a response to procure reward and learning the cue-drug association are indications that subjects become sensitized to such events throughout drug experience. Moreover, the process by which the reinforcing effects of drug transfer to environmental stimuli provides a likely reason as to why subjects attribute significance to drug-paired stimuli after extinction (Davis and Smith, 1976). Context-induced reinstatement offers a similar glance into relapse-related behavior but, instead of discrete cues, utilizes entire contexts.

Context-induced reinstatement is comparable to cue-induced reinstatement except that the effect of context on the resumption of drug seeking after extinction is examined. Context-induced reinstatement requires subjects to undergo self-administration in a distinguishable environment composed of unique stimuli (i.e., context A). Following self-administration in context A, subjects undergo extinction in a different context composed of its own set of distinct cues (i.e., context B). The subsequent reinstatement test is then performed in the drug-paired context (i.e., context A) and responding on the once-drug-paired lever (or nose poke port) is quantified. One study utilizing this procedure found that context A, but not B, reinstates responding, suggesting that the environment in which subjects previously took drug is an important determinant of drug-seeking behavior (Crombag et al., 2002).

Yet, the effects of drug experience clearly reach beyond modulating only those behaviors directly associated with drug. Indeed, drug experience has been found to induce broader changes

in both animal and human behavior, indicating more generalized changes in information processing and deficits in decision making. One clear example of this is seen in increased impulsivity (Roesch et al., 2007; Simon et al., 2007). Utilization of delay discounting tasks enables the study of the cognitive processes underlying the ability to compare values between small immediate and large delayed consumption of a commodity (e.g., reward outcome). The task is predicated on the idea that rewards lose value as the delay of delivery becomes longer, such that the delay discounts subjective value. Operationalized, such assays commonly require subjects to pick between immediately available small rewards or big rewards of varying delay. Subjects are thought to be impulsive when the imposition of delay has a greater negative effect on the subjective value of the outcome, and therefore exhibit steeper delay functions when several delays are tested. This assay can be used to examine how drugs of abuse impact impulsivity as it pertains to decision making. One study examined this by first administering rats cocaine (i.e., experimental rats) or saline (i.e., control rats) over two weeks and tested them on a variation of the delay discounting task following a three-week withdrawal period. The study found that while all rats demonstrated delay discounting, those with prior cocaine experience were less likely to choose the big, more delayed outcome compared to saline controls. Such a preference for small, more immediate rewards suggests that prior repeated cocaine exposure increases impulsive decision making (Simon et al., 2007).

Another example of how drugs of abuse cause generalized changes in information processing is seen in impaired Pavlovian reinforcer devaluation following repeated drug exposure. As discussed previously, devaluation can be used to examine the ability of subjects to learn to predict the current value of the expected outcome such that changing representations alters the value of the outcome, the ability of which is critical for adaptive decision making. Briefly, when a

neutral stimulus is paired with a reinforcer, establishing the association between the stimulus and the sensory / motivational representation of the outcome, then it stands to reason that devaluation of the outcome should reduce the conditioned response to the reinforcer-predicting cue. Importantly, reduced responding occurs quickly even though the cue is never overtly paired with the devalued reinforcer, suggesting that it is not new learning. Instead, it demonstrates that the subject utilizes the cue to evoke a representation regarding the outcome and its new value and then uses the updated information to guide decision-making behavior. In one study, experimenters administered rats with cocaine (i.e., experimental rats) or saline (i.e., control rats) over two weeks and, following a three-week withdrawal period, asked whether the repeated exposure to drug altered the associative processing important for outcome devaluation. To do this, all rats were first conditioned to associate a cue light with a food reward outcome. The food reward was then devalued for half of the rats of each group by pairing it with LiCl-induced malaise (i.e., devalued rats), while the other half received a control saline injection (i.e., non-devalued rats). On a subsequent session, rats underwent a 16-trial extinction probe test during which they were again presented with the cue and responding was compared across the four experimental groups. In control rats, responding to the cue light was sensitive to alterations in the value of the food reward outcome, but such modifications in responding were not present in the experimental cocaine-treated group, suggesting that prior drug prevented the reduced responding to the cue light caused by outcome devaluation. Such deficits in the ability to adjust behavioral responding reflect an inability to learn and further use information of the value of predicted outcomes, indicating that the maladaptive decision making observed following repeated drug experience could be, in part, due to a persistent deficit in the ability to represent and use the information about current outcome value that is often critical for guiding behavior (Schoenbaum and Setlow, 2005). A similar effect

was observed in amphetamine-sensitized rats that failed to show proper devaluation while behaving on a reinforced devaluation task (Nelson and Killcross, 2006).

The effects of drugs on the associative processing underlying decision-making behavior have also been repeatedly assessed using reversal learning paradigms. As discussed previously, reversal learning paradigms are useful for measuring the ability to adapt responding following changes in reward contingency, serving as a useful proxy by which to quantify behavioral flexibility. Briefly, such tasks are composed of repeated pairings of a stimulus (e.g., odor cue, visual cue, etc.) with an outcome (e.g., delivery of reward), enabling subjects to learn contingencies through sensory properties of cues that predict reward and stimuli associated with them. Accomplishing this necessitates subjects to first learn that cues carry information regarding whether a stimulus predicts a reward (e.g., presentation of odor cue 1 means respond to left, but not right, fluid delivery well to receive reward; presentation of triangle, but not circle, leads to reward). Once discrimination is observed, suggesting acquisition of associative rules, contingencies are reversed. Adaptive behavior following reversals requires suppression of the initially acquired association and expression of the previously unreinforced association. Reversal learning paradigms can be used to examine processes underlying flexible decision making and whether repeated drug experience causes persistent deficits in such processes. One seminal study examined this by first administering monkeys with cocaine (i.e., experimental) or saline (i.e., control) for 14 days. Following a 9 or 30-day withdrawal period, they then tested the monkeys on a variation of an object discrimination reversal learning task. The study found that although the experimental and control groups demonstrated similar acquisition of the initial discrimination, regardless of the withdrawal period, monkeys with prior cocaine exposure made more errors following the reversal as compared to the control subjects (Jentsch et al., 2002). Similar results

were found when monkeys were allowed to self-administer cocaine over several months (Porter et al., 2011). Furthermore, these deficits have been consistently demonstrated across several species (Schoenbaum et al., 2004; Fillmore and Rush, 2006; Calu et al., 2007; Ersche et al., 2008; Krueger et al., 2009), suggesting that repeated exposure to drugs of abuse impairs guidance of behavior when outcomes change.

Effects of repeated drug experience on neurophysiology

The aforementioned behavioral effects are likely the result of drug-induced neurophysiological changes. A great amount of literature delves into both the acute and persistent effects of drug exposure on brain activity. In humans, studying the effects of drug on brain activity has two important differences from preclinical research. First, drug-naïve individuals typically cannot participate in studies for which they would be receiving a drug of abuse for the first time. Thus, human studies generally compare the effects of acute drug versus saline administration in chronic drug users, making it difficult to directly compare the effects of drug in new users versus chronic users. Second, despite a few instances of in vivo electrophysiology in humans, assessing brain activity in humans is generally limited to non-invasive methods that do not entail implanted devices to record brain activity. Instead, humans can undergo functional magnetic resonance imaging (fMRI) positron-emitting tracer (PET) imaging, which give a broad sense of neural activity changes, but are low-resolution compared to methods available in preclinical studies.

One study using fMRI in humans to examine the acute effects of drug found distinct patterns of brain activation following experimenter administration of cocaine, revealing neural circuitry related to cocaine-induced craving and euphoria. In this study, cocaine-dependent subjects first underwent fMRI imaging for 5 minutes prior to receiving a treatment to establish baseline activity. Subsequently, individuals received either an infusion of cocaine or saline and

then entire brains were imaged for 13 minutes during which subjects reported their feeling on a scale of rush, high, low, and craving. Compared to saline, infusions of cocaine were found to increase signal in the NAc, VTA, putamen, caudate thalamus, insula, basal forebrain, cingulate, hippocampus, parahippocampal gyrus, temporal and lateral prefrontal cortices, parietal cortex, striate/extrastriate cortices, and pons, while decreasing signal in the medial frontal cortex, amygdala, and temporal pole. Further, while brain areas that showed early and transient signal maxima including regions of the lateral prefrontal cortex, VTA, pons, basal forebrain, caudate, and cingulate were correlated with self-reported rush ratings, regions showing early and persistent signal maxima including the right parahippocampal gyrus, regions of the lateral prefrontal cortex, and the NAc , as well as negative signal change in the amygdala, were more related to craving. Together, these finding demonstrate that drugs of abuse alter brain activity, which in certain areas relates to reward processing (Breiter et al., 1997). Another human study using blood oxygenation level-dependent (BOLD) fMRI responses to examine the acute effects of cocaine following a history of chronic cocaine use on brain activity found that administration of drug to cocaine users showed temporal negative or positive BOLD responses and activated mesolimbic and mesocortical dopaminergic projection areas. Further, cocaine infusions were found to elicit activity in hierarchical brain networks in the OFC, prefrontal cortex, and Brodmann area 10. Together, these results show that acute drug exposure leads to robust changes in brain activity, suggesting a role for both hierarchical networks and dopaminergic circuitry in facilitating the reward, memory, motivation, and associative learning processes underlying cocaine abuse in humans (Kufahl et al., 2005).

Similar examinations have been performed in non-human subjects, with the added benefit of the comparison of drug effect between naïve and experienced subjects One study in cocaine-

experienced monkeys examining the acute effects of cocaine on brain activity as a function of drug experience found that the effects of drug and impacted regions change dynamically over time. Briefly, PET imaging with F-18 labeled fluorodeoxyglucose (FDG) was used in rhesus macaques to characterize the metabolization of glucose, as a proxy for regional brain activity, following acute experimenter administered intramuscular (i.m.) cocaine infusions at different phases of intravenous (i.v.) cocaine self-administration history. Initially, FDG-PET measurement following acute experimenter-administered cocaine in cocaine-naïve monkeys revealed increased glucose metabolism in regions of the medial prefrontal cortex; however, as cocaine experience increased from limited to extended access self-administration, acute drug-induced increases in metabolism expanded to include the anterior cingulate, sensorimotor-related areas, as well as the striatum. Such extensions of increased activity suggest that while the initial acute effects of cocaine could be facilitated by the prefrontal cortex, the effects following chronic drug exposure impact higher-order regions and their related processes including attentiveness, impulsivity, compulsivity, and emotion. After four weeks of withdrawal from cocaine self-administration, acute infusions of cocaine no longer resulted in increased patterns of metabolic activity. Overall, these data show that the acute effects of drugs on brain activity are incredibly dynamic, changing as a function of prior drug history (Porrino and Lyons, 2000; Henry et al., 2010; Hanlon et al., 2013).

However, the effects of drugs on neurophysiology extend beyond those directly related with drug. In another human study examining the chronic effects of drug, researchers performed a coordinate-based meta-analysis using fMRI studies from individuals with addictive alcohol, cannabis, nicotine, cocaine, or unbiased substance use during tasks that engaged functional domains defined as “Cognition” with executive-control, emotion, and behavior and “Reward” composed of motivation, anticipation, or processing of reward tasks. Analysis revealed common

changes in frontal and striatal circuits involved in executive control, reward processing, and habit formation between different behavioral tasks and drugs. Further sub-analyses of paradigm-specific changes between substances showed pronounced striatal changes during processes related to reward and larger frontal changes during more cognitive processes. These findings suggest that prior substance use results in the dysregulation of frontostriatal circuits during cognitive and reward-related processes even when drug is not on board, likely contributing to the maladaptive decision making observed in SUDs which may contribute to substance-specific behavioral alterations (Klugah-Brown et al., 2020).

While gross imaging studies nicely capture the acute and persistent effects of drugs on the neurophysiology related to reward processing and decision making, they lack the resolution that in vivo electrophysiology offers, limiting experimental interpretations to relatively large brain areas, when in reality some effects may result from alterations in specific brain subnuclei or neuronal subpopulations. While *in vivo* electrophysiology is being pioneered in humans, historically these finer-grained investigations have been limited to application in non-human subjects. Moreover, human studies typically cannot enroll drug-naïve individuals to receive a drug for the first time. Thus, animal studies are also able to better capture the effects of drug on the brain at early experiential timepoints.

Single unit recordings during acute drug administration

Early studies looking at the acute effects of drug on striatal activity in awake freely moving cats found that experimenter-administered amphetamine increased single-unit activity. Specifically, the study recorded single units in the striatum of 6 cats while awake and mobile. Comparison of pre-treatment baseline firing rates of individual units with firing rates collected after cats received treatments of varying doses of amphetamine revealed that at high doses (5.0

mg/kg, intraperioneally (i.p.)) amphetamine produced exclusively excitatory responses in striatal units, while at lower doses (0.5 and 2.0 mg/kg, i.p.) it elicited both excitatory and inhibitory responses. Further, amphetamine treatment evoked a stereotypy and hyperactive behavioral response in the cats, which researchers described as tending to outlast unit firing rate changes. Such early studies in awake and freely moving animals were some of the first to demonstrate the acute effects of drug treatment on brain activity (Trulson and Jacobs, 1979).

Another, more recent study examined the effect of experimenter-administered cocaine on DMS activity in awake behaving mice. To examine the acute effects of cocaine, mice received an i.p. injection of either cocaine or saline and single units were recorded in the DMS as mice freely moved about an open field. Consistent with previous findings, acute administration of cocaine increased DMS firing compared to saline controls. However, calcium imaging, which in this study enabled the isolated assessment of either direct or indirect pathway MSNs, showed that cocaine interestingly did not differentially modulate these two opposing populations, but instead increased the activity of both. This contrasted with proposed models suggesting that stimulants drive hyperactivity via activating direct pathway neurons and inhibiting indirect pathway neurons. These changes were linked to a strengthening of upstream inputs to the DMS originating from the OFC, which importantly did not show differences in connectivity rate or strength when synapsing between the two DMS populations. Accordingly, weakening this OFC-DMS synapse with a high-frequency stimulation protocol blocked these cocaine-induced connectivity changes and inhibited the expression of cocaine-induced locomotor sensitization when delivered during a cocaine challenge test. These findings thus directly implicate the OFC-DMS pathway as a key modulator of the behavioral effects of cocaine (Bariselli et al., 2020).

Single unit recordings following repeated drug exposure

The effects of drug exposure on cortical activity independent of the DMS have also been investigated. One study in rats examined the gradual and persistent change in the activity of OFC and medial prefrontal cortex (mPFC) units to repeated acute amphetamine exposure. In the first experiment, researchers examined the effects of amphetamine on neural activity in a dose-dependent manner by injecting naïve rats with either a low (0.5 mg/kg) or high dose (2 mg/kg) of amphetamine and performing single unit recordings. Findings revealed that acute drug exposure produced region- and dose-dependent patterns of activity changes. Both doses of amphetamine increased unit activity in the OFC. In contrast, in the mPFC, while low dose was found to increase activity, high dose predominantly decreased unit activity. Researchers then went on to examine the dynamic neurophysiological changes that occurred over repeated exposure to drug. To do this, rats first underwent five recording sessions during which they received an experimenter administered injection of high dose amphetamine and freely locomoted as unit activity was recorded. Rats then underwent a 10 day "drug off" period followed by three subsequent sessions one week apart during which rats received challenge doses of amphetamine to examine whether prior drug exposure potentiated future responses to the high or low dose amphetamine. Although repeated amphetamine injections similarly increased the responsivity of single units in the OFC and mPFC to acute high dose amphetamine injections, it altered the distribution of response types differently across regions. While the proportion of units in the OFC with a drug-induced excitatory response increased with no change in the fraction of units with an inhibitory response, the opposite was true for the mPFC, with no change in the relative amount of units with an excitatory response but an increase in the proportion with an inhibitory response. Amphetamine challenges of the same high dose also revealed regional differences. While acute injections resulted in greater magnitudes of both excitatory and inhibitory responses in the OFC, increased magnitude in inhibitory

responses was only observed in the mPFC, suggesting that prior repeated drug exposure results in the potentiation of responsivity to the same dose of drug. Interestingly, the low dose amphetamine challenge produced effects resembling that of the higher dose across both regions. In the OFC, acute administration of low dose amphetamine increased the proportion of both excitatory and inhibitory units and increased the magnitude of their responsivity. In the mPFC, however, low dose challenge elicited an inhibitory effect. To test whether such amphetamine-induced neurophysiological changes in the OFC and mPFC were behaviorally relevant, the group examined single unit activity in these regions as rats performed an instrumental goal-directed task. Rats were trained on an operant task in which a nose poke earned the delivery of a food pellet, and the total number of food pellets was quantified as the measure of instrumental responding. Once trained, recording sessions began. To examine the effect of drug on behavior, rats intermittently performed the task during recordings prior to amphetamine injections to collect baseline unit activity, after which they were treated with amphetamine, returned to the task, and recording of unit activity resumed. Overall, data revealed that repeated amphetamine exposure elicited greater and more persistent excitatory responses in OFC units and inhibitory responses in mPFC units, and that such effects were coupled with the gradual impairment of behavior on the instrumental learning task with a strong correlation between task performance and the effects of drug on behavior, suggesting that such cortical regions could be important substrates for the effects of drug on goal-directed behavior (Homayoun and Moghaddam, 2006).

Another study recording single units in rats found that repeated methamphetamine self-administration dysregulated unit activity in the dorsomedial prefrontal cortex (dmPFC). Rats were trained to self-administer methamphetamine or receive yoked-saline infusions using a 14-day escalation procedure. Following extinction and reinstatement, rats were anesthetized and putative

pyramidal units in the dmPFC were recorded. Analysis of unit activity showed that prior repeated methamphetamine exposure increased both basal firing frequencies and the proportion of burst-firing units compared to yoked controls. These findings also correlated with behavioral deficits observed in the same study in which the previously methamphetamine-administering animals exhibited deficits during extradimensional shifts in the ASST. Together, these data show that prior chronic methamphetamine disrupts dmPFC unit activity and PFC-dependent behavior (Parsegian et al., 2011).

Single unit recordings following drug experience during normal information processing

As a slight distinction from the above studies, one investigation of the chronic effects of drug on OFC function found that prior drug exposure disrupts the processing of behaviorally relevant associative information during decision making. To examine this, rats first underwent 14 days of an experimenter-administered cocaine or saline procedure. Rats were then trained on a go, no-go odor discrimination paradigm during which they had to sample one of two odors on each trial and respond at a fluid well. While responses to the fluid well following a "positive" odor was deemed correct and resulted in the delivery of a 0.05 mL of a 10% sucrose solution; the same response following a "negative" odor was incorrect and resulted in the delivery of 0.02 mL of a quinine solution. Over training, rats were able to discriminate the odor cues such that they would consistently respond to the fluid well following a positive odor cue paired with sucrose and withhold the response (i.e., no-go) following the negative cue paired with quinine. After training, single units in the OFC of rats were recorded while they performed the task. It was found that while units recorded from cocaine-experienced rats responded normally to the expectation of quinine once the choice had been made to respond, they did not increase activity during the consideration of the quinine-predicting odor stimulus, suggesting that prior cocaine treatment

impaired the ability of OFC units to signal adverse outcomes. Additionally, the study found that following contingency reversal of cue-outcome associations, odor cue-selective unit responses were less adaptive in cocaine-experienced rats. Overall, these results demonstrate that prior chronic cocaine exposure can cause persistent, behaviorally meaningful alterations in the representation of the associative information about natural outcomes important for guiding behavior more generally (Stalnaker et al., 2006).

Summary

Impaired decision making is an important feature of SUDs. In this section, we have reviewed historical and recent findings supporting the central idea that drugs of abuse cause persistent changes on regional brain activity that results in these effects on decision making, contributing further to SUD. While drug-induced changes in human brain activity can be subject to somewhat limited interpretation, animal studies make it easier to directly compare drug-naïve and experienced subjects and offer higher-resolution approaches for measurement and analysis. Moreover, we reviewed animal models of decision making as well as substance abuse disorders. While it is clear that animal models of SUDs certainly do not entirely capture the nuance of human SUD, the common view in the field is that volitional drug intake paradigms are the current best approach for modeling SUD in animals. Furthermore, various animal models of decision making are clearly reliant on distinct brain regions, some of the most notable regions being the OFC, PL, and further downstream, the DMS, which are all involved in maintaining state representations to support optimal decision making. Impairments to the functions of these regions as a result of continuous drug exposure may in turn impair decision making, causing suboptimal decisions and perhaps contributing to continued drug seeking and taking despite negative consequences.

The overarching goal of this series of experiments was to examine how drug exposure affects decision making in rats. Since this necessitated a two-stage approach, we implemented procedures in each phase to maximize the degree to which drug effects reflected those occurring clinically, and draw nuanced conclusions about how decision making and information in discrete brain regions affected by drug exposure. Importantly, we did not necessarily anticipate drug effects that were maximally disruptive; as reviewed recently, even pathologies in human drug users are often overstated in clinical studies, where drug-naïve and –experienced subjects can often not be accurately identified by brain scans or task performance if their group labels are removed (Grifell and Hart, 2018). Instead, we sought to examine the most clinically relevant effects, which might be expressed as quite minute changes in brain activity and behavior, and otherwise be easily overlooked. As such, we leveraged the above-described advantages offered by the self-administration model of volitional drug intake and *in vivo* electrophysiological recordings during an odor-guided discrimination task of reversal learning. Cocaine self-administration permitted the important interaction between drug exposure and the choice to consume the drug. Subsequently, the odor-guided choice task presented a richly textured canvas upon which the performance of specific neuronal populations could be thoroughly characterized. Performing this sequence for high-order cortical regions in the PL and OFC as well as the behavioral output-proximal DMS gave us access to a systems-level understanding of potential drug-induced neural pathologies.

Chapter 2: Effects of Cocaine Experience on OFC and PL Representations During Decision Making

This chapter is in preparation for submission.

Mueller LE, Konya C, Sharpe MJ, Moghaddam B, Wikenheiser AM, Schoenbaum G.

Introduction

Dysfunctional decision making is a prominent component of substance use disorders (SUDs), demonstrated by continued drug use in the face of adverse consequences (APA, 2013; SAMHSA, 2016; Koffarnus and Kaplan, 2018). Previous work has shown that prior drug experience disrupts the associations necessary for proper behavioral adaptation when reward contingencies change, leading to impaired choice behavior. Impaired outcome-guided choice behavior has been repeatedly demonstrated on reversal tasks in humans (Fillmore and Rush, 2006; Verdejo-Garcia et al., 2007), rats (Schoenbaum et al., 2004; Calu et al., 2007), mouse (Krueger et al., 2009), and monkey studies (Jentsch et al., 2002; Porter et al., 2011), with prior drug experience resulting in marked perseveration following the contingency switch. Additionally, rats with prior drug histories demonstrate impairments on other assays used to probe aspects of decision making, such as outcome devaluation (Schoenbaum and Setlow, 2005; Nelson and Killcross, 2006), and contingency degradation (Dickinson et al., 2002; Miles et al., 2003), suggesting that drugs of abuse disrupt outcome-guided choice behavior. Such impairments in adaptive choice behavior could be related to why animals and humans continue drug use despite negative consequences.

Optimal decision making is contingent upon adaptive behavior, the ability to flexibly adjust actions or responses according to changes in the world. Such behavior will be facilitated if one can properly encode appropriate rules for each situation, requiring the interlacing of memories of past experiences and their outcomes with newly updated information. The ability to do so is achieved by assigning learning in different situations or contexts into compartmentalized "states"; whereby the information contained within each state defines the rules appropriate for guiding behavior under particular conditions (Bradfield et al., 2013; Schoenbaum et al., 2013; Stalnaker et al., 2016; Mueller et al., 2021; Stalnaker et al., 2021). The ability to develop and alternate between states speeds adaptive responding when circumstances change. Recent work suggests that suboptimal

behavior observed following drug experience might be related to the abnormal encoding of such states (Mueller et al., 2021). Specifically, it is possible that drugs affect the way the brain generalizes and separates the rules used to guide situationally-optimal behavior.

This previous work found that prior drug experience affects representations of state-related information in the dorsomedial striatum (DMS), preserving and strengthening the representation of information not normally maintained in naïve controls, changes which correlated with a failure to optimize behavior during valued-based decision making (Mueller et al., 2021). While such changes in state-related encoding may be due to direct drug action on striatal populations, it is also possible they reflect contributions from afferent cortical regions also key for adaptive behavior and impacted by drug, such as the orbitofrontal (OFC) and prelimbic cortices (PL). Drugs of abuse have been shown to cause persistent impairments in the ability of the OFC and PL to process associative information (Jentsch and Taylor, 1999; Volkow and Fowler, 2000; Gobin et al., 2019), effects evident at both the behavioral (Grant et al., 2000; Bechara et al., 2001; Schoenbaum and Setlow, 2005; Calu et al., 2007; Ersche et al., 2008; Gobin et al., 2019) and neurophysiological levels (Thorpe et al., 1983; Homayoun and Moghaddam, 2006; Stalnaker et al., 2006; Chen et al., 2013). As regions that project to the DMS, the OFC and PL may be crucial for shaping the encoding of state-related information required for adaptive decision making in DMS.

To address whether such drug-induced changes in the DMS are secondary to changes occurring in the OFC and PL during choice behavior, we recorded WS (wide spike) units in the OFC and PL of male rats that had previously self-administered cocaine as they performed a decision-making task similar to that used to characterize activity in DMS in prior work. In this task, the response to procure the most valuable reward switched across blocks of trials. By pseudo-randomly presenting both free- and forced-choices in our task, blocks of trials established unique

multi-layered states since each choice type within each state required a distinct rule necessary for optimal performance on the task. We examined whether neural representations of these generic (i.e., trial block) and specific states (i.e., choice type within trial block) in the OFC and PL would be affected by prior cocaine self-administration.

Materials and Methods

General methods

Subjects

Eight male Long-Evans rats (250-300 g, Charles River Labs), aged approximately three months, were subjects for this experiment. During the odor-guided choice task training and recording sessions, rats received water *ad libitum* for 10 minutes per day and *ad libitum* food. During self-administration, rats were food deprived to 85% of initial weight with *ad libitum* water. Self-administration and odor task testing were performed during the light phase. All experimental procedures complied with Institutional Animal Care and Use Committee of the US National Institute of Health guidelines.

Surgical methods for jugular catheter and electrode implantation

All surgeries were performed under aseptic conditions as previously described (Stalnaker et al., 2010). Rats were randomly assigned to cocaine or sucrose self-administration groups. Rats used for cocaine self-administration (n = 4) received chronic indwelling jugular catheter implants; rats used for sucrose self-administration (n = 4) received sham surgeries in which the jugular vein was exposed but no catheter was implanted. Rats recovered for seven days before self-administration began. During recovery and self-administration, catheters were flushed daily with

a cocktail of enrofloxacin and heparinized saline to maintain patency. Following self-administration and training on the decision-making task, recording electrodes composed of drivable bundles of sixteen 25-μm diameter NiCr wires electroplated to an impedance of ~200 kΩ were implanted into the PL of one hemisphere and the OFC of the opposite hemisphere in each rat (PL: 2.9 mm AP, ±0.7 mm ML, 3.5 mm V (Burgos-Robles et al., 2013); OFC: 3.0 mm AP, ±3.2 mm ML, 4.0 mm V (Wikenheiser et al., 2017); relative to bregma and dura). The hemispheres of the implanted regions were counterbalanced across rats. Rats were allowed two weeks to recover and then returned to the decision-making task for one week of reminder training and subsequent recording sessions. Electrodes were advanced following each recording session to capture new units.

Self-administration

The self-administration procedure was similar to that described previously (Lucantonio et al., 2014). Briefly, rats were trained to self-administer intravenous cocaine-HCl (0.75 mg/kg/infusion; n = 4) or oral sucrose (10%, wt/vol; n = 4) under a fixed ratio 1 schedule in 3 h sessions over 14 consecutive days.

Odor-guided choice task

The decision-making task was similar to that described previously (Stalnaker et al., 2016). Task training and recording was performed using chambers equipped with two fluid wells and an odor port. Self-paced sessions began with the illumination of a house light, with trials initiated by a nose poke in the odor port. Upon odor port entry, rats held for a 500-ms fixation period, which was followed by a 500-ms presentation of one of three instructional odors that remained constant

throughout the experiment. Following odor presentation, rats withdrew from the odor port and indicated their choice by entering either the left or right fluid well within 3000 ms. Upon fluid well entry, rats held for 500 ms before fluid reward delivery began. Once rats consumed the reward and withdrew from the fluid well, the house light was extinguished, and the trial ended. If rats withdrew prematurely at any point prior to fluid delivery, the trial was aborted, and the house light turned off. Two instructional odors specified a forced-choice reward delivery at either the left or the right fluid well. The third odor indicated a free choice trial. Presentation of odors occurred in a pseudorandom sequence with forced-choice right/left odors delivered equally on 65% of trials and free-choice odors delivered on 35% of trials.

Reward outcomes differed in flavor identity (vanilla or chocolate milk) and size (small [single 0.05 mL drop] or big [three 0.05 mL drops]), with opposite flavor and size outcomes being delivered at opposing fluid wells. Sessions consisted of a short initial block and followed by four blocks of approximately 60 trials each. Response-reward contingencies were consistent for blocks of trials but switched across blocks; block switches were not explicitly signaled and their timing varied randomly to prevent anticipation. Rats were trained before electrode implantation until they performed forced-choice trials with greater than 70% accuracy and completed all blocks. Following electrode implantation and recovery, rats received an additional week of reminder training to increase task performance and acclimation to recording cables.

Single-unit recording

Neural activity was recorded using Plexon Multichannel Acquisition Processor Systems (Plexon Inc.) interfaced with a training chamber. Electrode signals were amplified 20x via operational head stages on electrode arrays and then passed to differential preamplifiers where

they underwent an additional 50x amplification and filtration at 150-9000 Hz (Stalnaker et al., 2010; Stalnaker et al., 2016). From there, signals were passed to multichannel acquisition processors where they underwent an additional 250-8000 Hz filtration, 40 kHz digitization, and final 1-32 x amplification. Signal waves were derived from active channels with event timestamps documented by the behavioral program.

Experimental Design and Statistical Analyses

Behavioral epochs

Task trials were divided into six epochs either 0.5 s or 0.8 s in length (Figs. 3 and 4, top). The pre-odor port nose poke epoch started after house light illumination, signaling the availability of a new trial, and concluded after 0.8 s, immediately prior to odor port entry. The fixation epoch began at time of odor port entry, which required rats to hold their snouts in the odor port, and ended after 0.5 s, immediately prior to odor cue delivery. The odor sample epoch began at the time of odor onset and ended after 0.5 s when odor cue delivery ended. Once the odor cue presentation ended, rats chose between the two fluid delivery wells. The movement epoch comprised 0.5 s immediately prior to entry into the fluid delivery well. The anticipation epoch began at the time of fluid delivery well entry and ended after 0.5 s, immediately prior to outcome delivery. The consumption epoch began at the time of outcome delivery and ended after 0.8 s.

Spike sorting and unit type classification

Neural activity was stored and manually sorted into putative single units using Offline Sorter (Plexon Inc.). Files were processed using NeuroExplorer (Nex Technologies) to extract timestamps and further analyzed using MATLAB (The MathWorks, Inc.). Orbitofrontal and

prelimbic cortical populations were classified as wide waveform units (wide spike; WS) and narrow waveform units (narrow spike; NS) using a hierarchical unsupervised cluster analysis as previously described (Quirk et al., 2009; Letzkus et al., 2011; Burgos-Robles et al., 2013). Scatter plots of spike waveform widths (μs) and firing rates (Hz) show resulting clusters per region and treatment group (OFC Fig. 2b; PL: Fig. 2c). Of the 1737 units recorded from the OFC of sucrose-experienced rats, 1545 were classified as WS and 192 were classified as NS. Of the 1124 units recorded from the OFC of cocaine-experienced rats, 970 were classified as WS and 154 were classified as NS. Of the 1080 units recorded from the PL of sucrose-experienced rats, 985 were classified as WS and 95 were classified as NS. Of the 1313 units recorded from the PL of cocaine-experienced rats, 1209 were classified as WS and 98 were classified as NS.

Neural firing rate dynamics

To analyze the general response properties of units, firing rates for each unit were computed in 50 ms bins, averaged across completed free- and forced-choice trials in all blocks, and peak normalized (Figs. 3a, 4a). Units were counted as maximally active for a trial epoch if, for at least one bin during that period, the unit's peak-normalized average firing rate exceeded 95% of its absolute maximum value (Figs. 3b, 4b). Raw, un-normalized population firing rates were computed across all units of a given population and trials in 50 ms bins (Figs. 3c, 4c).

Regression analyses

Dynamics of neural selectivity for task variables were analyzed using linear regression models fit to each unit's firing rate over the course of completed free- and forced-choice trials as previously described (Wikenheiser et al., 2017; MATLAB function: fitglm; Fig. 5a,c). Within each

epoch, firing rates were calculated in 50-ms bins, and for each bin, an individual regression model was computed. Outcome flavor, size, response direction and choice type served as categorical predictors. Within each unit, p-values were corrected for multiple comparisons, and the percent of the total population significant for each predictor was calculated for individual time bins (corrected p-value < 0.05). To look at selectivity during the odor sample epoch we implemented a similar approach using one 0.5-s time bin. Units were deemed selective for a dimension if corrected regression model p values were less than 0.05 (Fig. 5b,d).

Multi-dimensional representational similarity analyses

To measure the degree to which multiple task dimensions were represented by cortical population ensembles, WS units per region and treatment group were pooled and median unit firing rates were z-scored across each of the 16 unique *choice type* x *direction* x *size* x *flavor* combinations, creating population vectors for each combination. Similar to methods previously described, the relationship between ensemble representations were then measured by computing the Pearson's correlation coefficient between population vectors for every pair of trial-related variable combinations (Kriegeskorte et al., 2008; McKenzie et al., 2014; Figs. 6a, 7a). Population level hierarchical representations were visualized using dendrograms (MATLAB functions: *linkage* and *dendrogram*). Population vectors were created as described above and cluster trees were computed using the unweighted average distance (UPGMA) between pairs of vectors and correlation as the distance metric (Figs. 6b, 7b).

Quantification of dimension representations – choice type, direction, size and flavor – was calculated using the d' metric in 100 unit pseudoensembles. Pseudoensembles were generated by randomly pulling subsets of 100 units for inclusion from the entire classified WS population per

region of recorded units across all sessions. Pseudoensemble generation and d' calculations were performed independently for sessions recorded from sucrose- and cocaine-experienced rats. Pseudoensembles were used to create population vectors as described above, correlations were drawn between vectors associated with each trial-related dimension combination, and d' metrics were computed measuring the separation of distributions of the correlations from within-dimension conditions (e.g., pairs of identical choice type combinations) versus the distribution of the correlations from between-dimension conditions (e.g., pairs of opposing choice type combinations). For each pseudosensemble, we performed a shuffling procedure to confirm that d' of randomized data was at the theoretical chance level of zero. To do this, we calculated d' values using trial-related combination labels randomly re-assigned to within and between-dimension correlation values. D' values of each dimension were averaged across 250 runs, comparisons were made between dimension (choice type- vs. direction- vs. size- vs. flavor), treatment groups (sucrose- vs. cocaine-experienced), and time (100 ms bins), and effects were analyzed using two-way ANOVAs (Figs. 8 and 9).

To examine whether the strength of OFC WS choice representations during odor sample were related to between-choice type transition errors, we generated pseudoensembles by binning every two recording sessions for which each rat was run and the pooled unit population exceeded 20 units (n = 8 bins; i.e., first bin session: all rats' sessions 2-3, second bin session: all rat's sessions: 4-5, etc.) and pulled subsets of 20 classified OFC WS units per treatment group. Population vectors were created using pseudoensembles for each of the eight bin sessions and d' of choice type was calculated for each bin as described above. Corresponding between-choice transition errors were calculated by averaging over the same sessions included in each bin. Coefficients of correlation were computed between d' of choice values and associated transition

errors and an ANCOVA was used to determine the relationship of best-fit line slopes between treatment groups. Because slopes were similar, treatment group data was combined and one correlation was drawn from pooled data (Fig. 10b).

Statistical Analyses

All analyses were conducted using MATLAB (The MathWorks Inc., Natick, MA) and GraphPad Prism 8 software (GraphPad, La Jolla, CA). One- and two-way ANOVAs, ANCOVAs, and t-tests were used to analyze all data, as reported in results and figure legends. The Bonferroni-Holm procedure was used to correct for multiple comparisons. P-values were significant if they fell below 0.05. Sample sizes were determined from previous studies utilizing similar behavioral and recording procedures (Stalnaker et al., 2006; Stalnaker et al., 2016; Wikenheiser et al., 2017; Mueller et al., 2021). Error bars or shading on plots represent the standard error of the mean.

Results

Rats were shaped on a decision-making task in which blocks of trials defined unique states containing conflicting associative information (Fig. 1a,b; Stalnaker et al., 2016). The paradigm consisted of a short warm-up block, followed by four full length trial blocks across a single session. On each trial within the same block, one of three unique odor cues was presented at the odor port, requiring rats to respond to the left or to the right fluid well on forced-choice trials, or in either direction on free-choice trials, to receive milk rewards that differed in size (big: three 0.05 ml drops; small: one 0.05 ml drop) and flavor per well (vanilla or chocolate milk). At the unsignaled block switches, the reward solutions available in each fluid well changed in either flavor or size, establishing distinct sets of response-outcome contingencies per trial block. After acquiring this

basic task, rats were trained to self-administer either oral sucrose (n = 4) or cocaine (n = 4) using a fixed ratio one schedule of reinforcement for 3 h/d for 14 d (Fig. 1a,c). Two-way repeated-measures ANOVAs revealed significant session × lever interactions in both treatment groups (sucrose: *p < 0.0001, F(13,39) = 7.348; cocaine: *p < 0.0001, F(13,39) = 20.21). Following self-administration, rats received some additional training on the decision-making paradigm, after which two drivable bundles of electrodes were implanted into each rat, one in the PL and one in the OFC, with the hemisphere of the implanted regions counterbalanced across rats (i.e., left PL, right OFC or left OFC, right PL; Fig. 1a). After recovery, rats were given seven days of reminder training for acclimation to recording cables and then neural recording sessions began (Fig. 1a); electrodes were advanced after each session to capture new neurons. To control for potentially interacting effects of time and self-administration experience, recording sessions were run concurrently for both treatment groups.

Prior cocaine self-administration slowed decision making

During recording sessions, both sucrose- and cocaine-experienced rats demonstrated a strong preference for the big reward on free-choice trials, quickly adjusting their choice for the well producing the big reward over the first 20 trials following a value block switch (Fig. 1d). The rate of change was similar across treatment groups. A two-way ANOVA on free choice behavior following a value block switch revealed a significant main effect of trial number but no treatment group or interaction (trial number: *p < 0.0001, F(5,8832) = 242.48); treatment group: p = 0.421, F(1,8832) = 0.65; int: p = 0.4257, F(5,8832) = 0.98; Fig. 1d), and a direct comparison of early versus late trials failed to show a significant interaction between these two factors (p = 0.5627, F(1,230) = 0.3360, Fig. 1d, inset). The strong preference for the big reward was also similar

between treatment groups on free-choice trials following flavor identity block switches (Fig. 1e). A two-way ANOVA on free choice behavior following flavor identity block switches revealed a subtle but significant main effect of trial number, but there was neither main effect nor interaction with treatment group (trial number: *p = 0.0145, F(5,8885) = 2.84); treatment group: p = 0.2533, F(1,8885) = 1.31; int: p = 0.6866, F(5,8885) = 0.62; Fig. 1e), and there were no group differences when early and late trials were directly compared (p = 0.9936, F(1,230) < 0.1, Fig. 1e, inset). Both treatment groups also exhibited enhanced accuracy on big forced-choice trials, again with no effect of cocaine on this measure (Fig. 1f). A two-way ANOVA revealed a significant main effect of size but not treatment group or interaction (size: *p < 0.0001, F(1,230) = 292.0; treatment group: p = 0.3271, F(1,230) = 0.9643; int.: p = 0.6433, F(1,230) = 0.2151; Fig. 1f). Consistent with previous findings (Mueller et al., 2021), cocaine caused a general slowing of responding to the fluid well after odor sample (Fig. 1g). A two-way ANOVA revealed significant main effects of treatment group, size, and interaction (treatment group: *p < 0.0001, F(1,230) = 484.1; size: *p < 0.0001, F(1,230) = 148.4; int.: *p < 0.0001, F(1,230) = 18.11). Bonferroni-Holm post hoc testing showed that the cocaine group had significantly slower reaction times for both small and big rewards (small: *p < 0.0001, t(460) = 23.2; big: *p < 0.0001, t(460) = 20.92; Fig. 1g).

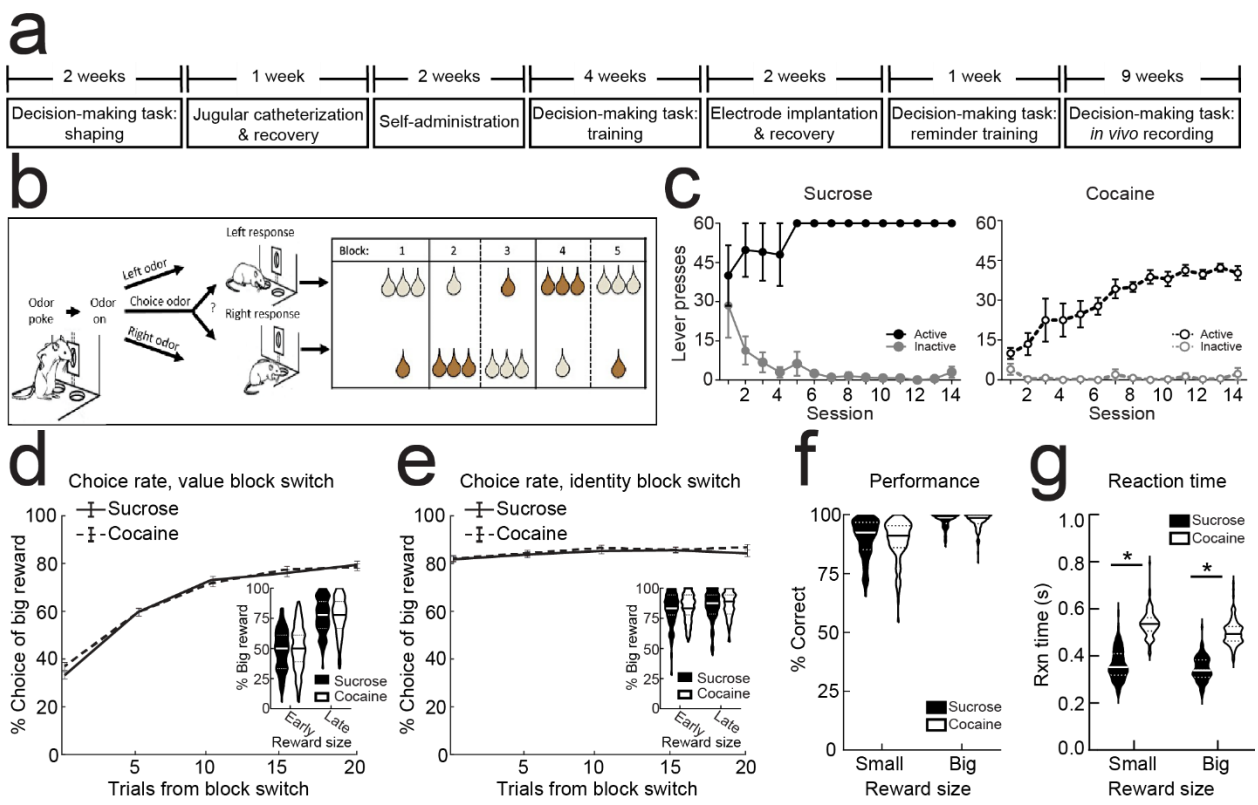

Figure 1. Self-administration, task design, and behavior.
a) Experimental timeline. b) The decision-making task consisted of four unique blocks of trials defined by distinct response-reward contingencies. Instructional odors delivered to the central odor port indicated which action (go left, go right, go either direction) would be reinforced on each trial. On forced-choice trials, correct responses to the left or the right fluid well were reinforced with fluid delivery. On free-choice trials, responses to either fluid well resulted in fluid delivery. Fluid reinforcers delivered at each well differed in size and flavor and alternated following unsignaled block switches as shown in the example block sequence. c) Number of active (black circles) and inactive (grey circles) lever presses during each session of sucrose (left; n = 4) and cocaine (right; n = 4) self-administration. Error bars indicate SEM in lever presses. d) The percent of big rewards chosen on free-choice trials was computed and aligned to value block switches or (e) identity block switches (sucrose: solid black line; cocaine: perforated black line). Error bars indicate SEM in the percent choice of the big reward. Choice rate behavior for early and late trials following block switches (insets). Solid lines indicate medians and dots represent upper and lower quartiles. f) Correct performance on forced-choice trials for small (left) or big rewards (right). Solid lines indicate medians and dots represent upper and lower quartiles g) Reaction times for small (left) and big rewards (right) on free- and forced-choice trials. Solid lines indicate medians and dots represent upper and lower quartiles. *$p < 0.05$, Bonferroni-Holm post hoc testing.

Prior cocaine self-administration disrupted response dynamics of OFC and PL WS units

To investigate whether prior cocaine experience altered neural activity in OFC and PL cortices, units were recorded in the OFC and the PL (Fig. 2a). A previously established unsupervised hierarchical clustering method was used to divide unit populations in each region and treatment group into two clusters, one consisting of wide spikes (WS) and the other narrow spikes (NS; Quirk et al., 2009; Letzkus et al., 2011; Sotres-Bayon et al., 2012). OFC recordings in sucrose-experienced rats yielded a total of 1737 units of which 1545 were classified WS and 192 NS, and 1124 total units in cocaine-experienced rats with 970 WS and 154 NS (Fig. 2b). Session recordings in the PL of sucrose-experienced rats yielded 1080 units with 985 defined WS and 95 NS, and 1307 units in cocaine-experienced rats with 1209 defined WS and 98 NS (Fig. 2c). Proportions of classified WS and NS units per region were similar across treatment groups and were consistent with previous findings (Quirk et al., 2009; Sotres-Bayon et al., 2012; Insel and Barnes, 2015). Studies indicate that NS are likely putative interneurons, consisting of both fast-spiking and non-fast-spiking units, and WS are likely putative pyramidal units (McCormick et al., 1985; Kawaguchi and Kubota, 1997; Cauli et al., 2000; Quirk et al., 2009; Sotres-Bayon et al., 2012). In this study, we focus on WS unit populations.

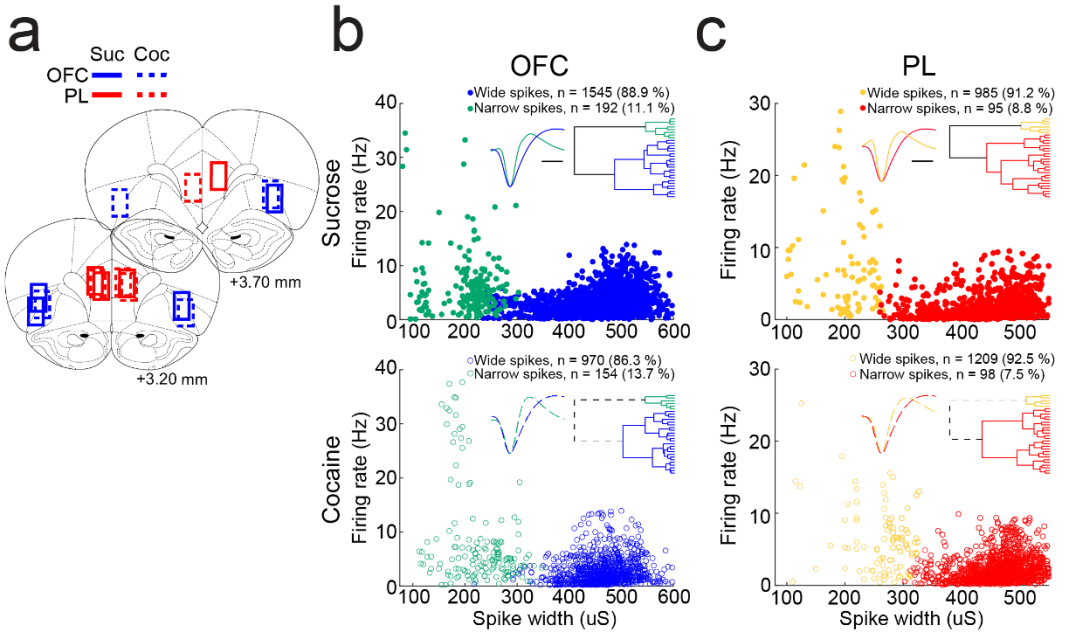

Figure 2. Histology and unit classification.
a) Approximate location of neural recordings in the PL and OFC are indicated by boxes (sucrose: solid line; cocaine: perforated line). b) Mean firing rates and spike widths for each OFC unit recorded from sucrose- (top; NS, 6.51 Hz, 205.70 µs; WS, 3.25 Hz, 461.17 µs) and cocaine-experienced rats (bottom; NS, 8.05 Hz, 229.69 µs; WS, 2.7 Hz, 465.08 µs) yielding two unique populations consistent with putative narrow-spiking interneurons (NS, green) and wide-spiking pyramidal cells (WS, blue). Dendrogram insets depict population clusters identified by the unsupervised cluster analysis. Wave insets show the average waveforms for classified unit populations. Scale bar; 100 µs. c) Mean firing rates and spike widths for each PL unit recorded from sucrose- (top; NS, 7.27 Hz, 207.50 µs; WS, 1.84 Hz, 446.22 µs) and cocaine-experienced rats (bottom; NS, 6.18 Hz, 265.1 µs; WS, 2.33 Hz, 468 µs) yielding two unique populations consistent with putative narrow-spiking interneurons (NS, yellow) and wide-spiking pyramidal cells (WS, red). Dendrogram insets depict population clusters identified by the unsupervised cluster analysis. Wave insets show the average waveforms for classified unit populations. Scale bar; 100 µs.

To examine whether prior cocaine experience altered the general response properties of OFC and PL WS units during performance of the decision-making task, completed free- and forced-choice trials were divided into five epochs (Figs. 3 and 4, top), and units captured throughout the entire recording experiment were sorted according to the time of maximum peak-normalized response (Figs. 3a and 4a). This revealed a significant increase in the proportion of OFC WS units with a peak response during the fixation and odor sample epochs, while significantly decreasing the proportion during the anticipation and consumption epochs in cocaine-

as compared to sucrose experienced-rats (Z-test for population proportions, corrected for multiple comparisons: fixation, *p < 0.0001, z(2514) = 10.2535; odor sample, *p < 0.0001, z(2514) = 6.92; anticipation, *p < 0.0001, z(2514) = -5.9984; consumption, *p < 0.0001, z(2514) = -8.8052; Figs. 3a,b, asterisks). Additionally, cocaine significantly decreased the raw, un-normalized average firing rate for OFC WS (Fig. 3c). A two-way ANOVA revealed a significant main effect of treatment (*p < 0.0001; F(1,140728) = 1506) along with significant main effects and interactions with time (time: *p < 0.0001, F(55,140728) = 7.81; int.: *p < 0.0001, F(1,55) = 12.02). Bonferroni-Holm post hoc tests revealed between-group differences reaching significance during the fixation, odor sample, and movement epochs (Fig. 3c, asterisks).

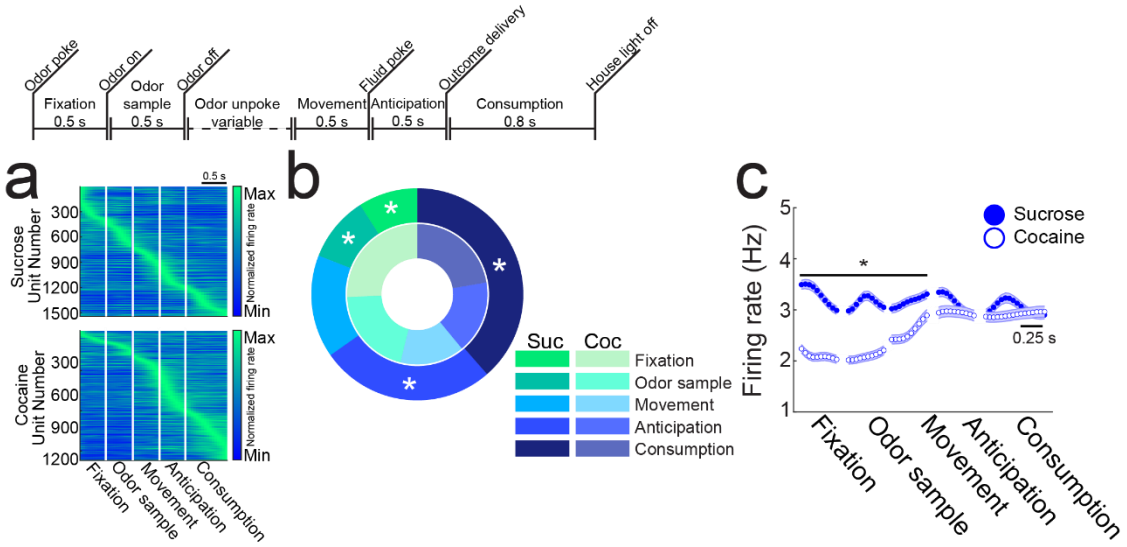

Figure 3. Time course of OFC WS neural responses across treatment groups.
Completed free- and forced-choice trials were divided into five epochs bounded by important task events. a) Peri-event time histograms (PETHs) of peak-normalized firing rates in units recorded from sucrose- (top) and cocaine-experienced rats (bottom) aligned to each of the five task epochs. Each row shows the firing rate of a single unit using a color scale ranging from blue (zero) to green (peak) with time on the x-axis (each panel = 0.5 s or 0.8 s; bin size was 50 ms). b) Proportion of units maximally active per epoch in units recorded from sucrose- and cocaine-experienced rats (outer and inner rings, respectively). *p < 0.05, z-test for population proportions corrected for multiple comparisons. c) PETHs of raw, un-normalized population-average firing rates in units recorded from sucrose- and cocaine-experienced rats (closed and open circles, respectively) aligned to each of the five task epochs. Bin size was 50 ms; shading indicates SEM. *p < 0.05, Bonferroni-Holm post-hoc testing.

Prior cocaine experience had the opposite, though weaker, effect on firing rate dynamics in the PL, significantly decreasing the proportion of PL WS units with a peak response during the odor sample and movement epochs, while increasing the proportion during the consumption epoch

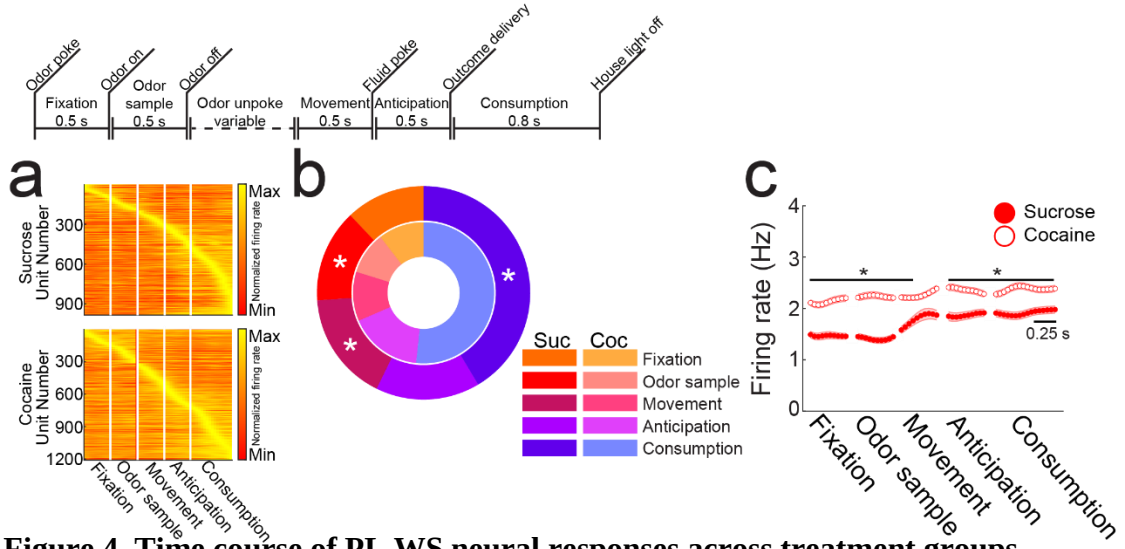

Figure 4. Time course of PL WS neural responses across treatment groups.
Completed free- and forced-choice trials were divided into five epochs bounded by important task events. a) Peri-event time histograms (PETHs) of peak-normalized firing rates in units recorded from sucrose- (top) and cocaine-experienced rats (bottom) aligned to each of the five task epochs. Each row shows the firing rate of a single unit using a color scale ranging from red (zero) to yellow (peak) with time on the x-axis (each panel = 0.5 s or 0.8 s; bin size was 50 ms). b) Proportion of units maximally active per epoch in units recorded from sucrose- and cocaine-experienced rats (outer and inner rings, respectively). *p < 0.05, z-test for population proportions corrected for multiple comparisons. c) PETHs of raw, un-normalized population-average firing rates in units recorded from sucrose- and cocaine-experienced rats (closed and open circles, respectively) aligned to each of the five task epochs. Bin size was 50 ms; shading indicates SEM. *p < 0.05, Bonferroni-Holm post-hoc testing.

(Z-test for population proportions, corrected for multiple comparisons: odor sample, *p = 0.0104, z(2193) = -3.08; movement, *p = 0.0018, z(2193) = -3.5729; consumption, *p < 0.0001, z(2193); Figs. 4a,b, asterisks) and generally increased the raw, un-normalized population average firing rate (Fig. 4c). A two-way ANOVA revealed a significant main effect of treatment (*p < 0.0001; F(1,122752) = 2016.68) along with significant main effects and interactions with time (time: *p < 0.0001, F(55,122752) = 10.32; int.: *p < 0.0001, F(1,55) = 2.69). Bonferroni-Holm post hoc tests revealed between-group differences reaching significance in the trial epochs (Fig. 4c, asterisks).

Prior cocaine self-administration reduced OFC selectivity for choice type

While the prior analysis demonstrated that cocaine experience disrupted the response properties of OFC and PL WS units, it did not address whether the information they encoded was impacted. Similar to previous studies (Feierstein et al., 2006; Roesch et al., 2006; Tsujimoto et al., 2009; Stalnaker et al., 2014; Wikenheiser et al., 2017), we found that units from these cortical regions represented single-dimension features related to the trial. All four variables defining the trial and its outcomes – choice type (free or forced), direction (left or right fluid delivery well), size (big or small), and flavor (chocolate or vanilla) – were represented by individual OFC and PL WS units (Fig. 5).

To compute selectivity for single-dimension task variables across all WS units per cortical region recorded throughout performance of the task, we fit generalized linear models to each units' firing rate at every time bin throughout completed free- and forced-choice trials and plotted the percent of units selective for each individual variable (following Bonferroni-Holm correction for multiple comparisons) at each point.(Figs. 5a,c) This showed that selectivity for most of the variables increased either during or following the odor sample epoch, the period in a trial that rats had knowledge about the choice type of the trial they were performing on, as well as the available

actions and outcomes. However, while the proportion of OFC WS units selective for size and flavor was similar between unit populations from both treatment groups, the proportion selective for choice type and direction was significantly lower in the population recorded from cocaine-experienced rats (Fig. 5a). Closer examination of the effect of prior cocaine on the ability of units to represent information during the odor sample epoch revealed a significant decrease in the proportion of WS OFC units selective for choice type and direction (Z-test for population proportions: choice type: $*p < 0.0001$, $z(2514) = 5.1774$; direction: $*p < 0.0001$, $z(2514) = 4.5907$; Fig. 5b). However, a different effect was observed in PL. While the fraction of units selective for choice type, direction and flavor were similar across treatment groups, the fraction selective for size was significantly higher in the unit population recorded from cocaine-experienced rats (Fig. 5c). Calculation of the fraction of PL WS units that selectively encoded size during the odor sample epoch revealed a significant increase in the amount of size-selective units following cocaine experience (Z-test for population proportions: $*p < 0.0001$, $z(2193) = -3.3243$; Fig. 5d).

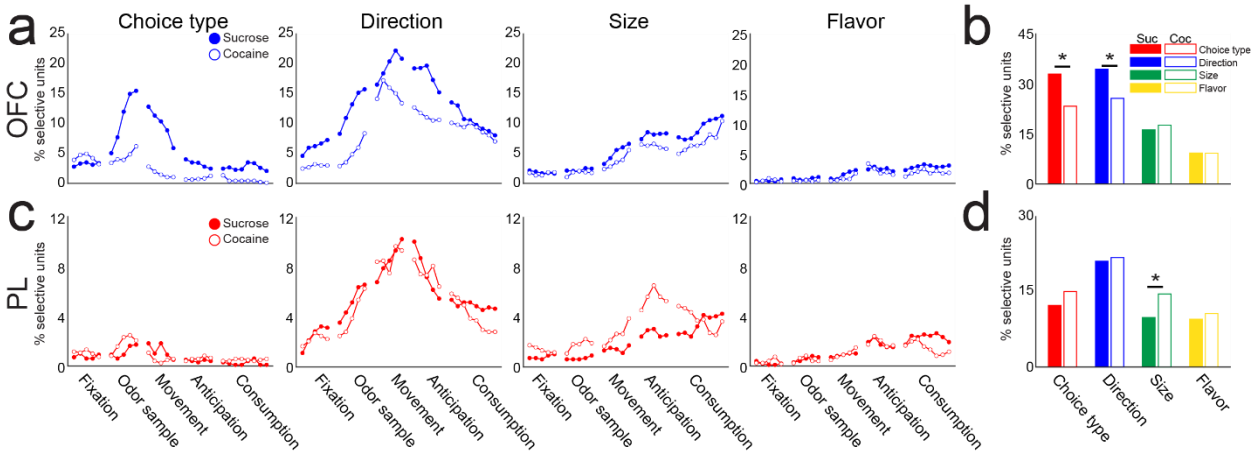

Figure 5. Prior cocaine self-administration reduced OFC WS unit selectivity for choice type. Regression models fit to neural data examined how well the firing rates of individual WS units were explained by dimensions of trial choice type, response direction, outcome size, and outcome flavor. a) Proportion of OFC WS units recorded from sucrose- and cocaine-experienced rats (closed and open circles, respectively) with activity significantly changed by trial-related dimensions throughout the trial. Bin size used to fit models was 100 ms. b) Proportion of OFC WS units recorded from sucrose- and cocaine-experienced rats (closed and open bars, respectively) with activity significantly modulated by dimensions during the odor sample epoch. Bin size used to fit models was 500 ms. *$p < 0.05$, z-test for population proportions. c and d) Same as (a and b) but for PL WS unit populations.

Prior cocaine self-administration altered the organization OFC WS multi-dimensional hierarchical representations

The preceding analyses revealed that activity of individual WS units in the OFC and PL reflect single dimension features characterizing task trials – choice type, location, size, and flavor – and that prior cocaine experience altered representations of such variables. However, these analyses did not address whether such representations spanned multiple dimensions of the task and how information was integrated and structured at the population level. To examine this, we used representational similarity analysis (RSA) (Lucantonio et al., 2012; McKenzie et al., 2014) to look at the geometry of the neural activity space across the four dimensions characterizing task trials. To compute the similarity of population level ensemble representations of different trial dimensions for each cortical region, we calculated the average z-normalized firing rate for each

WS unit for all 16 unique choice type x direction x size x flavor trial combinations during the odor sample epoch, and constructed a population vector for each combination based on normalized rates.

To visualize the pattern of combination-wise similarity between population vectors across all trial dimensions during the odor sample epoch, we constructed similarity matrices for WS unit populations per cortical region in which trial variables were sorted by choice type, direction, size, and flavor, with exemplar matrices of possible similarity patterns for comparison (Figs. 6 top, 6a and 7a). Further, to characterize the format, or hierarchy, of the information represented on the matrices, dendrograms were constructed by grouping correlation values into clusters (Figs. 6b and 7b). Matrices show ensemble similarity patterns that reflect trial-related dimensions (Figs. 6a and 7a).

The matrix generated using the OFC WS unit population recorded from sucrose-experienced rats revealed ensemble similarity patterns of choice type on free-choice trials (Fig. 6a, top), akin to the lower right quadrant of the choice type exemplar matrix (Fig. 6 top, choice type pattern). The same matrix also revealed ensemble similarity patterns of direction on forced-choice trials (Fig. 6a, top), like what is seen in the upper left quadrant of the direction exemplar matrix (Fig. 6 top, direction pattern). A follow-up dendrogram showed that the most separable representation encoded by the OFC WS unit population recorded from sucrose-experienced rats is choice type, and that depending upon the choice type (free- or forced-choice) the hierarchical structure of information differs. On forced-choice trials, when rats are required to respond to the fluid delivery well signaled by the instructional odor cue, a structured format of information was encoded, with direction being the most separable representation, followed by size and then flavor. However, on free-choice trials, when rats chose their preferred fluid delivery well, a different and

more ambiguously-organized hierarchical structure was encoded, with trial types first grouped based on location relative to the value block switch, followed by combinations of size and opposing outcome identities within a block, and then flavor (Fig. 6b, top).

Prior cocaine experience altered ensemble representation patterns in the OFC WS unit population, causing them to exhibit less similarity (Fig. 6a, bottom). For example, less ensemble similarity of choice type was observed on free-choice trials, as well as less ensemble similarity of direction on forced-choice trials. Visual comparison of the matrix with exemplar patterns did not reveal a striking similarity but rather a multiplexed representation of many of the matrices, suggesting that prior cocaine experience diminishes ensemble representations (Fig. 6 top, and 6a, bottom). A follow-up dendrogram revealed that prior cocaine experience altered the format of information encoded by the OFC WS unit population, resulting in location relative to value block switch being the most separable feature, followed by an ambiguous division of size and opposing outcome identities within a block, and then choice type and flavor (Fig. 6b, bottom).

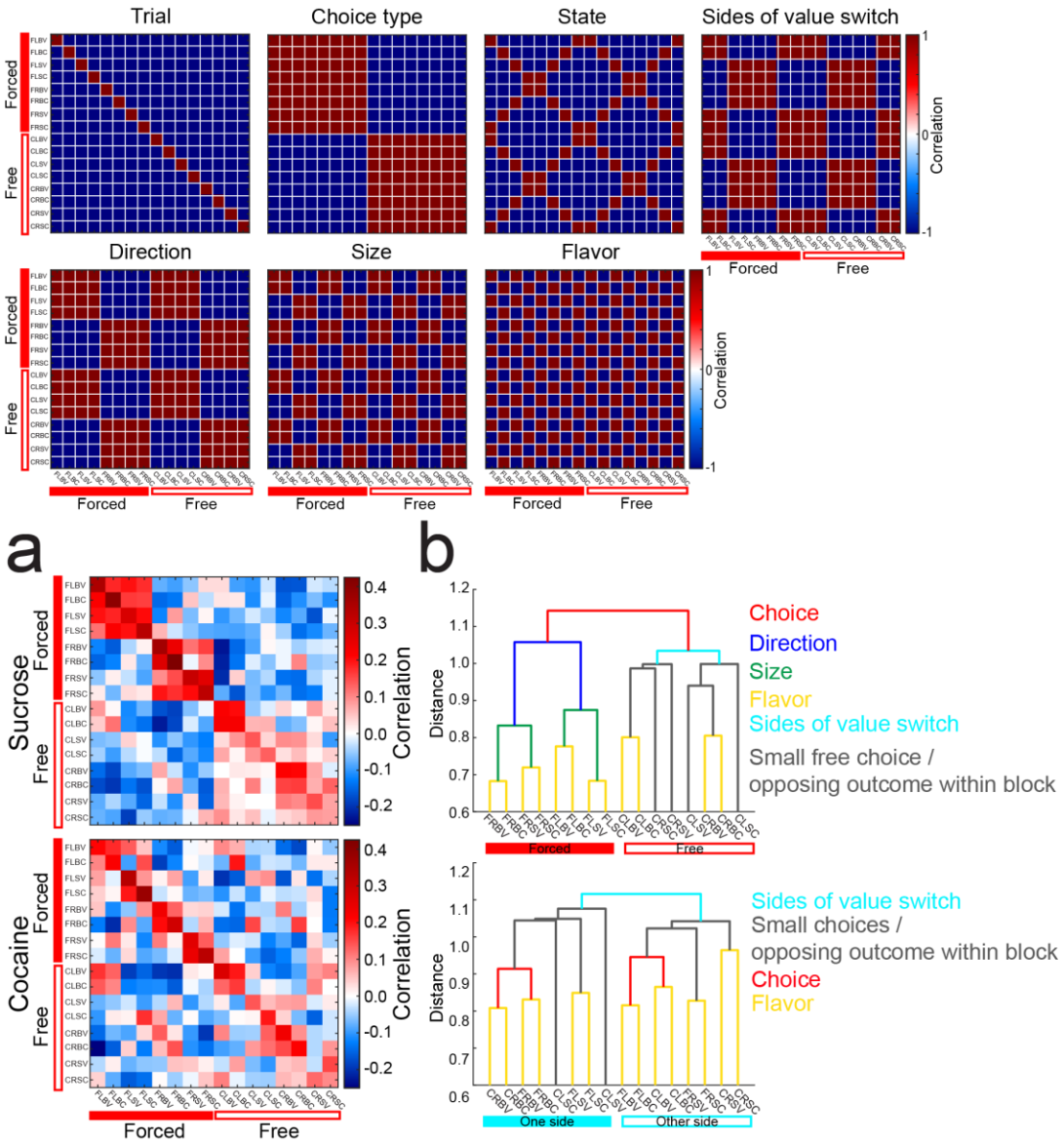
Trial
Choice type
State
Sides of value switch
Direction
Size
Flavor
Forced
Free
Correlation
a
b
Sucrose
Cocaine
Choice
Direction
Size
Flavor
Sides of value switch
Small free choice /
opposing outcome within block
Sides of value switch
Small choices /
opposing outcome within block
Choice
Flavor
Distance
One side
Other side

Figure 6. OFC WS population ensemble hierarchical representation during odor sampling. Correlation matrices illustrate the seven exemplar representations ("Trial", "Choice type", "State", "Sides of value switch", "Direction", "Size", and "Flavor"). The axes represent the correlation coefficients from between- and within-dimension variables, ordered by choice type, direction, size and flavor (FLBV = forced-left-big-vanilla; FLBC = forced-left-big-chocolate; FLSV = forced-left-small-vanilla; FLSC = forced-left-small-chocolate; FRBV = forced-right-big-vanilla; FRBC = forced-right-big-chocolate; FRSV = forced-right-small-vanilla; FRSC = forced-right-small-chocolate; CLBV = free-left-big-vanilla; CLBC = free-left-big-chocolate; CLSV = free-left-small-vanilla; CLSC = free-left-small-chocolate; CRBV = free-right-big-vanilla; CRBC = free-right-big-chocolate; CRSV = free-right-small-vanilla; CRSC = free-right-small-chocolate). a) Similarity matrices from OFC WS population ensembles recorded from sucrose- and cocaine-experienced rats (top and bottom, respectively) showing the correlation coefficients from between- and within-dimension variables of choice type, direction, size, and flavor by color scale during the odor sample epoch (right bar). b) Dendrograms from OFC WS population ensembles recorded from sucrose- and cocaine-experienced rats (top and bottom, respectively) showing the format of information represented during the odor sample epoch as a function of correlation of population vectors along various trial-related dimensions.

The matrix produced using the PL WS unit population recorded from sucrose-experienced rats showed ensemble similarity patterns of opposing trial dimensions of direction, size, and flavor (i.e., the two trial types occurring within the same state on both free- and forced-choice trials (Fig. 7a, top), like the pattern observed on the state exemplar matrix (Fig. 6 top, state pattern). A follow-up dendrogram revealed a less obvious hierarchical structure compared to the OFC, with the greatest separable representation being location relative to value block switch, followed by combinations of size and opposing outcome identities within the block, and then flavor and choice type (Fig. 7b, top).

Interestingly, prior cocaine experience did not have an obvious effect on PL WS similarity patterns (Fig. 7a, bottom). Accordingly, a follow-up dendrogram showed that prior cocaine experience did not alter the format of information represented by the PL WS unit population (Fig. 7b, bottom).

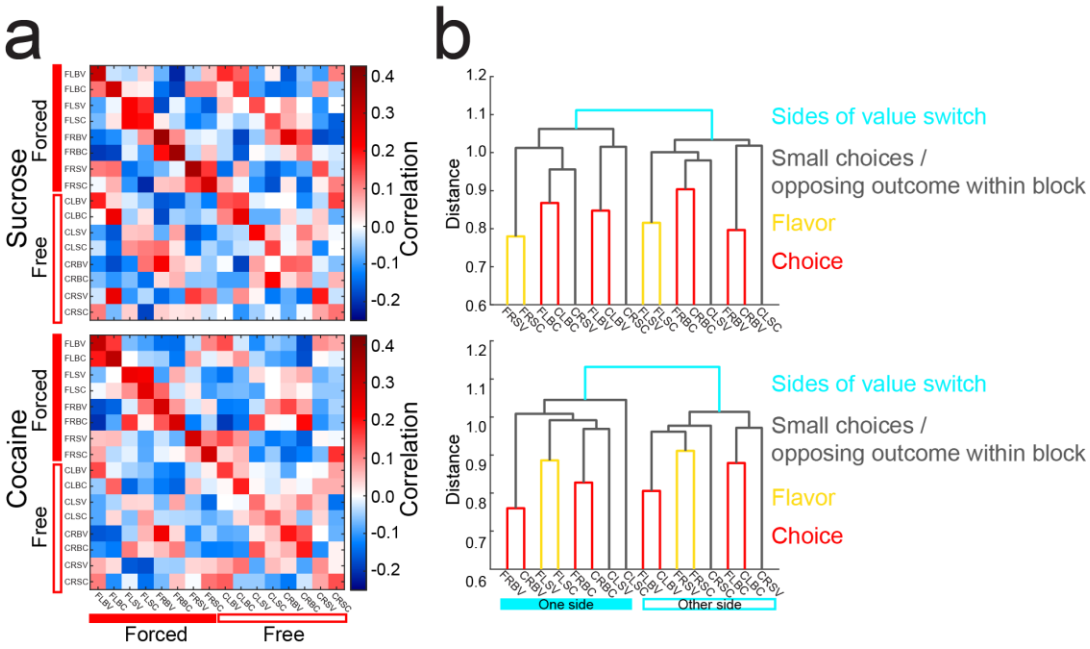

Figure 7. PL WS population ensemble hierarchical representation during odor sampling. a) Similarity matrices from PL WS population ensembles recorded from sucrose- and cocaine-experienced rats (top and bottom, respectively) showing the correlation coefficients from between and within dimension variables of choice type, direction, size, and flavor by color scale during the odor sample epoch (right bar). b) Dendrograms from PL WS population ensembles recorded from sucrose- and cocaine-experienced rats (top and bottom, respectively) showing the format of information represented during the odor sample epoch as a function of correlation of population vectors along various trial-related dimensions.

Prior cocaine self-administration decreased strength of OFC WS choice type representation

Although the preceding analyses illustrated the type and format of information encoded by these regions, it did not quantify the strength of each dimension – choice type, direction, size, and flavor. To examine the degree to which each dimension was represented during the behavioral task, we used d' to compute the separation of the distributions of correlation coefficients from within- (e.g., pairs of the same variables) and between-dimension conditions (e.g., pairs of opposing variables). That is, for each dimension of interest (e.g., choice type), we subtracted the average between-dimension distribution of correlation coefficients (e.g., each forced-choice-

associated population vector correlated with each other free-choice-associated population vector) from the average within-dimension distribution of correlation coefficients (e.g., each free-choice-associated population vector correlated with each free-choice-associated population vector, or forced-choice with forced-choice) and then divided the values by the spread of the distribution (i.e., standard deviation). Synthetic pseudoensembles were created by randomly choosing subsets of 100 units with replacement from classified populations across all sessions. Pseudoensembles were used to create population vectors associated with each of the 16 unique trial variable combinations, which were then used to compute d' for each of the four trial-related variables – choice type, direction, size, and flavor – this was repeated 250 times for each region and treatment group.

The results showed that OFC WS d' values of choice type, direction, and size exceeded the shuffled chance level throughout particular epochs of the trial (Fig. 8a). D' values prior to odor delivery were low; however, once the odor was presented, revealing the trial type, d' of choice increased, reflecting the increased importance of information about the type of choice rats were performing on during the odor sample epoch. D' of choice gradually decreased and direction increased during the movement epoch, indicating the importance of reward location as rats responded to the fluid delivery well. D' of direction gradually decreased during the anticipation and consumption epochs as d' of size increased suggesting increased representation of outcome value as rats expected and consumed the reward. Prior cocaine self-administration decreased the strength of choice type representations (Fig. 8b). Two-way ANOVA revealed significant trial-related dimension, treatment group, and interaction effects (dim.: $*p < 0.0001$, $F(3,216) = 61.87$; grp.: $*p = 0.0195$, $F(1,216) = 5.540$; int.: $*p = 0.0146$, $F(3,216) = 3.584$), post-tests showed significant between-treatment group differences for the choice type dimension ($*p = 0.0023$). A

follow-up two-way ANOVA on d' of choice type revealed significant time, group, and interaction effects (time: *p < 0.0001, F(27,13944) = 689.5; grp.: *p < 0.0001, F(1,13944) = 4461; int.: *p < 0.0001, F(27,13944) = 162.9). Bonferroni-Holm post-hoc testing on time bins found that d' of choice type was significantly decreased in OFC WS pseudoensembles from the cocaine group during most time points (Fig. 8c, red, asterisks), suggesting that prior cocaine self-administration reduces the amount of choice type information being encoded.

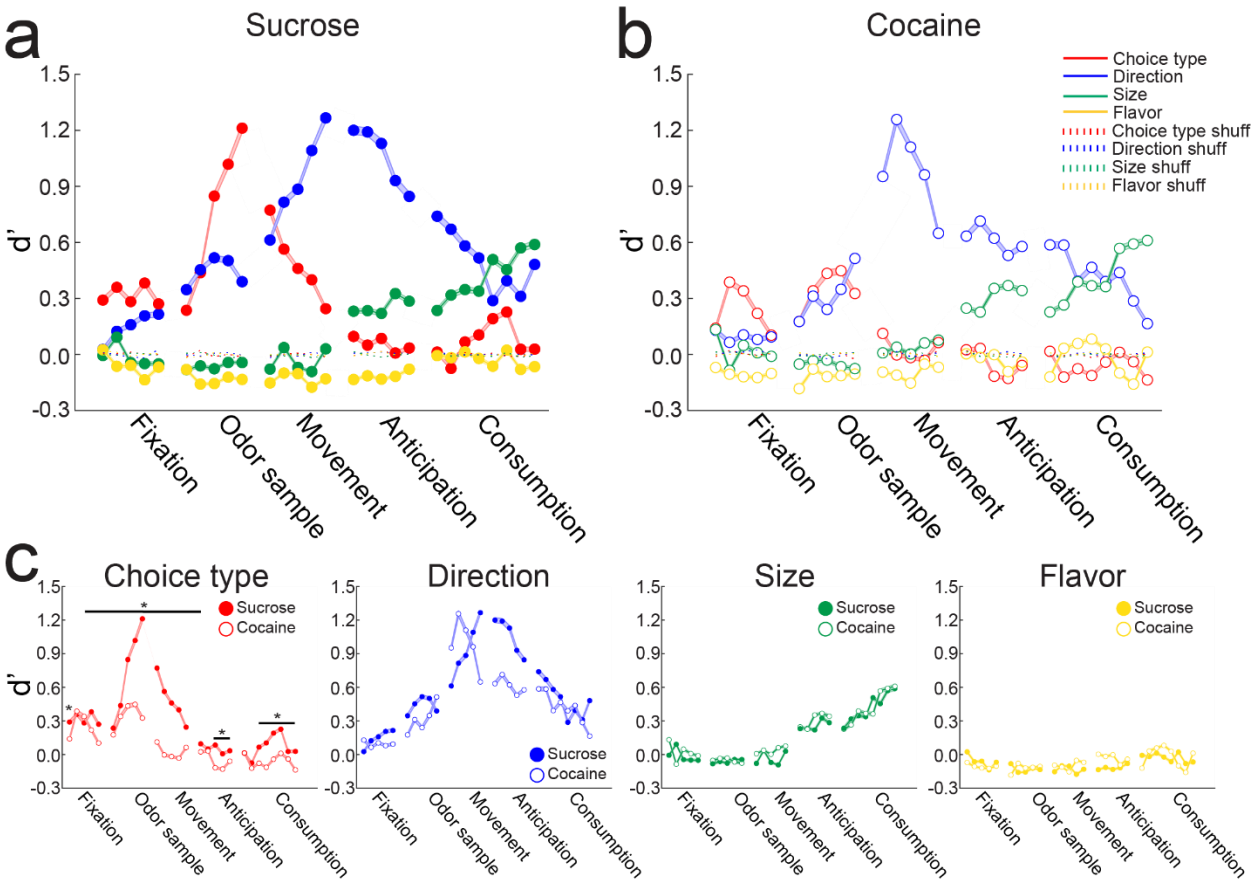

Figure 8. Strength of OFC WS representations of trial dimensions over time.
a) Mean d' of trial-related variables using pseudoensembles of 100 OFC WS units randomly selected from pooled units recorded across all sessions from sucrose- and (b) cocaine-experienced rats (filled and open circles, respectively) compared to respective shuffled values (perforated lines). c) Between-treatment group comparisons of d' values for choice (red), direction (blue), size (green), and flavor (yellow). Bin size was 100 ms; shading indicates the SEM in d'. *p < 0.05, Bonferroni-Holm post-hoc testing of d' values between pseudoensembles drawn from sucrose- and cocaine-experienced rats.

The same analysis performed using pseudoensembles of PL WS units show that the only dimension in which the d' value exceeded the shuffled chance level is direction (Fig. 9a). The d' value of direction became particularly high at the end of the odor sample epoch and through the movement epoch, likely reflecting the increased importance of reward location as rats decided and chose the action necessary to respond to the fluid delivery well. Prior cocaine experience altered the strength of direction representations (Fig. 9b). Two-way ANOVA showed significant trial-related dimension, treatment group, and interaction effects (dim.: *$p < 0.0001$, $F(3,216) = 59.51$; grp.: *$p = 0.0330$, $F(1,216) = 4.604$; int.: *$p < 0.0001$, $F(3,216) = 13.58$), post-tests showed significant between-treatment group differences for the direction dimension (dir.:*$p < 0.0001$). A follow-up two-way ANOVA on d' of direction revealed significant time, group, and interaction effects (time: *$p < 0.0001$, $F(27,13944) = 396.4$; grp.: *$p < 0.0001$, $F(1,13944) = 4035$; int.: *$p < 0.0001$, $F(27,13944) = 126.7$). Bonferroni-Holm post-hoc testing on time points found that d' of direction was significantly altered in PL WS pseudoensembles from the cocaine group during most time points (Fig. 9c, blue, asterisks), suggesting that prior cocaine self-administration reduces the amount of direction information being encoded.

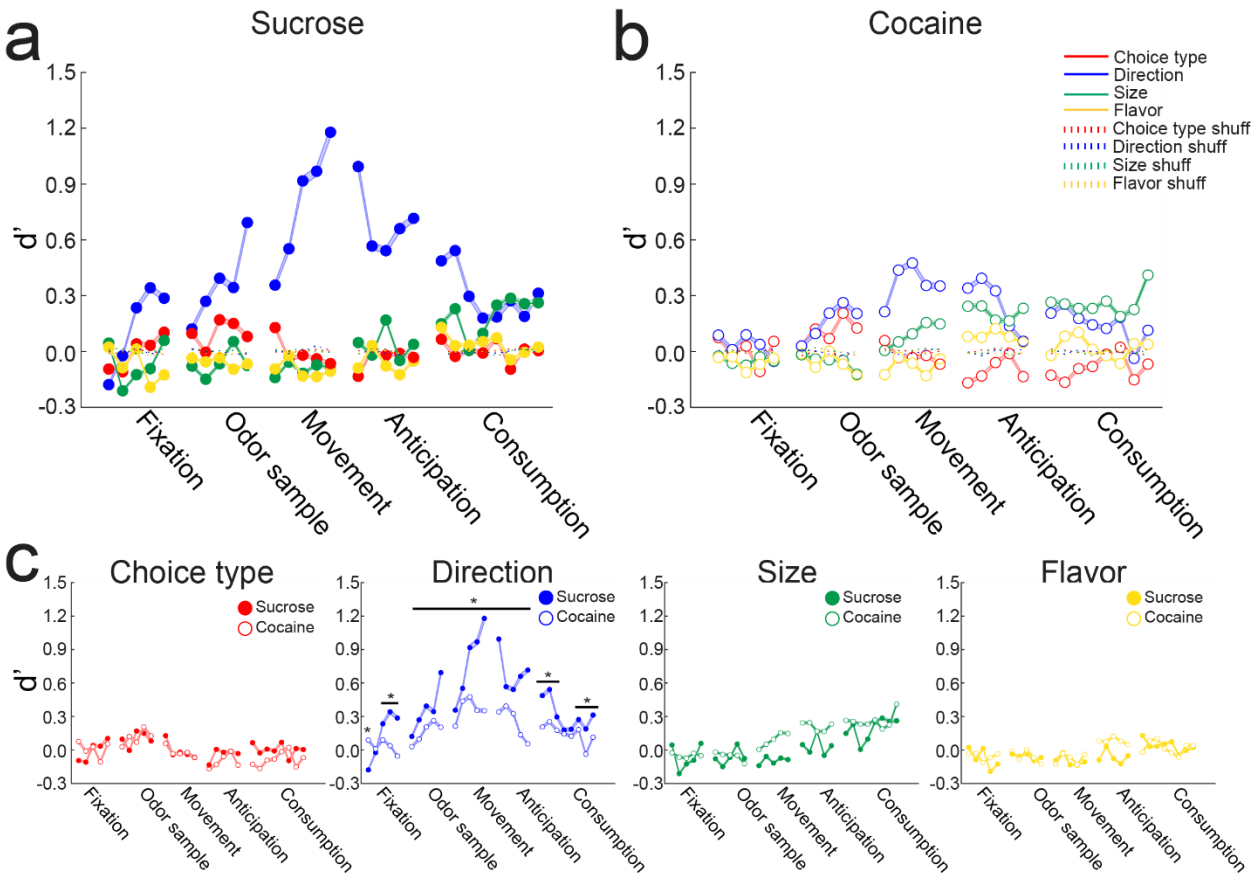

Figure 9. Strength of PL WS representations of trial dimensions over time.
a) Mean d' of trial-related variables using pseudoensembles of 100 PL WS units randomly selected from pooled units recorded across all sessions from sucrose- and (b) cocaine-experienced rats (filled and open circles, respectively) compared to respective shuffled values (perforated lines). c) Between-treatment group comparisons of d' values for choice (red), direction (blue), size (green), and flavor (yellow). Bin size was 100 ms; shading indicates the SEM in d'. *$p < 0.05$, Bonferroni-Holm post-hoc testing of d' values between pseudoensembles drawn from sucrose- and cocaine-experienced rats.

Representation of choice in OFC WS unit population correlates with task performance

The strong encoding of choice type in the OFC during the odor sample epoch and the weakening of such encoding in ensembles recorded from cocaine-experienced rats, led us to question whether differences in representational strength were reflected in differences in behavior across treatment groups. We postulated that if OFC encoding of choice type was behaviorally relevant then perhaps there would be behavioral differences when transitioning between free- and

forced-choice trials (i.e., choice-type transition), such that weaker representations when alternating between choice types could correspond to less optimal behavior. To examine whether the ability of OFC WS unit population to encode choice type representations was behaviorally relevant to performance on the decision-making task, we examined the ability of rats to make optimal responses following choice type transitions (Fig. 10a). For this analysis, we quantified the percent of transition errors whereby errors were defined by either (1) an incorrect response on a forced-choice trial following a free choice or (2) choice for the small reward on a free-choice trial following a forced choice. Comparison of the performance on choice transitions across treatment groups revealed that prior cocaine experience led to greater transition errors (*$p = 0.0003$, $t(150) = 3.691$; Fig. 10a); further examination of choice transition errors showed significant between-group differences on free to forced as well as forced to free transitions (*$p < 0.0001$, $t(150) = 4.826$; *$p = 0.0391$, $t(150) = 2.082$). Investigation of errors within choice type did not reveal significant group differences (forced to forced: $p = 0.2412$, $t(150) = 1.177$; free to free: $p = 0.9778$, $t(150) = 0.02794$).

To explore whether transitions involving trial types of conflicting information (e.g., left forced-choice small to a right free-choice big) played a particular role in the observed behavioral deficit, we calculated the amount of errors for all combinations of size / direction across choice transitions but found no significant differences (free big to forced big: $p = 0.1473$, $t(150) = 1.457$; free small to forced big: 0.6338, $t(140) = 0.4775$; free big to forced small: $p = 0.9691$, $t(150) = 0.03874$; free small to forced small: $p = 0.4544$, $t(140) = 0.7501$; forced big to free: $p = 0.8656$, $t(150) = 0.1696$; forced small to free: $p = 0.4579$, $t(150) = 0.7442$). Additionally, we looked at whether perseveration played a role in the increased number of errors. Analysis of the probability of repeating the same action as one, two, three, or four trials back did not reveal significant group

differences (p = 0.2564, t(150) = 1.139; p = 0.4777, t(150) = 0.7119; p = 0.9434, t(150) = 0.07106; p = 0.8331, t(150) = 0.2111; respectively). Taken together, these results show that prior cocaine self-administration generally decreases the ability to perform optimally across choice type transitions.

To test whether neural and behavioral findings might be related we calculated the coefficient of correlation between d' of choice and choice transition errors at different points throughout the recording experiment. A preliminary ANCOVA showed that the slopes of sucrose and cocaine treatments were not significantly different (p = 0.4005, F(1,12) = 0.7599), thus a correlation was calculated from pooled data which showed that d' of choice and choice transition errors were significantly related (*p = 0.0087, r = -0.6316; Fig. 10b). However, comparisons within d' of choice type and choice type transition errors showed significant between group differences (d': p = 0.0360, t(14) = 2.319; transition errors: p = 0.0368, t(14) = 2.308). This suggests that prior cocaine self-administration decreases the strength of choice encoding while increasing behavioral errors.

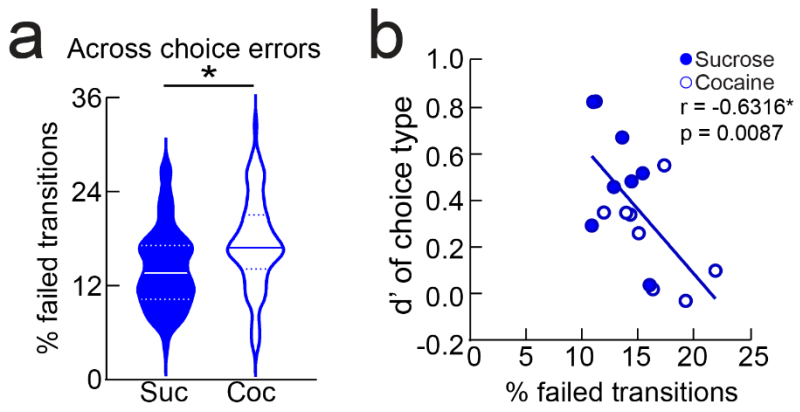

Figure 10. Strength of OFC WS choice representations correlate with choice transition behavior. a) Fraction of between-choice transition errors defined as either (1) an incorrect response on a forced-choice trial following a free choice or (2) choice for the small reward on a free-choice trial following a forced choice. Solid lines indicate medians and dots represent upper and lower quartiles. b) Correlation between OFC WS d' of choice computed using pseudoensembles drawn from sucrose- and cocaine-experienced rats (filled and open circles, respectively) during the odor sample epoch in bins of two recordings sessions across the experiment with choice transition errors averaged over corresponding sessions. Line indicates the best-fit line of the correlation for pooled sucrose and cocaine session data. *p < 0.05.

Discussion

Abnormal decision making is a key feature of SUDs. The ability to appropriately adapt choice behavior in response to changes in the world requires the interlacing of memories about past experiences with updated information about altered contingencies. This can be accomplished by assigning learning in different contexts into compartmentalized "states", whereby the information contained within each state defines the rules appropriate for guiding decision making under those conditions. A growing body of literature suggests that adaptive behavior is facilitated by proper state representations (Bradfield et al., 2013; Schoenbaum et al., 2013; Wilson et al., 2014; Stalnaker et al., 2016). Further, recent work suggests that the abnormal decision making observed in sufferers of SUDs could result from a failure to appropriately represent such states due to drug-induced neural alterations. One previous study found that prior drug experience disrupts state encoding in the DMS (Mueller et al., 2021), whereas separate work showed that the DMS receives inputs from higher order regions involved in decision making such as the OFC and PL (Stalnaker et al., 2016; Sharpe et al., 2019). Together, these findings suggest that drug-induced changes in DMS state encoding may be secondary to changes occurring in the OFC and PL. Here, we recorded OFC and PL neural activity in cocaine-experienced male rats during a decision-making task where blocks of free- and forced-choice trials represented distinct states to examine whether the representation of state information is altered by prior drug exposure. We found that OFC and PL WS units encoded state-related information, albeit in distinct manners, with OFC representing specific free- and forced-choice states within trial blocks and PL more generally representing trial blocks, and that prior cocaine experience altered these representations. Specifically, analysis of the hierarchical representations encoded by OFC and PL WS unit ensembles during the odor sample epoch revealed that prior cocaine exposure changed the format of information represented in the OFC but not the PL. Further, examination of the strength of

information encoded by OFC and PL WS unit ensembles revealed that prior cocaine weakened choice type and direction representations, respectively, and that weaker OFC WS choice type representations were related to more choice type transition errors. These data show that prior cocaine experience alters the format of state-related information being represented when rats are deciding their course of action and suggest a model in which disorganized and inappropriately weak representations of state- related information in the OFC disrupts choice behavior in well-trained subjects.

Individual OFC and PL WS units and ensembles recorded from control rats, which had previously self-administered sucrose and received extensive training on the decision-making task, encoded distinct information during the odor sample epoch as rats were presented with an odor cue and deciding on a course of action. Initial single-unit-level regression analyses looking at the encoding of four single-dimension trial-related variables – choice type, direction, size, and flavor – during the odor sample epoch showed that while the largest proportion of OFC units encoded either choice type or direction, the largest proportion of PL units did not encode choice type but rather only direction. Further analysis at the population ensemble level also revealed choice type encoding in the OFC. RSA looking at the encoding of multiple dimensions of trial-related variables during the odor sample epoch revealed that while both regions represent multi-dimensional hierarchical structures, informational rankings are quite distinct. While the highest tier in the OFC is choice type, with subsequent tiers being organized seemingly based on the demands of whether rats are performing on a free- or forced-choice trial, the structure in the PL is more ambiguous, with the highest tier being location relative to the value block switch, with lower tiers ranked based on combinations of size and opposing outcome identities within a block, with flavor and choice type at the lowest rank. An additional analysis examining the strength of representations during

the odor sample epoch further exemplified OFC choice-type representation. Specifically, during the odor epoch, encoding of choice type became increasingly high and peaked as the presentation of the odor cue ended, as if the odor cue allowed for the evolution of the choice-type information regarding the full rule necessary to guide behavior (i.e., odor 1, forced choice, go left, get big reward, etc.). The same analysis examining the strength of PL representations during the odor sample epoch did not reveal strong encoding of choice type, instead direction was most strongly encoded and seemed to rapidly increase as presentation of the odor was ending. Taken together, these results suggest that when the trial type signaling odor cue is presented, while the OFC encodes very specific information about where the rat is performing in state space (i.e., exact trial defining features of choice type, direction, size, flavor, etc.), the PL encodes more general information about the location (i.e., location relevant to value block switch, etc.), perhaps playing a more supportive background role in the ability of the OFC to define its specific location.

Prior cocaine self-administration had significant effects on the organization and strength of representations across the trial, particularly within the OFC. Regression analyses during the odor sample epoch showed that prior cocaine experience significantly decreased single-unit encoding of both choice type and direction in the OFC, while decreasing the encoding of size in the PL, although the proportion of PL units encoding size was quite small even in the control group. Disrupted unit encoding of single-dimension trial-related variables in the OFC was accompanied by changes in multi-dimensional representations at the population ensemble level. RSA revealed that prior cocaine experience altered the hierarchy of information represented in the OFC, with the dendrogram revealing an ambiguously tiered structure with location relative to value block switch at the top and lower tiers ranked based on combinations of size and opposing outcome identities within a block, choice type, and flavor at the lowest rank. In accordance with

the marginal effect it had on encoding at the level of the individual unit, prior cocaine did not obviously change the hierarchical organization of information represented in the PL with the highest tier still being location relative to the value block switch, with lower tiers ranked based on combinations of size and opposing outcome identities within a block, and flavor and choice type at the lowest rank. Further examination of the effect of cocaine on the strength of representations revealed a significant weakening of choice-type representation in the OFC and direction representation in the PL. These data suggest that prior cocaine experience disrupts both the structure and strength of state-related information encoding, particularly in the OFC.

The finding of choice type representations in the OFC during the odor sample epoch and the weakening of such representations following cocaine self-administration led us to ask whether differences in representational strength were related to behavioral differences. Our behavioral analysis was predicated on the idea that if OFC encoding of choice was behaviorally relevant, then perhaps representation of choice type would be related to task performance when alternating or transitioning between free- and forced-choice trials. Indeed, examination of the relationship between behavior on choice-type transitions and strength of OFC choice-type representations revealed a significant correlation. Furthermore, prior cocaine experience significantly worsened choice-type-transition behavior and decreased the strength of OFC choice-type representations. That is, weakened state-related representations of choice type in the OFC related to impaired value-based choice behavior.

A limitation of the current study is that population vector analyses, made possible by combining all units recorded across all rats and all sessions per region and treatment group, did not allow us to study between-rat differences in neural encoding and how it relates to behavior. To this end, ongoing work using reinforcement learning models to examine drug-naïve rat behavior

on a similar task suggests that there are individual differences in representational learning (Song et al., In prep.), warranting a close examination of individual differences on the task and, further, whether cocaine intake impacts it. More subjects in combination with the ability to record higher numbers of simultaneous units in the future will likely allow for these detailed investigations, perhaps providing finer resolution to the relationship of choice type representations and behavior in control animals and how this relationship is altered by prior cocaine. Another limitation to the study is that simultaneous recordings of the two regions were not performed in the same hemisphere, preventing us from performing the most accurate OFC-PL pairwise analysis. Future recording studies performed within the same hemisphere utilizing genetic tools to identify specific cell populations within each region should be performed to provide the most detailed view of how the OFC and PL and the cell types within them interact during decision making and the potential effect of prior cocaine experience.

The findings of this study, in combination with previous work, suggest a picture of altered state representation in the prefrontal cortex, particularly in the OFC, after cocaine experience. Prior work proposes that the OFC functions to recognize and track specific locations within "hidden" or latent states, likely sending information downstream to the DMS where it maintains real-time representations of the current state so that optimal information can be used to influence behavior (Bradfield et al., 2013; Wilson et al., 2014; Schuck et al., 2016; Stalnaker et al., 2016; Baltz et al., 2018; Stalnaker et al., 2021). By contrast, the PL functions in a distinct but complementary way to the OFC, perhaps not encoding a specific location within the state, but rather a more generic representation of the overall state space (Sharpe et al., 2019). Accordingly, results from this study suggest that the disruption of the hierarchical structure and representational strength of choice type in the OFC following cocaine experience led to greater behavioral errors, whereas the weaker

effect of cocaine on the PL allowed the rats to still make generally appropriate decisions. To draw an analogy, previous and current drug users often exhibit largely normal cognitive and decision-making abilities, except perhaps in specific situations that might be analogous to the OFC-dependent components of the task used here. Since the OFC lies upstream of the DMS, the drug-induced changes in DMS state representations and behavior that were previously observed on a similar task may be secondary to changes in the OFC (Mueller et al., 2021). While drug-induced changes in representational strength were paradoxical, with cocaine weakening encoding in the OFC and strengthening encoding in the DMS, it is possible that this is simply a function of over-compensation in the DMS for the information that has been disorganized or lost upstream. The smaller effect of cocaine on PL representations could be because the task in this study was not heavily PL-dependent. Commonly assessed in attentional set-shifting tasks (ASST), PL function (Dias et al., 1996; Birrell and Brown, 2000) and activity (Rich and Shapiro, 2009) is typically implicated in alternating between extradimensional response strategies. That is, if a task requires subjects to simply attend to the odor cues and adjust behavior when contingencies are reversed, as it does in our odor-guided choice task, then PL function might not be as important, compared to if subjects needed to first attend to the odor cue and then switch to a light cue. In the future, utilizing a task that employs both intra- and extra-dimensional strategy shifts may be better suited to examine the effects of cocaine experience on PL state representations.

Overall, these findings indicate that the neural representations related to adaptive value-based choice behavior in OFC and PL are disrupted by prior cocaine exposure. Cocaine-induced disruptions in the hierarchical organization and strength of state-related encoding may contribute to the dysfunctional decision making observed in SUDs (Jentsch et al., 2002; Schoenbaum and Setlow, 2005; SAMHSA, 2016; Koffarnus and Kaplan, 2018). In the next section, we examined

how similar drug experience affected state encoding in the DMS in a slightly modified version of this odor-guided choice task. The DMS receives inputs from the PL and OFC and is important for behavioral/motor output, representing a likely site at which to examine neural changes arising either from changes in upstream cortical processing or direct drug action, given the importance of local dopaminergic processing in the striatum. We present findings wherein cocaine self-administration produced long-term changes on state-related encoding of two neuronal populations, FSIs and MSNs, in the DMS.

Chapter 3: Effects of Cocaine Experience on DMS Representations During Decision Making

This chapter is adapted from a manuscript published in the *Journal of Neuroscience.*

Mueller LE, Sharpe MJ, Stalnaker TA, Wikenheiser AM, Schoenbaum G (2021). Prior cocaine use alters the normal evolution of information coding in striatal ensembles during value-based decision-making. *Journal of Neuroscience*. 41(2):342-353.

Introduction

Behavioral inflexibility is a core feature of substance use disorders (SUDs), exemplified by persistent drug use despite negative consequences (Association, 2013; SAMHSA, 2016). Within the lab setting, prior drug experience impairs the ability of subjects to adapt behavior following changes in reward contingencies. Rat (Jentsch and Taylor, 2001; Egerton et al., 2005; Stalnaker et al., 2009; Sokolic et al., 2011; McCracken and Grace, 2013; Zhukovsky et al., 2019), monkey (Jentsch et al., 2002; Wright et al., 2013; Seip-Cammack and Shapiro, 2014), and human studies have found that prior drug experience impairs reversal learning, resulting in marked perseveration following the switch (Fillmore and Rush, 2006; Verdejo-Garcia et al., 2007; Ersche et al., 2008; however see Patzelt et al., 2014). Rodents previously exposed to drugs also fail to demonstrate proper contingency degradation or outcome devaluation, suggesting that drugs of abuse disrupt flexible value-based decision-making (Dickinson et al., 2002; Miles et al., 2003; Schoenbaum and Setlow, 2005; Nelson and Killcross, 2006; Corbit et al., 2014; however see Hogarth et al., 2019). Such inflexibility may be related to why animals and humans continue to seek drugs despite adverse consequences (Vanderschuren and Everitt, 2004; Pelloux et al., 2007).

Proper behavioral flexibility, the ability to adjust actions according to changes in environmental contingencies, is optimal if one can properly represent the rules appropriate for each situation, interlacing existing memories with new updated information in changing environments. This can be easily achieved by assigning learning in each context into compartmentalized "states" (Bradfield et al., 2013; Stalnaker et al., 2016). The information contained within each state defines the rules to be recalled for guiding behavior under particular circumstances, and the ability to develop and alternate between states facilitates adaptive responding when circumstances change. Though not often framed this way, the maladaptive behavior observed in sufferers of SUDs may be at least partly the result of a failure to properly encode states due to drug-induced alterations in

the responsible neural mechanisms. That is, drugs may affect the manner in which the brain segregates and generalizes between distinct circumstances that govern which rules should be recalled and applied to guide situationally-appropriate behavior.

The dorsomedial striatum (DMS) is important for adaptive behavior, functioning as a mediator of choice between specific courses of action (Yin et al., 2005; Balleine et al., 2009; Balleine and O'Doherty, 2010; Corbit and Janak, 2010). Prior studies found that interference with DMS activity impairs adaptive behavior by disrupting reversals during discrimination tasks (Ragozzino et al., 2002; Castane et al., 2010), preventing proper outcome devaluation and contingency degradation (Yin et al., 2005; Corbit and Janak, 2010), and obstructing goal-directed learning when action-outcome contingencies change (Bradfield et al., 2013). These results suggest that this region, which is comprised of medium spiny neurons (MSNs), fast-spiking interneurons (FSIs), and cholinergic interneurons (CINs), is a crucial substrate for the encoding of state relevant information that is required for behavioral flexibility. Given that drugs of abuse cause maladaptive behavior, further investigations are needed to determine the role of the DMS as a site affected by drugs of abuse, particularly studies that examine DMS activity in drug-experienced subjects during decision making.

Here, we recorded neurons in the DMS of male rats with prior self-administration experience as they performed a value-guided decision-making task in which the response to obtain the more valuable reward changed across blocks of trials. The trial blocks established unique states since each required a distinct set of rules necessary for optimal performance on the task. We examined whether neural representations of these states and related information about the trials in the DMS would be affected by prior experience self-administering cocaine.

Materials and Methods

General Methods

Subjects

Nine male Long-Evans rats (250-300 g, Charles River Labs), aged approximately three months, were subjects for this experiment. During the odor-guided choice task training and recording sessions rats received water *ad libitum* for 10 minutes per day and *ad libitum* food, during self-administration rats were food deprived to 85% of initial weight with *ad libitum* water. Self-administration and odor task testing were performed during the light phase. All experimental procedures complied with Institutional Animal Care and Use Committee of the US National Institute of Health guidelines.

Surgical methods for jugular catheter and electrode implantation

All surgeries were performed in aseptic conditions as previously described (Stalnaker et al., 2010). Rats were randomly assigned to cocaine or sucrose self-administration groups. Rats used for cocaine self-administration (n =5) received chronic indwelling jugular catheter implants; rats used for sucrose self-administration (n = 4) received sham surgeries in which the jugular vein was exposed but no catheter was implanted. Rats recovered for seven days before self-administration began. During recovery and self-administration, catheters were flushed daily with a cocktail of enrofloxacin and heparinized saline to maintain patency. Following self-administration and training on the decision-making task, recording electrodes composed of driveable bundles of sixteen 25-µm diameter NiCr wires electroplated to an impedance of ~200 kΩ were bilaterally implanted above the DMS (-0.4 mm AP, ±2.6 mm ML, and 3.5-4.0 mm V; relative to bregma and dura). Rats were allowed two weeks to recover, after which they were returned to the decision-making task for one week of reminder training and subsequent recording

sessions. Electrodes were advanced 40 µm following each recording session to capture new neurons.

Self-administration

The self-administration procedure was similar to that described previously (Lucantonio et al., 2014). Rats were trained to self-administer intravenous cocaine-HCl (0.75 mg/kg/infusion; n = 5) or oral sucrose (10%, wt/vol; n = 4) under a fixed ratio 1 schedule in 3 h sessions over 14 consecutive days.

Odor-guided choice task

An adaptation of a previously-described decision-making task was used (Stalnaker et al., 2016). Task training and recording was performed using chambers equipped with two fluid wells and an odor port. Self-paced sessions began with the illumination of a house light, with trials initiated by a nose poke in the odor port. Upon odor port entry, rats held for a 500 ms fixation period, which was followed by a 500 ms presentation of one of three instructional odors that remained constant throughout the experiment. Following odor presentation, rats withdrew from the odor port and indicated their choice by entering either the left or right fluid well within 3000 ms. Upon fluid well entry, rats held for 500 ms before fluid reward delivery began. Once rats consumed the reward and withdrew from the fluid well, the house light was extinguished and the trial ended. If rats withdrew prematurely at any point prior to fluid delivery, the trial was aborted, and the house light turned off. Two instructional odors specified a forced-choice reward delivery at either the left or the right fluid well. The third odor indicated a free choice trial. Presentation of

odors occurred in a pseudorandom sequence with forced-choice right/left odors delivered equally on 65% of trials and free-choice odors delivered on 35% of trials.

Reward outcomes were constant in flavor identity (vanilla milk) but differed in size, small (single 0.05 mL drop) or big (three 0.05 mL drops), with small and big size outcomes delivered at opposing fluid wells. Sessions began with an initial short block and were followed by two blocks of approximately 120 trials each. Response-reward contingencies were consistent for blocks of trials but switched across blocks; block switches were not explicitly signaled and their timing varied randomly to prevent anticipation. Rats were trained before electrode implantation until they performed forced-choice trials with greater than 70% accuracy and completed all blocks. Following electrode implantation and recovery, rats received an additional week of reminder training to increase task performance and acclimation to recording cables.

Single-unit recording

Neural activity was recorded using Plexon Multichannel Acquisition Processor Systems (Plexon Inc.) interfaced with a training chamber. Electrode signals were amplified 20x via operational head stages on electrode arrays and then passed to differential preamplifiers where they underwent an additional 50x amplification and filtration at 150-9000 Hz (Stalnaker et al., 2010; Stalnaker et al., 2016). From there, signals were passed to multichannel acquisition processors where they underwent an additional 250-8000 Hz filtration, 40 kHz digitization, and final 1-32 x amplification. Signal waves were derived from active channels with event timestamps documented by the behavioral program.

Experimental Design and Statistical Analyses

Behavioral epochs

Task trials were divided into five epochs either 0.5 s or 0.8 s in length (Fig. 13a). The fixation epoch began at time of odor port entry, which required rats to hold their snouts in the odor port, and ended after 0.5 s, immediately prior to odor cue delivery. The odor sample epoch began at the time of odor onset and ended after 0.5 s when odor cue delivery ended. Once the odor cue presentation ended, rats chose between the two fluid delivery wells. The movement epoch comprised 0.5 s immediately prior to entry into the fluid delivery well. The anticipation epoch began at the time of fluid delivery well entry and ended after 0.5 s, immediately prior to outcome delivery. The consumption epoch began at the time of outcome delivery and ended after 0.8 s.

Spike sorting and cell type classification

Neural activity was stored and manually sorted into putative single units using Offline Sorter (Plexon Inc.). Files were processed using NeuroExplorer (Nex Technologies) to extract timestamps and further analyzed using MATLAB (The MathWorks, Inc.). Striatal units were classified as a putative fast-spiking interneuron (FSI), medium spiny neuron (MSN), or unidentified (UID), based on three distinct clusters formed in scatter plot of two measurements of mean spike waveform duration: (1) the peak duration of one-half maximum (Sucrose: FSI, 60 - 130 μs, MSN, 105 – 250 μs, UID, 97.5 – 195 μs; Cocaine: FSI, 60 – 142.5 μs, MSN, 97.5 – 250 μs, UID, 77.5 – 215 μs), and (2) the valley to peak duration (Sucrose: FSI, 97.5 – 282.5 μs, MSN, 412.5 – 670 μs, UID, 237.5 – 442.5 μs; Cocaine: FSI, 90 – 272.5 μs, MSN, 410 – 677.5 μs, UID, 170 – 442.5 μs; Figs. 12b - d). Of the 784 units recorded from sucrose-experienced rats, 432 were classified as putative MSNs and 160 were classified as putative FSIs. Of the 1253 single units

recorded from cocaine-experienced rats, 679 were classified as putative MSNs and 307 were classified as putative FSIs.

Neuron firing rate dynamics

To analyze the general response properties of neurons, firing rates for each unit were computed in 50 ms bins, averaged across correctly-performed forced-choice trials in both blocks, and peak normalized. Neurons were counted as maximally active for a trial epoch if, for at least one bin during that period, the cell's peak-normalized average firing rate exceeded 95% of its absolute maximum value (Fig. 13b). Raw, un-normalized population firing rates were computed across all neurons of a given population and trials in 50 ms bins (Fig. 13c).

Classification analyses

Support vector machines (MATLAB function: *fitcecoc*) were used to decode information about outcome direction and size in 100 neuron pseudoensembles. The observation data used to train and test the binary classifier consisted on trial-by-trial spike counts from each neuron during the five distinct trial epochs. Data was limited to the counterbalanced set of correct-performed, forced-choice trials. Recording sessions with a minimum of 30 completed trials of each trial type per block were analyzed. Spike patterns across units included in pseudoensembles were classified into four unique *direction* x *size* categories. Pseudoensembles were generated by randomly pulling a subset of 100 neurons for inclusion from an entire classified population of recorded units across all sessions. Pseudoensemble generation and classification accuracy testing was performed independently for sessions recorded from sucrose- and cocaine-experienced rats. All but one trial was used to fit the classifier (training set), and the remaining trial was used to evaluate

classification accuracy (test set). The random pulling and testing of pseudosensembles was repeated 20 times. Since there was a slight variation in the amount of trials per outcome, to avoid overrepresentation of particular outcomes in the training and testing sets, we controlled for the number of trials for each outcome. To do this, we found the smallest number of trials across all neurons in each pseudoensemble over all of the outcomes and set the trial number to that value. For each pseudosensemble, we performed a shuffling procedure to confirm that accuracy of randomized data is at the theoretical chance level of ¼ correct. To do this we trained classifiers using trial type labels randomly re-assigned to training set firing rates. Classification accuracies were averaged across the 20 runs, comparisons were made between cell types (MSNs vs. FSIs), treatment groups (sucrose- vs. cocaine-experienced), and time (100 ms bins), and effects were analyzed using repeated-measures ANOVAs (Fig. 13d). Confusion matrices show the proportion of test samples classified correctly (along the diagonal) and incorrectly (off of the diagonal) averaged across the 20 repetitions (Figs. 14a, 15a). Confusion matrices were generated and the coefficient of correlation was computed between each matrix and each exemplar pattern. Effects of cocaine on such representational patterns were analyzed using multi-way ANOVAs (Figs. 14b, 15b). Since *trial* emerged as the dominant pattern in matrices generated using pseudoensembles pulled from cocaine-experienced rats, Bonferroni post-hoc tests were performed on *trial* pattern coefficients of the correlation between the treatment groups across each trial epoch. To better visualize the dominant information patterns of confusion matrices, we binarized raw confusion matrices with different thresholds (10% - 80%). For each threshold (e.g., 10%), any value in a confusion matrix that was less than or equal to this threshold was set to zero (black), and other values were set to one (white; Figs. 14c, 15c). To examine how trial and direction information evolved in FSIs over the course of the experiment we generated pseudoensembles by binning every

two recording sessions across the experiment. To do this we pulled subsets of six classified FSIs per treatment group across two consecutive sessions, resulting in 10 bins (i.e., first bin session: sessions 1-2, second bin session: sessions 3-4, etc.). Confusion matrices were created for each of the 10 bin sessions as described above. Coefficients of correlation between each of the 10 bin session matrices with the exemplar trial pattern and the exemplar direction pattern were calculated and plotted over bin sessions (Fig. 16a). Corresponding reaction times were calculated over forced-choice trials by averaging reaction times over the same sessions included in each bin (Fig. 16b).

Statistical Analyses

All analyses were conducted using MATLAB (The MathWorks Inc., Natick, MA) and GraphPad Prism 8 software (GraphPad, La Jolla, CA). One-, two- and three-way ANOVAs, ANCOVAs, and t-tests were used to analyze all data, as reported in results and figure legends. The Bonferroni procedure was used to correct for multiple comparisons. P-values were significant if they fell below 0.05. Sample sizes were determined from previous studies utilizing similar behavioral and recording procedures (Stalnaker et al., 2006; Stalnaker et al., 2016; Wikenheiser et al., 2017). Error bars on plot represent the standard error of the mean.

Results

The experimental timeline is shown in Figure 11a. At the start of the experiment, the rats were shaped on a decision-making paradigm similar to a task used previously in which blocks of trials define states (Stalnaker et al., 2016). For this task (Fig. 11c), one of three different odor cues was presented on each trial, signaling the rat to respond left or right on forced choice trials or in either direction on free choice trials to receive big (three 0.05 mL drops) or small (one drop) amounts of vanilla-flavored milk solution. Response-reward contingencies were constant over

each of three blocks of ~120 trials, switching at unsignaled transitions between each block. Subsequently, male rats were trained to self-administer either oral sucrose (n = 4) or cocaine (n = 5) using a fixed ratio 1 schedule of reinforcement for 3 h per day for 14 days (Figs. 11a and 11b). Two-way repeated-measures ANOVAs revealed significant session × lever interactions in both groups (sucrose: *$p < 0.0001$, $F(13,39) = 16.67$; cocaine: *$p < 0.0001$, $F(13,52) = 13.94$). Rats were then given additional training on the decision-making task, after which drivable bundles of electrodes were bilaterally implanted in the DMS (Fig. 11a). Following recovery, rats received an additional week of reminder training for acclimation to recording cables and then neural recording sessions began (Fig. 11a); electrodes were advanced 40 μm after each recording session to capture new neurons. To control for potentially interacting effects of time and self-administration experience, recording sessions occurred in parallel for sucrose and cocaine groups.

Prior cocaine self-administration disrupted changes in task performance

During the recording sessions, each group of rats demonstrated a strong preference for the big reward on free-choice trials, learning to select the well producing the big reward quickly over the first 20 trials after a block switch (Fig. 11d, left). The rate of this change was slightly but significantly slower in cocaine-experienced rats. A two-way ANOVA on free choice behavior produced significant main effects of trial number (*$p < 0.0001$, $F(6,1134) = 471.8$), treatment group (*$p < 0.0001$, $F(1,1134) = 15.31$, Fig. 11d, left), and interaction (*$p = 0.0072$, $F(6,1134) = 2.959$), and direct comparison of early versus late trials revealed a significant interaction between these two factors (*$p = 0.0075$, $F(1,161) = 7.336$, Fig. 11d, right). Bonferroni post-hoc testing found significant group differences in the second and fourth trial bins following the block transition (*$p < 0.05$; Fig. 11d, left) and on early but not late trials (early: *$p = 0.0077$, $t(322) =$

2.912; late: $p > 0.99$, $t(322) = 0.04409$; Fig. 11d, right). Both groups also exhibited enhanced accuracy on big forced-choice trials; however there was no effect of cocaine on this measure (Fig. 11e). Two-way repeated-measures ANOVAs revealed significant main effects of size (Early: $^*p = 0.0006$, $F(1,161) = 12.21$; Late: $^*p < 0.0001$, $F(1,161) = 150.9$), but no main effects of treatment group or interactions with treatment group (Early, Group: $p = 0.9214$, $F(1,161) = 0.0098$, Int.: $p = 0.0773$, $F(1,161) = 3.161$; Late, Group: $p = 0.7982$, $F(1,161) = 0.06556$, Int.: $p = 0.3192$, $F(1,161) = 0.9985$). Cocaine also caused a general slowing of responding to the fluid well after odor sampling on free-choice trials (Fig. 11f). A three-way repeated-measures ANOVA revealed significant main effects of treatment group and trial early/late (Treatment group: $^*p = 0.0004$, $F(1,161) = 13.16$; Trial early/late: $^*p < 0.0001$, $F(1,161) = 17.84$) but not size ($p = 0.052$, $F(1,161) = 3.769$), with significant interactions of size and treatment group ($^*p = 0.0043$, $F(1,150) = 8.396$) and size and trial early/late ($^*p = 0.0212$, $F(1,161) = 5.414$) but not of treatment group and trial early/late ($p = 0.1656$, $F(1,161) = 1.940$) or a three-way interaction between size and treatment group and trial early/late ($p = 0.2586$, $F(1,150) = 1.286$). Bonferroni post-hoc testing showed that the cocaine group had significantly slower reaction times for both small and big rewards early (Small: $^*p = 0.0274$, $t(311) = 3.328$; Big: $^*p < 0.0008$, $t(311) = 4.244$) and for big rewards late ($^*p = 0.0018$, $t(311) = 4.053$). A similar effect was observed on forced-choice trials (Fig. 11g). A three-way repeated-measures ANOVA revealed significant main effects of treatment group and trial early/late (Treatment group: $^*p < 0.0001$, $F(1,161) = 34.79$; Trial early/late: $^*p < 0.0001$, $F(1,161) = 21.82$) but not size (p 0.7973, $F(1,161) = 0.06617$), with a significant interaction of size and trial early/late ($p = 0.0351$, $F(1,161) = 4.516$) but not of size and treatment group ($p = 0.5072$, $F(1,161) = 0.4419$), treatment group and trial early/late ($p = 0.4344$, $F(1,161) = 0.6140$), or three-way interaction between size and treatment group and trial early/late ($p = 0.2810$, $F(1,161) = 1.170$).

Bonferroni post-hoc testing showed the cocaine group has significantly slower reaction times for both small and big rewards early (Small: *$p < 0.0001$, $t(644) = 4.822$; Big: *$p < 0.0001$, $t(644) = 5.854$) and late (Small: *$p < 0.0001$, $t(644) = 5.064$; Big: *$p < 0.0001$, $t(644) = 4.902$). Notably, fitting correlations to these data revealed that the more self-administration responses cocaine-experienced rats made the slower their forced-choice trial reaction times, whereas the more self-administration responses the sucrose rats made the faster their reaction times (Fig. 11h). ANCOVA revealed significant difference between slopes of best-fit lines (*$p < 0.0001$, $F(1,159) = 53.13$; Sucrose: *$p < 0.0001$, $r = -0.6252$; Cocaine: *$p < 0.0001$, $r = 0.3906$). Furthermore, the effect of cocaine on forced-choice trial reaction time emerged with greater experience on the task during the course of recording; reaction times in the cocaine group failed to show the same decrease that was observed in sucrose-experienced rats suggesting that prior cocaine experience prevents optimization of decision-making (Fig. 11i). ANCOVA revealed significant differences between the slopes of best-fit lines (*$p = 0.0075$, $F(1,159) = 7.341$; Sucrose: *$p < 0.0001$, $r = -0.4506$; Cocaine: $p = 0.7844$, $r = -0.0283$).

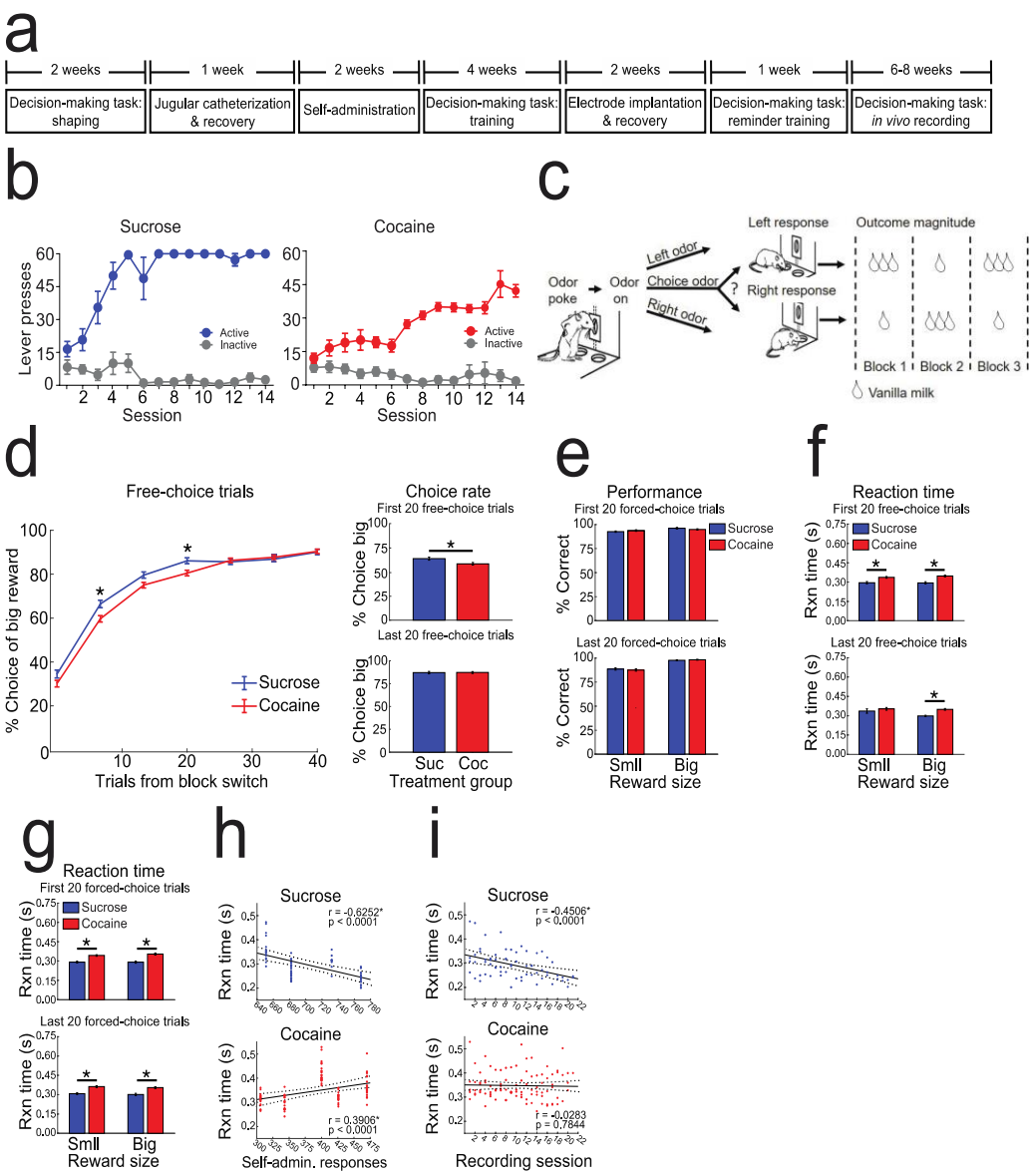
a
2 weeks
1 week
2 weeks
4 weeks
2 weeks
1 week
6-8 weeks
Decision-making task: shaping
Jugular catheterization & recovery
Self-administration
Decision-making task: training
Electrode implantation & recovery
Decision-making task: reminder training
Decision-making task: in vivo recording
b
Sucrose
Cocaine
Lever presses
Active
Inactive
Session
c
Odor poke
Odor on
Left odor
Choice odor
Right odor
?
Left response
Right response
Outcome magnitude
Block 1
Block 2
Block 3
Vanilla milk
d
Free-choice trials
% Choice of big reward
Sucrose
Cocaine
Trials from block switch
Choice rate
First 20 free-choice trials
Last 20 free-choice trials
% Choice big
Suc
Coc
Treatment group
e
Performance
First 20 forced-choice trials
Last 20 forced-choice trials
% Correct
Sucrose
Cocaine
SmlI
Big
Reward size
f
Reaction time
First 20 free-choice trials
Last 20 free-choice trials
Rxn time (s)
Sucrose
Cocaine
SmlI
Big
Reward size
g
Reaction time
First 20 forced-choice trials
Last 20 forced-choice trials
Rxn time (s)
Sucrose
Cocaine
SmlI
Big
Reward size
h
Sucrose
Cocaine
Rxn time (s)
r = -0.6252*
p < 0.0001
r = 0.3906*
p < 0.0001
Self-admin. responses
i
Sucrose
Cocaine
Rxn time (s)
r = -0.4506*
p < 0.0001
r = -0.0283
p = 0.7844
Recording session

Figure 11. Self-administration, task design, and behavior.
a) Experimental timeline. b) Number of active (colored circles) and inactive (grey circles) lever presses during each session of sucrose (left) (n = 4), and cocaine (right) (n = 5) self-administration. c) The decision-making task consisted of two unique blocks of trials defined by distinct response-reward contingencies. Instructional odors delivered to the central odor port indicated which action (go left, go right, go either direction) would be reinforced on each trial. On forced-choice trials, correct responses to the left or the right fluid well were reinforced with fluid delivery. On free-choice trials, responses to either fluid well resulted in fluid delivery. Fluid reinforcers delivered at each well differed in size and alternated following unsignaled block switches as shown in the example block sequence. d) The percent of big rewards chosen on free-choice trials was computed and aligned to block switches (left). Choice rate behavior for early (right, top) and late trials (right, bottom) following block switches. *$p < 0.05$, Bonferroni post-hoc testing. e) Correct performance on forced-choice trials for small (left) and big (right) rewards on early (top) and late trials (bottom) following block switches. f) Reaction times for small (left) and big (right) rewards on early (top) and late free-choice trials (bottom) following block switches. *$p < 0.05$, Bonferroni post-hoc testing. g) Reaction times for small (left) and big (right) rewards on early (top) and late forced-choice trials (bottom) following block switches. *$p < 0.05$, Bonferroni post-hoc testing. h) Correlation of reaction times for the first 20 forced-choice trials of each block and the cumulative number of lever presses for sucrose (top) and cocaine (bottom) self-administration. Solid line indicates best-fit line of the correlation between the reaction time and self-administration responses. Dotted lines show 95% confidence intervals of the best-fit lines. * $p < 0.05$, Bonferroni post-hoc. i) Correlation of reaction times for the first 20 forced-choice trials of each block and the recording session in sucrose- (top) and cocaine-experienced rats (bottom). Solid line indicates best-fit line of the correlation between the reaction time and recording session. Dotted lines show 95% confidence intervals of the best-fit lines. * $p < 0.05$, Bonferroni post-hoc testing.

Time course of DMS MSN and FSI neural responses across treatment groups

To investigate whether prior cocaine experience altered activity of dorsomedial striatal (DMS) neuron populations, units were recorded in the DMS (Fig. 12a). Recordings yielded a total of 784 and 1253 single units from recording sessions performed by sucrose- and cocaine-experienced rats, respectively. Examination of spike waveforms revealed three distinct clusters (Fig. 12b) (Berke et al., 2004; Gage et al., 2010; Gittis et al., 2011), which were proportionally similar across recordings performed by sucrose- and cocaine-experienced rats (Fig. 12c). The largest cluster of cells (sucrose, n = 432; cocaine, n = 679) had long-duration waveforms and low firing rates (Fig. 12c, d), characteristics of MSNs. A second cluster of cells (sucrose, n = 160;

cocaine, n = 307) displayed very brief waveforms and higher firing rates (Fig. 12c, d), features typical of FSIs. These neurons were classified as putative MSNs and putative FSIs, respectively.

Though FSIs are believed to only represent 1% of the total striatal cell population (Luk and Sadikot, 2001), here we found a higher proportion of FSIs comparable to findings of previous reports (Berke et al., 2004; Schmitzer-Torbert and Redish, 2008; Wiltschko et al., 2010; Benhamou et al., 2014; Stalnaker et al., 2016); the distinct waveform and firing properties of FSIs make them easily extracted from other activity. A third cluster of cells (sucrose, n = 192; cocaine, n = 267) had intermediate waveform durations and could not be identified as belonging to a particular population with any certainty (Fig. 12c); these were classified as unidentified (UIDs) and excluded from analysis due to their unknown phenotype. The CINs likely fell within the UID cluster; however, waveform metrics alone proved insufficient to adequately define them.

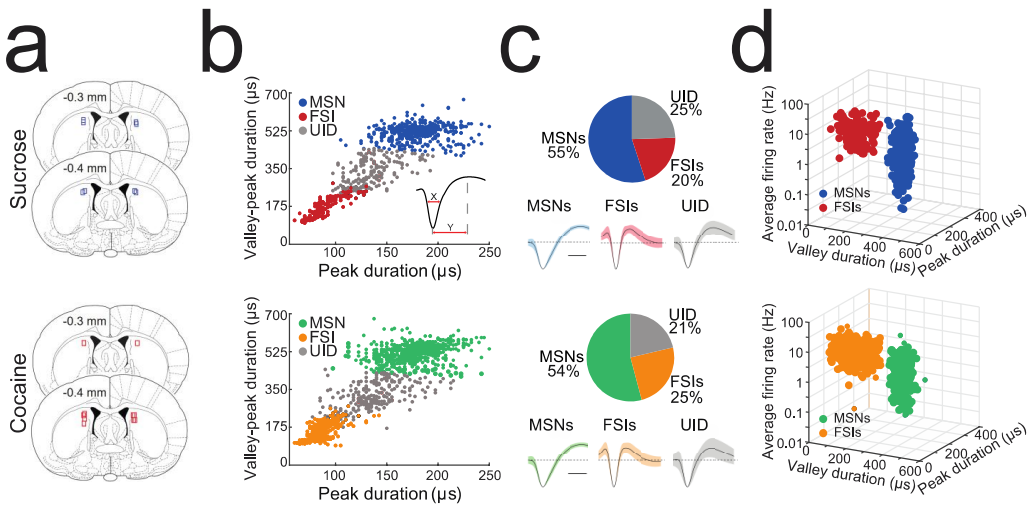

Figure 12. Histology and neuron classification.
a) Approximate locations of neural recordings in the DMS, are indicated by colored boxes (Sucrose: top; Cocaine: bottom). b) Scatter plots of average spike waveform durations (x, peak duration at half maximum; y, valley-peak duration) for each single-unit recorded in sucrose- (top) and cocaine-experienced rats (bottom). MSN, medium spiny neuron; FSI, fast-spiking interneuron; UID, unidentified. c) The approximate proportions and average waveforms superimposed to show inter-wave variability for classified cell types recorded from sucrose- (top) and cocaine-experienced rats (bottom). Dotted line, zero voltage; scale bar, 100 μs. d) Mean peak duration, valley duration, and session-wide firing rate for MSN and FSI populations recorded from sucrose- (Top panel; MSNs, 180 μs, 362 μs, 3.09 Hz; FSIs, 91 μs, 197 μs, 13.48 Hz) and cocaine-experienced rats (Bottom panel; MSNs, 177 μs, 362 μs, 2.33 Hz; FSIs, 90 μs, 177 μs, 10.94 Hz).

To determine whether prior cocaine exposure altered the general response properties of DMS MSNs and FSIs during performance of the decision-making task, correctly performed forced-choice trials were divided into five epochs (Fig. 13a), and units collected throughout the entire recording experiment were sorted according to the time of maximum peak-normalized response (Fig. 13b). This revealed a significant increase in the proportion of MSNs with a peak response during the odor sampling epoch in cocaine- as compared to sucrose-experienced rats (Z-test for population proportions, corrected for multiple comparisons: odor sample, *$p < 0.0001$, $z(1110) = -4.3059$, Fig. 13b, top). Cocaine had no significant effect on the time course of DMS FSI neural responses (Fig. 13b, bottom). Additionally, cocaine significantly decreased the raw, un-normalized population-average firing rates for both MSNs and FSIs (Fig. 13c). Two-way ANOVAs revealed significant main effects of treatment for both cell types (MSN: *$p < 0.0001$; $F(1,62104) = 230.46$; FSI: *$p < 0.0001$; $F(1,26040) = 522.47$) along with significant main effects and interactions with time in the MSNs (time: *$p < 0.0001$; $F(55,62104) = 11.31$; int: *$p = 0.0017$; $F(55,62104) = 1.65$). Bonferroni post-hoc tests revealed between-group differences reaching significance in the reward anticipation and consumption epochs (Fig. 13c, asterisks).

Prior cocaine self-administration disrupted the normal evolution of representations in DMS MSNs and FSIs

To examine the information represented by these populations during the behavioral task we used support vector machines to test how well trial type, defined by particular combinations of outcome size and direction, could be decoded from the firing patterns of DMS MSNs and FSIs over the time course of the trial. Synthetic DMS MSN and FSI pseudoensembles were created by randomly choosing subsets of 100 units with replacement from classified populations of cells

recorded across all sessions; the subsequent analysis used a leave-one-out cross validation procedure to classify 30 trials of each of the four trial types. This was repeated 20 times for each group and cell type.

The results showed that classification accuracy of both MSN and FSI pseudoensembles exceeded the 25% chance level throughout the trial (Fig. 13d). Classification accuracy was above chance even prior to odor delivery, presumably reflecting knowledge of the trials available in the particular trial block. Once the odor was presented, revealing the trial type, classification accuracy improved substantially, remaining high until after the response, when it declined slightly, perhaps reflecting the reduced importance of information about the trial, once the choice was made and the reward was delivered. Classification accuracy was generally higher in FSIs than MSNs and better in neurons from the cocaine than from the sucrose treatment group. Accordingly, a two-way repeated-measures ANOVA revealed significant main effects of cell type ($^*p < 0.0001$, $F(1,81) = 156.05$) and treatment group ($^*p < 0.0001$, $F(1,81) = 35$), as well as a significant interaction between the two ($^*p < 0.0001$, $F(1,81) = 18.56$). One-way repeated-measures ANOVAs comparing decoding within each cell type showed that improved decoding was evident for both MSNs and FSIs after cocaine (MSNs: $^*p = 0.0049$, $F(1,27) = 9.38$; FSIs: $^*p < 0.0001$, $F(1,27) = 61.56$;). Bonferroni post-hoc testing on time bins found that decoding was significantly more accurate for MSNs from the cocaine group during fluid well entry and reward consumption and for FSIs during all epochs except consumption (Fig. 13d, asterisks), suggesting that prior cocaine experience increases the amount of trial-related information being encoded by pseudoensembles.

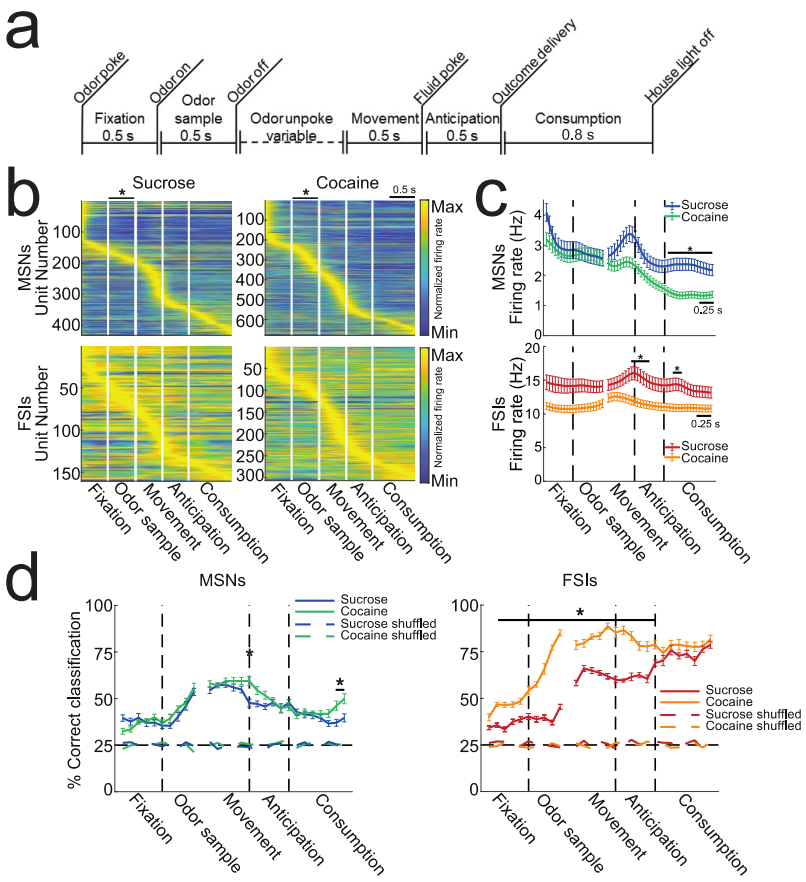

Figure 13. Time course of DMS MSN and FSI neural responses across treatment groups. a) Correctly performed forced-choice trials were divided into five epochs bounded by important task events. b) Peri-event time histograms (PETHs) of peak-normalized firing rates in MSNs (top) and FSIs (bottom) recorded from sucrose- (left) and cocaine-experienced rats (right) aligned to each of the five task epochs. Each row shows the firing rate of a single cell using a color scale ranging from blue (zero) to yellow (peak) with time on the x-axis (each panel = 0.5 s or 1 s; bin size was 50 ms). *$p < 0.05$, z-test for maximally active population proportions per epoch corrected for multiple comparisons. c) PETHs of raw, un-normalized population-average firing rates in MSNs (top) and FSIs (bottom) recorded from sucrose- and cocaine-experienced rats aligned to each of the five task epochs. Bin size was 50 ms; error bars indicate SEM. *$p < 0.05$, Bonferroni post-hoc testing. d) Proportion of test samples correctly classified from each trial type using pseudoensembles of 100 MSNs (left) or FSIs (right) randomly selected from units recorded across all sessions from sucrose- and cocaine-experienced rats compared to respective shuffled data (dotted lines). Dashed horizontal lines indicate chance levels for classification. Bin size was 100 ms; error bars indicate the SEM in classification accuracy. *$p < 0.05$, Bonferroni post-hoc testing of classification performance between pseudoensembles drawn from sucrose- and cocaine-experienced rats.

To gain insight into the informational basis of this improved decoding, we compared changes in the patterns of classification over trial epochs in each group and cell type using

confusion matrices, which illustrate how trials were classified (or misclassified). The pseudoensembles used in this analysis were drawn from classified units recorded across all sessions. The overall pattern of classification is informative because each trial was contained within a block in which two different size rewards were available following the execution of two different responses. Encoding of any combination of this information could support above chance classification accuracy (Fig. 13d), however each gives rise to different patterns of classification when plotted in a confusion matrix (Figs. 14 and 15, exemplar patterns).

A comparison of the actual classification patterns across epochs in pseudoensembles recorded from sucrose-experienced rats showed that the MSNs and FSIs followed a characteristic trajectory through each trial. This started with the fixation epoch, in which both MSN and FSI pseudoensembles showed a pattern consistent with encoding of trial block – that is, the most dominant pattern of encoding during fixation, evident in the raw, line, and filtered confusion plots (Figs. 14a-c and 15a-c), was most similar to that of *block* (Figs. 14 and 15, block pattern). During the odor sampling epoch, the *block* and diagonal *trial* pattern were both evident, after which the checkerboard *direction* pattern became most prominent during the movement and anticipation epochs, followed by a return to the diagonal *trial* pattern during reward consumption.

To quantify this evolution, the coefficient of correlation was computed between the confusion matrices and each exemplar pattern (Figs. 14b and 15b). A comparison confirmed the trajectory described above and identified significant changes as a result of cocaine use. Three-way ANOVAs comparing patterns within each cell type across trial epochs revealed significant main effects of treatment group (MSNs: $*p = 0.0141$, $F(1,760) = 6.06$; FSIs: $*p < 0.0001$, $F(1,760) = 155.04$), trial epoch (MSNs: $*p < 0.0001$, $F(4,760) = 35.64$; FSIs: $*p < 0.0001$, $F(4,760) = 140.61$), and pattern (MSNs: $*p < 0.0001$, $F(3,760) = 160.26$; FSIs: $*p < 0.0001$, $F(3,760) = 420.49$), as

well as significant interactions between treatment group and trial epoch (MSNs: *p = 0.0001, F(4,760) = 6.07; FSIs: *p = 0.0058, F(4,760) = 3.66), treatment group and pattern (MSNs: *p = 0.0405, F(3,760) = 2.77; FSIs: *p < 0.0001, F(3,760) = 46.38), trial epoch and pattern (MSNs: *p < 0.0001, F(12,760) = 38.76; FSIs: *p < 0.0001, F(12,760) = 124.29), and significant three-way interactions between treatment group, trial epoch and pattern (MSNs: *p < 0.0001, F(12,760) = 5.56; FSIs: *p < 0.0001, F(12,760) = 4.75). Two-way ANOVAs comparing *trial* patterns within each cell type across epochs revealed that representations were altered after cocaine (MSNs: treatment group, p = 0.1302, F(1,38) = 2.393; trial epoch, *p < 0.0001, F(4,152) = 18.94; int., *p = 0.0143, F(4,152) = 3.223; FSIs: treatment group, * p < 0.0001, F(1,38) = 71.05; trial epoch, * p < 0.0001, F(4,152) = 54.92; int., *p = 0.0034, F(4,152) = 4.113). Consistent with subtle differences between treatment groups in MSN classification accuracy (Fig. 13d, left), the only difference in MSNs from the cocaine group was stronger encoding of trial type information during the anticipation epoch (*p = 0.0045, t(190) = 3.376, Figs. 14a-c, "Anticipation"). By contrast, pseudoensembles composed of FSIs, which exhibited much larger differences in classification accuracy (Fig. 13d, right), showed prominent changes in classification patterns in the cocaine group, exhibiting significantly stronger representation of trial type across all epochs except consumption (Fixation: *p = 0.004, t(190) = 4.017; Odor sample: *p < 0.0001; t(190) = 4.736; Movement: * p = 0.0146, t(190) = 3.014; Anticipation: *p < 0.0001, t(190) = 6.976; Figs. 15a-c).

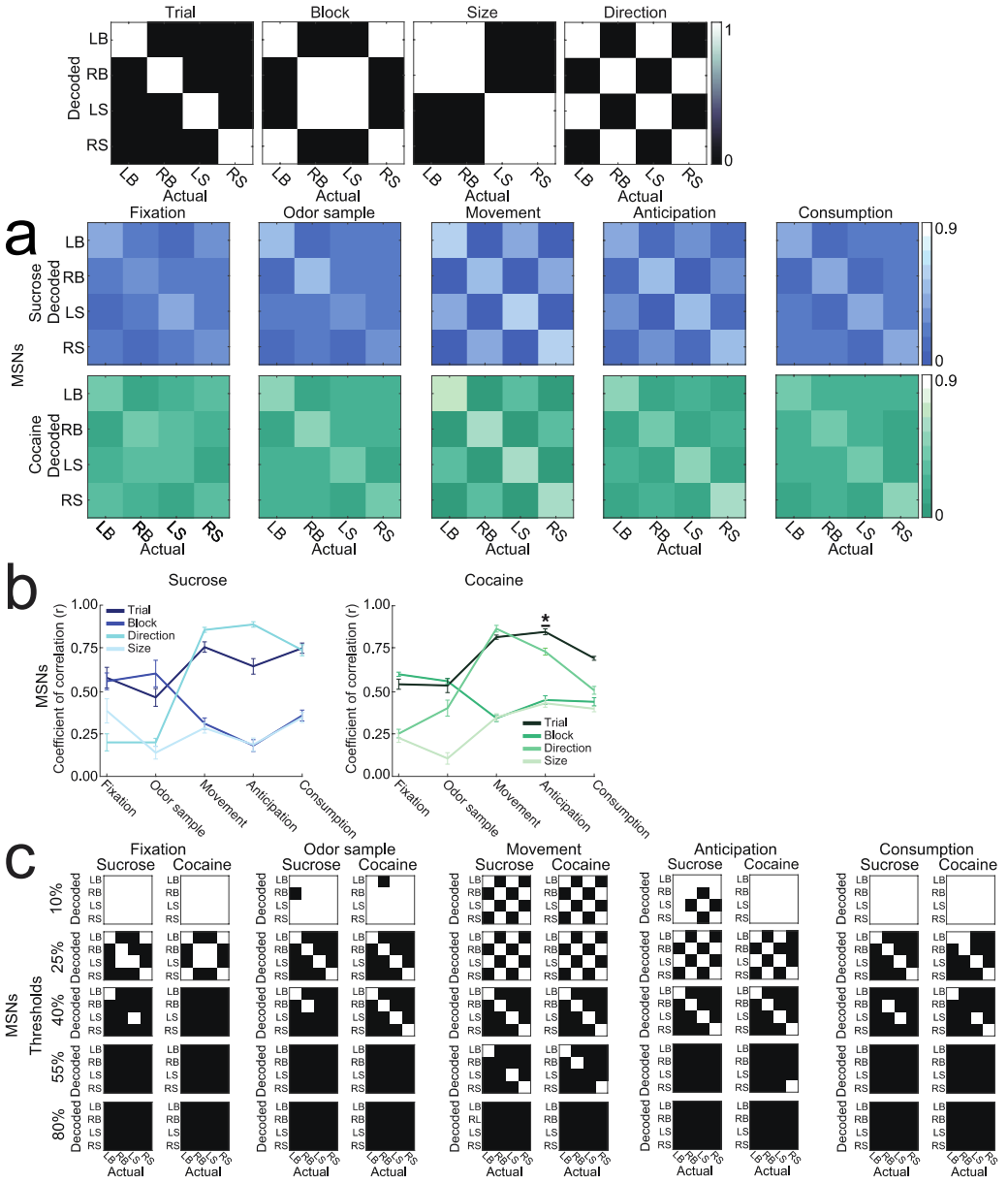
Trial
Block
Size
Direction
Decoded
LB
RB
LS
RS
Actual
1
0
a
Fixation
Odor sample
Movement
Anticipation
Consumption
Sucrose Decoded
Cocaine Decoded
MSNs
0.9
0
b
Sucrose
Cocaine
Coefficient of correlation (r)
1.00
0.75
0.50
0.25
0.00
Trial
Block
Direction
Size
Fixation
Odor sample
Movement
Anticipation
Consumption
c
Sucrose
Cocaine
MSNs
Thresholds
10%
25%
40%
55%
80%

Figure 14. Evolution of representations in DMS MSNs.
Confusion matrices illustrate the four exemplar representations ("Trial", "State", "Size", and "Direction"). The y axis represents the ground-truth four trial types, ordered by direction and size (LB = left-big; RB = right-big; LS = left-small; RS = right-small). The x axis represents how these four predicted trial types would be classified according to hypothesized information (e.g., trial, state, size or direction). a) Confusion matrices depict the proportion of test samples correctly classified (along the diagonal) and incorrectly (off the diagonal) during the five trial epochs using pseudoensembles of 100 categorized MSNs in sucrose (top) and cocaine sessions (bottom). b) Coefficient of the correlation between exemplar patterns and raw MSN confusion matrices generated using pseudoensembles drawn from sucrose- (left) and cocaine-experienced rats (right) during each trial epoch. Error bars indicate SEM in coefficient of the correlation. *$p < 0.05$, Bonferroni post-hoc testing revealed higher coefficient of the correlation of trial pattern with raw confusion matrices generated using pseudoensembles drawn from cocaine- than sucrose-experienced rats. c) Raw MSN confusion matrices generated for each epoch were filtered at graded thresholds (10-80%) and binarized to extract patterns of representations (values exceeding thresholds are colored white and the rest black).

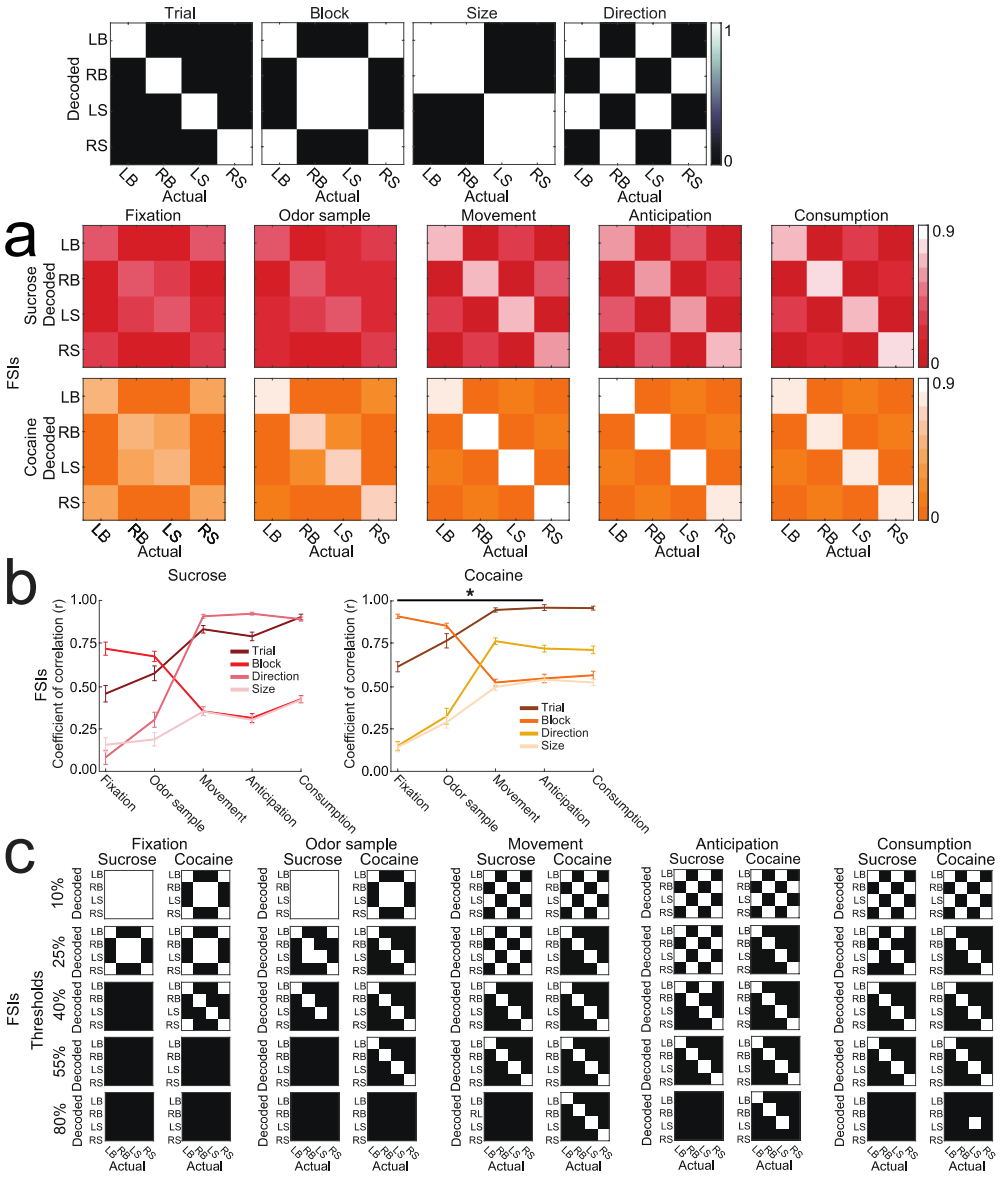

Figure 15. Evolution of representations in DMS FSIs.
Confusion matrices illustrate the four exemplar representations (See Fig. 4 legend). a) Confusion matrices depict the proportion of test samples correctly classified (along the diagonal) and incorrectly (off the diagonal) during the five trial epochs using pseudoensembles of 100 categorized FSIs in sucrose (top) and cocaine sessions (bottom). b) Coefficient of the correlation between exemplar patterns and raw FSI confusion matrices generated using pseudoensembles drawn from sucrose- (left) and cocaine-experienced rats (right) during each trial epoch. Error bars indicate SEM in coefficient of the correlation. *$p < 0.05$, Bonferroni post-hoc testing revealed higher coefficient of the correlation of trial pattern with raw confusion matrices generated using pseudoensembles drawn from cocaine- than sucrose-experienced rats. c) Raw FSI confusion matrices generated for each epoch were filtered at graded thresholds (10-80%) and binarized to extract patterns of representations (values exceeding thresholds are colored white and the rest black).

Prior cocaine's effects on representations in DMS parallels effects on task performance

The above analysis suggests that prior cocaine self-administration paradoxically results in improved representation of specific information about the unique trial types. At the same time, cocaine-experienced rats showed deficits in switching their behavior early in trial blocks after rewards changed compared to sucrose-experienced rats, and failed to become faster at performing the task over the course of the entire recording experiment, an effect which was observed in sucrose-experienced rats. To test whether these neural and behavioral findings might be related, we calculated the coefficient of correlation between the FSI confusion matrices from the movement epoch at different points throughout the recording experiment and the exemplar direction and trial patterns, which were the dominant patterns during the movement epoch in FSIs recorded from sucrose- and cocaine-experienced rats, respectively. This analysis showed that, in pseudoensembles drawn from sucrose-experienced rats, information related to direction increased and information related to trial decreased as the recording experiment progressed (Fig. 16a, top), whereas the representation of information about trial and direction did not change in pseudoensembles drawn from cocaine-experienced rats (Fig. 16a, bottom). ANCOVAs supported these observations, revealing significant differences between the slopes of best-fit lines in FSIs recorded from sucrose- but not cocaine-experienced rats (Sucrose: $*p < 0.0001$, $F(1,16) = 44.79$; Cocaine: $p = 0.5689$, $F(1,16) = 0.3383$; best-fit lines for sucrose rats: Direction: $*p = 0.0005$, $r = 0.8915$; Trial: $*p = 0.0054$, $r = -0.8004$; best-fit lines for cocaine rats: Direction: $p = 0.1953$, $r = 0.4469$; Trial: $p = 0.0942$, $r = 0.5573$). Consistent with the relationship between recording session number and reaction time (Fig. 11), we found similar results when the neural data was plotted against forced-choice reaction time. ANCOVAs revealed significant differences between the slopes of best-fit lines in FSIs (Sucrose: $*p = 0.0002$, $F(1,16) = 22.11$; Cocaine: $p = 0.5877$, $F(1,16)$

= 0.3061), while forced-choice reaction times were negatively correlated with the representation of direction and positively correlated with representation of trial in sucrose-experienced rats (Direction: *p = 0.0065, r = -0.7903; Trial: *p = 0.0087, r = 0.7736; Fig. 16b, top), they were not significantly correlated in cocaine-experienced rats (Direction: p = 0.5715, r = -0.2042; Trial: p = 0.9698, r = 0.0138; Fig. 16b, bottom). This suggests that a shift from trial to direction information leads to faster reaction times in sucrose-experienced rats, and that prior cocaine experience prevents this shift from occurring over the course of the recording experiment, ultimately blocking optimization of decision-making in rats.

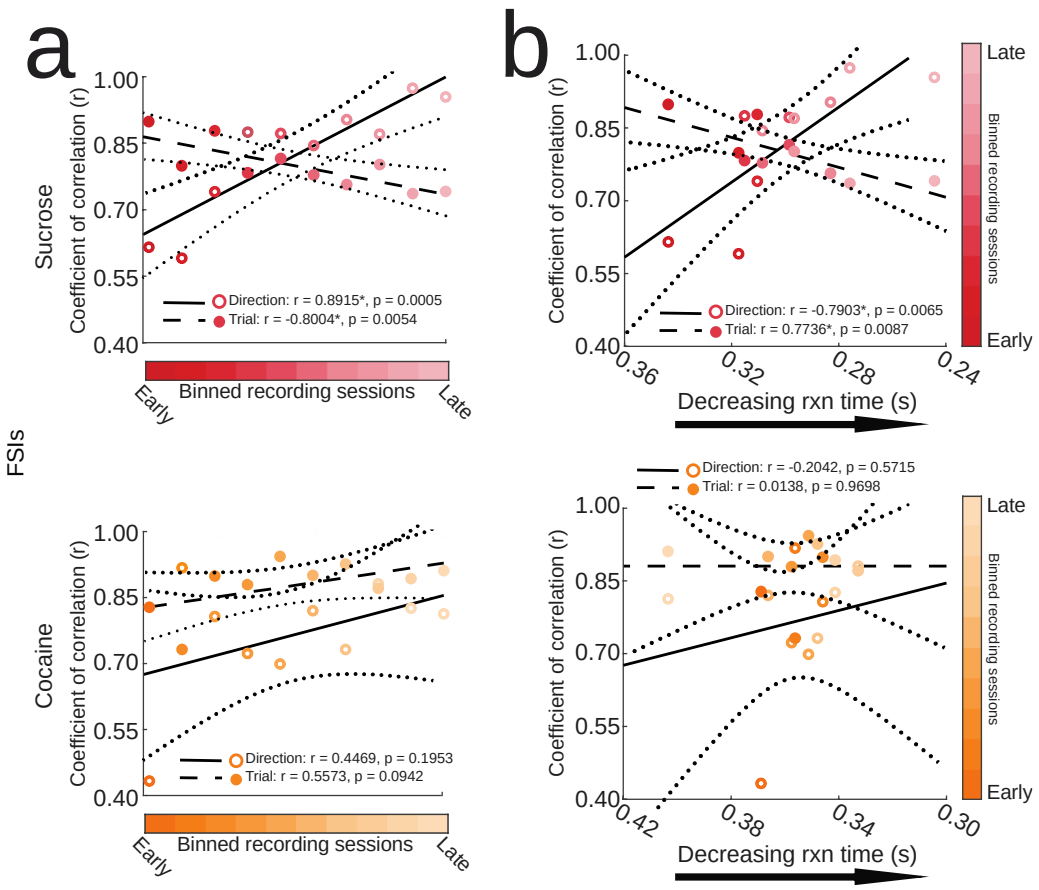

Figure 16. Evolution of representations in DMS FSIs across recording sessions correlates with behavior.
a) Coefficient of the corrleation between raw FSI confusion matrices generated using pseudoensembles drawn from sucrose- (top) and cocaine-experienced rats (bottom) during the movement epoch in bins of two recording sessions across the experiment and examplar direction (open circles) and trial patterns (closed circles). Solid line indicates best-fit line of the correlation between the coefficient of correlation with direction pattern and recording session. Dashed line indicates best-fit line of the correlation between the coefficient of correlation with trial pattern and recording session. Dotted lines show 95% confidence intervals of the best-fit lines. b) Correlation of the coefficients of the correlation between raw FSI confusion matrices generated using pseudoensembles drawn from sucrose- (top) and cocaine-experienced rats (bottom) during the movement epoch in bins of two recording sessions across the experiment with exemplar direction (open circles) and trial patterns (closed circles) with forced-choice reaction times averaged over corresponding sessions. Solid line indicates best-fit line of the correlation between the coefficient of correlation with direction pattern and forced-choice reaction times. Dashed line indicates best-fit line of the correlation between the coefficient of correlation with trial pattern and forced-chocie reaction times. Dotted lines show 95% confidence intervals of the best-fit lines.

Discussion

SUDs are characterized by behavioral inflexibility. Though not often conceptualized this way, such an effect may be the result of a failure to appropriately interlace existing memories or rules with new information about the changed environment, due to drug-induced changes in the neural mechanisms governing the compartmentalization of context or "states". Thus, drugs may affect the way the brain separates and generalizes between different circumstances that govern which rules should be recalled to appropriately guide behavior. Previously, neurons in DMS, particularly CINs, were found to compartmentalize state information and encode real-time representations of state, the ability of which depended on orbitofrontal cortex (OFC) input (Bradfield et al., 2013; Stalnaker et al., 2016). However, whether other DMS populations encode state and state-relevant information and how these may be affected by cocaine experience was unknown. Here, we recorded neural activity in DMS in cocaine-experienced rats during a decision-making paradigm where blocks of trials represented distinct states in order to test whether the

compartmentalization of such information might be affected by drug use. We found that pseudoensembles of DMS MSNs and FSIs encode information relevant to state and that prior cocaine experience disrupts the evolution of such representations, suggesting that prior cocaine experience does alter DMS state and rule encoding. While addiction is a complex multi-hit process and these results cannot account for all of its many facets, these findings describe early effects of drug exposure on decision making that may facilitate the progression to SUD.

MSN and FSI pseudoensembles recorded from control rats, which had previously self-administered sucrose and received extensive training on the decision-making task, carried very similar information. Closer examination of the type of information being encoded in DMS in these well-trained rats revealed an interesting evolution of these representations over the course of the trial, which reflected the information logically required to guide behavior most efficiently in each epoch. During the fixation epoch, prior to presentation of the informative odor cue, pseudoensembles encoded the trial block, suggesting that rats were maintaining information about the current trial block or state, preparatory to the trial. During the odor sample epoch, when the informative cue was presented, pseudoensembles represented both state and trial information, as if the odor cue allowed encoding to shift from state to trial-specific information regarding the full rule necessary to guide behavior (i.e. odor 1, go left, get big reward). During the movement and anticipation epochs, as the rats physically responded to and waited in the fluid delivery well, the pseudoensembles in controls rapidly switched to encoding only direction, shedding information irrelevant to either the physical response or location. However, during the consumption epoch, pseudoensembles again encoded the trial type, which was signaled by the presence of the particular outcome in a given well. Additionally, over the course of the recording experiment, reaction times decreased in controls, and this decrease was associated with a decline in the representation of trial-

specific information and an increase in the representation of direction, particularly in pseudoensembles composed of FSIs. Thus, optimization of behavior appeared to reflect the pruning or minimization of information beyond response direction in these neurons.

Prior cocaine self-administration had significant effects on the evolution of these representations, both within trials and across recording sessions. Within individual sessions, MSN and FSI pseudoensembles recorded from cocaine-experienced rats exhibited stronger trial-specific representations compared to those recorded from sucrose-experienced rats. While prior cocaine self-administration did not alter MSN and FSI pseudoensemble encoding appreciably during the fixation epoch, trial information was the primary information represented in all subsequent epochs, dominating the raw confusion matrices and persisting at higher filtering thresholds in binarized matrices. The over-encoding of trial-specific information persisted throughout the recording experiment, relative to pseudoensembles recorded from sucrose-experienced rats, which switched to simpler response representations during some epochs. This persistence was associated with slightly slower switching of choice behavior at the start of blocks throughout experimental sessions, and a failure to develop faster reaction times on the task over multiple sessions, as if the stronger representations were slowing normal flexibility and responding.

This drug-induced over-coding of trial type was most prominent in FSI pseudoensembles, which displayed representations of specific information about trial type throughout each trial in cocaine-experienced rats. These findings are consistent with other studies that suggest a role for FSIs in information processing during the performance of striatum-dependent behaviors. Previously striatal FSIs have been reported to increase firing during choice execution on a striatum-dependent sequence learning task, activity proposed to suppress prepotent but situationally-inappropriate responses (Gage et al., 2010). Further, FSIs are thought to be important

for the acquisition of striatum-dependent action selection strategies and that impairments in FSI signaling result in learning deficits which can be overcome with prolonged training (Lee et al., 2017; Owen et al., 2018). These ideas nicely complement our results.

A limitation of the current study is that pseudoensembles were formed from units recorded from different subjects, which did not allow us to study between-rat differences in the effects of cocaine intake on neural encoding. The ability to record higher numbers of simultaneous units in the future will likely allow for these detailed investigations. Moreover, while our classifications of putative MSNs and FSIs are supported by previous literature (Berke et al., 2004; Gage et al., 2010; Gittis et al., 2011), future studies utilizing genetic tools to identify specific cell populations should be performed to confirm and expand upon our findings, as well as clarify the role of CINs in modulating other DMS populations.

The results of this study, added to prior work, suggest a picture of altered state and rule encoding in DMS after cocaine experience. Specifically, although we were not able to identify a population of CINs in the current study, prior work suggests this population is actively engaged in regulating proper representation of the current state, aiding the recall of the rules to guide behavior in settings such as ours (Stalnaker et al., 2016). Failure of this function led to behavioral confusion between shifting rules (Bradfield et al., 2013). These state representations have been found to depend on both thalamic and orbitofrontal (and likely prefrontal) input (Bradfield et al., 2013; Stalnaker et al., 2016). Given that cocaine exposure similar to what rats experienced here disrupts orbitofrontal function (Jentsch et al., 2002; Schoenbaum et al., 2004; Stalnaker et al., 2006; Calu et al., 2007; Porter et al., 2011; Lucantonio et al., 2012), one possible explanation of the current results lies in dysregulation of this mechanism for adaptive behavior. Specifically, if proper state representation by the OFC-CIN circuit were altered, perhaps becoming too rigid and less able to

recognize and respond to state changes, then this might result in the observed effects. For example, slower switching behavior on free-choice trials at the start of new blocks would be a logical result of a failure of neural representations to recognize or flexibly adapt to the new state, and the failure to develop faster and more efficient responses on forced choice trials could also reflect an over-representation of trial-type specific information instead of more general information about response direction. Interestingly, this idea accords with the proposal that a core contribution of the OFC is to recognize and track such "hidden" or latent states (Wilson et al., 2014; Schuck et al., 2016; Stalnaker et al., 2016; Baltz et al., 2018). Notably, both changes were associated with persistent representation of highly specific information about the trial type as defined by the odor cue, which was only briefly encoded by DMS neurons recorded from sucrose-experienced rats. While these changes were most prominent in FSIs, they were also apparent in MSNs to a lesser degree, and thus could be impacting DMS output and behavior.

Overall, these results indicate that neural representations related to adaptive value-based decision-making in DMS are altered by prior cocaine experience; cocaine-induced alterations in the evolution of encoding in DMS cell populations may contribute to the behavioral inflexibility observed in individuals suffering from SUDs (Jentsch et al., 2002; Schoenbaum and Setlow, 2005). Moreover, while direct drug action in the DMS could account for some of these changes, these results also link drug-induced changes in higher cortical areas like the OFC and PL to changes in behavior through the output hub of the DMS. These and other considerations as well as limitations of both Chapters 2 and 3 are further presented in Chapter 4.

Chapter 4: General Discussion

Summary of findings

Decision making requires the ability to adapt actions according to changes in environmental contingencies, the optimization of which can be achieved by properly representing the rules appropriate for each situation, interlacing existing memories with new updated information in changing environments. Organisms can achieve this by assigning learning in each context into compartmentalized "states" (Bradfield et al., 2013; Stalnaker et al., 2016). The associative information contained within each state defines the rules and information to be recalled for optimally guiding behavior under particular circumstances, and the ability to develop and transition between states facilitates adaptive decision making. Though not often conceptualized in this manner, the maladaptive decision making that is a hallmark of SUDs may be the result of an impairment to properly encode such representations that support decision making due to drug-induced alternations in the responsible neural mechanisms. That is, drugs may alter the way in which the brain segregates and generalizes between distinct circumstances that govern which rules should be recalled and applied to guide situationally-appropriate choices. The experiments included in this dissertation investigated whether chronic drug experience altered representations that support choice behavior in three brain regions previously shown to be involved in various aspects of the decision-making process.

The study presented in Chapter 2 used behavioral and in vivo electrophysiological approaches to examine the effect of prior cocaine experience on representations supporting decision making in the OFC and PL. To examine whether the encoding of such state-related information was altered by prior drug exposure, single unit activity was recorded in the OFC and PL of cocaine-experienced rats during a decision-making task where blocks of free- and forced-choice trials characterized experimenter-defined states. Overall, the experiment found that prior cocaine experience disrupted the ability of the OFC to represent *cognitive*-based state-related

information (e.g., choice type), that is, subjects' specific location within the task state (i.e., free- and forced-choice trials), but did not disrupt the more general representations of state by the PL. Analysis of the structural hierarchy of information encoded by these cortical regions during the odor sample epoch showed that prior cocaine experience significantly altered the organization of the state-related representation in the OFC while resulting in only minimal changes in the PL. Furthermore, investigation of the strength of representations encoded by these regions during the same epoch showed that cocaine experience weakened choice type and direction information, respectively, and that weaker OFC choice-type representations by wide-spiking (WS) neurons were related to increased behavioral errors during decision making. Together, the study presented in Chapter 2 shows that prior cocaine experience alters the format of state-related information being represented when rats decide their course of action and suggest a model in which disorganized and inappropriately weak representations of *cognitive*-based state-related information in the OFC disrupts choice behavior in well-trained subjects. These findings, in combination with previous work showing that DMS state correlates are OFC input-dependent (Stalnaker et al., 2016), suggest that prior cocaine experience disrupts representations supporting choice behavior in the DMS, which may, in turn, account for the changes in behavior. The potential effect of cocaine on DMS encoding of representations was thus explored in Chapter 3.

Chapter 3 also used behavioral and in vivo electrophysiological approaches to investigate whether representations supporting decision making in the DMS were altered following cocaine. To examine whether the encoding and compartmentalization of state-related information was altered by prior drug exposure, single unit activity was recorded in the DMS of cocaine-experienced rats during a decision-making paradigm, slightly modified from the paradigm used in Chapter 2, where blocks of trials represented experimenter-defined states. Overall, we found that

prior cocaine experience also disrupted the ability of DMS MSNs and FSIs to properly encode state-related information, causing significant changes in the evolution of DMS MSN and FSI *action*-based state-related representations (e.g., direction), both within trials and across recording sessions. Specifically, within recording sessions, MSN and FSI pseudoensembles exhibited stronger trial-specific representations compared to those recorded from sucrose-experienced rats. The over-representation of trial-specific information persisted throughout the recording relative to pseudoensembles recorded from sucrose-experienced rats, which switched to simpler response representations. The over-encoding of state-related information was associated with slower switching of choice behavior for the large reward, and a failure to develop quicker reaction times on the task, suggesting that prior cocaine exposure results in abnormal state representations that prevent optimal behavior on a decision-making task.

Together, these data suggest a model in which higher-order cortical regions encode *cognitive*-based state-related representations (e.g., choice type) that possibly get relayed downstream where information is translated into *action*-based state-related representations (e.g., direction) before being outputted for behavior (Sharpe et al., 2019). Our finding that the OFC encodes highly-structured state-related information agrees with findings from several studies, supporting that the role of the OFC during decision making is to track the current location within the cognitive map of the current task space (Wilson et al., 2014). A prior human neuroimaging study utilizing pattern classification on data collected during a 16-state decision-making task found that OFC activity decoded unobservable task states and that the accuracy of decoding was related to decision-making performance. Additionally, behavioral accuracy following state transitions correlated with the similarity between the neural representations of those states (Schuck et al., 2016). Previous work in rats showed that following the acquisition of spatial-specific object-

reward associations, OFC neural ensembles encode value-based schemas that represented objects in the context and location where they are associated with reward or non-reward. Similarly, OFC neural ensembles signal distinct spatial contexts that delineate the mapping of specific stimuli to actions and outcomes in a highly-structured fashion (Farovik et al., 2015; McKenzie et al., 2016). Our recent work in combination with these data support the idea that the OFC provides an essential representation of the current state of the environment which guides outcome-oriented actions.

Our finding that the DMS encodes response-based state-related information agrees with previous findings suggesting that the region maintains representations important for guiding behavior on a decision-making task in an OFC-dependent manner. Prior rat work found that the DMS encodes the associative information required for optimal behavior, functioning as a mediator of choice between specific courses of action (Yin et al., 2005; Balleine et al., 2009; Balleine and O'Doherty, 2010). Additionally, a rat behavior study implicated neurons within the DMS in the specific organization of associative information and have been shown to track information about outcomes that could be used to guide appropriate decision making (Bradfield et al., 2013). Furthermore, a single-unit recording study in rats found that units within the DMS represent state in real-time, dependent on OFC-input (Stalnaker et al., 2016). Our work, in addition to these data, suggest that the DMS uses input from the OFC to encode action-based state-related information and that such representations guide behavior.

Our recent data also suggest a model in which the maladaptive decision making observed in SUDs could be, in part, related to drug-induced changes in the brain's ability to encode state-related information due to changes in prefrontal-striatal circuits. Specifically, we showed that prior cocaine experience alters the structure and strength of information represented in the OFC during decision making. Furthermore, since the DMS encodes state-related information in an OFC-

dependent manner (Stalnaker et al., 2016), the alterations in DMS state representations following cocaine experience could be secondary to changes in OFC representations. Our work is consistent with previous studies that found that drugs of abuse cause persistent impairments in the ability of prefrontal regions to function properly (Jentsch and Taylor, 1999; Volkow and Fowler, 2000; Homayoun and Moghaddam, 2006; West et al., 2014; Gobin et al., 2019). Cocaine exposure was found to impair response latency during discrimination learning and slower acquisition of serial reversals, mirroring findings in OFC-lesioned animals (Graham JH, 1995). In a recording study, a discrimination task was used to show that OFC correlates in rats with cocaine histories failed to encode expected outcomes. While OFC neurons in both control and cocaine-exposed animals fired for the expectation of outcome during learning, only cells in control animals exhibited selective activation by predictive odors after originally learned contingencies were reversed, whereas neurons in cocaine rats fired non-selectively (Stalnaker et al., 2007). These findings, in addition to our recent work, suggest that prior cocaine self-administration disrupts state-related neural representations in the OFC. Nonetheless, these findings do not preclude a potentially important effect of direct drug action in the DMS. Prior studies found that interference with DMS activity impairs adaptive behavior by impairing reversals in discrimination tasks, outcome devaluation and contingency degradation, and goal-directed learning when action-outcome contingencies change (Ragozzino et al., 2002; Yin et al., 2005; Castane et al., 2010; Corbit and Janak, 2010; Bradfield et al., 2013). These findings clearly indicate that the DMS itself is crucial for tracking state-related information required for behavioral flexibility, perhaps offering a more output-proximal region for drug exposure to induce a pathological behavioral phenotype.

Limitations

One great limitation, as with many in vivo electrophysiological studies, is that the results provided by these studies are merely correlative, not causal. For instance, the neural and behavioral deficits do not necessarily indicate that one was caused by the other. Moreover, even though the cortical regions we studied both project to the DMS, it is still unclear whether the any cortical deficits did, in fact, produce the changes we observed within the DMS. Thus, future causal studies used specific neural perturbation approaches are warranted. Relatedly, rescue studies examining whether the drug-induced deficits behavior can be alleviated by bolstering neural activity in areas important for maintaining state representations could also identify a causal relationship and a potential therapeutic approach.

A second limitation of these studies is that putative unit types were pooled across subjects within a treatment group for the pseudoensemble analyses, which did not permit us to examine between-rat differences in the effects of cocaine intake on neural representations of behavior. Technical advances allowing the ability to record higher numbers of simultaneous units in the future will likely allow for these more detailed investigations.

A third limitation is that while our classifications of putative DMS MSNs and FSIs, as well as OFC and PL WS and NS units (believed to be putative pyramidal and fast-spiking interneurons, respectively), are supported by previous literature (Berke et al., 2004; Quirk et al., 2009; Gage et al., 2010; Gittis et al., 2011; Letzkus et al., 2011), future studies utilizing genetic tools to identify specific cell populations should be performed to confirm and expand upon our findings, as well as clarify the role of region unit populations modulating one another.

Due to the length and complexity of the studies included in Chapters 2 and 3, we only included male rats in our experiments. We recognize the importance of including female subjects and acknowledge that it is a limitation of our work, particularly because sex-dependent differences

in choice behavior following chronic drug exposure have been previously observed (Aguirre et al., 2020). Subsequent studies should focus on examining the effects of cocaine on state representations and decision making in female rats.

Future directions

Although previous work has shown that the ability of the DMS to encode state representations is OFC-dependent (Stalnaker et al., 2016), and we observed drug-induced alterations in state-related representations in both the OFC and DMS on similar decision making tasks, it is not necessarily the case that OFC and DMS deficits, nor either neural deficit and behavior, are related. To more closely examine whether neural changes are related, large scale simultaneous DMS and OFC recordings in the same hemisphere while rats perform the decision-making task would be a great step toward this end. This would provide a direct comparison across the regions on the exact same task, and large-scale recordings will provide the power necessary to examine unit pairs, with a particular focus on how pairs function during choice behavior in control rats and further whether prior cocaine experience impacts it. A better understanding of how the two regions interact would facilitate subsequent optogenetic neural manipulation studies to identify a causal link between the areas. Moreover, because certain cell types within these regions are thought to play particular roles in contributing to state representations (Stalnaker et al., 2016; Sharpe et al., 2019), it will be necessary to use genetic tools to get the most accurate view of how the regions and cell populations within them interact to encode state representation that support choice behavior.

The studies presented in Chapters 2 and 3 revealed that prior cocaine experience disrupts representations of state-related information in frontostriatal regions that are thought to support

good decision making and that such changes may be related to suboptimal choice behavior. As an input to these circuits further upstream, the HPC projects to the OFC (Barbas and Blatt, 1995), and OFC state representations require intact HPC-output, acting as a facilitator by incorporating information regarding variable external features into stable internal representations of state space in the OFC (Wikenheiser et al., 2017). Thus, an obvious follow-up investigation would be to examine whether the drug-induced changes in OFC state representations are related to changes in representations encoded by the HPC. Not only would the recording experiment shed light on the effects of chronic drug experience on HPC function during decision making, for which only sparse literature exists, it would also shed light on the effects of drugs on the HPC-OFC-DMS pathway, and be a useful study for understanding how drugs of abuse change representations underlying choice behavior.

References

Aguirre CG, Stolyarova A, Das K, Kolli S, Marty V, Ray L, Spigelman I, Izquierdo A (2020) Sex-dependent effects of chronic intermittent voluntary alcohol consumption on attentional, not motivational, measures during probabilistic learning and reversal. PLoS One 15:e0234729.

APA (2013) Diagnostic and statistical manual of mental disorders, 5th Edition. Arlington, VA.

Association AP (2013) Diagnostic and statistical manual of mental disorders, 5th Edition. Arlington, VA.

Balleine BW, Dickinson A (1998) Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. Neuropharmacology 37:407-419.

Balleine BW, O'Doherty JP (2010) Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. Neuropsychopharmacology 35:48-69.

Balleine BW, Liljeholm M, Ostlund SB (2009) The integrative function of the basal ganglia in instrumental conditioning. Behav Brain Res 199:43-52.

Baltz ET, Yalcinbas EA, Renteria R, Gremel CM (2018) Orbital frontal cortex updates state-induced value change for decision-making. Elife 7.

Barbas H, Blatt GJ (1995) Topographically specific hippocampal projections target functionally distinct prefrontal areas in the rhesus monkey. Hippocampus 5:511-533.

Bariselli S, Miyazaki NL, Creed MC, Kravitz AV (2020) Orbitofrontal-striatal potentiation underlies cocaine-induced hyperactivity. Nat Commun 11:3996.
Bechara A, Dolan S, Denburg N, Hindes A, Anderson SW, Nathan PE (2001) Decision-making deficits, linked to a dysfunctional ventromedial prefrontal cortex, revealed in alcohol and stimulant abusers. Neuropsychologia 39:376-389.
Benhamou L, Kehat O, Cohen D (2014) Firing pattern characteristics of tonically active neurons in rat striatum: context dependent or species divergent? J Neurosci 34:2299-2304.
Berke JD, Okatan M, Skurski J, Eichenbaum HB (2004) Oscillatory entrainment of striatal neurons in freely moving rats. Neuron 43:883-896.
Birrell JM, Brown VJ (2000) Medial frontal cortex mediates perceptual attentional set shifting in the rat. J Neurosci 20:4320-4324.
Bobadilla AC, Heinsbroek JA, Gipson CD, Griffin WC, Fowler CD, Kenny PJ, Kalivas PW (2017) Corticostriatal plasticity, neuronal ensembles, and regulation of drug-seeking behavior. Prog Brain Res 235:93-112.
Bradfield LA, Bertran-Gonzalez J, Chieng B, Balleine BW (2013) The thalamostriatal pathway and cholinergic control of goal-directed action: interlacing new with existing learning in the striatum. Neuron 79:153-166.
Breiter HC, Gollub RL, Weisskoff RM, Kennedy DN, Makris N, Berke JD, Goodman JM, Kantor HL, Gastfriend DR, Riorden JP, Mathew RT, Rosen BR, Hyman SE (1997) Acute effects of cocaine on human brain activity and emotion. Neuron 19:591-611.
Burgos-Robles A, Bravo-Rivera H, Quirk GJ (2013) Prelimbic and infralimbic neurons signal distinct aspects of appetitive instrumental behavior. PloS one 8:e57575-e57575.
Calu DJ, Stalnaker TA, Franz TM, Singh T, Shaham Y, Schoenbaum G (2007) Withdrawal from cocaine self-administration produces long-lasting deficits in orbitofrontal-dependent reversal learning in rats. Learn Mem 14:325-328.
Castane A, Theobald DE, Robbins TW (2010) Selective lesions of the dorsomedial striatum impair serial spatial reversal learning in rats. Behav Brain Res 210:74-83.
Cauli B, Porter JT, Tsuzuki K, Lambolez B, Rossier J, Quenet B, Audinat E (2000) Classification of fusiform neocortical interneurons based on unsupervised clustering. Proc Natl Acad Sci U S A 97:6144-6149.
Chen BT, Yau HJ, Hatch C, Kusumoto-Yoshida I, Cho SL, Hopf FW, Bonci A (2013) Rescuing cocaine-induced prefrontal cortex hypoactivity prevents compulsive cocaine seeking. Nature 496:359-362.
Chudasama Y, Kralik JD, Murray EA (2007) Rhesus monkeys with orbital prefrontal cortex lesions can learn to inhibit prepotent responses in the reversed reward contingency task. Cereb Cortex 17:1154-1159.
Corbit LH, Janak PH (2010) Posterior dorsomedial striatum is critical for both selective instrumental and Pavlovian reward learning. Eur J Neurosci 31:1312-1321.
Corbit LH, Chieng BC, Balleine BW (2014) Effects of repeated cocaine exposure on habit learning and reversal by N-acetylcysteine. Neuropsychopharmacology 39:1893-1901.
Crombag HS, Grimm JW, Shaham Y (2002) Effect of dopamine receptor antagonists on renewal of cocaine seeking by reexposure to drug-associated contextual cues. Neuropsychopharmacology 27:1006-1015.

Davis WM, Smith SG (1976) Role of conditioned reinforcers in the initiation, maintenance and extinction of drug-seeking behavior. Pavlov J Biol Sci 11:222-236.
Dias R, Robbins TW, Roberts AC (1996) Dissociation in prefrontal cortex of affective and attentional shifts. Nature 380:69-72.
Dickinson A, Wood N, Smith JW (2002) Alcohol seeking by rats: action or habit? Q J Exp Psychol B 55:331-348.
Egerton A, Brett RR, Pratt JA (2005) Acute delta9-tetrahydrocannabinol-induced deficits in reversal learning: neural correlates of affective inflexibility. Neuropsychopharmacology 30:1895-1905.
Ersche KD, Roiser JP, Robbins TW, Sahakian BJ (2008) Chronic cocaine but not chronic amphetamine use is associated with perseverative responding in humans. Psychopharmacology (Berl) 197:421-431.
Everitt BJ, Morris KA, O'Brien A, Robbins TW (1991) The basolateral amygdala-ventral striatal system and conditioned place preference: further evidence of limbic-striatal interactions underlying reward-related processes. Neuroscience 42:1-18.
Everitt BJ, Parkinson JA, Olmstead MC, Arroyo M, Robledo P, Robbins TW (1999) Associative processes in addiction and reward. The role of amygdala-ventral striatal subsystems. Ann N Y Acad Sci 877:412-438.
Farovik A, Place RJ, McKenzie S, Porter B, Munro CE, Eichenbaum H (2015) Orbitofrontal cortex encodes memories within value-based schemas and represents contexts that guide memory retrieval. J Neurosci 35:8333-8344.
Feierstein CE, Quirk MC, Uchida N, Sosulski DL, Mainen ZF (2006) Representation of spatial goals in rat orbitofrontal cortex. Neuron 51:495-507.
Fillmore MT, Rush CR (2006) Polydrug abusers display impaired discrimination-reversal learning in a model of behavioural control. J Psychopharmacol 20:24-32.
Floresco SB, Block AE, Tse MT (2008) Inactivation of the medial prefrontal cortex of the rat impairs strategy set-shifting, but not reversal learning, using a novel, automated procedure. Behav Brain Res 190:85-96.
Gage GJ, Stoetzner CR, Wiltschko AB, Berke JD (2010) Selective activation of striatal fast-spiking interneurons during choice execution. Neuron 67:466-479.
Gallagher M, McMahan RW, Schoenbaum G (1999) Orbitofrontal cortex and representation of incentive value in associative learning. J Neurosci 19:6610-6614.
Gittis AH, Leventhal DK, Fensterheim BA, Pettibone JR, Berke JD, Kreitzer AC (2011) Selective inhibition of striatal fast-spiking interneurons causes dyskinesias. J Neurosci 31:15727-15731.
Gobin C, Shallcross J, Schwendt M (2019) Neurobiological substrates of persistent working memory deficits and cocaine-seeking in the prelimbic cortex of rats with a history of extended access to cocaine self-administration. Neurobiol Learn Mem 161:92-105.
Graham JH PL (1995) Neuroanatomical substrates of cocaine self-administration. Boca Raton, FL: CRC Press.
Grant S, Contoreggi C, London ED (2000) Drug abusers show impaired performance in a laboratory test of decision making. Neuropsychologia 38:1180-1187.
Grifell M, Hart CL (2018) Is Drug Addiction a Brain Disease? American Scientist 106:160.

Hanlon CA, Beveridge TJ, Porrino LJ (2013) Recovering from cocaine: insights from clinical and preclinical investigations. Neurosci Biobehav Rev 37:2037-2046.
Heilig M, Goldman D, Berrettini W, O'Brien CP (2011) Pharmacogenetic approaches to the treatment of alcohol addiction. Nat Rev Neurosci 12:670-684.
Henry PK, Murnane KS, Votaw JR, Howell LL (2010) Acute brain metabolic effects of cocaine in rhesus monkeys with a history of cocaine use. Brain Imaging Behav 4:212-219.
Hikosaka K, Watanabe M (2000) Delay activity of orbital and lateral prefrontal neurons of the monkey varying with different rewards. Cereb Cortex 10:263-271.
Hogarth L, Lam-Cassettari C, Pacitti H, Currah T, Mahlberg J, Hartley L, Moustafa A (2019) Intact goal-directed control in treatment-seeking drug users indexed by outcome-devaluation and Pavlovian to instrumental transfer: critique of habit theory. Eur J Neurosci 50:2513-2525.
Homayoun H, Moghaddam B (2006) Progression of cellular adaptations in medial prefrontal and orbitofrontal cortex in response to repeated amphetamine. J Neurosci 26:8025-8039.
Hunt WA, Barnett LW, Branch LG (1971) Relapse rates in addiction programs. J Clin Psychol 27:455-456.
Insel N, Barnes CA (2015) Differential Activation of Fast-Spiking and Regular-Firing Neuron Populations During Movement and Reward in the Dorsal Medial Frontal Cortex. Cereb Cortex 25:2631-2647.
Izquierdo A, Jentsch JD (2012) Reversal learning as a measure of impulsive and compulsive behavior in addictions. Psychopharmacology (Berl) 219:607-620.
Izquierdo A, Belcher AM (2012) Rodent models of adaptive decision making. Methods Mol Biol 829:85-101.
Izquierdo A, Suda RK, Murray EA (2004) Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. J Neurosci 24:7540-7548.
Jentsch JD, Taylor JR (1999) Impulsivity resulting from frontostriatal dysfunction in drug abuse: implications for the control of behavior by reward-related stimuli. Psychopharmacology (Berl) 146:373-390.
Jentsch JD, Taylor JR (2001) Impaired inhibition of conditioned responses produced by subchronic administration of phencyclidine to rats. Neuropsychopharmacology 24:66-74.
Jentsch JD, Olausson P, De La Garza R, 2nd, Taylor JR (2002) Impairments of reversal learning and response perseveration after repeated, intermittent cocaine administrations to monkeys. Neuropsychopharmacology 26:183-190.
Kalivas PW (2004) Glutamate systems in cocaine addiction. Curr Opin Pharmacol 4:23-29.
Kalivas PW, McFarland K, Bowers S, Szumlinski K, Xi ZX, Baker D (2003) Glutamate transmission and addiction to cocaine. Ann N Y Acad Sci 1003:169-175.
Kawaguchi Y, Kubota Y (1997) GABAergic cell subtypes and their synaptic connections in rat frontal cortex. Cereb Cortex 7:476-486.
Klugah-Brown B, Di X, Zweerings J, Mathiak K, Becker B, Biswal B (2020) Common and separable neural alterations in substance use disorders: A coordinate-based meta-analyses of functional neuroimaging studies in humans. Hum Brain Mapp 41:4459-4477.

Koffarnus MN, Kaplan BA (2018) Clinical models of decision making in addiction. Pharmacol Biochem Behav 164:71-83.
Kriegeskorte N, Mur M, Bandettini P (2008) Representational similarity analysis - connecting the branches of systems neuroscience. Front Syst Neurosci 2:4.
Krueger DD, Howell JL, Oo H, Olausson P, Taylor JR, Nairn AC (2009) Prior chronic cocaine exposure in mice induces persistent alterations in cognitive function. Behav Pharmacol 20:695-704.
Kufahl PR, Li Z, Risinger RC, Rainey CJ, Wu G, Bloom AS, Li SJ (2005) Neural responses to acute cocaine administration in the human brain detected by fMRI. Neuroimage 28:904-914.
Lee K, Holley SM, Shobe JL, Chong NC, Cepeda C, Levine MS, Masmanidis SC (2017) Parvalbumin Interneurons Modulate Striatal Output and Enhance Performance during Associative Learning. Neuron 93:1451-1463.e1454.
Leshner AI (1997) Addiction is a brain disease, and it matters. Science 278:45-47.
Letzkus JJ, Wolff SB, Meyer EM, Tovote P, Courtin J, Herry C, Lüthi A (2011) A disinhibitory microcircuit for associative fear learning in the auditory cortex. Nature 480:331-335.
Lucantonio F, Stalnaker TA, Shaham Y, Niv Y, Schoenbaum G (2012) The impact of orbitofrontal dysfunction on cocaine addiction. Nat Neurosci 15:358-366.
Lucantonio F, Takahashi YK, Hoffman AF, Chang CY, Bali-Chaudhary S, Shaham Y, Lupica CR, Schoenbaum G (2014) Orbitofrontal activation restores insight lost after cocaine use. Nat Neurosci 17:1092-1099.
Luk KC, Sadikot AF (2001) GABA promotes survival but not proliferation of parvalbumin-immunoreactive interneurons in rodent neostriatum: an in vivo study with stereology. Neuroscience 104:93-103.
Lüscher C (2013) Drug-evoked synaptic plasticity causing addictive behavior. J Neurosci 33:17641-17646.
Lüscher C, Ungless MA (2006) The mechanistic classification of addictive drugs. PLoS Med 3:e437.
Marquis JP, Killcross S, Haddon JE (2007) Inactivation of the prelimbic, but not infralimbic, prefrontal cortex impairs the contextual control of response conflict in rats. Eur J Neurosci 25:559-566.
McCormick DA, Connors BW, Lighthall JW, Prince DA (1985) Comparative electrophysiology of pyramidal and sparsely spiny stellate neurons of the neocortex. J Neurophysiol 54:782-806.
McCracken CB, Grace AA (2013) Persistent cocaine-induced reversal learning deficits are associated with altered limbic cortico-striatal local field potential synchronization. J Neurosci 33:17469-17482.
McKenzie S, Frank AJ, Kinsky NR, Porter B, Rivière PD, Eichenbaum H (2014) Hippocampal representation of related and opposing memories develop within distinct, hierarchically organized neural schemas. Neuron 83:202-215.
McKenzie S, Keene CS, Farovik A, Bladon J, Place R, Komorowski R, Eichenbaum H (2016) Representation of memories in the cortical-hippocampal system: Results from the application of population similarity analyses. Neurobiol Learn Mem 134 Pt A:178-191.
Miles FJ, Everitt BJ, Dickinson A (2003) Oral cocaine seeking by rats: action or habit? Behav Neurosci 117:927-938.

Mueller LE, Sharpe MJ, Stalnaker TA, Wikenheiser AM, Schoenbaum G (2021) Prior Cocaine Use Alters the Normal Evolution of Information Coding in Striatal Ensembles during Value-Guided Decision-Making. J Neurosci 41:342-353.
Mueller LE, Kausch MA, Markovic T, MacLaren DA, Dietz DM, Park J, Clark SD (2015) Intra-ventral tegmental area microinjections of urotensin II modulate the effects of cocaine. Behav Brain Res 278:271-279.
Nelson A, Killcross S (2006) Amphetamine exposure enhances habit formation. J Neurosci 26:3805-3812.
Nestler EJ, Carlezon WA, Jr. (2006) The mesolimbic dopamine reward circuit in depression. Biol Psychiatry 59:1151-1159.
NIDA (2020a) The Science of Drug Use and Addiction: The Basics.
NIDA (2020b) Costs of substance abuse.
O'Brien CP (2005) Anticraving medications for relapse prevention: a possible new class of psychoactive medications. Am J Psychiatry 162:1423-1431.
O'Brien CP, Childress AR, McLellan AT, Ehrman R (1992) Classical conditioning in drug-dependent humans. Ann N Y Acad Sci 654:400-415.
Owen SF, Berke JD, Kreitzer AC (2018) Fast-Spiking Interneurons Supply Feedforward Control of Bursting, Calcium, and Plasticity for Efficient Learning. Cell 172:683-695.e615.
Pardo-Garcia TR, Garcia-Keller C, Penaloza T, Richie CT, Pickel J, Hope BT, Harvey BK, Kalivas PW, Heinsbroek JA (2019) Ventral Pallidum Is the Primary Target for Accumbens D1 Projections Driving Cocaine Seeking. J Neurosci 39:2041-2051.
Parsegian A, Glen WB, Jr., Lavin A, See RE (2011) Methamphetamine self-administration produces attentional set-shifting deficits and alters prefrontal cortical neurophysiology in rats. Biol Psychiatry 69:253-259.
Patzelt EH, Kurth-Nelson Z, Lim KO, MacDonald AW, 3rd (2014) Excessive state switching underlies reversal learning deficits in cocaine users. Drug Alcohol Depend 134:211-217.
Pelloux Y, Everitt BJ, Dickinson A (2007) Compulsive drug seeking by rats under punishment: effects of drug taking history. Psychopharmacology (Berl) 194:127-137.
Pierce RC, Kumaresan V (2006) The mesolimbic dopamine system: the final common pathway for the reinforcing effect of drugs of abuse? Neurosci Biobehav Rev 30:215-238.
Porrino LJ, Lyons D (2000) Orbital and medial prefrontal cortex and psychostimulant abuse: studies in animal models. Cereb Cortex 10:326-333.
Porter JN, Olsen AS, Gurnsey K, Dugan BP, Jedema HP, Bradberry CW (2011) Chronic cocaine self-administration in rhesus monkeys: impact on associative learning, cognitive control, and working memory. J Neurosci 31:4926-4934.
Quirk MC, Sosulski DL, Feierstein CE, Uchida N, Mainen ZF (2009) A defined network of fast-spiking interneurons in orbitofrontal cortex: responses to behavioral contingencies and ketamine administration. Front Syst Neurosci 3:13.
Ragozzino ME, Ragozzino KE, Mizumori SJ, Kesner RP (2002) Role of the dorsomedial striatum in behavioral flexibility for response and visual cue discrimination learning. Behav Neurosci 116:105-115.
Ragozzino ME, Kim J, Hassert D, Minniti N, Kiang C (2003) The contribution of the rat prelimbic-infralimbic areas to different forms of task switching. Behav Neurosci 117:1054-1065.

Rahman S, McBride WJ (2000) Feedback control of mesolimbic somatodendritic dopamine release in rat brain. J Neurochem 74:684-692.
Rescorla RA (1990) Instrumental responses become associated with reinforcers that differ in one feature. Animal Learning & Behavior 18:206-211.
Rich EL, Shapiro M (2009) Rat prefrontal cortical neurons selectively code strategy switches. J Neurosci 29:7208-7219.
Roesch MR, Taylor AR, Schoenbaum G (2006) Encoding of time-discounted rewards in orbitofrontal cortex is independent of value representation. Neuron 51:509-520.
Roesch MR, Takahashi Y, Gugsa N, Bissonette GB, Schoenbaum G (2007) Previous cocaine exposure makes rats hypersensitive to both delay and reward magnitude. J Neurosci 27:245-250.
Russo SJ, Nestler EJ (2013) The brain reward circuitry in mood disorders. Nat Rev Neurosci 14:609-625.
SAMHSA (2015) Substance use disorder.
SAMHSA (2016) Mental and Substance Use Disorders.
Schmitzer-Torbert NC, Redish AD (2008) Task-dependent encoding of space and events by striatal neurons is dependent on neural subtype. Neuroscience 153:349-360.
Schoenbaum G, Setlow B (2005) Cocaine makes actions insensitive to outcomes but not extinction: implications for altered orbitofrontal-amygdalar function. Cereb Cortex 15:1162-1169.
Schoenbaum G, Chiba AA, Gallagher M (1999) Neural encoding in orbitofrontal cortex and basolateral amygdala during olfactory discrimination learning. J Neurosci 19:1876-1884.
Schoenbaum G, Stalnaker TA, Niv Y (2013) How did the chicken cross the road? With her striatal cholinergic interneurons, of course. Neuron 79:3-6.
Schoenbaum G, Nugent SL, Saddoris MP, Setlow B (2002) Orbitofrontal lesions in rats impair reversal but not acquisition of go, no-go odor discriminations. Neuroreport 13:885-890.
Schoenbaum G, Setlow B, Nugent SL, Saddoris MP, Gallagher M (2003) Lesions of orbitofrontal cortex and basolateral amygdala complex disrupt acquisition of odor-guided discriminations and reversals. Learn Mem 10:129-140.
Schoenbaum G, Saddoris MP, Ramus SJ, Shaham Y, Setlow B (2004) Cocaine-experienced rats exhibit learning deficits in a task sensitive to orbitofrontal cortex lesions. Eur J Neurosci 19:1997-2002.
Schuck NW, Cai MB, Wilson RC, Niv Y (2016) Human Orbitofrontal Cortex Represents a Cognitive Map of State Space. Neuron 91:1402-1412.
Seip-Cammack KM, Shapiro ML (2014) Behavioral flexibility and response selection are impaired after limited exposure to oxycodone. Learn Mem 21:686-695.
Shaham Y, Shalev U, Lu L, de Wit H, Stewart J (2003) The reinstatement model of drug relapse: history, methodology and major findings. Psychopharmacology (Berl) 168:3-20.
Sharpe MJ, Stalnaker T, Schuck NW, Killcross S, Schoenbaum G, Niv Y (2019) An Integrated Model of Action Selection: Distinct Modes of Cortical Control of Striatal Decision Making. Annu Rev Psychol 70:53-76.
Simon NW, Mendez IA, Setlow B (2007) Cocaine exposure causes long-term increases in impulsive choice. Behav Neurosci 121:543-549.

Smith JE, Koves TR, Co C (2003) Brain neurotransmitter turnover rates during rat intravenous cocaine self-administration. Neuroscience 117:461-475.
Sokolic L, Long LE, Hunt GE, Arnold JC, McGregor IS (2011) Disruptive effects of the prototypical cannabinoid Δ^9-tetrahydrocannabinol and the fatty acid amide inhibitor URB-597 on go/no-go auditory discrimination performance and olfactory reversal learning in rats. Behav Pharmacol 22:191-202.
Song M, Takahashi YK, Burton AC, Roesch MR, Schoenbaum G, Niv Y, Langdon AJ (In prep.) Evaluating generalization in learning the representation of an odor guided choice task.
Sotres-Bayon F, Sierra-Mercado D, Pardilla-Delgado E, Quirk GJ (2012) Gating of fear in prelimbic cortex by hippocampal and amygdala inputs. Neuron 76:804-812.
Stalnaker TA, Raheja N, Schoenbaum G (2021) Orbitofrontal State Representations Are Related to Choice Adaptations and Reward Predictions. J Neurosci 41:1941-1951.
Stalnaker TA, Takahashi Y, Roesch MR, Schoenbaum G (2009) Neural substrates of cognitive inflexibility after chronic cocaine exposure. Neuropharmacology 56 Suppl 1:63-72.
Stalnaker TA, Berg B, Aujla N, Schoenbaum G (2016) Cholinergic Interneurons Use Orbitofrontal Input to Track Beliefs about Current State. J Neurosci 36:6242-6257.
Stalnaker TA, Roesch MR, Franz TM, Burke KA, Schoenbaum G (2006) Abnormal associative encoding in orbitofrontal neurons in cocaine-experienced rats during decision-making. Eur J Neurosci 24:2643-2653.
Stalnaker TA, Calhoon GG, Ogawa M, Roesch MR, Schoenbaum G (2010) Neural correlates of stimulus-response and response-outcome associations in dorsolateral versus dorsomedial striatum. Front Integr Neurosci 4:12.
Stalnaker TA, Roesch MR, Calu DJ, Burke KA, Singh T, Schoenbaum G (2007) Neural correlates of inflexible behavior in the orbitofrontal-amygdalar circuit after cocaine exposure. Ann N Y Acad Sci 1121:598-609.
Stalnaker TA, Cooch NK, McDannald MA, Liu TL, Wied H, Schoenbaum G (2014) Orbitofrontal neurons infer the value and identity of predicted outcomes. Nat Commun 5:3926.
Stefani MR, Moghaddam B (2005) Systemic and prefrontal cortical NMDA receptor blockade differentially affect discrimination learning and set-shift ability in rats. Behav Neurosci 119:420-428.
Stefani MR, Groth K, Moghaddam B (2003) Glutamate receptors in the rat medial prefrontal cortex regulate set-shifting ability. Behav Neurosci 117:728-737.
Teitelbaum H (1964) A COMPARISON OF EFFECTS OF ORBITOFRONTAL AND HIPPOCAMPAL LESIONS UPON DISCRIMINATION LEARNING AND REVERSAL IN THE CAT. Exp Neurol 9:452-462.
Thorpe SJ, Rolls ET, Maddison S (1983) The orbitofrontal cortex: neuronal activity in the behaving monkey. Exp Brain Res 49:93-115.
Tremblay L, Schultz W (1999) Relative reward preference in primate orbitofrontal cortex. Nature 398:704-708.
Trulson ME, Jacobs BL (1979) Effects of d-amphetamine on striatal unit activity and behavior in freely moving cats. Neuropharmacology 18:735 - 738.
Tsujimoto S, Genovesio A, Wise SP (2009) Monkey orbitofrontal cortex encodes response choices near feedback time. J Neurosci 29:2569-2574.

Vanderschuren LJ, Everitt BJ (2004) Drug seeking becomes compulsive after prolonged cocaine self-administration. Science 305:1017-1019.
Verdejo-Garcia AJ, Perales JC, Perez-Garcia M (2007) Cognitive impulsivity in cocaine and heroin polysubstance abusers. Addict Behav 32:950-966.
Volkow ND, Fowler JS (2000) Addiction, a disease of compulsion and drive: involvement of the orbitofrontal cortex. Cereb Cortex 10:318-325.
Wallis JD, Miller EK (2003) Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. Eur J Neurosci 18:2069-2081.
West EA, Saddoris MP, Kerfoot EC, Carelli RM (2014) Prelimbic and infralimbic cortical regions differentially encode cocaine-associated stimuli and cocaine-seeking before and following abstinence. Eur J Neurosci 39:1891-1902.
Wikenheiser AM, Marrero-Garcia Y, Schoenbaum G (2017) Suppression of Ventral Hippocampal Output Impairs Integrated Orbitofrontal Encoding of Task Structure. Neuron 95:1197-1207.e1193.
Wilson RC, Takahashi YK, Schoenbaum G, Niv Y (2014) Orbitofrontal cortex as a cognitive map of task space. Neuron 81:267-279.
Wiltschko AB, Pettibone JR, Berke JD (2010) Opposite effects of stimulant and antipsychotic drugs on striatal fast-spiking interneurons. Neuropsychopharmacology 35:1261-1270.
Wise RA (2004) Dopamine, learning and motivation. Nat Rev Neurosci 5:483-494.
Wolf ME (2016) Synaptic mechanisms underlying persistent cocaine craving. Nat Rev Neurosci 17:351-365.
Wright MJ, Jr., Vandewater SA, Parsons LH, Taffe MA (2013) Δ(9)Tetrahydrocannabinol impairs reversal learning but not extra-dimensional shifts in rhesus macaques. Neuroscience 235:51-58.
Yin HH, Ostlund SB, Knowlton BJ, Balleine BW (2005) The role of the dorsomedial striatum in instrumental conditioning. Eur J Neurosci 22:513-523.
Zhukovsky P, Puaud M, Jupp B, Sala-Bayo J, Alsio J, Xia J, Searle L, Morris Z, Sabir A, Giuliano C, Everitt BJ, Belin D, Robbins TW, Dalley JW (2019) Withdrawal from escalated cocaine self-administration impairs reversal learning by disrupting the effects of negative feedback on reward exploitation: a behavioral and computational analysis. Neuropsychopharmacology 44:2163-2173.