Behavioral/Cognitive

Prior Cocaine Use Alters the Normal Evolution of Information Coding in Striatal Ensembles during Value-Guided Decision-Making

Lauren E. Mueller,[1,2] Melissa J. Sharpe,[3] Thomas A. Stalnaker,[1] Andrew M. Wikenheiser,[3*] and Geoffrey Schoenbaum[1*]

[1]Cellular Neurobiology Research Branch, Behavioral Neurophysiology Research Section, National Institute on Drug Abuse Intramural Research Program, Baltimore, Maryland 21224, [2]Department of Behavioral Neuroscience, Oregon Health and Science University, Portland, Oregon 97239, and [3]Department of Psychology, University of California, Los Angeles, Los Angeles, California 90095

Substance use disorders (SUDs) are characterized by maladaptive behavior. The ability to properly adjust behavior according to changes in environmental contingencies necessitates the interlacing of existing memories with updated information. This can be achieved by assigning learning in different contexts to compartmentalized "states." Though not often framed this way, the maladaptive behavior observed in individuals with SUDs may result from a failure to properly encode states because of drug-induced neural alterations. Previous studies found that the dorsomedial striatum (DMS) is important for behavioral flexibility and state encoding, suggesting the DMS may be an important substrate for these effects. Here, we recorded DMS neural activity in cocaine-experienced male rats during a decision-making task where blocks of trials represented distinct states to probe whether the encoding of state and state-related information is affected by prior drug exposure. We found that DMS medium spiny neurons (MSNs) and fast-spiking interneurons (FSIs) encoded such information and that prior cocaine experience disrupted the evolution of representations both within trials and across recording sessions. Specifically, DMS MSNs and FSIs from cocaine-experienced rats demonstrated higher classification accuracy of trial-specific rules, defined by response direction and value, compared with those drawn from sucrose-experienced rats, and these overly strengthened trial-type representations were related to slower switching behavior and reaction times. These data show that prior cocaine experience paradoxically increases the encoding of state-specific information and rules in the DMS and suggest a model in which abnormally specific and persistent representation of rules throughout trials in DMS slows value-based decision-making in well trained subjects.

Key words: addiction; cocaine; rat; single unit; striatum

Significance Statement

Substance use disorders (SUDs) may result from a failure to properly encode rules guiding situationally appropriate behavior. The dorsomedial striatum (DMS) is thought to be important for such behavioral flexibility and encoding that defines the situation or "state." This suggests that the DMS may be an important substrate for the maladaptive behavior observed in SUDs. In the current study, we show that prior cocaine experience results in over-encoding of state-specific information and rules in the DMS, which may impair normal adaptive decision-making in the task, akin to what is observed in SUDs.

Introduction

Behavioral inflexibility is a core feature of substance use disorders (SUDs), exemplified by persistent drug use despite negative consequences (American Psychiatric Association, 2013; Substance Abuse and Mental Health Services Administration, 2016). Within the laboratory 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-García et al., 2007; Ersche et al., 2008; but 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; but 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 individuals with SUDs may be at least partly the result of a failure to properly encode states because of 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; Castañé 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 composed 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
Subjects
Nine male Long–Evans rats (weight, 250–300 g; age, ~3 months; Charles River Laboratories) were subjects for this experiment. During

the odor-guided choice task training and recording sessions, rats received water *ad libitum* for 10 min/d and food *ad libitum*, and during self-administration rats were food deprived to 85% of initial weight with *ad libitum* access to 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 Institutes 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 7 d 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 16 25-µm-diameter NiCr wires electroplated to an impedance of ~200 kΩ were bilaterally implanted above the DMS (anteroposterior, −0.4 mm; mediolateral, ±2.6 mm; ventral, 3.5–4.0 mm; relative to bregma and dura). Rats were allowed 2 weeks to recover, after which they were returned to the decision-making task for 1 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%, w/v; $n = 4$) 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 before 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 ~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 >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 interfaced with a training chamber. Electrode signals were amplified 20× via operational head stages on electrode arrays and then passed to differential preamplifiers, where they underwent an additional 50× amplification and filtration at 150–9000 Hz (Stalnaker et al., 2010, 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× amplification. Signal waves were derived from active channels with event time stamps documented by the behavioral program.

Experimental design and statistical analyses

Behavioral epochs. Task trials were divided into five epochs either 0.5 or 0.8 s in length (see Fig. 3a). 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 before 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 before 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 before 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). Files were processed using NeuroExplorer (Nex Technologies) to extract time stamps and further analyzed using MATLAB (MathWorks). Striatal units were classified as a putative FSI, MSN, or unidentified (UID), based on the following three distinct clusters formed in a 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; see Fig. 2b–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 peak-normalized average firing rate of the cell exceeded 95% of its absolute maximum value (see Fig. 3b). Raw, un-normalized population firing rates were computed across all neurons of a given population and trials in 50 ms bins (see Fig. 3c).

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 of trial-by-trial spike counts from each neuron during the five distinct trial epochs. Data were limited to the counterbalanced set of correctly 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 × 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 were 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 pseudoensembles was repeated 20 times. Since there was a slight variation in the number 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 pseudoensemble, we performed a shuffling procedure to confirm that the accuracy of randomized data are at the theoretical chance level of one-quarter correct. To do this, we trained classifiers using trial-type labels randomly reassigned to training-set firing rates. Classification accuracies were averaged across the 20 runs, comparisons were made between cell types (MSN vs FSI), treatment groups (sucrose experienced vs cocaine experienced), and time (100 ms bins), and the effects were analyzed using repeated-measures ANOVAs (see Fig. 3d). Confusion matrices show the proportion of test samples classified correctly (along the diagonal) and incorrectly (off of the diagonal) averaged across the 20 repetitions (see Figs. 4a, 5a). 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 multiway ANOVAs (Figs. 4b, 5b). 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 0 (black), and other values were set to 1 (white; see Figs. 4c, 5c). 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 (e.g., first bin session, sessions 1–2; second bin session, sessions 3–4). 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 (see Fig. 6a). Corresponding reaction times were calculated over forced-choice trials by averaging reaction times over the same sessions included in each bin (see Fig. 6b).

Statistical analyses. All analyses were conducted using MATLAB (MathWorks) and GraphPad Prism 8 software (GraphPad Software). 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. The *p* values were significant if they fell to <0.05. Sample sizes were determined from previous studies using similar behavioral and recording procedures (Stalnaker et al., 2006, 2016; Wikenheiser et al., 2017). Error bars on the plot represent the SEM.

Results

The experimental timeline is shown in Figure 1a. 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. 1c), 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/d for 14 d (Fig. 1a,b). 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. 1a). Following recovery, rats received an additional week of
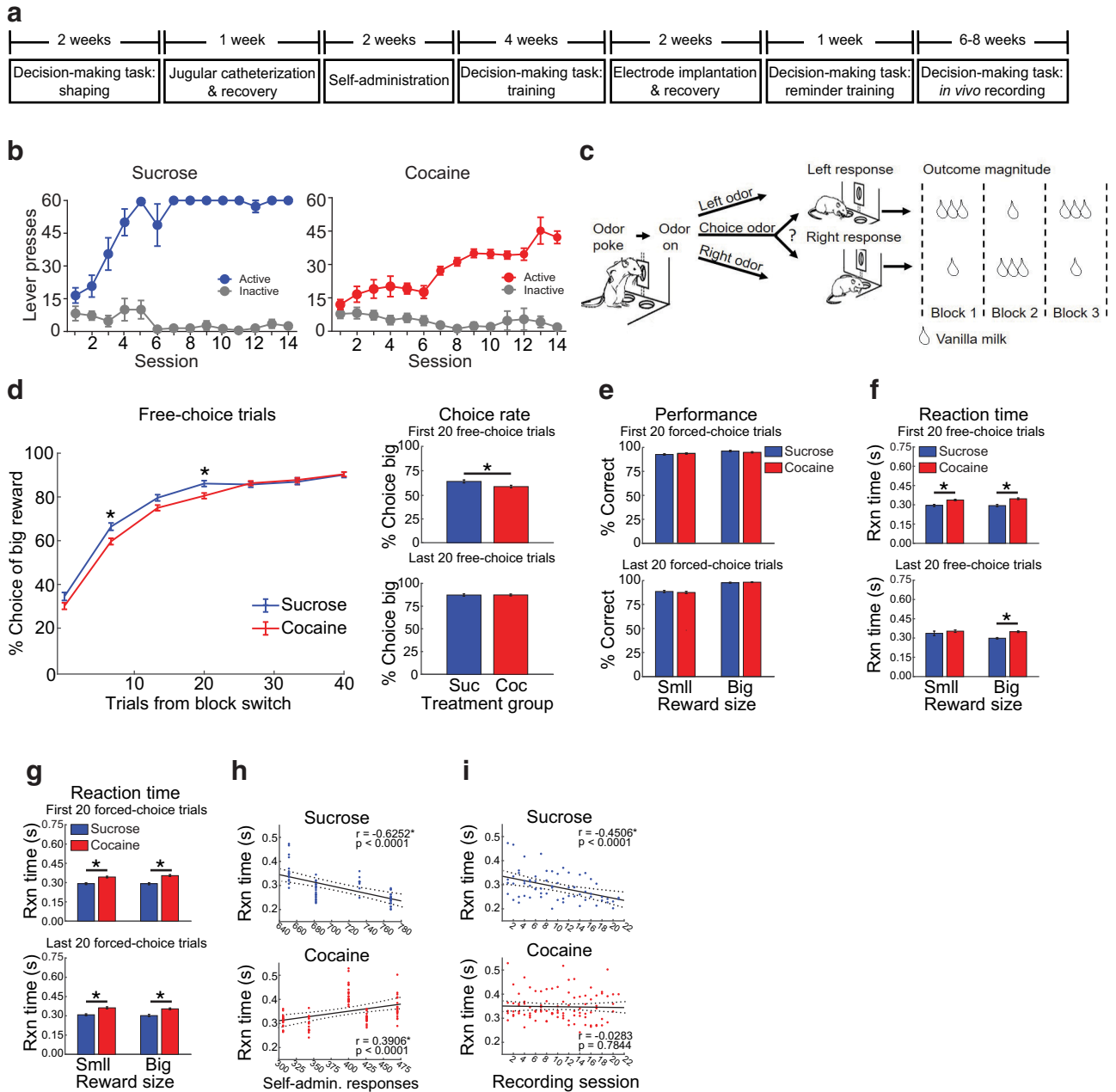
Figure 1. Self-administration, task design, and behavior. **a**, Experimental timeline. **b**, Number of active (colored circles) and inactive (gray circles) lever presses during each session of sucrose (left; $n = 4$) and cocaine (right; $n = 5$) self-administration. **c**, The decision-making task 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 percentage 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 (bottom) trials 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 (bottom) forced-choice trials 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-experienced rats (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.

reminder training for acclimation to recording cables and then neural recording sessions began (Fig. 1*a*); 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. 1*d*, left). The rate of this
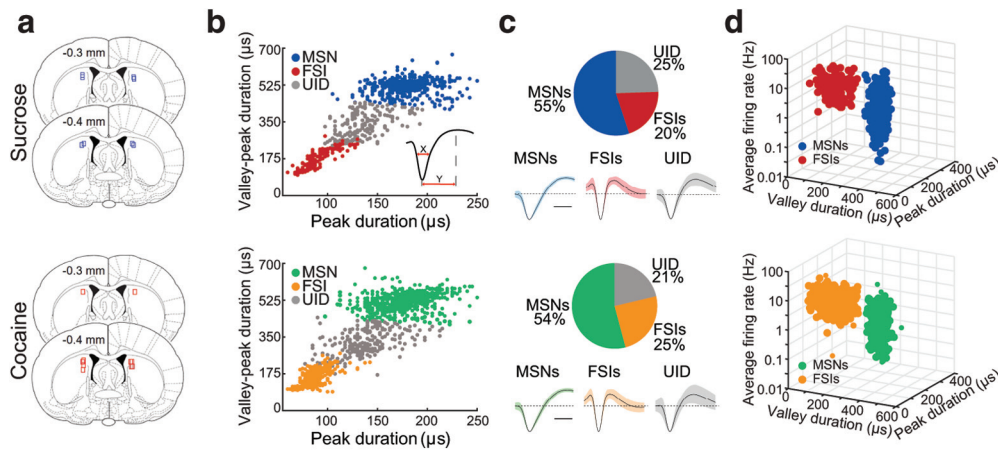
Figure 2. Histology and neuron classification. ***a***, Approximate locations of neural recordings in the DMS are indicated by colored boxes (sucrose, top; cocaine, bottom). ***b***, Scatter plots of average spike waveform durations (*x*, peak duration at half-maximum; *y*, valley–peak duration) for each single unit recorded in sucrose-experienced rats (top) and cocaine-experienced rats (bottom). ***c***, The approximate proportions and average waveforms superimposed to show interwave variability for classified cell types recorded from sucrose-experienced rats (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-experienced rats (top; MSNs: 180 μs, 362 μs, 3.09 Hz; FSIs: 91 μs, 197 μs, 13.48 Hz) and cocaine-experienced rats (bottom; MSNs: 177 μs, 362 μs, 2.33 Hz; FSIs: 90 μs, 177 μs, 10.94 Hz).

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. 1*d*, 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. 1*d*, right). Bonferroni *post hoc* testing found significant group differences in the second and fourth trial bins following the block transition (*$p < 0.05$; Fig. 1*d*, left) and on early but not late trials (early: *$p = 0.0077$, $t_{(322)} = 2.912$; late: $p > 0.99$, $t_{(322)} = 0.04,409$; Fig. 1*d*, right). Both groups also exhibited enhanced accuracy on big forced-choice trials; however, there was no effect of cocaine on this measure (Fig. 1*e*). 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: $p = 0.9214$, $F_{(1,161)} = 0.0098$; interaction: $p = 0.0773$, $F_{(1,161)} = 3.161$; late, group: $p = 0.7982$, $F_{(1,161)} = 0.06,556$; interaction: $p = 0.3192$, $F_{(1,161)} = 0.9985$). Cocaine also caused a general slowing of responding to the fluid well after odor sampling on free-choice trials (Fig. 1*f*). 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 among size and treatment group and trial early/late ($p = 0.2586$, $F_{(1150)} = 1.286$). 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. 1*g*). 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 among size and treatment group and trial early/late ($p = 0.2810$, $F_{(1,161)} = 1.170$). Bonferroni *post hoc* testing showed that the cocaine group had 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. 1*h*). ANCOVA revealed significant differences 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. 1*i*). 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: $p = 0.7844$, $r = -0.0283$).

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

To investigate whether prior cocaine experience altered activity of DMS neuron populations, units were recorded in the DMS (Fig. 2*a*). 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. 2*b*; 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. 2*c*). The largest cluster of cells (sucrose, *n* = 432; cocaine, *n* = 679) had long-duration waveforms and low firing
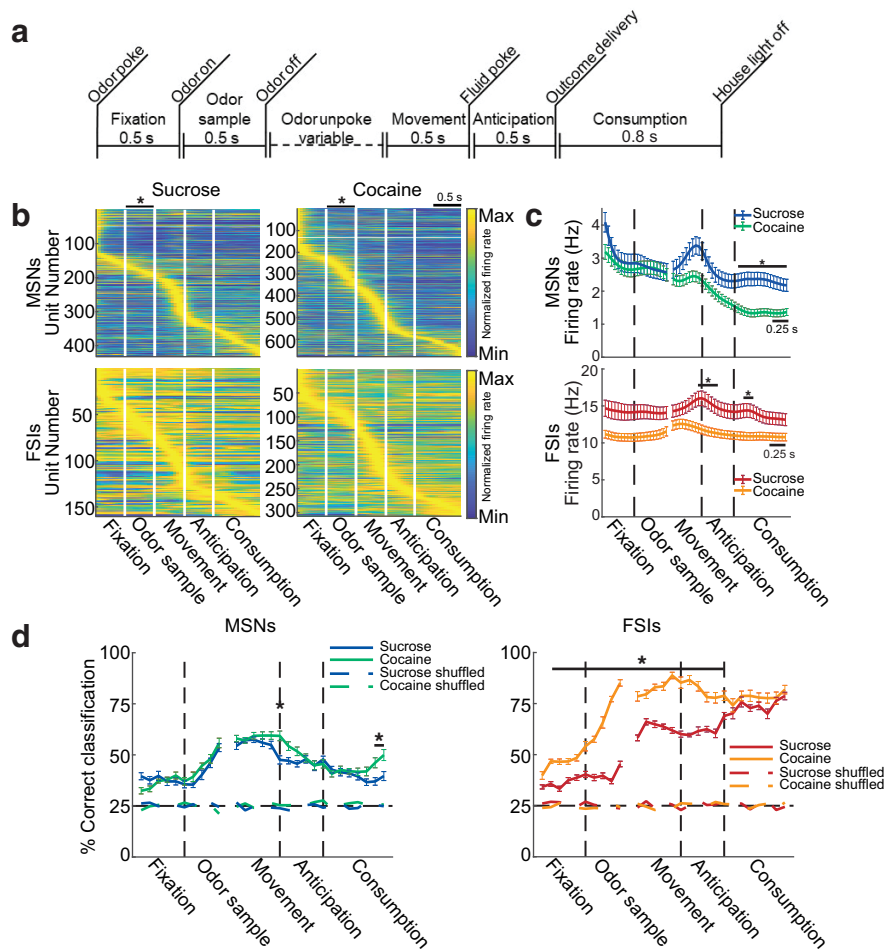
Figure 3. Time course of DMS MSN and FSI neural responses across treatment groups. ***a***, Correctly performed forced-choice trials were divided into five epochs bounded by important task events. ***b***, Perievent time histograms (PETHs) of peak-normalized firing rates in MSNs (top) and FSIs (bottom) recorded from sucrose-experienced rats (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 or 1 s; bin size, 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-averaged 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 with 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. Min, Minimum; Max, maximum.

rates (Fig. 2c,d), which are characteristics of MSNs. A second cluster of cells (sucrose, $n = 160$; cocaine, $n = 307$) displayed very brief waveforms and higher firing rates (Fig. 2c,d), features that are 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 that were 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. 2c); these were classified as UIDs and excluded from analysis because of their unknown phenotype. The CINs

likely fell within the UID cluster; however, waveform metrics alone proved insufficient to adequately define them.

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. 3a), and units collected throughout the entire recording experiment were sorted according to the time of maximum peak-normalized response (Fig. 3b). This revealed a significant increase in the proportion of MSNs with a peak response during the odor sampling epoch in cocaine-experienced rats compared with sucrose-experienced rats (z-test for population proportions, corrected for multiple comparisons: odor sample, ∗$p < 0.0001$, $z_{(1110)} = -4.3059$; Fig. 3b, top). Cocaine had no significant effect on the time course of DMS FSI neural responses (Fig. 3b, bottom). Additionally, cocaine significantly decreased the raw, un-normalized population-average firing rates for both MSNs and FSIs (Fig. 3c). Two-way ANOVAs revealed significant main effects of treatment for both cell types (MSN: ∗$p < 0.0001$; $F_{(1,62,104)} = 230.46$; FSI: ∗$p < 0.0001$; $F_{(1,26040)} = 522.47$) along with significant main effects and interactions with time in the MSNs (time: ∗$p < 0.0001$; $F_{(55,62104)} = 11.31$; interaction: ∗$p = 0.0017$; $F_{(55,62,104)} = 1.65$). Bonferroni *post hoc* tests revealed between-group differences reaching significance in the reward anticipation and consumption epochs (Fig. 3c, 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 the classification accuracy of both MSN and FSI pseudoensembles exceeded the 25% chance level throughout the trial (Fig. 3d). Classification accuracy was above chance even before 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 was 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. 3d, asterisks), suggesting that prior cocaine experience increases the amount of trial-related information being encoded by pseudoensembles.

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. 3d); however, each gives rise to different patterns of classification when plotted in a confusion matrix (Figs. 4, 5, 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. 4a–c, 5a–c), was most similar to that of block (Figs. 4, 5, block pattern). During the odor-sampling epoch, the block and diagonal trial patterns 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. 4b, 5b). 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 among 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, $*p < 0.0001$, $F_{(4,152)} = 18.94$; interaction, $*p = 0.0143$, $F_{(4,152)} = 3.223$; FSIs: treatment group, $*p < 0.0001$, $F_{(1,38)} = 71.05$; trial epoch, $*p < 0.0001$, $F_{(4,152)} = 54.92$; interaction, $*p = 0.0034$, $F_{(4,152)} = 4.113$). Consistent with subtle differences between treatment groups in MSN classification accuracy (Fig. 3d, 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. 4a–c, anticipation). By contrast, pseudoensembles composed of FSIs, which exhibited much larger differences in classification accuracy (Fig. 3d, 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. 5a–c).

Prior cocaine 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 with sucrose-experienced rats, and failed to become faster at performing the task over the course of the entire recording experiment, an effect that 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. 6a, top), whereas the representation of information about trial and direction did not change in pseudoensembles drawn from cocaine-experienced rats (Fig. 6a, bottom). ANCOVAs supported these observations, revealing significant differences between the slopes of best-fit lines in FSIs recorded from sucrose-experienced 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. 1), we found similar results when the neural data were 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. 6b, top), they were not significantly correlated in cocaine-experienced rats (direction: $p = 0.5715$, $r = $
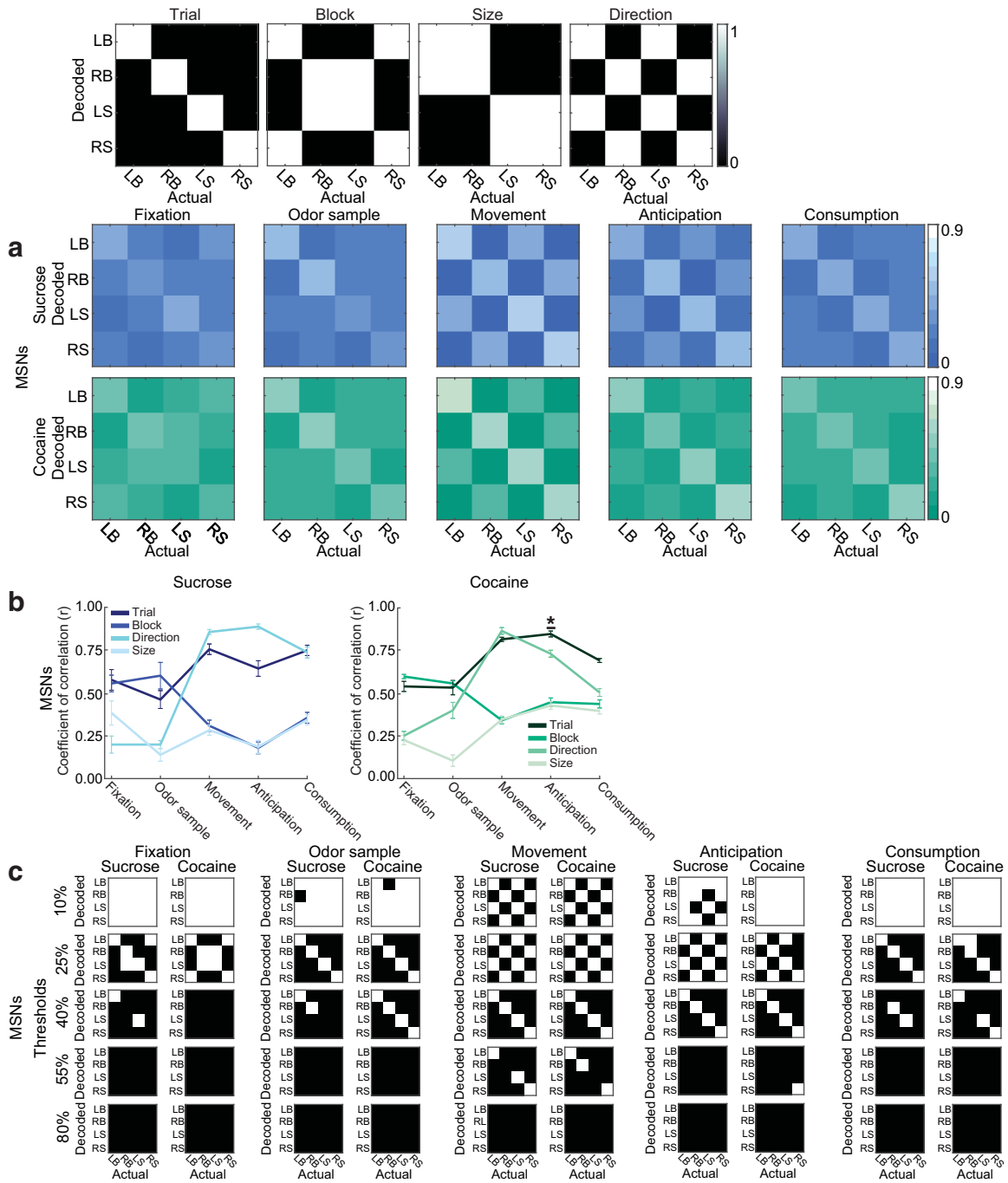
Figure 4. Evolution of representations in DMS MSNs. Confusion matrices illustrate the four exemplar representations (Trial, State, Size, and Direction). The y-axis represents the ground-truth four trial types, ordered by direction and size (LB, left-big; RB, right-big; LS, left-small; RS, right-small). 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-experienced rats (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 a higher coefficient of the correlation of trial pattern with raw confusion matrices generated using pseudoensembles drawn from cocaine-experienced rats 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).

−0.2042; trial: $p = 0.9698$, $r = 0.0138$; Fig. 6*b*, 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.

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 because of drug-induced changes in the neural mechanisms governing
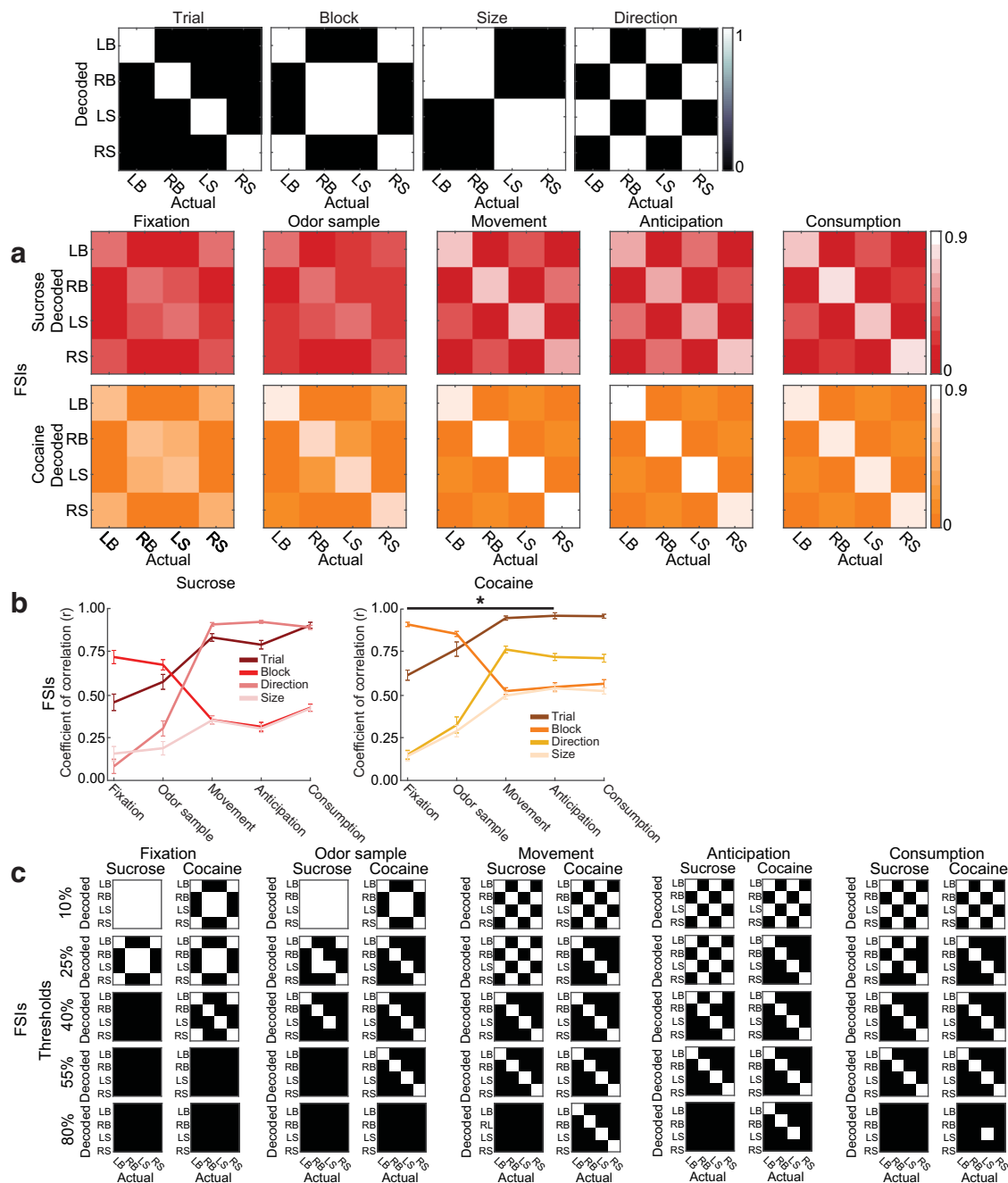
Figure 5. Evolution of representations in DMS FSIs. Confusion matrices illustrate the four exemplar representations (Fig. 4 legend). **a**, Confusion matrices depict the proportion of test samples classified correctly (along the diagonal) and incorrectly (off the diagonal) during the five-trial epochs using pseudoensembles of 100 categorized FSIs in sucrose sessions (top) and cocaine sessions (bottom). **b**, Coefficient of the correlation between exemplar patterns and raw FSI confusion matrices generated using pseudoensembles drawn from sucrose-experienced rats (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-experienced rats than those from 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).

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 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
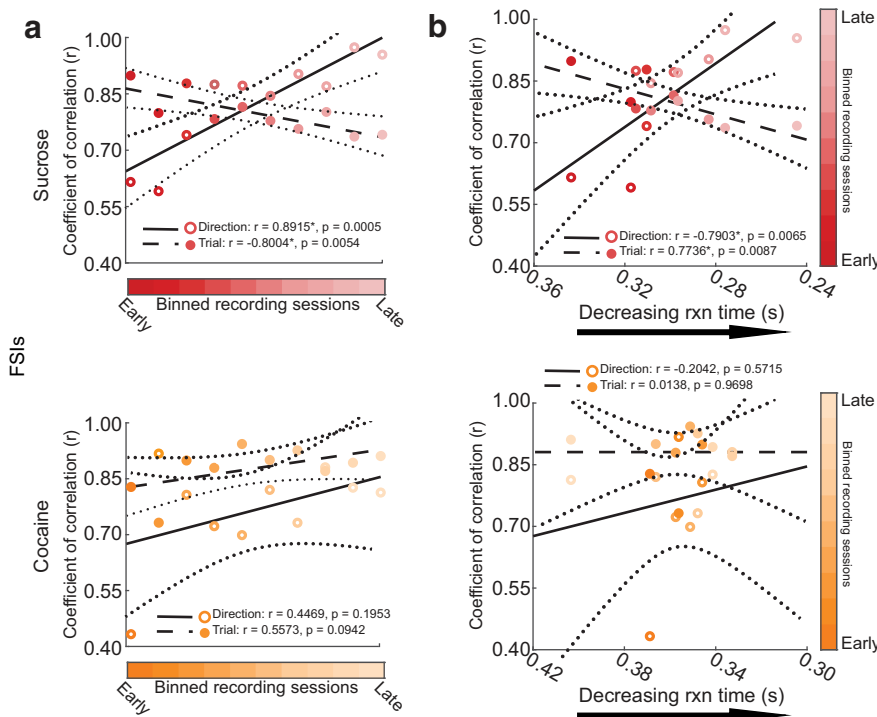
Figure 6. Evolution of representations in DMS FSIs across recording sessions correlates with behavior. ***a***, Coefficient of the correlation between raw FSI confusion matrices generated using pseudoensembles drawn from sucrose-experienced rats (top) and cocaine-experienced rats (bottom) during the 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 the 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-experienced rats (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.

rule encoding. While addiction is a complex multihit 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, before 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 (e.g., 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 with 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 overencoding 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-encoding 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, it is thought that FSIs are important for the acquisition of striatum-dependent action selection strategies and that impairments in FSI signaling result in learning deficits that 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 using genetic tools to identify specific cell populations should be performed to confirm and expand on our findings, as well as to 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 that 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 overrepresentation 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 with SUDs (Jentsch et al., 2002; Schoenbaum and Setlow, 2005).

References

American Psychiatric Association (2013) Diagnostic and statistical manual of mental disorders, Ed 5. Arlington, VA: American Psychiatric Association.

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:e35988.

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.

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.

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.

Castañé 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.

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.

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.

Fillmore MT, Rush CR (2006) Polydrug abusers display impaired discrimination-reversal learning in a model of behavioural control. J Psychopharmacol 20:24–32.

Gage GJ, Stoetzner CR, Wiltschko AB, Berke JD (2010) Selective activation of striatal fast-spiking interneurons during choice execution. Neuron 67:466–479.

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.

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.

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.

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.e4.

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.

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.

Miles FJ, Everitt BJ, Dickinson A (2003) Oral cocaine seeking by rats: action or habit? Behav Neurosci 117:927–938.

Nelson A, Killcross S (2006) Amphetamine exposure enhances habit formation. J Neurosci 26:3805–3812.

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.e15.

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.

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.

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.

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, 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.

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.

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, Takahashi Y, Roesch MR, Schoenbaum G (2009) Neural substrates of cognitive inflexibility after chronic cocaine exposure. Neuropharmacology 56 Suppl 1:63–72.

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, Berg B, Aujla N, Schoenbaum G (2016) Cholinergic interneurons use orbitofrontal input to track beliefs about current state. J Neurosci 36:6242–6257.

Substance Abuse and Mental Health Services Administration (2016) Mental and substance use disorders. Rockville, MD: Substance Abuse and Mental Health Services Administration.

Vanderschuren LJ, Everitt BJ (2004) Drug seeking becomes compulsive after prolonged cocaine self-administration. Science 305:1017–1019.

Verdejo-García AJ, Perales JC, Pérez-García M (2007) Cognitive impulsivity in cocaine and heroin polysubstance abusers. Addict Behav 32:950–966.

Wikenheiser AM, Marrero-Garcia Y, Schoenbaum G (2017) Suppression of ventral hippocampal output impairs integrated orbitofrontal encoding of task structure. Neuron 95:1197–1207.e3.

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.

Wright MJ Jr, Vandewater SA, Parsons LH, Taffe MA (2013) Δ(9) Tetrahydrocannabinol impairs reversal learning but not extradimensional 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, Alsiö 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.