Ventral tegmental dopamine inputs to the nucleus accumbens mediates cue-triggered motivation but not reward expectancy

Briac Halbout[1*], Andrew T. Marshall[1], Ali Azimi[1], Mimi Liljeholm[2], Stephen V. Mahler[3], Kate M Wassum[4], Sean B. Ostlund[1*]

1. *Department of Anesthesiology and Perioperative Care and Irvine Center for Addiction Neuroscience, University of California, Irvine, CA, USA*
2. *Department of Cognitive Sciences, University of California, Irvine, CA, USA*
3. *Department of Neurobiology & Behavior and Irvine Center for Addiction Neuroscience, University of California, Irvine, CA, USA*
4. *Department of Psychology and Brain Research Institute University of California, Los Angeles, CA, USA.*

Corresponding authors (*):
Briac Halbout
University of California, Irvine
837 Health Science Road, 2216 GNRF
Irvine, CA 92697
Email: halboutb@uci.edu
Tel: 949-824-9826

Sean Ostlund
University of California, Irvine
837 Health Science Road, 3111GNRF
Irvine, CA 92697
Email: sostlund@uci.edu
Tel: 949-824-1347

Running Title: VTA-dopamine neurons mediate cue-triggered reward seeking

Abstract 148

Article: 5234
Introduction: 764
Materials and Methods: 2081
Results: 2038
Discussion: 1657
Figures count: 5
References count: 68
Supplementary Materials: 1

Reward-paired cues stimulate reward-seeking behavior through a combination of motivational and cognitive processes. We used pathway-specific chemogenetic inhibition of dopamine neurons to determine their role in the cue-elicited reward seeking, and applied a novel analytic approach to assay cue-induced changes in reward expectancy. Inhibiting ventral tegmental area dopamine neurons abolished cue-induced reward seeking but not reward retrieval, indicating that this pathway modulates response vigor but not reward expectancy. Locally inhibiting dopamine inputs to nucleus accumbens also disrupted cue-elicited reward seeking but not retrieval. Interestingly, the suppression produced by this treatment was greater in rats that responded to cues with exploratory reward seeking, without attempting to retrieve reward, than in rats exhibiting complete bouts of seeking with retrieval, indicating that individuals differ in their reliance on a dopamine-mediated motivational process. These findings shed new light on the behavioral and neural mechanisms of adaptive and compulsive forms of cue-motivated behavior.

Cues that signal reward availability elicit vigorous reward-seeking behavior by engaging independent motivational and cognitive processes (Balleine and Ostlund, 2007; Liljeholm and O'Doherty, 2011; Nadler *et al*, 2011). This is normally an adaptive response to reward-paired cues, but it may also contribute to compulsive reward seeking in addiction and related disorders (Berridge and Robinson, 2016; Cartoni *et al*, 2016; Everitt *et al*, 2001; Hogarth *et al*, 2013; LeBlanc *et al*, 2012). While the mesocorticolimbic dopamine system is an important mediator of cue-elicited reward seeking (Dickinson *et al*, 2000; Lex and Hauber, 2008; Wassum *et al*, 2011; Wyvell and Berridge, 2000), much remains to be learned about its specific contributions to the motivational and cognitive processes underlying this multifaceted behavior.

Pavlovian-to-instrumental transfer (PIT) refers to the ability of a conditioned stimulus (CS+) that signals passive reward delivery to invigorate performance of a separately trained instrumental reward-seeking action (e.g., lever pressing). While the PIT effect depends partly on the motivational properties of the CS+ (Bindra, 1974; Dickinson and Balleine, 1994; Rescorla and Solomon, 1967), cognitive processes also make important contributions, influencing *which* actions are selected (Colwill and Motzkin, 1994; Kruse *et al*, 1983) and *when* they are performed (Delamater and Oakeshott, 2007; Marshall and Ostlund, 2018; Matell and Della Valle, 2018). One popular account of PIT proposes that the CS+ invigorates reward-seeking actions by temporarily raising expectations that such behavior will be effective in producing reward (Cartoni *et al*, 2016; Cartoni *et al*, 2013; Hogarth *et al*, 2014; Hogarth and Troisi II, 2015; Rescorla, 1994). This view predicts that rats will

have a stronger expectancy of instrumental (i.e., response-contingent) reward delivery in the presence of the CS+ than in its absence.

While dopamine plays an important role in regulating PIT performance, the scope and nature of its involvement remains unclear. For instance, blocking dopamine receptor activation disrupts expression of the PIT effect (Dickinson *et al*, 2000; Ostlund and Maidment, 2011; Wassum *et al*, 2011), but leaves intact the capacity for reward-paired cues to guide action selection based on the identity of the anticipated reward (Ostlund *et al*, 2011). While this suggests that cognitive processes influencing PIT performance may not strongly depend on dopamine transmission, other findings suggest that dopamine supports relevant executive functions, such as applying cognitive flexibility (Floresco *et al*, 2005; Radke *et al*, 2018) or using complex cognitive reward representations (Sharpe and Schoenbaum, 2018). Dopamine signaling also underlies cognitive distortions in reward expectancy brought about by 'near-misses' on a rodent "slot machine" task (Cocker *et al*, 2016). Thus, it is plausible that dopamine contributes to PIT by mediating cue-induced inflation of instrumental reward expectancy. Alternatively, dopamine may selectively mediate the response-invigorating, motivational influence of reward-paired cues. Testing these predictions has been a challenge given the lack of an established method for measuring cue-induced changes in reward expectancy. However, in a recent PIT study we found that exposing rats to a food-paired CS+ not only increased the rate of lever pressing, it also increased the likelihood that rats would follow these reward-*seeking* actions with an immediate attempt to *retrieve* reward by approaching the food cup (Marshall *et al*, 2018). Thus, the CS+ caused rats to behave as if they had

newly heightened expectations that their lever-press responses would be effective in delivering food reward. We hypothesize that bouts of lever pressing immediately followed by reward retrieval reflect the use of a deliberate reward-seeking strategy based on cognitive reward expectancy, whereas lever pressing without reward retrieval represents a more exploratory approach to reward seeking that may contribute to persistent, compulsive-like behavior (Joel and Avisar, 2001; Joel and Doljansky, 2003; Joel *et al*, 2008).

In the current study, we investigated how chemogenetic inhibition of ventral tegmental area (VTA) dopamine neurons, or their inputs to the nucleus accumbens (NAc) or medial prefrontal cortex (mPFC), impacts the expression of PIT. Our analysis includes assessment of cue-induced changes in instrumental reward expectancy using the reward-retrieval measure, which was validated in an initial experiment. We hypothesized that if dopamine's role in PIT relates to cue-induced inflation of instrumental reward expectancy, then inhibiting dopamine neurons would disrupt the influence of the CS+ on reward seeking as well as that cue's tendency to increase the likelihood of response-contingent reward retrieval. However, if dopamine's role in PIT relates primarily to a motivational function, then inhibiting dopamine neurons should disrupt cue-induced reward seeking but not reward retrieval. As an additional test of cognitive control over reward seeking, we also investigated the effects of inhibiting VTA dopamine neurons on rats' ability to select instrumental actions based on changes in expected reward value (Balleine and Dickinson, 1998).

Materials and Methods

Animals: In total, 89 male and female Long-Evans Tyrosine hydroxylase (Th):Cre+ rats (hemizygous Cre+) (Mahler *et al*, *in press;* Mahler *et al*, 2014; Witten *et al*, 2011) and wildtype (WT) littermates were used for this study. Subjects were at least 3 months of age at the start of the experiment and were single- or paired-housed in standard Plexiglas cages on a 12h/12h light/dark cycle. Animals were maintained at ~85% of their free-feeding weight during behavioral procedures. All experimental procedures that involved rats were approved by the UC Irvine Institutional Animal Care and Use Committee and were in accordance with the National Research Council Guide for the Care and Use of Laboratory Animals.

Apparatus: Behavioral procedures took place in sound- and light-attenuated Med Associates chambers (St Albans, VT, USA). Each chamber was equipped with two retractable levers. Grain-based dustless precision pellets (45 mg, BioServ, Frenchtown, NJ, USA) were delivered into a food cup positioned in a recessed magazine between the levers. A photobeam detector was used to record food-cup approaches. Chambers were illuminated during all sessions.

Surgery: Th:Cre+ rats were anesthetized using isoflurane and placed in a stereotaxic frame for microinjections of a Cre-dependent (DIO) serotype 2 adeno-associated virus (AAV) vectors to induce dopamine neuron-specific expression of the inhibitory designer receptor exclusively activated by designer drug (DREADD) hM4Di fused to mCherry (AAV-hSyn-DIO-hM4Di-mCherry), or mCherry alone (AAV-hSyn-DIO-mCherry) (University of North Carolina Chapel Hill vector Core, Chapel Hill, NC, USA / Addgene, Cambridge, MA, USA; Experiment 2 was replicated with both sources)(Mahler

et al, *in press;* Witten *et al*, 2011). The AAV was injected bilaterally into the VTA (-5.5 mm AP, +/- 0.8 mm ML, -8.15 mm DV; 1uL / side). Experiment 3 rats were bilaterally implanted with guide cannulae (22 gauge, Plastic One) 1 mm dorsal to NAc (+1.3 AP, +/-1.8 ML, -6.2 DV) or mPFC (+3.00 AP, +/-0.5 ML, -3.0 DV) for subsequent clozapine-n-oxide (CNO) microinjections. Animals were randomly assigned to virus (hM4Di or mCherry) and cannula location (NAc or mPFC) groups. Animals were allowed at least 5 days of recovery before undergoing food restriction and behavioral training. Testing occurred at least 25 days after surgery to allow adequate time for viral expression of hM4Di throughout dopamine neurons, including in terminals within the NAc and mPFC.

Experiment 1: Microstructural analysis of reward seeking and reward retrieval under conditions of varying response-contingent feedback about reward delivery.

Instrumental learning. WT rats (n = 9) underwent 2 d of magazine training (40 pellet deliveries in each 1-h daily session) before receiving 9 d of instrumental training. In each session, rats had continuous access to a single lever, which could be pressed to deliver food pellets into the food cup. The schedule of reinforcement was adjusted over days from continuous reinforcement (CRF) to increasing random intervals (RI), with one day each of CRF, RI-15s, RI-30s, and 6 days of RI-60s. Each session was terminated after 30 min or after 20 rewards deliveries.

Varying response-contingent feedback. Following training, rats were given a series of tests to assess the influence of response-contingent

feedback signaling reward delivery on instrumental reward-seeking (lever presses) and reward-retrieval responses (food-cup approach after lever pressing), allowing us to determine if the latter measure reliably tracks changes in instrumental reward expectancy. Rats were given 3 tests (30-min each, pseudorandom order over days) during which lever pressing caused: 1) activation of the pellet dispenser to deliver a pellet into the food cup (RI-60s schedule; *Food and Cues Test*), 2) activation of the pellet dispenser to deliver a pellet into an external cup not accessible to the rats, producing associated sound and tactile cues but no reward (also RI-60s schedule; *Cues Only Test*), or 3) no dispenser activation (i.e., extinction; *No Food or Cues Test*).

Experiments 2 and 3: Role of mesocorticolimbic dopamine in cue-motivated reward seeking and retrieval:

Pavlovian conditioning. Th:Cre+ rats (n = 60) underwent 2 d of magazine training, as in Experiment 1, before they received 8 daily Pavlovian conditioning sessions, each of which consisted of a series of 6 presentations of a two-min audio cue (CS+; either 10Hz tone or white noise; 80dB), with trials separated by a variable 3-min interval (range 2-4-min). During each CS+ trial, pellets were delivered on a 30-s random time schedule, resulting in an average of 4 pellets per trial. Rats were separately habituated to an unpaired auditory stimulus (CS-; alternative auditory stimulus).

Instrumental training. Following Pavlovian conditioning, all rats were given 9 d of instrumental training with lever pressing ultimately reinforced on an RI-60s schedule, as in Experiment 1.

Pavlovian-to-instrumental transfer test. After a session of Pavlovian conditioning, rats were given a 30-min extinction session, during which lever presses were recorded but had no consequence (i.e., no food or cues). On the next day, rats were given a PIT test, during which the lever was continuously available but produced no rewards. Following 8 min of extinction, the CS+ and CS− were each presented 4 times (2 min per trial) in pseudorandom order and separated by a 3-min fixed interval. Before each new round of testing, rats were given two sessions of instrumental retraining (RI-60s), one session of CS+ retraining, and one 30-min extinction session, as described above.

Experiment 2: Th:Cre+ rats expressing hM4Di (n = 18) or mCherry only (n = 14) in VTA dopamine neurons were used to assess the effects of system-wide inhibition of the mesocorticolimbic dopamine system on PIT performance. These groups were run together and received CNO (5mg/kg, i.p.) or vehicle (5% DMSO in saline) injection 30 min prior to testing. They were subjected to a second test following retraining, during which they were given the alternative drug pretreatment.

Experiment 3: In Experiment 3a, Th:Cre+ rats expressing hM4Di in VTA dopamine neurons were used to assess the impact of locally inhibiting dopaminergic terminals in the NAc (n = 7) or mPFC (n = 9) on PIT performance. Because microinjection procedures produced additional variability in task performance, rats in this experiment underwent a total of 4 tests. Rats received either CNO microinfusions (1mM, 0.5uL/side or 0.3uL/side, for NAc and mPFC respectively) or vehicle (DMSO 5% in aCSF) 5 min before the start of each test, and were given two rounds of testing each

with CNO and vehicle (test order counterbalanced across other experimental conditions). To determine if the effects of CNO microinjections depended on hM4Di expression, we ran a separate control study (Experiment 3b) using Th:Cre+ rats expressing mCherry only in VTA dopamine neurons. Experiments 3a and 3b were run and analyzed separately.

Experiment 4: *Role of mesocorticolimbic dopamine in goal-directed decision making*:

Instrumental Training. Th:Cre+ rats (n = 20) began with 2 d of magazine training, during which they received 20 grain-pellets and 20 liquid sucrose rewards (0.1 mL of 20% sucrose solution, w/v) on a random 30-s schedule. This was followed by 10 d of instrumental training with two distinct action–outcome contingencies (e.g., left lever press → grain; right lever press→ sucrose) on a reinforcement schedule that was gradually shifted from continuous reinforcement (CRF) to random ratio 20 (RR-20). The left and right lever-press responses were trained in separate sessions, at least 2 h apart, each day. The specific action-outcome arrangements were counterbalanced across subjects. Sessions were terminated after 30 min elapsed or 20 pellets were earned.

Devaluation Testing. To selectively devalue one of the food outcomes prior to testing, all rats were satiated on grain pellets or sucrose solution by providing them with 90 min of unrestricted access to that food in the home cage. After 60 min of feeding, rats received CNO (5mg/kg, i.p.) or vehicle injections. After an additional 30 minutes of feeding, rats were placed in the chamber for a test in which they had continuous access to both levers. The

test began with a 5-min extinction phase (no food or cues), which was immediately followed by a 15-min reinforced phase, during which each action was reinforced with its respective outcome (CRF for the first 5 rewards, then RR-20). Rats were given a total of 4 devaluation tests, 2 after CNO and 2 after vehicle, alternating the identity of the devalued outcome across the 2 tests in each drug condition (test order counterbalanced across training and drug conditions).

Histology: Rats were deeply anesthetized with a lethal dose of pentobarbital and perfused with 1x PBS followed by 4% paraformaldehyde. Brains were postfixed in 4% paraformaldehyde, cryoprotected in 20% sucrose and sliced at 40 µm on a cryostat. To visualize hM4Di expression, we performed immunohistochemistry for Th and mCherry tag. Tissue was first incubated in 3% normal donkey serum PBS plus Triton X-100 (PBST; 2 h) and then in primary antibodies in PBST at 4°C for 48 hours using rabbit anti-DsRed (mCherry tag; 1:500; Clontech; 632496), and mouse anti-TH (1:1,000, Immunostar; 22941) antibodies. Sections were incubated for 4 h at room temperature in fluorescent conjugated secondary antibodies (Alexa Fluor 488 goat anti-mouse (TH; 1:500; Invitrogen; A10667) and Alexa Fluor 594 goat anti-rabbit (DsRed; 1:500; Invitrogen; A11037).

Drugs: CNO was obtained from NIMH (Experiments 2 and 4) or Sigma-Aldrich (St. Louis, MO, USA; Experiment 3), and dissolved in 5% DMSO in saline, or aCSF for microinjection.

Statistical Analysis: Data were analyzed using general(ized) linear mixed-effects models (Pinheiro and Bates, 2000), which allows for simultaneous parameter estimation as a function of condition (fixed effects) and the individual rat (random effects) (Boisgontier and Cheval, 2016; Bolker *et al*, 2008; Pinheiro *et al*, 2000). Analyses on count data (e.g., press frequency) incorporated a Poisson response distribution and a log link function (Coxe *et al*, 2009). Fixed-effects structures included an overall intercept and the full factorial of all primary manipulations (Experiment 2: Group, Drug, CS Type, CS Period; Experiment 3: Site, Drug, CS Type, CS Period; Experiment 4: Group, Drug, Lever), and the random-effects structures included by-subjects uncorrelated intercepts adjusted for the within-subjects manipulations (i.e., Experiments 2 and 3: Drug, CS Type, and CS Period; Experiment 4: Drug, Lever). “CS Type” refers to the distinction between the CS+ and CS-, while “CS Period” refers to the distinction between the 120-s CS duration and the 120-s period preceding its onset. *Reward retrieval* was quantified as the proportion of all lever presses that were followed within 2.5 sec by a food-cup approach. These data were square-root transformed prior to analysis to correct positive skew, but are plotted in non-transformed space for ease of interpretation. Reward retrieval data were collapsed across pre-CS+ and pre-CS- periods, such that the factor “CS Period” had 3 levels (CS+, CS-, and Pre-CS). The fixed- and random-effects structures of this analysis was identical to the frequency analysis above with the exception that CS Type was not included in the analysis, and the random-effects structure only included by-subjects intercepts.

All statistical analyses were conducted using the Statistics and Machine Learning Toolbox in MATLAB (The MathWorks; Natick, MA, USA). The alpha level for all tests was .05. As all predictors were categorical in the mixed-effects analysis, effect size was represented by the unstandardized regression coefficient (Baguley, 2009), reported as *b* in model output tables. Mixed-effects models provide *t*-values to reflect the statistical significance of the coefficient relative to the population mean (i.e., simple effects). These simple effects are indicative of main effects and interactions when a factor has only two levels. For factors with at least 3 levels, *F*-tests were conducted to reveal the overall significance of the effect or interaction(s) involving this factor. The source of significant interactions was determined by secondary mixed-effects models identical to those described above but split by the relevant factor of interest. For analyses in which a significant main effect had more than two levels, post-hoc tests of main effects employed MATLAB's *coefTest* function, and interactions were reported in-text as the results of ANOVA *F*-tests (i.e., whether the coefficients for each fixed effect were significantly different from 0).

PIT Scores (CS+ minus Pre-CS+) were calculated for more focused analyses of CS+ elicited reward seeking. One-sample *t*-tests were used to compare PIT Scores to 0 for each group in the vehicle condition. As a metric of CNO-induced suppression of CS+ elicited reward seeking, the PIT Score from the vehicle condition was subtracted from the corresponding score in the CNO condition. For each group, this was compared to 0 via a one-sample *t*-test. We defined lever presses that were not followed within 2.5 s by a food-cup approach response as *exploratory reward-seeking actions*. In Experiment

3, we assessed the correlation between individual differences in CS+ elicited exploratory reward seeking in the vehicle condition against differences in CNO-induced suppression of PIT.

Results

Effects of response-contingent feedback about reward delivery on reward retrieval

In a recent study (Marshall *et al*, 2018), we proposed that rats' tendency to approach the food cup after lever pressing (reward seeking) represents an attempt to retrieve response-contingent reward and may therefore serve as a behavioral assay of instrumental reward expectancy. An initial experiment (see Figure 1A for illustration) was run to evaluate this interpretation of reward-retrieval behavior. Rats were trained to lever press on a RI-60s schedule, such that reward seeking was often nonreinforced but was on occasionally reinforced with a food pellet. As seen in Figure 1B and Figure1-Suplement 1, rats were more likely to approach the food cup within a few seconds of performing a lever-press response than they were at other times. Not surprisingly, rats were much more likely to approach the food cup when a lever press was reinforced than when it was not ($t(8) = 19.33$, $p < .001$), suggesting that they could detect when pellets were delivered based on sound and tactile cues that were produced when the dispenser was activated. However, food-cup approaches also increased following nonreinforced presses (Figures 1B and 1C), indicating that rats made sporadic attempts to retrieve rewards that were not actually delivered. To determine how response-contingent cues signaling reward delivery influence reward-retrieval behavior,

rats were given a series of tests in which lever pressing produced either 1) pellet dispenser cues and actual pellet delivery (Food and Cues), 2) pellet dispenser cues only (Cues Only), or 3) no pellet dispenser cues or pellet delivery (No Food or Cues). Reward-retrieval actions were defined as bouts of food-cup approach occurring within 2.5 s of a lever press (Figure 1E) based on the typical range of inter-response-intervals between these actions (Figure 1B and Figure 1-Supplement 1). Because reward retrieval opportunities are contingent on reward-seeking behavior (Figure 1D), we focused our analysis on a normalized measure of reward retrieval – the proportion of lever presses that were followed by food-cup approach (Figure 1F).

We found that reward-retrieval actions were much more likely to occur after reinforced than nonreinforced lever-press responses (Figure 1F), regardless of whether pellet dispenser cues were presented alone or together with actual food delivery (Figure 1F; ts(8) ≥ 13.74, ps < .001). When lever presses had no consequence at all (e.g., No Food or Cues Test), rats continued to seek reward at a high rate (Figure 1D), but rarely followed such actions with an attempt to retrieve food from the cup (Figure 1F). Thus, without feedback that their behavior was effective in producing reward, rats continued to *explore* the instrumental contingency but did not behave as if they expected their lever pressing to produce reward. These findings show that the tendency to approach the food cup immediately after pressing is not a fixed response sequence, but instead tracks changes in instrumental reward expectancy signaled by response-contingent cues.

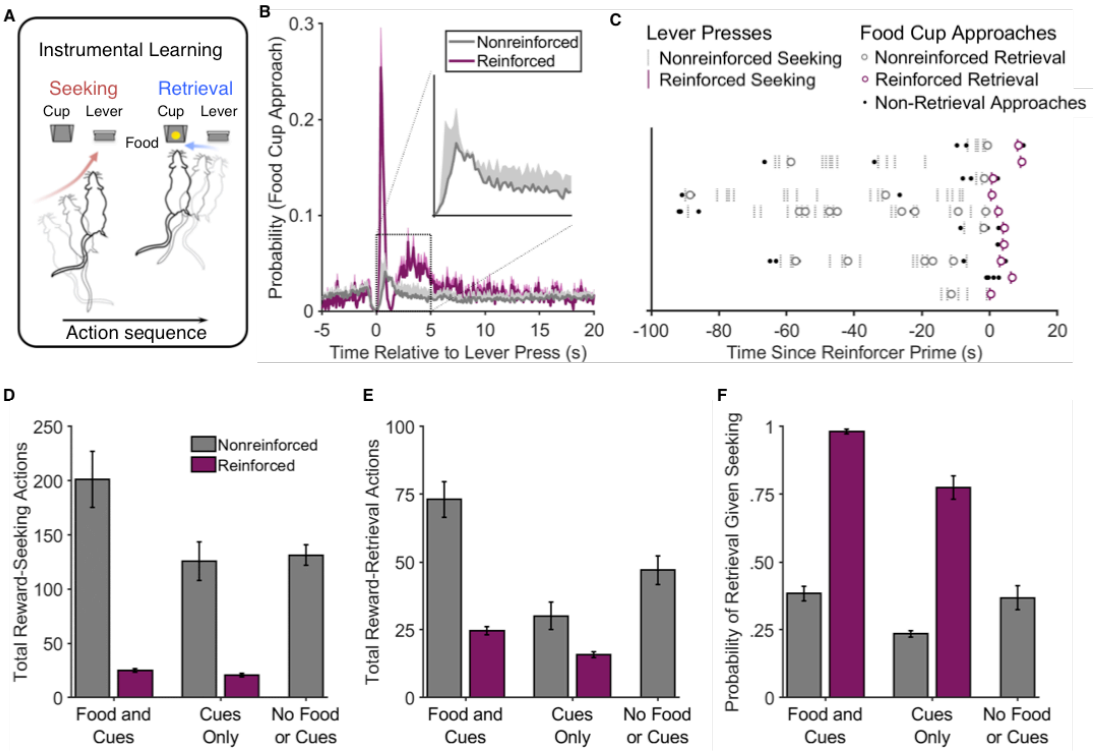

Figure 1. Microstructural organization of instrumental behavior. **A.** Hungry rats were trained to perform a self-paced task (RI-60s schedule of reinforcement) in which "reward-seeking" lever presses could be followed by immediate food-cup approaches to collect the food (i.e., "reward retrieval"). **B.** Probability of food-cup approaches as a function of time surrounding a reinforced or nonreinforced lever press. Shaded regions above each thick line represent +1 between-subjects SEM. **C.** Representative organization of instrumental behaviors for an individual rat showing reward-seeking and –retrieval actions for each reinforced or nonreinforced press. Each line represents a reinforced trial during RI-60s training. **D, E.** Effects of the manipulation of reinforcement contingency on the organization of instrumental actions. Total reward-seeking (D) or –retrieval (E) actions at tests during which lever pressing was either reinforced under an RI-60s schedule of reinforcement with pellet delivery (Food and Cues), cues associated with food pellets delivery (Cues Only; i.e., activation of the pellet dispenser), or not reinforced (No Food or Cues). **F.** The probability of reward retrieval given a reward-seeking action is highly dependent on the reinforcement contingency condition. Attempted reward retrieval was much more likely to occur after reinforced than nonreinforced trials, regardless of whether reinforcement was with Food and Cues, or Cues Only. Rats rarely followed lever press with an attempt to collect food from the cup in nonreinforced (including No Food or Cues) conditions. Error bars in D-F represent ± 1 between-subjects SEM.

Inhibiting dopamine neurons during Pavlovian-to-instrumental transfer preferentially disrupts cue-motivated reward seeking, but not reward retrieval

Our previous findings suggest that noncontingent exposure to reward-predictive cues *both* invigorate reward-seeking behavior (i.e., the PIT effect) and increase the likelihood that such actions are followed by an attempt to

retrieve reward from the food cup (Marshall *et al*, 2018). Together with findings from Experiment 1, this supports the hypothesis that the response-invigorating influence of reward-paired cues may at least partly depend on their ability to inflate instrumental reward expectancy. Using the new behavioral index of reward expectancy, we next asked whether mesocorticolimbic dopamine signaling plays a role in this cognitive component of PIT or whether it instead preferentially mediates the motivational influence of reward-paired cues on reward seeking.

Rats with VTA dopamine neuron-specific expression of hM4Di or mCherry (Figure 2) were trained on the PIT task (Figure 3A) consisting of a Pavlovian conditioning phase, in which two different auditory cues were paired (CS+) or unpaired (CS-) with food pellets, and a separate instrumental training phase, in which rats were trained to lever press for pellets. During PIT testing, we noncontingently presented the CS+ and CS- while rats were free to lever press and check the food cup without response-contingent food or cue delivery.

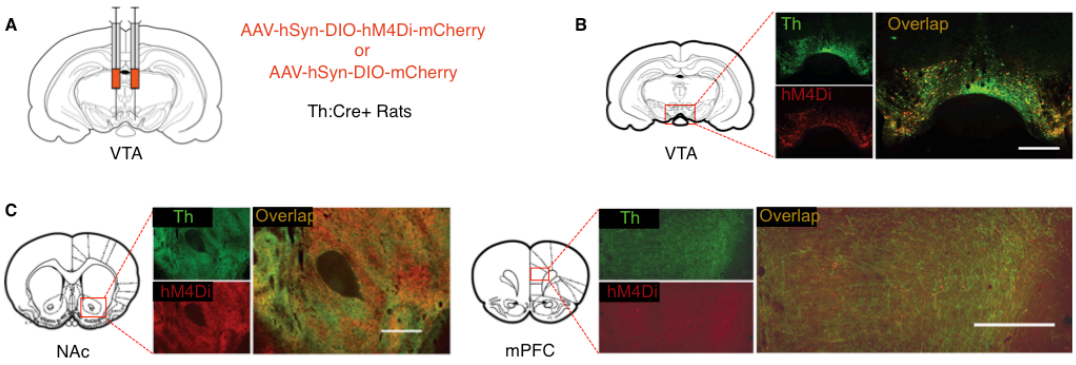

Figure 2. DREADD expression in Th:Cre+ rats. **A.** Th:Cre+ rats received bilateral injections of AAV-hSyn-DIO-hM4Di-mCherry or AAV-hSyn-DIO-mCherry in the VTA. **B.** Representative expression of the mCherry-tagged inhibitory DREADD hM4Di (red) in VTA Th positive neurons (green) of Th:Cre+ rats, as well as in neuronal terminals (**C.**) projecting to the nucleus accumbens (NAc) and medial prefrontal cortex (mPFC). Scale bar is 500um.

We found that pretreatment with the DREADD ligand CNO suppressed reward seeking specifically during CS+ periods in hM4Di but not mCherry rats (Figure 3B; significant Group * Drug * CS Period * CS Type interaction, $t(240) = -3.14$, $p = .002$; see Supplementary File Table 1 for full generalized linear mixed-effects model output). Control rats in the mCherry group exhibited elevated lever pressing during CS+ trials (relative to pre-CS response rates; CS Period * CS Type interaction, $p < .001$), and this effect was not altered by CNO (Drug * CS Period * CS Type interaction, $p = .780$). In contrast, CNO pretreatment disrupted expression of CS+ induced reward seeking in the hM4Di group (Drug * CS Period * CS Type interaction, $p < .001$). Specifically, hM4Di rats showed a CS+ specific elevation in lever pressing when pretreated with vehicle (CS Period * CS Type interaction, $p < .001$) but not CNO (CS Period * CS Type interaction, $p = .684$). The increase in lever pressing during the CS+ (PIT score: CS+ - pre-CS) was significant for both vector groups when pretreated with vehicle (one-sample t-tests; *p*s < .001; Figure 3C), but was significantly suppressed by CNO in the hM4Di group ($t(17) = -3.83$, $p = .001$), and not in the mCherry group ($t(13) = -1.21$, $p = .249$; Figure 3C).

These findings indicate that inhibiting VTA dopamine neurons disrupts cue-induced invigoration of reward seeking. We then investigated if the influence the CS+ on instrumental reward expectancy was also sensitive to dopamine neuron inhibition (Figures 3E and 3F). We found that the CS+ ($p < .001$) but not the CS- ($p = .501$) increased the probability that rats would attempt to retrieve a reward after lever pressing, even though no rewards were delivered at test. Importantly, there was no evidence that inhibiting DA

neurons impacted this CS+ induced increase in reward retrieval (interactions involving Drug or Group, *p*s > .109; Supplementary File Table 3; see Figure 3-Supplement 1 and Supplementary File Table 2 for analysis of reward retrieval *frequency*). Therefore, although noncontingent CS+ presentations cause rats to act as if they expect their reward-seeking behavior to be effective in producing reward, this influence did not depend on VTA dopamine neuron activity.

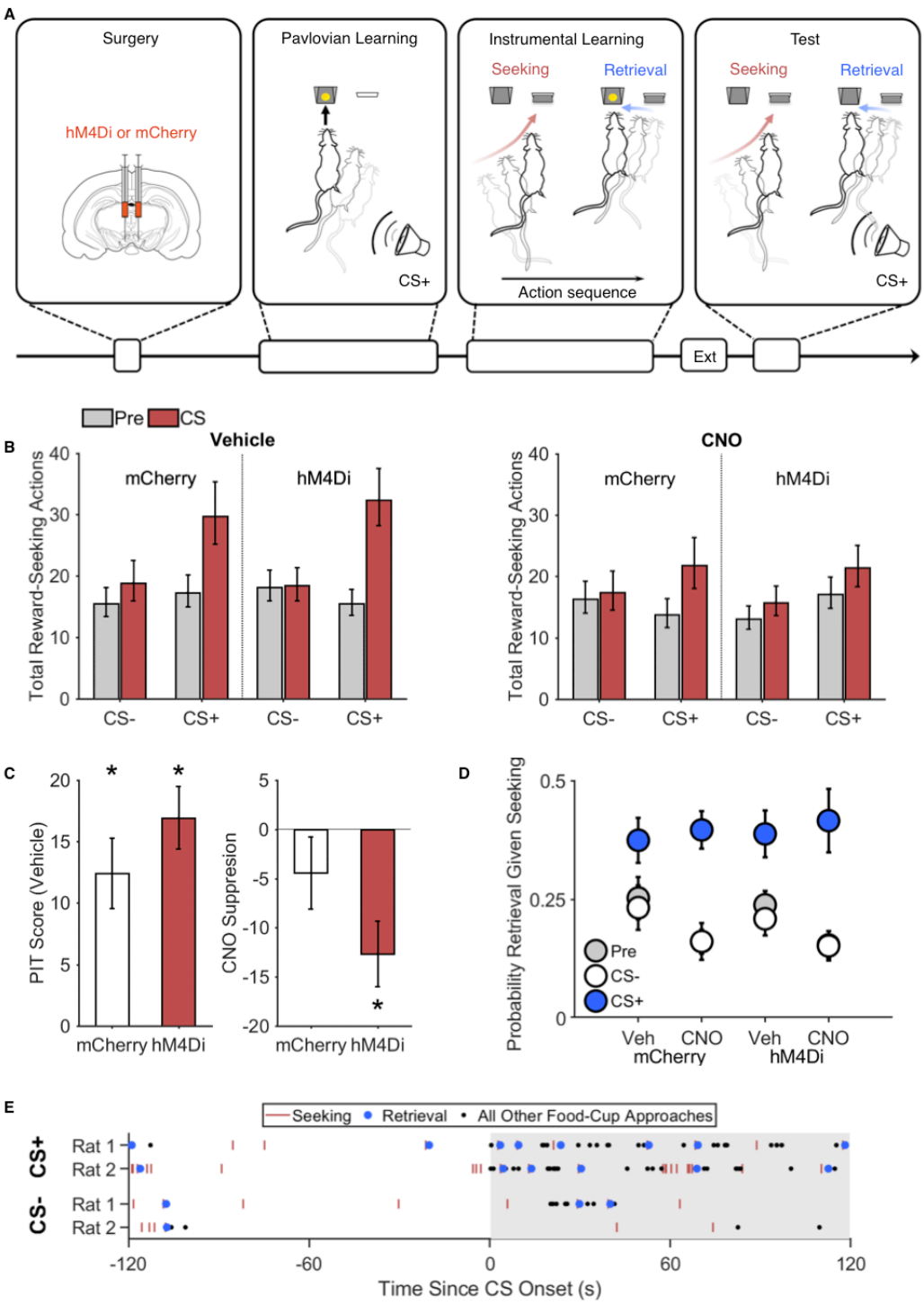
A
Surgery
hM4Di or mCherry
Pavlovian Learning
CS+
Instrumental Learning
Seeking
Retrieval
Action sequence
Test
Seeking
Retrieval
CS+
Ext
B
Pre
CS
Vehicle
CNO
mCherry
hM4Di
Total Reward-Seeking Actions
CS-
CS+
C
PIT Score (Vehicle)
CNO Suppresion
mCherry hM4Di
D
Probability Retrieval Given Seeking
Pre
CS-
CS+
Veh CNO Veh CNO
mCherry hM4Di
E
Seeking
Retrieval
All Other Food-Cup Approaches
CS+
CS-
Rat 1
Rat 2
Time Since CS Onset (s)

Figure 3. Chemogenetic inhibition of dopamine neurons in Pavlovian to instrumental transfer (PIT). **A.** Experimental design: Following viral vectors injections and recovery, rats received Pavlovian training, during which they learned to associate an auditory cue (CS+) with food pellet delivery. During instrumental learning rats learned to lever press according to the same task used in Experiment 1. Lever pressing was extinguished (Ext) before rats were submitted to a Pavlovian to Instrumental Transfer test, during which CS+ (or a control CS-) were randomly presented. We studied the effect of CNO (0 or 5mg/kg within-subject design) pre-treatment on reward-seeking and reward-retrieval at test. **B.** Chemogenetic inhibition of VTA dopamine neurons disrupts cue-motivated reward seeking. Reward-seeking actions in rats expressing the inhibitory DREADD hM4Di or mCherry following vehicle (left) or CNO (5mg/kg, right) treatment prior to test. Grey bars represent instrumental actions during the pre-CS period, red bars represent reward-seeking actions during CSs presentation. Error bars represent ±1 standard error of the estimated marginal means from the corresponding fitted generalized linear mixed-effects model. **C.** PIT expression is specifically impaired in hM4Di expressing Th:Cre+ rats. Left panel, analysis of PIT scores after vehicle treatment (Reward seeking during CS minus reward seeking during Pre-CS) show a significant elevation of reward-seeking behaviors during CS+ period for both groups; right panel, analysis of the CNO suppression score (PIT score following CNO treatment minus PIT score following vehicle treatment) show a significant effect of CNO only in hM4Di expressing rats. *$p<0.05$ vs 0. Error bars refer to ±1 between-subjects SEM. **D.** Proportions of reward-seeking actions that were immediately followed by an attempt to retrieve the reward during different PIT periods do not differ as a function of treatment or group, but were increased during CS+ as opposed to CS- or pre-CS periods. Error bars are as in C. **E.** Representative organization of instrumental behaviors in reward-retrieving and -seeking actions in PIT. Data show responses for two individual Th:Cre+ rats expressing mCherry and receiving vehicle treatment during pre-CS and CS presentation periods at test.

Pathway-specific inhibition of dopamine projections to NAc, but not mPFC, disrupts cue-motivated reward seeking in Pavlovian instrumental transfer

As previously reported (Mahler *et al*, *in press*), hM4Di expression in VTA dopamine neurons resulted in transport of DREADDs to axonal terminals in the NAc and mPFC (Figure 2). Therefore, we took advantage of this to investigate the roles of these two pathways in PIT performance, again distinguishing between the influence of reward-paired cues on reward seeking and reward retrieval. Guide cannulae were aimed at the NAc or mPFC in rats expressing hM4Di in VTA dopamine neurons (Experiment 3a; Figure 4A and Figure 4-Supplement 1). These rats underwent training and testing for PIT, as described above (Figure 4B), but were pretreated with intra-NAc or mPFC injections of CNO (1mM) or vehicle to achieve local inhibition of neurotransmitter release (Lichtenberg *et al*, 2017; Mahler *et al*, 2014;

Stachniak *et al*, 2014), an approach previously shown to be effective in inhibiting these specific dopamine projections (Mahler *et al*, *in press*). Figure 4C shows that, in hM4Di expressing rats, the CS+ specific increase in reward seeking was disrupted by CNO in a manner that depended on the microinjection site (Drug * CS Period * CS Type * Site interaction, $t(240) = -2.99$, $p = .003$; Supplementary File Table 4). After intracranial vehicle injections, rats showed a CS+ specific elevation in reward seeking (CS Period * CS Type interaction, $p < .001$), which did not differ significantly across vehicle injection sites (CS Period * CS Type * Site interaction, $p = .151$). Unlike with systemic CNO, the CS+ remained effective in eliciting reward seeking after CNO microinjection into the mPFC (CS Type * CS Period interaction, $p < .001$) or NAc (CS Type * CS Period interaction, $p < .001$). However, this effect was significantly attenuated after CNO in NAc, relative to mPFC (CS Period * CS Type * Site interaction, $p = .012$). A more focused analysis (Fig 4D) of the net change in pressing on CS+ trials (PIT score) confirmed that both NAc and mPFC rats showed significant elevations when pretreated with vehicle (one-sample t-tests; *p*s ≤ .001), and that this effect was suppressed after intra-NAc ($t(6) = -2.49$, $p = .047$), but not intra-mPFC ($t(8) = 0.34$, $p = .746$) CNO.

In a separate group of rats expressing mCherry only in VTA dopamine neurons we examined if these behavioral effects of CNO microinfusions were hM4Di-dependent (Experiment 3b). We found that there was no significant unconditional effects of injecting CNO into either the NAc or mPFC on CS+ induced changes in reward seeking (Figure 4-Supplement 2).

As in the previous experiment, instrumental reward-retrieval behavior was also influenced by noncontingent cue exposure (Figure 4E). Specifically, we found that the CS+ ($p < .001$) increased the probability that reward-seeking actions were followed by an attempt to retrieve reward, which was not affected by CNO at either site (*p*s > .187; Supplementary File Table 6; see Figure 4 –Supplement 3 and Supplementary File Table 5 for analysis of reward retrieval frequency). The CS- produced a decrease in the probability of reward retrieval compared to both pre-CS baseline ($p = .040$) and the CS+ ($p < .001$).

Given that inhibiting dopamine terminals in the NAc resulted in only a partial suppression of CS+ motivated reward seeking, but not its ability to promote reward retrieval, we hypothesized that variability across rats in sensitivity to the suppressive effects of this treatment may relate to their tendency to exhibit a motivational rather than cognitive form of PIT. We assumed that rats applying a motivational PIT strategy would respond to the CS+ by engaging in *exploratory* bouts of reward seeking (i.e., presses that were not followed by an immediate attempt to retrieve reward). Consistent with this, we found that for NAc rats, CS+ elicited increases in exploratory reward seeking (CS+ - pre-CS+) during the vehicle test were negatively correlated with CNO-induced suppression of reward seeking (all presses; $r = -0.81$, $p = .027$; Figure 4F). No such relationship was seen between exploratory seeking on vehicle day and sensitivity to CNO in mPFC rats ($r = -0.19$, $p = .618$).

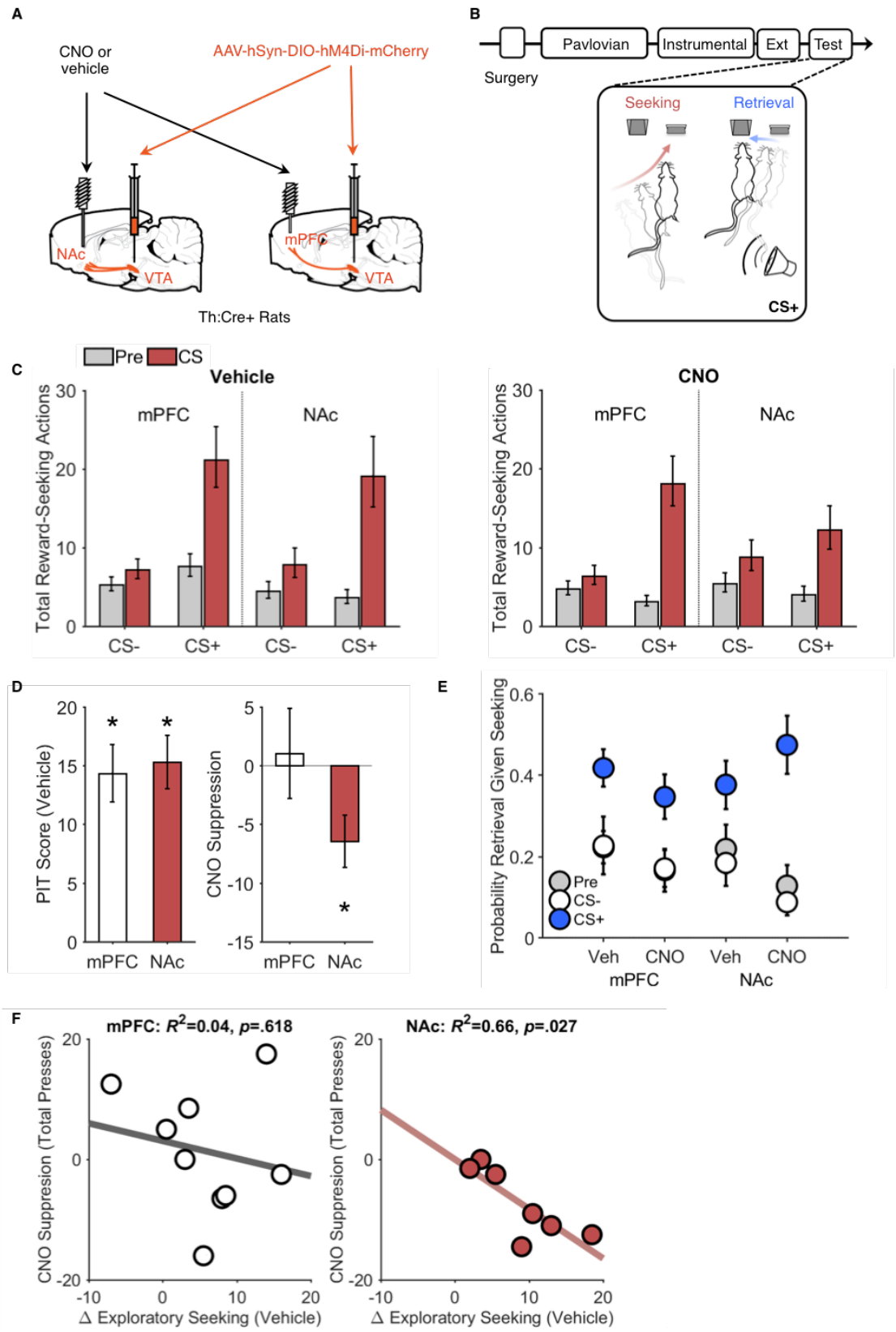
A
CNO or vehicle
AAV-hSyn-DIO-hM4Di-mCherry
NAc
VTA
mPFC
VTA
Th:Cre+ Rats
B
Pavlovian
Instrumental
Ext
Test
Surgery
Seeking
Retrieval
CS+
C
Pre
CS
Vehicle
CNO
mPFC
NAc
Total Reward-Seeking Actions
CS-
CS+
D
PIT Score (Vehicle)
CNO Suppression
mPFC
NAc
E
Probability Retrieval Given Seeking
Pre
CS-
CS+
Veh
CNO
mPFC
NAc
F
mPFC: R^2=0.04, p=.618
NAc: R^2=0.66, p=.027
CNO Suppresion (Total Presses)
Δ Exploratory Seeking (Vehicle)

Figure 4. Pathway specific chemogenetic inhibition of dopamine in PIT. **A.** Th:Cre+ rats initially received VTA AAV-hSyn-DIO-hM4Di-mCherry injections and were implanted with guide cannulas aimed at the nucleus accumbens (NAc) or medial prefrontal cortex (mPFC) for microinjection of CNO (1mM) or vehicle (veh) to inhibit dopamine release at terminals at test. **B.** Following surgery, rats underwent PIT training and testing as described above (Pavlovian: Pavlovian Learning; Instr: Instrumental Learning; Ext: Extinction). We analyzed the microstructural organization of behavior (reward-seeking and reward-retrieval actions) at test. **C.** Pathway specific inhibition of dopamine release in the nucleus accumbens (NAc) but not the medial prefrontal cortex (mPFC) disrupts reward seeking in PIT. Reward-seeking actions during PIT in rats expressing the inhibitory DREADD hM4Di and receiving CNO or vehicle microinfusions in either the NAc or mPFC. Grey bars represent instrumental actions during the pre-CS period, red bars represent reward-seeking during CS presentation. Error bars represent ±1 standard error of the estimated marginal means from the corresponding fitted generalized linear mixed-effects model. **D.** PIT expression is specifically impaired following NAc CNO treatment. Left panel, analysis of PIT scores followingy vehicle treatment (Reward seeking during CS minus reward seeking during Pre-CS) show a significant elevation of reward-seeking behaviors during CS+ period for both groups; right panel, analysis of the CNO suppression score (PIT score following CNO treatment minus PIT score following vehicle treatment) show a significant effect of CNO only when injected in the NAc. *p<0.05 vs 0. Error bars refer to ±1 between-subjects SEM.**E.** Proportions of reward-seeking actions that were immediately followed by an attempt to retrieve the reward during different PIT periods do not differ as a function of treatment or group, but were increased during CS+ as opposed to CS- or pre-CS periods. Error bars are as in D. **F.** Scatter plot showing the relationship of individual differences in the effect of CS+ on exploratory reward-seeking actions in the vehicle condition (Δ Exploratory Seeking; i.e., presses non followed by an attempt to collect reward) and the suppressive effect of CNO on reward seeking. Data points represent individual rats receiving intra-mPFC (left panel) or –NAc (right panel) CNO microinjections.

Inhibiting dopamine neurons spares use of outcome value when choosing between reward-seeking actions

The results of the previous experiments indicate that the mesolimbic dopamine system mediates the motivational influence of reward-paired cues but does not play a major role in using or modifying response-contingent reward expectancies. As an addition test of this hypothesis, we investigated dopamine's role in another important aspect of cognitive control over reward seeking – the ability to choose actions based on expected outcome value. Sensitivity to changes in the value of food rewards, induced by sensory-specific satiety, was assessed in a separate set of rats expressing mCherry or hM4Di in VTA dopamine neurons, following administration of CNO (5mg/kg) or vehicle (Figure 5A). We found that rats pressed significantly less for the devalued food than for the non-devalued food ($t(148) = -5.41$, $p < .001$; Figure

5B; Supplementary File Table 7). CNO treatment did not significantly affect sensitivity to devaluation in either hM4Di or mCherry rats (*p*s ≥ .146), indicating that this aspect of goal-directed decision making is not disrupted by dopamine neuron inhibition. Inhibiting VTA dopamine neurons also failed to disrupt sensitivity to devaluation during reinforced testing (see Figure 5-Supplement 1).

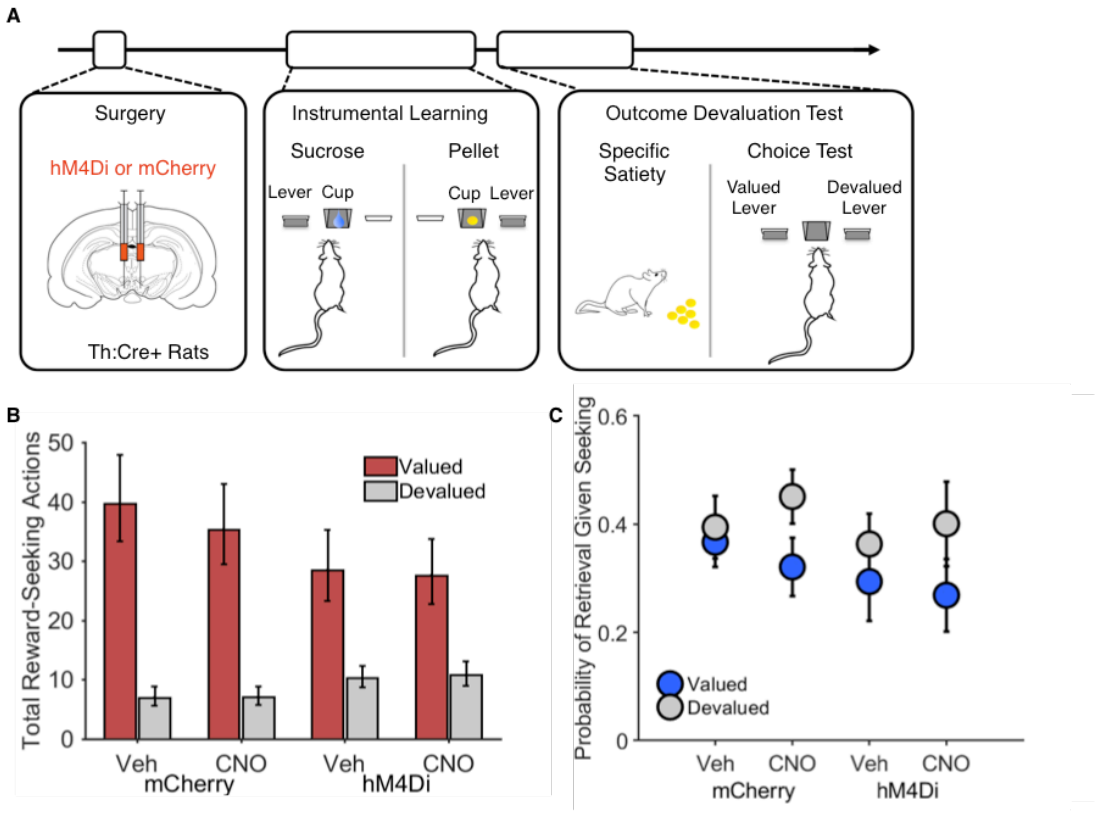

Figure 5. Chemogenetic inhibition of dopamine neurons in outcome specific devaluation test. **A.** Th:Cre+ rats received VTA injections of AAV-hSyn-DIO-hM4Di-mCherry or AAV-hSyn-DIO-mCherry. Following recovery, rats were trained on two distinct lever press actions for two different rewards (Instrumental Learning). Rats then underwent specific outcome devaluation testing following treatment with CNO (5mg/kg) or vehicle. **B.** Chemogenetic VTA dopamine inhibition does not affect outcome specific devaluation. Total reward seeking actions at test on the valued (red bars) and devalued (grey) levers in hM4di or mCherry expressing Th:Cre+ rats, following CNO (5 mg/kg) or vehicle treatments. Error bars represent ±1 standard error of the estimated marginal means from the corresponding fitted generalized linear mixed-effects model. **C.** Probability of engaging in a reward retrieval action given a seeking action for the different treatments and virus expression conditions for the valued and devalued levers. Error bars refer to ±1 between-subjects SEM.

We found VTA dopamine neuron inhibition also had no significant effect on rats' tendency to retrieve reward after pressing during devaluation

testing (Figure 5C; *p*s ≥ .109; Supplementary File Table 9; see Figure 5-Supplement 2 and Supplementary File Table 8 for analysis of reward retrieval *frequency*). Interestingly, the proportion of reward-seeking actions that were followed by an attempt to retrieve reward was slightly but significantly *higher* when rats were pursuing the devalued vs. the valued reward, $p = .040$. Thus, once a reward-seeking action was performed, expecting a devalued reward did not deter rats from attempting to collect it.

Discussion

The current study investigated the involvement of the mesocorticolimbic dopamine system in cognitive and motivational effects of reward-paired cues during expression of PIT. It is believed that reward-paired cues trigger reward seeking partly by raising expectations that reward-seeking actions will be effective in producing reward (Cartoni *et al*, 2016; Cartoni *et al*, 2013; Hogarth *et al*, 2014; Hogarth *et al*, 2015; Rescorla, 1994). We applied a novel analytic approach for assaying cue-induced changes in instrumental (response-contingent) reward expectancy and dissociating this influence from the response-invigorating effects of reward-paired cues on reward seeking. We show that rats are more likely to attempt to retrieve reward after lever pressing when they are given feedback that their behavior was effective in producing reward, relative to when no such feedback is given, bolstering the view that post-seeking reward retrieval reflects instrumental reward expectancy. We also confirmed our previous finding (Marshall *et al*, 2018) that noncontingent CS+ presentations during PIT testing both invigorated reward

seeking and increased the likelihood that individual lever-press responses were followed by an attempt to retrieve reward, suggesting that the CS+ was indeed able to inflate rats' expectations of response-contingent reward.

We found that chemogenetic inhibition of VTA dopamine neurons attenuated the ability of a CS+ to increase reward seeking, even though this cue continued to increase the likelihood that the few remaining presses were followed by attempts to retrieve reward. This suggests that the mesocorticolimbic dopamine system is not required for the cognitive process through which reward-paired cues alter instrumental reward expectancy, but does play an important and dissociable role in the motivational process through which such cues invigorate reward seeking. We observed a similar, albeit less pronounced, deficit in cue-motivated reward seeking when we inhibited dopamine terminals in the NAc, and, again, the ability of the CS+ to increase reward retrieval was spared.

It is worth noting that researchers often distinguish between general (nonspecific) vs. outcome-specific PIT effects (Balleine *et al*, 2007; Cartoni *et al*, 2016; Corbit and Balleine, 2016; Holmes *et al*, 2010). The ability of a CS+ to nonspecifically invigorate reward seeking, regardless of the identity of the outcome of the cue or instrumental response, is referred to as *general PIT*, and is believed to result from a state of heightened arousal or motivation (Corbit *et al*, 2016; Rescorla *et al*, 1967). In contrast, rats trained with multiple stimulus-outcome and action-outcome contingencies exhibit an outcome-specific form of PIT, in which cue-induced response invigoration is greater for actions that share the same outcome as the CS+ than for actions trained with a different outcome (Colwill *et al*, 1994; Kruse *et al*, 1983). *Specific PIT,*

therefore, requires use of cognitive, sensory-specific reward representations to guide response selection. Perhaps not surprisingly, it has been suggested that specific PIT depends more heavily than general PIT on cue-induced inflation of instrumental reward expectancy, or response efficacy (Cartoni *et al*, 2013). This is notable because previous PIT research has shown that dopamine receptor blockade disrupts the general, response-invigorating effects of reward-paired cues but leaves intact their ability to bias response selection in an outcome-specific manner (Ostlund *et al*, 2011). Likewise, the current findings indicate that dopamine signaling is not required for cue-induced inflation of instrumental reward expectancy, which helps strengthen the link between these two cognitive features of PIT. However, it would be a mistake to assume that cognitive processes only contribute to PIT in situations in which the outcome-specific nature of this effect is being assessed. For instance, using a nonspecific (single-reward type) PIT task, we have shown that rats use previous stimulus-reward interval learning to regulate the temporal patterning of reward seeking during CS+ presentations (Marshall *et al*, 2018). Moreover, we have found that reward-paired cues increase instrumental reward expectancy, reflected in post-seeking reward retrieval, regardless of whether specific (unpublished findings) or nonspecific PIT protocols are used (current findings, and Marshall *et al*, 2018). Such findings suggest that PIT is a multifaceted behavioral phenomenon that is not defined solely by the presence or absence of outcome-specificity.

Our findings are also compatible with previous studies showing that dopamine supports a motivational process that facilitates the expression of preparatory, or exploratory, reward-seeking behavior, but is relatively

unimportant for subsequent responses required for reward consumption (Fibiger and Phillips, 1986; Ikemoto and Panksepp, 1996; Veeneman *et al*, 2012). Fast-scan cyclic voltammetry studies also indicate that the mesolimbic dopamine system is strongly engaged during the initiation but not the completion of reward-seeking action sequences (Cacciapaglia *et al*, 2012; Collins *et al*, 2016; Klanker *et al*, 2015; Wassum *et al*, 2012). Likewise, prior studies using the PIT task have found that phasic dopamine release in the NAc is correlated with the vigor of cue-motivated lever pressing (Ostlund *et al*, 2014; Wassum *et al*, 2013) and not food cup approaches (Aitken *et al*, 2016), and that individual transient dopamine release events are temporally correlated with the execution of discrete lever-press responses (Ostlund *et al*, 2014). Such findings suggest that dopamine's role in regulating the pursuit of rewards varies as a function of reward proximity, contributing predominantly to the initiation of reward seeking. Our findings generally support this view, but also provide a critical test of dopamine's role in cue-motivated reward seeking. By focusing our analysis on a microstructurally-defined subset of reward-retrieval actions, we show that inhibiting VTA dopamine neurons attenuates cue-motivated reward seeking in a direct manner, rather than by disrupting cue-induced inflation of instrumental reward expectancy. Previous studies have shown that food-paired cues elicit dopamine release in the mPFC (Bassareo and Di Chiara, 1997; Feenstra *et al*, 1999), and electrophysiological findings directly implicate this region in expression of the PIT effect (Homayoun and Moghaddam, 2009). It is therefore notable that inhibiting dopamine projections to the mPFC had no effect on the ability of the CS+ to increase either reward seeking or reward retrieval. However, mPFC

lesion studies have typically found little evidence that this structure is necessary for PIT performance (Cardinal *et al*, 2003; Corbit and Balleine, 2003). Future research is therefore warranted to investigate whether the mPFC plays a causal role in PIT expression or is merely engaged during this task.

Our finding that inhibiting VTA dopamine neurons did not impact rats' sensitivity to outcome devaluation is consistent with previous findings (Dickinson *et al*, 2000; Lex and Hauber, 2009; Lex and Hauber, 2010; Wassum *et al*, 2011), even though regions innervated by this dopamine system, including the NAc and mPFC, are known to make important contributions to this feature of goal-directed decision making (Bradfield and Balleine, 2017; Sharpe *et al*, 2019). However, these results do not rule out the possibility that the mesolimbic dopamine system plays a more substantial role when goal-directed decisions are more complex and/or require greater cognitive resources (Cools, 2015; Floresco, 2013; Westbrook and Braver, 2016), which is an issue that deserves further investigation.

The present results may also have implications for understanding the role of dopamine in pathologies of behavioral control such as obsessive-compulsive disorder (OCD). In the signal attenuation model of OCD (Joel *et al*, 2001), rats learn that they can no longer rely on response-contingent cues to signal whether or not their reward-seeking behavior has been successful. When this happens, logical organization of their reward-seeking and reward-retrieval behavior disintegrates, and rats come to exhibit excessive reward seeking, typically without attempting to collect reward from the food cup. Importantly, blocking D1-dopamine receptors disrupts expression of these

compulsive (but incomplete) bouts of reward seeking, without affecting the production of complete bouts of seeking followed by reward retrieval, which continue to be infrequently performed on some test trials even though rats no longer have reason to be confident in the efficacy of their reward-seeking behavior (Joel *et al*, 2003). Considered in this light, our findings suggest that the mesolimbic dopamine system may mediate the tendency for reward-paired cues to promote this exploratory, or even compulsive, form of reward seeking, rather than more purposeful attempts to pursue and collect a reward that is expected. Consistent with this interpretation, we found that rats showing the greatest propensity to respond to the CS+ by engaging in exploratory seeking, without attempting to retrieve reward, showed the greatest suppressive effect of NAc dopamine inhibition. This link between mesolimbic dopamine signaling and exploratory cue-motivated reward seeking deserves further research, as it may reflect a biobehavioral marker useful for understanding and treating compulsive disorders like OCD and addiction (Joel *et al*, 2008; Robinson *et al*, 2014).

In summary, our findings indicate that the mesolimbic dopamine system selectively mediates the influence of reward-paired cues on response vigor but not instrumental reward expectancy, with dopaminergic inputs to the NAc playing a crucial role. This study also raises several new questions for future research. For instance, the current study shows that that cue-induced changes in reward expectancy do not dependent on VTA dopamine neuron function, but the neural circuitry underlying this influence remains to be elucidated. Further research is also needed to evaluate how cue-induced changes in instrumental reward expectancy relate to the cue-induced changes

in reward seeking. While we show that these effects are dissociable in terms of their dependence on mesolimbic dopamine signaling, this does not mean that reward expectations play no role in the expression of PIT. Indeed, we found that the PIT effect was relatively insensitive to NAc dopamine inhibition for rats that tended to respond to the CS+ with complete bouts of reward seeking and retrieval. This behavioral measure of reward expectancy therefore seems to be linked to processes responsible for controlling the execution of reward-*seeking* actions. Furthermore, as noted above, the connection between instrumental reward expectancy and reward seeking is believed to be even stronger for the outcome-specific PIT task (Cartoni *et al*, 2016; Cartoni *et al*, 2013). Finally, given growing evidence that independent neural circuits mediate the cognitive and motivational effects of reward-paired cues, it is important for future research to determine if they differentially contribute to the pathological forms of cue-elicited reward-seeking behavior apparent in food and drug addiction and related disorders.

Funding and Disclosure
Our work is supported by National Institutes of Health (NIH) grants MH106972, DK098709, AG045380, and DA029035. The authors declare no competing interests.

Acknowledgment
The authors acknowledge the assistance of Christy N. Munson in the acquisition of behavioral data.

References

Aitken TJ, Greenfield VY, Wassum KM (2016). Nucleus accumbens core dopamine signaling tracks the need‐based motivational value of food‐paired cues. *Journal of neurochemistry* **136**(5): 1026-1036.

Baguley T (2009). Standardized or simple effect size: what should be reported? *British Journal of Psychology* **100**: 603-617.

Balleine BW, Dickinson A (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* **37**(4-5): 407-419.

Balleine BW, Ostlund SB (2007). Still at the Choice-Point. *Annals of the New York Academy of Sciences* **1104**(1): 147-171.

Bassareo V, Di Chiara G (1997). Differential influence of associative and nonassociative learning mechanisms on the responsiveness of prefrontal and accumbal dopamine transmission to food stimuli in rats fed ad libitum. *J Neurosci* **17**(2): 851-861.

Berridge KC, Robinson TE (2016). Liking, wanting, and the incentive-sensitization theory of addiction. *American Psychologist* **71**(8): 670.

Bindra D (1974). A motivational view of learning, performance, and behavior modification. *Psychological review* **81**(3): 199.

Boisgontier MP, Cheval B (2016). The anova to mixed model transition. *Neuroscience & Biobehavioral Reviews* **68**: 1004-1005.

Bolker BM, Brooks ME, Clark CJ, Geange SW, Poulsen JR, Stevens MHH*, et al* (2008). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology & Evolution* **24**(3): 127-135.

Bradfield L, Balleine B (2017). The Learning and Motivational Processes Controlling Goal-Directed Action and Their Neural Bases. *Decision Neuroscience*. Elsevier, pp 71-80.

Cacciapaglia F, Saddoris MP, Wightman RM, Carelli RM (2012). Differential dopamine release dynamics in the nucleus accumbens core and shell track distinct aspects of goal-directed behavior for sucrose. *Neuropharmacology* **62**(5-6): 2050-2056.

Cardinal RN, Parkinson JA, Marbini HD, Toner AJ, Bussey TJ, Robbins TW*, et al* (2003). Role of the anterior cingulate cortex in the control over behavior by Pavlovian conditioned stimuli in rats. *Behav Neurosci* **117**(3): 566-587.

Cartoni E, Balleine B, Baldassarre G (2016). Appetitive Pavlovian-instrumental Transfer: A review. *Neurosci Biobehav Rev* **71**: 829-848.

Cartoni E, Puglisi-Allegra S, Baldassarre G (2013). The three principles of action: a Pavlovian-instrumental transfer hypothesis. *Frontiers in Behavioral Neuroscience* **7**: 153.

Cocker PJ, Hosking JG, Murch WS, Clark L, Winstanley CA (2016). Activation of dopamine D4 receptors within the anterior cingulate cortex enhances the erroneous expectation of reward on a rat slot machine task. *Neuropharmacology* **105**: 186-195.

Collins AL, Greenfield VY, Bye JK, Linker KE, Wang AS, Wassum KM (2016). Dynamic mesolimbic dopamine signaling during action sequence learning and expectation violation. *Scientific reports* **6**.

Colwill RM, Motzkin DK (1994). Encoding of the unconditioned stimulus in Pavlovian conditioning. *Animal Learning & Behavior* **22**(4): 384-394.

Cools R (2015). The cost of dopamine for dynamic cognitive control. *Current Opinion in Behavioral Sciences* **4**: 152-159.

Corbit LH, Balleine BW (2003). The role of prelimbic cortex in instrumental conditioning. *Behav Brain Res* **146**(1-2): 145-157.

Corbit LH, Balleine BW (2016). Learning and Motivational Processes Contributing to Pavlovian-Instrumental Transfer and Their Neural Bases: Dopamine and Beyond. *Curr Top Behav Neurosci* **27**: 259-289.

Coxe S, West SG, Aiken LS (2009). The analysis of count data: A gentle introduction to Poisson regression and its alternatives. *Journal of Personality Assessment* **91**(2): 121-136.

Delamater AR, Oakeshott S (2007). Learning about multiple attributes of reward in Pavlovian conditioning. *Annals of the New York Academy of Sciences* **1104**(1): 1-20.

Dickinson A, Balleine B (1994). Motivational control of goal-directed action. *Animal Learning & Behavior* **22**(1): 1-18.

Dickinson A, Smith J, Mirenowicz J (2000). Dissociation of Pavlovian and instrumental incentive learning under dopamine antagonists. *Behavioral neuroscience* **114**(3): 468.

Everitt BJ, Dickinson A, Robbins TW (2001). The neuropsychological basis of addictive behaviour. *Brain Research Reviews* **36**(2-3): 129-138.

Feenstra MG, Teske G, Botterblom MH, De Bruin JP (1999). Dopamine and noradrenaline release in the prefrontal cortex of rats during classical aversive and appetitive conditioning to a contextual stimulus: interference by novelty effects. *Neurosci Lett* **272**(3): 179-182.

Fibiger H, Phillips A (1986). Reward, motivation, cognition: psychobiology of mesotelencephalic dopamine systems. *Comprehensive Physiology*.

Floresco SB (2013). Prefrontal dopamine and behavioral flexibility: shifting from an "inverted-U" toward a family of functions. *Front Neurosci* **7**: 62.

Floresco SB, Magyar O, Ghods-Sharifi S, Vexelman C, Tse MTL (2005). Multiple Dopamine Receptor Subtypes in the Medial Prefrontal Cortex of the Rat Regulate Set-Shifting. *Neuropsychopharmacology* **31**: 297.

Hogarth L, Balleine BW, Corbit LH, Killcross S (2013). Associative learning mechanisms underpinning the transition from recreational drug use to addiction. *Annals of the New York Academy of Sciences* **1282**(1): 12-24.

Hogarth L, Retzler C, Munafò MR, Tran DMD, Troisi JR, Rose AK*, et al* (2014). Extinction of cue-evoked drug-seeking relies on degrading hierarchical instrumental expectancies. *Behaviour Research and Therapy* **59**: 61-70.

Hogarth L, Troisi II JR (2015). A Hierarchical Instrumental Decision Theory of Nicotine Dependence. In: Balfour DJK, Munafò MR (eds). *The Neurobiology and Genetics of Nicotine and Tobacco*. Springer International Publishing: Cham, pp 165-191.

Holmes NM, Marchand AR, Coutureau E (2010). Pavlovian to instrumental transfer: A neurobehavioural perspective. *Neuroscience & Biobehavioral Reviews* **34**(8): 1277-1295.

Homayoun H, Moghaddam B (2009). Differential representation of Pavlovian–instrumental transfer by prefrontal cortex subregions and striatum. *European Journal of Neuroscience* **29**(7): 1461-1476.

Ikemoto S, Panksepp J (1996). Dissociations between appetitive and consummatory responses by pharmacological manipulations of reward-relevant brain regions. *Behavioral neuroscience* **110**(2): 331.

Joel D, Avisar A (2001). Excessive lever pressing following post-training signal attenuation in rats: a possible animal model of obsessive compulsive disorder? *Behavioural brain research* **123**(1): 77-87.

Joel D, Doljansky J (2003). Selective Alleviation of Compulsive Lever-Pressing in Rats by D 1, but not D 2, Blockade: Possible Implications for the Involvement of D 1 Receptors in Obsessive–Compulsive Disorder. *Neuropsychopharmacology* **28**(1): 77.

Joel D, Stein DJ, Schreiber R (2008). CHAPTER 6 - Animal Models of Obsessive–Compulsive Disorder: From Bench to Bedside via Endophenotypes and Biomarkers A2 - McArthur, Robert A. In: Borsini F (ed). *Animal and Translational Models for CNS Drug Discovery*. Academic Press: San Diego, pp 133-164.

Klanker M, Sandberg T, Joosten R, Willuhn I, Feenstra M, Denys D (2015). Phasic dopamine release induced by positive feedback predicts individual differences in reversal learning. *Neurobiology of Learning and Memory* **125**: 135-145.

Kruse JM, Overmier JB, Konz WA, Rokke E (1983). Pavlovian conditioned stimulus effects upon instrumental choice behavior are reinforcer specific. *Learning and motivation* **14**(2): 165-181.

LeBlanc KH, Ostlund SB, Maidment NT (2012). Pavlovian-to-instrumental transfer in cocaine seeking rats. *Behavioral neuroscience* **126**(5): 681.

Lex A, Hauber W (2008). Dopamine D1 and D2 receptors in the nucleus accumbens core and shell mediate Pavlovian-instrumental transfer. *Learn Mem* **15**(7): 483-491.

Lex B, Hauber W (2009). The role of dopamine in the prelimbic cortex and the dorsomedial striatum in instrumental conditioning. *Cerebral cortex* **20**(4): 873-883.

Lex B, Hauber W (2010). The role of nucleus accumbens dopamine in outcome encoding in instrumental and Pavlovian conditioning. *Neurobiol Learn Mem* **93**(2): 283-290.

Lichtenberg NT, Pennington ZT, Holley SM, Greenfield VY, Cepeda C, Levine MS*, et al* (2017). Basolateral Amygdala to Orbitofrontal Cortex Projections Enable Cue-Triggered Reward Expectations. *The Journal of Neuroscience* **37**(35): 8374-8384.

Liljeholm M, O'Doherty JP (2011). Subcortical contributions to the motivational and cognitive control of instrumental performance by Pavlovian and discriminative stimuli. *Neural basis of motivational and cognitive control Cambridge (Massachusetts): The MIT Press*: 149-162.

Mahler SV, Brodnik ZD, Cox BM, Buchta WC, Bentzley BS, Quintanilla J*, et al* (*in press*). Chemogenetic manipulations of ventral tegmental area dopamine neurons reveal multifaceted roles in cocaine abuse. . *bioRxiv*.

Mahler SV, Vazey EM, Beckley JT, Keistler CR, McGlinchey EM, Kaufling J*, et al* (2014). Designer receptors show role for ventral pallidum input to ventral tegmental area in cocaine seeking. *Nat Neurosci* **17**(4): 577-585.

Marshall AT, Ostlund SB (2018). Repeated cocaine exposure dysregulates cognitive control over cue-evoked reward-seeking behavior during Pavlovian-to-instrumental transfer. *Learning & Memory* **25**(9): 399-409.

Matell MS, Della Valle RB (2018). Temporal specificity in Pavlovian-to-instrumental transfer. *Learning & Memory* **25**(1): 8-20.

Nadler N, Delgado MR, Delamater AR (2011). Pavlovian to instrumental transfer of control in a human learning task. *Emotion* **11**(5): 1112.

Ostlund SB, LeBlanc KH, Kosheleff AR, Wassum KM, Maidment NT (2014). Phasic mesolimbic dopamine signaling encodes the facilitation of incentive motivation produced by repeated cocaine exposure. *Neuropsychopharmacology* **39**(10): 2441-2449.

Ostlund SB, Maidment NT (2011). Dopamine Receptor Blockade Attenuates the General Incentive Motivational Effects of Noncontingently Delivered Rewards and Reward-Paired Cues Without Affecting Their Ability to Bias Action Selection. *Neuropsychopharmacology* **37**: 508.

Pinheiro J, Bates D (2000). *Mixed-effects models in S and S-Plus* Springer: New York.

Radke AK, Kocharian A, Covey DP, Lovinger DM, Cheer JF, Mateo Y*, et al* (2018). Contributions of nucleus accumbens dopamine to cognitive flexibility. *European Journal of Neuroscience*.

Rescorla RA (1994). Control of instrumental performance by Pavlovian and instrumental stimuli. *Journal of Experimental Psychology: Animal Behavior Processes* **20**(1): 44.

Rescorla RA, Solomon RL (1967). Two-process learning theory: Relationships between Pavlovian conditioning and instrumental learning. *Psychological review* **74**(3): 151.

Robinson TE, Yager LM, Cogan ES, Saunders BT (2014). On the motivational properties of reward cues: individual differences. *Neuropharmacology* **76**: 450-459.

Sharpe MJ, Schoenbaum G (2018). Evaluation of the hypothesis that phasic dopamine constitutes a cached-value signal. *Neurobiology of learning and memory* **153**: 131-136.

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. *Annual Review of Psychology* **70**(1): null.

Stachniak TJ, Ghosh A, Sternson SM (2014). Chemogenetic synaptic silencing of neural circuits localizes a hypothalamus-->midbrain pathway for feeding behavior. *Neuron* **82**(4): 797-808.

Veeneman MMJ, van Ast M, Broekhoven MH, Limpens JHW, Vanderschuren LJMJ (2012). Seeking–taking chain schedules of cocaine and sucrose self-administration: effects of reward size, reward omission, and α-flupenthixol. *Psychopharmacology* **220**(4): 771-785.

Wassum KM, Ostlund SB, Balleine BW, Maidment NT (2011). Differential dependence of Pavlovian incentive motivation and instrumental incentive learning processes on dopamine signaling. *Learning & memory* **18**(7): 475-483.

Wassum KM, Ostlund SB, Loewinger GC, Maidment NT (2013). Phasic mesolimbic dopamine release tracks reward seeking during expression of Pavlovian-to-instrumental transfer. *Biological psychiatry* **73**(8): 747-755.

Wassum KM, Ostlund SB, Maidment NT (2012). Phasic mesolimbic dopamine signaling precedes and predicts performance of a self-initiated action sequence task. *Biological psychiatry* **71**(10): 846-854.

Westbrook A, Braver TS (2016). Dopamine Does Double Duty in Motivating Cognitive Effort. *Neuron* **89**(4): 695-710.

Witten IB, Steinberg EE, Lee SY, Davidson TJ, Zalocusky KA, Brodsky M*, et al* (2011). Recombinase-driver rat lines: tools, techniques, and optogenetic application to dopamine-mediated reinforcement. *Neuron* **72**(5): 721-733.

Wyvell CL, Berridge KC (2000). Intra-accumbens amphetamine increases the conditioned incentive salience of sucrose reward: enhancement of reward "wanting" without enhanced "liking" or response reinforcement. *J Neurosci* **20**(21): 8122-8130.

SUPPLEMENTARY RESULTS

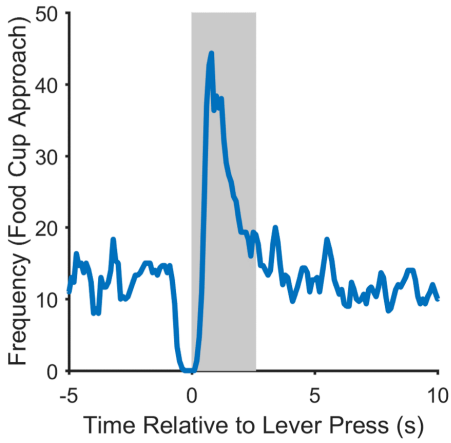

Figure 1-Supplement 1. Frequency of food-cup approaches as a function of time surrounding a lever press during the non CSs periods for the vehicle condition of Experiment 2. Shewhart process control chart analyses were used to determine the times when food-cup approach behavior was elevated with respect to constant background rates. This value approximated 2.5 s following each lever press, as shown by the shaded box.

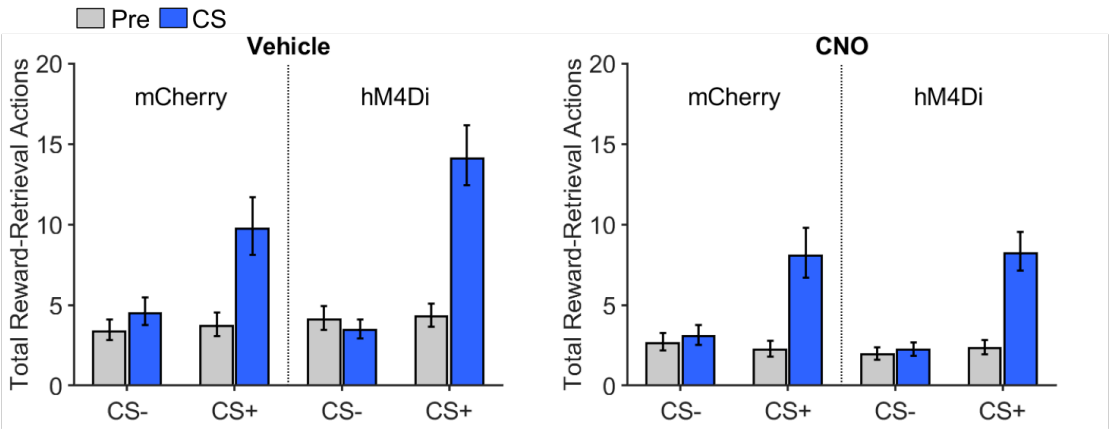

Figure 3-Supplement 1. Microstructural organization of behavior during PIT. Figure shows reward-retrieval actions in rats expressing the inhibitory DREADD hM4Di or mCherry following vehicle or CNO (5mg/kg) treatment prior to test (Experiment 2). Grey bars represent instrumental actions during the pre-CS period and blue bars represent reward-retrieval actions during CS presentation. Error bars represent ±1 standard error of the estimated marginal means from the corresponding fitted generalized linear mixed-effects model. We found that the CS+ but not the CS- strongly increased the frequency of reward-retrieval actions (CS Period * CS Type interaction, $t(240) = 7.84$, $p < .001$), an effect that was not dependent on the vector group or CNO

pretreatment (interactions involving these factors, *p*s ≥ .145; see Supplementary File Table 2 for full model output).

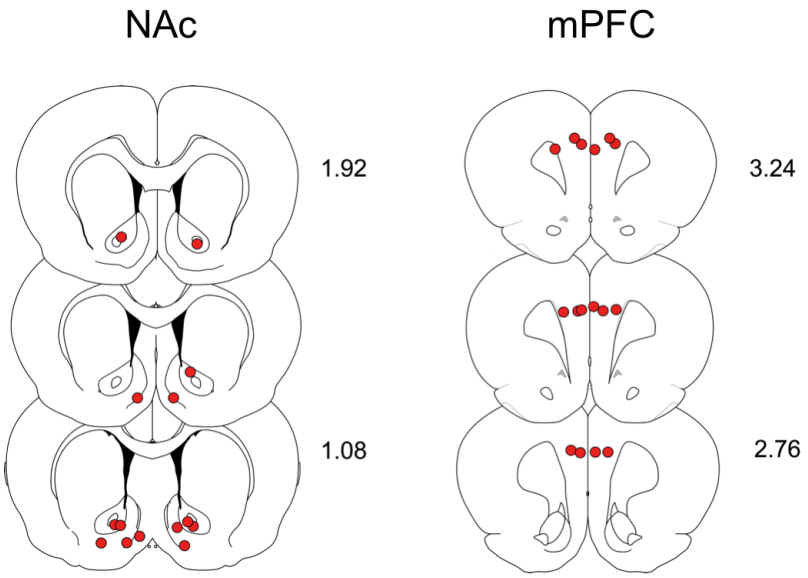

Figure 4-Supplement 1. Cannulae placements for Experiment 3a hM4Di expressing rats. Individual placements in nucleus accumbens (NAc) and medial prefrontal cortex (mPFC).

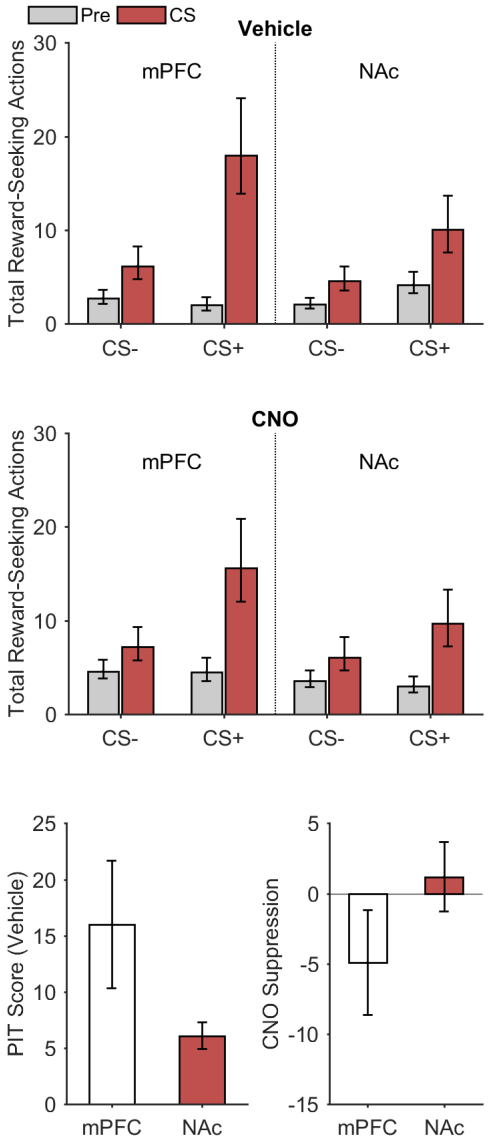

Figure 4-Supplement 2. Experiment 3b PIT Testing in mCherry expressing Th:Cre+ rats with localized (NAc, and mPFC; ns = 6) CNO micro-infusions. (A,B) Grey bars represent instrumental actions during the pre-CS period and red bars represent reward-taking lever presses during CS presentation. There was an overall CS Type * CS Period interaction, $t(176) = 4.51$, $p < .001$, which was not significantly moderated by site or drug, *p*s ≥ .060. The CS+ was much more likely to invigorate reward-seeking behavior compared to the CS-, especially in the mPFC [$t(88) = 4.37$, $p < .001$; NAc: $t(88) = 1.88$, $p = .063$], interactions that were not moderated by drug, *p*s ≥ .189. Error bars represent ±1 standard error of the estimated marginal means from the corresponding fitted generalized linear mixed-effects model. (C) While both groups showed positive PIT scores under the vehicle condition, CNO did not suppress behavior in this NAc control group as it did in the hM4Di group (see main text). Therefore, suppression by local CNO micro-infusions in the NAc was

dependent on whether the DREADD virus was expressed in that region. Error bars in D-F represent ± 1 between-subjects SEM.

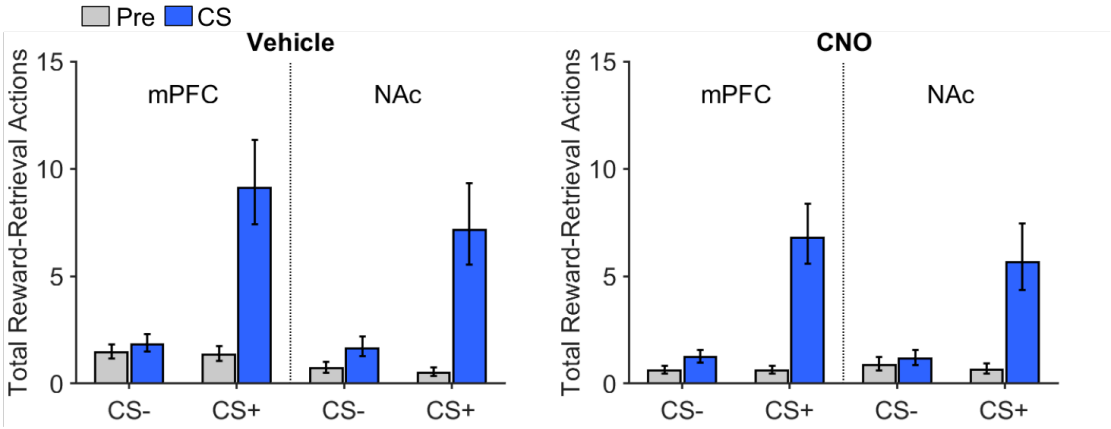

Figure 4-Supplement 3. Reward-retrieval actions during PIT test in rats expressing the inhibitory DREADD hM4Di and receiving CNO or vehicle microinfusions in either the NAc or mPFC (Experiment 3a). Grey bars represent instrumental actions during the pre-CS period and blue bars represent reward-taking lever presses during CS presentation. Error bars represent ±1 standard error of the estimated marginal means from the corresponding fitted generalized linear mixed-effects model. The cue-induced elevation in reward taking frequency was specific to the CS+ (CS Type x CS Period interaction, $t(240) = 7.45$, $p < .001$), and was not significantly modulated by CNO at either injection site (*p*s > .570; See Supplementary File Table 5).

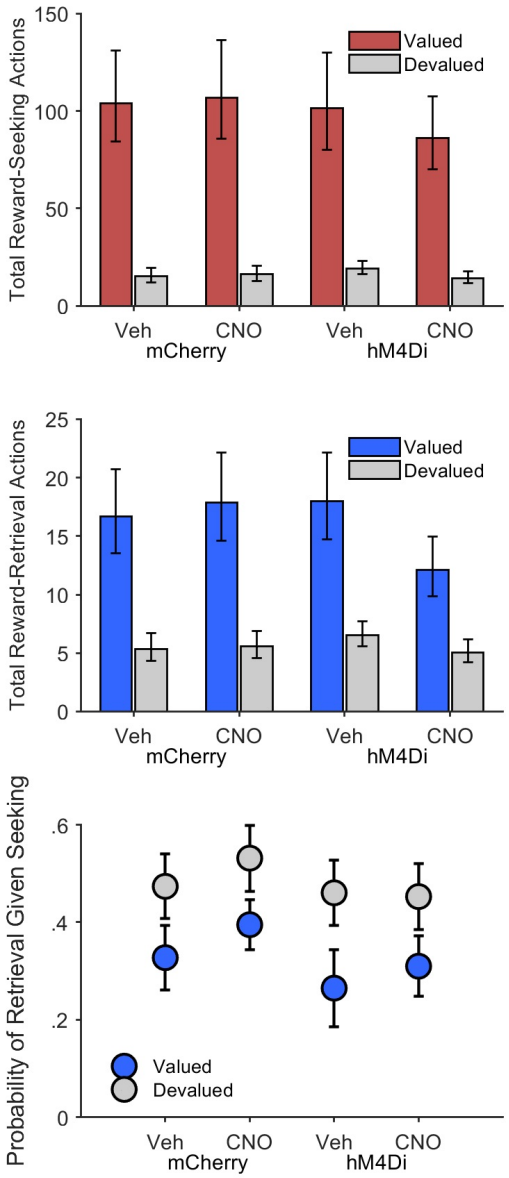

Figure 5-Supplement 1. Reinforced phase of outcome devaluation testing expressing the inhibitory DREADD hM4Di or mCherry following vehicle or CNO (5mg/kg) treatment. (A) There was significantly greater reward-seeking behavior directed at the valued lever compared to the devalued lever, $t(148) = -5.55$, $p < .001$, which was not moderated by group or drug, ps ≥ .095. (B) Similarly, reward-retrieval behavior was elevated following reward-seeking behaviors on the valued lever compared to the devalued lever, $t(148) = -5.46$, $p < .001$, which also did not depend on group or drug, ps ≥ .128. In A-B, error bars represent ±1 standard error of the estimated marginal means from the corresponding fitted generalized linear mixed-effects model. (C) The probability of performing a reward-retrieval behavior given a reward-seeking action was significantly elevated for devalued-lever actions, $t(131) = 4.11$, $p <$

.001, which also was not dependent on group and/or drug condition, *p*s ≥ .679.

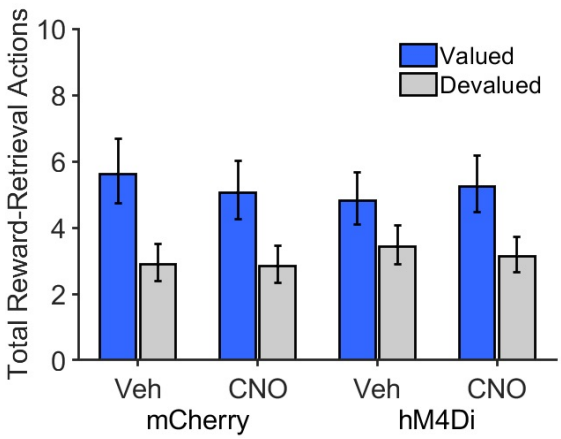

Figure 5-Supplement 2. During specific outcome devaluation test, rats from Experiment 4 exhibited fewer attempts to retrieve the devalued than the non-devalued reward ($t(148) = -4.40$, $p < .001$), and CNO failed to suppress retrieval attempts in either hM4Di or mCherry rats (*p*s ≥ .417; See Supplementary File Table 8). Error bars represent ±1 standard error of the estimated marginal means from the corresponding fitted generalized linear mixed-effects model.