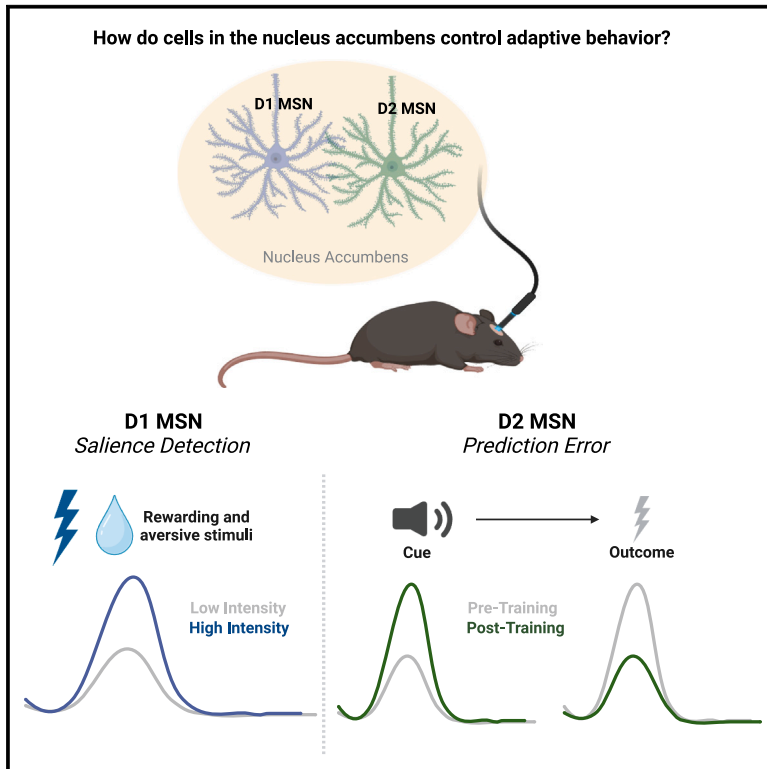


# D1 and D2 medium spiny neurons in the nucleus accumbens core have distinct and valence-independent roles in learning

## Graphical abstract



## Highlights

- D1 and D2 MSNs in the NAc of mice do not signal reward or valence
- D1 and D2 MSNs do not have opposing activity patterns during behavior
- D1 MSN responses are evoked by stimuli regardless of valence or contingency
- D2 MSNs track valence-free prediction error at the individual neuron level

## Authors

Jennifer E. Zachry, Munir Gunes Kutlu, Hye Jean Yoon, ..., P. Jeffrey Conn, Brad A. Grueter, Erin S. Calipari

## Correspondence

erin.calipari@vanderbilt.edu

## In brief

A large body of work has focused on how the nucleus accumbens (NAc) is critical to reward encoding. Zachry, Kutlu, et al. show that medium spiny neurons (MSNs) within the NAc of mice do not signal reward. Rather, D1 MSN responses are evoked by salient stimuli, while D2 MSNs track prediction errors.

Article

# D1 and D2 medium spiny neurons in the nucleus accumbens core have distinct and valence-independent roles in learning

Jennifer E. Zachry,<sup>1,2,8</sup> Munir Gunes Kutlu,<sup>1,2,7,8</sup> Hye Jean Yoon,<sup>1,2</sup> Michael Z. Leonard,<sup>1,2</sup> Maxime Chevée,<sup>1,2</sup> Dev D. Patel,<sup>1,2</sup> Anthony Gaidici,<sup>1,2</sup> Veronika Kondev,<sup>1,2</sup> Kimberly C. Thibeault,<sup>1,2</sup> Rishik Bethi,<sup>1,2</sup> Jennifer Tat,<sup>1,2</sup> Patrick R. Melugin,<sup>1,2</sup> Atagun U. Isiktas,<sup>1,2,3</sup> Max E. Joffe,<sup>1,2,4</sup> Denise J. Cai,<sup>5</sup> P. Jeffrey Conn,<sup>1,2</sup> Brad A. Grueter,<sup>1,2,6</sup> and Erin S. Calipari<sup>1,2,9,\*</sup>

<sup>1</sup>Department of Pharmacology, Vanderbilt University, Nashville, TN 37232, USA

<sup>2</sup>Vanderbilt Center for Addiction Research, Vanderbilt University, Nashville, TN 37232, USA

<sup>3</sup>Department of Neuroscience, Yale University, New Haven, CT 06520, USA

<sup>4</sup>Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA 15219, USA

<sup>5</sup>Nash Family Department of Neuroscience, Icahn School of Medicine, Mount Sinai, New York, NY 10029, USA

<sup>6</sup>Department of Anesthesiology, Vanderbilt University Medical Center, Nashville, TN 37232, USA

<sup>7</sup>Present address: Department of Cell Biology and Neuroscience, Rowan-Virtua School of Osteopathic Medicine, Rowan-Virtua School of Translational Biomedical Engineering and Sciences, Rowan University, Stratford, NJ 08084, USA

<sup>8</sup>These authors contributed equally

<sup>9</sup>Lead contact

\*Correspondence: [erin.calipari@vanderbilt.edu](mailto:erin.calipari@vanderbilt.edu)

<https://doi.org/10.1016/j.neuron.2023.11.023>

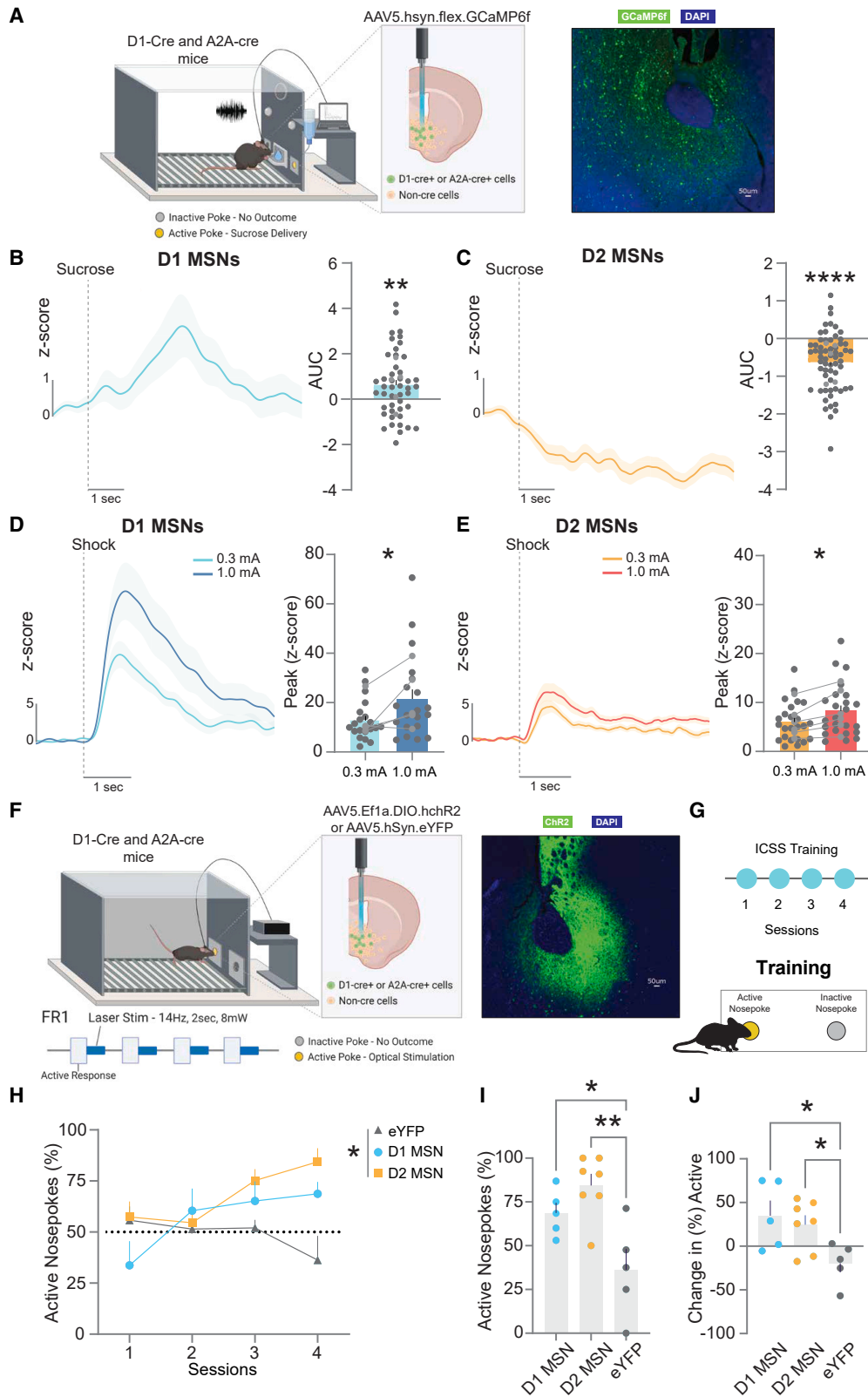
## SUMMARY

At the core of value-based learning is the nucleus accumbens (NAc). D1- and D2-receptor-containing medium spiny neurons (MSNs) in the NAc core are hypothesized to have opposing valence-based roles in behavior. Using optical imaging and manipulation approaches in mice, we show that neither D1 nor D2 MSNs signal valence. D1 MSN responses were evoked by stimuli regardless of valence or contingency. D2 MSNs were evoked by both cues and outcomes, were dynamically changed with learning, and tracked valence-free prediction error at the population and individual neuron level. Finally, D2 MSN responses to cues were necessary for associative learning. Thus, D1 and D2 MSNs work in tandem, rather than in opposition, by signaling specific properties of stimuli to control learning.

## INTRODUCTION

The ability to adaptively navigate an environment relies on associative learning, where animals form associations between predictive cues and external stimuli.<sup>1,2</sup> This process is at the core of nearly all adaptive behavior, and its dysregulation is also a key feature of a wide range of psychopathologies.<sup>3–9</sup> Associative learning is dependent on both the ability to identify the valence of external stimuli, as well as various other factors such as salience, prediction, and prediction error.<sup>10–16</sup> These valence-independent factors are important for helping animals navigate changing environments, where the same stimulus can often elicit opposite behavioral responses depending on the context in which it is encountered.<sup>17,18</sup> For example, an unavoidable footshock will induce freezing, but an avoidable footshock will elicit an escape response in rodents—even though the footshock in both cases has negative valence.<sup>18–21</sup> Without understanding the relationship between stimulus processing and behavioral action, our understanding of neural circuit control of behavior and its dysregulation in disease is incomplete.

The nucleus accumbens (NAc) is a hub for learning, selecting, and executing goal-oriented behaviors associated with both appetitive and aversive stimuli.<sup>22–26</sup> The NAc is a heterogeneous region primarily composed of medium spiny neurons (MSNs) that are classified based on their expression of D1- (which signal through  $G_{\alpha s}$ ) or D2-type dopamine receptors (which signal through  $G_{\alpha i}$ ).<sup>27–30</sup> Critically, these populations integrate information from dopaminergic inputs from the midbrain and glutamatergic inputs from across the brain, which combine to control their activity patterns.<sup>31,32</sup> The current understanding of D1 and D2 MSNs has been largely derived from work in the dorsal striatum showing that these two populations are segregated based on their projection targets into the direct (D1 MSNs) and indirect (D2 MSNs) pathways that drive action initiation and inhibition, respectively.<sup>33–36</sup> However, this pathway segregation is more nuanced in the NAc, and both D1 and D2 MSNs project to overlapping downstream areas, such as the ventral pallidum.<sup>30</sup> These fundamental organizational differences preclude our ability to infer MSN subtype-specific behavioral control based on this previous work.



(legend on next page)

In the NAc, much of our understanding of the role these populations play in behavioral control has been based on their role in reward-based behaviors and in response to drugs of abuse. Together these studies that span physiology, transcriptional plasticity, and pharmacological/optical manipulations have led to the hypothesis that these populations have opposing roles in behavior with D1 MSNs promoting reward and D2 MSNs promoting aversion/preventing reward seeking.<sup>37–45</sup> In this framework, these MSN populations function in opposition to one another, with D1 MSNs increasing and D2 MSNs decreasing activity in response to reward-based stimuli. However, emerging evidence is beginning to show that the functions of these neurons extend beyond simple reward-based coding,<sup>46–53</sup> and it is currently unclear exactly what environmental factors elicit responses in these populations and how these temporal dynamics are linked to behavioral control. By combining these recording/manipulation approaches with behaviors that include both operant and Pavlovian contingencies, we show the precise valence-independent role that these populations play in learning.

## RESULTS

### Aversive stimuli evoke both D1 and D2 MSN responses

To record temporally defined neural activity in awake and behaving animals, the genetically encoded calcium indicator GCaMP6f was selectively expressed in either D1 or D2 MSNs in the NAc core (using D1-cre or A2A-Cre mice, respectively). Because D2 receptors are expressed in some interneuron populations in the NAc, A2A-Cre mice were used to select for D2 MSNs, specifically.<sup>54</sup> Using fiber photometry, population-level calcium transients were monitored in awake and behaving animals (Figure 1A; see Figure S1A). First, D1 and D2 MSNs were recorded during a positive reinforcement task where mice emitted an operant response during an auditory cue for a sucrose reinforcer. In line with previous valence-based predictions,<sup>43,44</sup> at the time of sucrose collection D1 MSNs showed a positive response (Figure 1B), while D2 MSNs showed a small decrease (Figure 1C). However, when mice were presented with unsignaled footshocks at random intervals of varying intensity (0.3 or 1.0 mA), both D1 MSNs (Figure 1D) and D2

MSNs (Figure 1E) showed positive shock-evoked responses that increased with increasing shock intensity (Figures S1D and S1E).

### Stimulation of D1 and D2 MSNs supports reinforcement

Next, to test whether the stimulation of D1 MSNs or D2 MSNs supported reinforcement, the excitatory opsin (channelrhodopsin; ChR2) was expressed in each MSN population (Figures S1B and S1F–S1K). Mice nose poked for optical stimulation of either D1 or D2 MSNs (470 nm, 2 s, 14 Hz, 8 mW) (Figures 1F and 1G). Stimulation of either D1 or D2 MSNs supported reinforcement, while enhanced yellow fluorescent protein (eYFP) controls performed at chance levels over training (Figure 1H). Mice showed higher response rates when behavior was reinforced by either D1 or D2 MSN optical stimulation as compared with the unreinforced, eYFP condition (Figure 1I). Finally, mice in both the D1 MSN and D2 MSN stimulation groups increased their responses with additional training sessions as compared with eYFP controls (Figure 1J). Together, these data showed that there is no clear demarcation between D1 and D2 MSN functions in the NAc core based on valence. D1 MSN responses (in the positive direction) were evoked by both appetitive and aversive stimuli, rather than decreasing to aversive stimuli as would be predicted for a reward-based signal.<sup>55,56</sup> Further, D2 MSN stimulation supported reinforcement, suggesting that these neurons were not transmitting an aversive signal or preventing motivated responding.

### D2, but not D1, MSNs scale with learning

One of the difficulties with defining the precise role of neural populations in behavioral control is the ability to dissociate multiple behavioral factors. Thus, we first utilized a behavioral task developed in our laboratory (MCOAT<sup>17</sup>) to delineate the relationship between multiple task parameters (e.g., valence, action initiation, and prediction) and the resulting neural signals. Mice were first trained in a positive reinforcement task where an auditory cue predicted that an operant response would result in the delivery of sucrose (Figures 2A–2C, S2A, and S2B). Next, mice were trained in a negative reinforcement task where a distinct auditory cue indicated that a response on a second *operanda* prevented the delivery of a series of footshocks (Figures 2D, 2F, S2C, and

### Figure 1. D1 and D2 MSNs do not track stimulus valence

(A) Cre-dependent GCaMP6f (AAV5.hsyn.flex.CGAMP6f) was expressed in D1 MSNs (D1-cre mice) or D2 (A2A-cre mice) MSNs. (Right) Example of GCaMP6f expression in NAc core.

(B) D1 MSNs showed a positive response to sucrose retrieval in a positive reinforcement operant task (two-tailed independent sample t test,  $t_{45} = 2.897$ ,  $p = 0.0058$ ,  $n = 5$  mice). Dark gray dots are individual trials across all animals, light gray dots are averaged responses for each animal.

(C) D2 MSNs showed a decrease to sucrose retrieval in the same task (two-tailed independent sample t test,  $t_{60} = 6.29$ ,  $p < 0.0001$ ,  $n = 5$  mice).

(D) D1 MSNs showed an intensity-dependent positive response to unsignaled shock (nested ANOVA,  $F_{(1,39)} = 6.53$ ,  $p = 0.016$ ,  $n = 5$  mice).

(E) D2 MSNs showed an intensity-dependent positive response to unsignaled shock (nested ANOVA,  $F_{(1,47)} = 5.04$ ,  $p = 0.031$ ,  $n = 6$  mice).

(F) Intracranial self-stimulation (ICSS) task design. An excitatory opsin (ChR2; AAV5.Ef1a.DIO.hChR2) or a control vector (eYFP; AAV5.hSyn1.eYFP) was expressed in D1 MSNs or D2 MSNs in the NAc core. Nose pokes resulted in laser illumination (14 Hz, 2 s, 8 mW, 470 nm). Viral expression of ChR2 in the NAc core.

(G) Mice were trained to nose poke for optical stimulation of either D1 MSNs or D2 MSNs over 4 days.

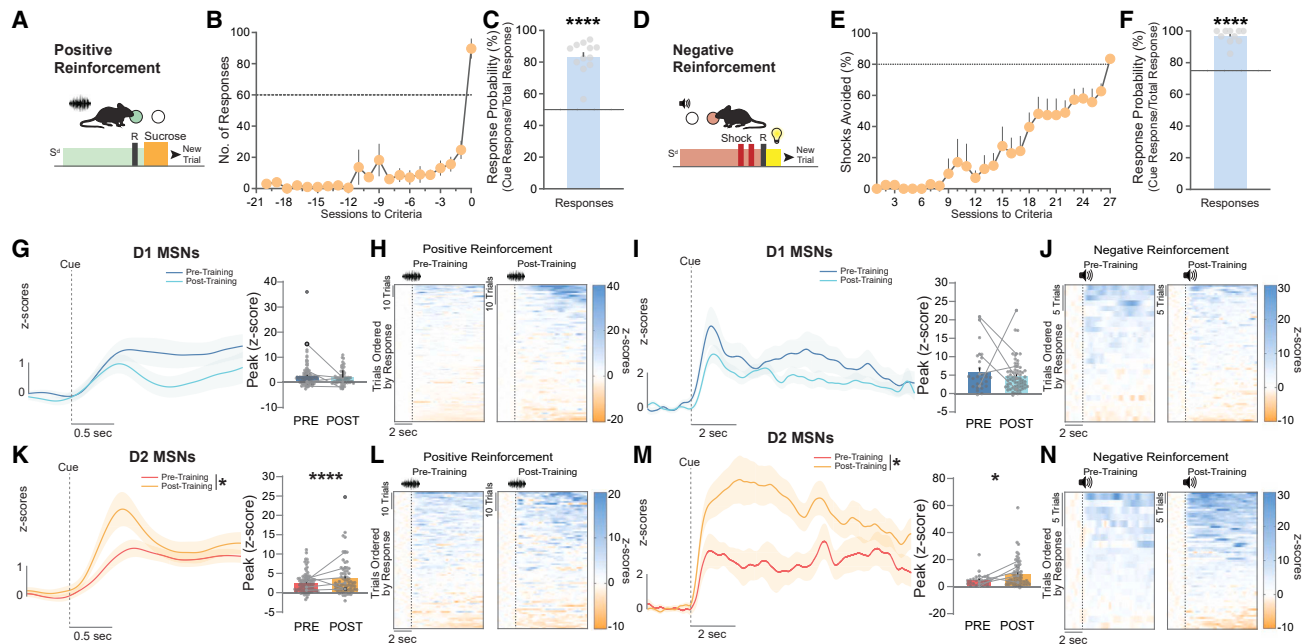
(H) D1-Cre (D1 MSN) and A2A-Cre (D2 MSN) mice showed a preference for the active nose poke as compared with eYFP controls (repeated measures ANOVA, trial  $\times$  group interaction  $F_{(6,42)} = 3.17$ ,  $p = 0.0118$ ).

(I) D1-cre ( $n = 5$  mice) and A2A-Cre (D2 MSNs,  $n = 7$  mice) showed a greater percentage of total responses on the active *operanda* as compared with the eYFP controls ( $n = 5$  mice, one-way ANOVA,  $F_{(2,14)} = 8.96$ ,  $p = 0.0031$ ; Dunnett's post-hoc eYFP versus D1,  $p = 0.036$ ; eYFP versus D2,  $p = 0.0016$ ).

(J) Training-dependent increase in responses in D1-Cre and A2A-Cre mice as compared with eYFP (one-way ANOVA,  $F_{(2,14)} = 4.60$ ,  $p = 0.0291$ ; Dunnett's post-hoc eYFP versus D1,  $p = 0.025$ ; eYFP versus A2A,  $p = 0.049$ ).

Data represented as mean  $\pm$  SEM; \* $p < 0.05$ ; \*\* $p < 0.01$ , \*\*\*\* $p < 0.0001$ . Scale bars = 50  $\mu$ m.





**Figure 2. D2 MSN responses to predictive cues scale with learning, while D1 MSN responses do not change**

(A) A discriminative cue ( $S^d$ ) indicated that responses on a fixed-ratio 1 schedule resulted in sucrose delivery.

(B) Mice acquired this task (>60 responses on the active operanda).

(C) Nearly all responses were made during the cue period, indicating that mice learned the value of this cue (two-tailed independent sample t test,  $t_{11} = 11.23$ ,  $p < 0.0001$ ,  $n = 12$  mice; chance = 50%).

(D) Mice responded during an  $S^d$  to prevent shock presentation.

(E) Mice avoided almost all possible shocks.

(F) Almost all responses were made during the cue period, indicating that mice learned the cue value (two-tailed independent sample t test,  $t_9 = 14.68$ ,  $p < 0.0001$ ,  $n = 10$  mice; critical value = 75.1%).

(G) D1 MSNs showed a response to the  $S^d$  that signaled positive reinforcement. This response did not change with training (nested ANOVA,  $F_{(1,176)} = 0.63$ ,  $p = 0.427$ ,  $n = 6$  mice). Dark gray dots are individual trials across all animals, light gray dots are the first response in each session for each animal.

(H) Heatmap of D1 MSN responses pre-training and post-training during positive reinforcement. All heatmaps are individual trials ordered by response magnitude to the cue.

(I) D1 MSNs showed a positive response to the  $S^d$  signaling negative reinforcement that did not change with experience (nested ANOVA  $F_{(1,91)} = 1.21$ ,  $p = 0.2747$ ,  $n = 5$  mice).

(J) Heatmap of D1 MSN responses pre-training and post-training during negative reinforcement.

(K) D2 MSNs showed an increase in response to the  $S^d$  signaling positive reinforcement between pre- and post-training (nested ANOVA  $F_{(1,166)} = 16.38$ ,  $p < 0.0001$ ,  $n = 6$  mice).

(L) Heatmap of D2 MSN responses.

(M) D2 MSNs showed a learning-dependent increase to the  $S^d$  signaling negative reinforcement (nested ANOVA  $F_{(1,88)} = 5.35$ ,  $p = 0.023$ ,  $n = 5$  mice).

(N) Heatmap of trial responses pre-training and post-training during negative reinforcement.

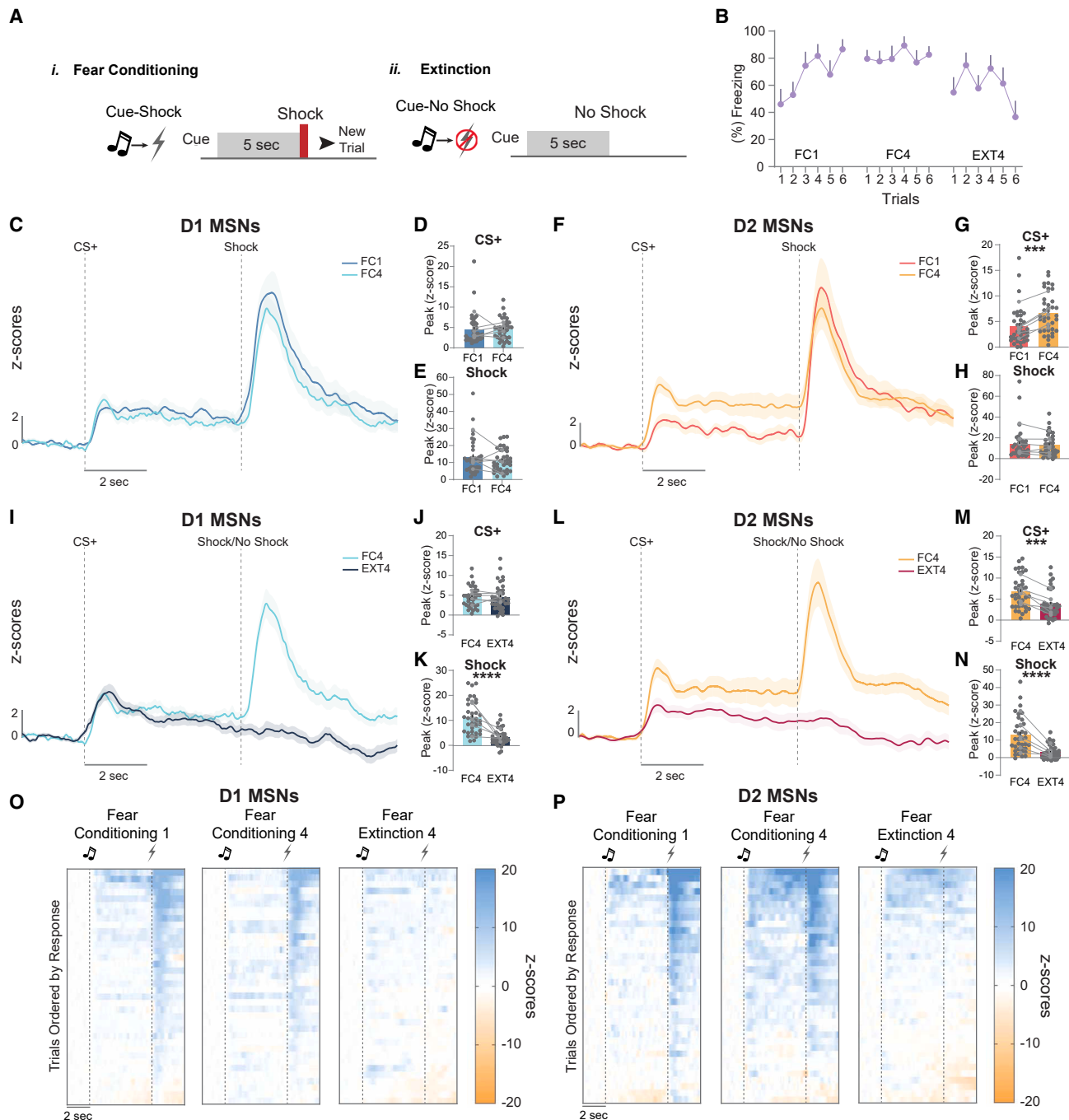
Data represented as mean  $\pm$  SEM. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\*\* $p < 0.0001$ .

S2D). Importantly, the operant response in the two phases is the same (nose poke), and both outcomes (sucrose retrieval or removal of shock) are positive; however, the maintaining stimulus has opposite valence (shock, negative; sucrose, positive). Thus, this behavioral task allows for the dissociation of motivational responding (the same between conditions) from the valence of the stimulus maintaining the behavior (the opposite).

In the positive reinforcement phase, mice learned quickly to respond on the active nose poke (Figure 2B) and responded with greater probability during the cue than during the inter-trial interval (Figures 2C, S2A, and S2B). During negative reinforcement, mice avoided greater than 80% of potential shocks within a session and had a higher probability of responding during the cue than during the inter-trial interval (Figures 2E, 2F, S2C, and

S2D). Calcium transients to the discriminative cues that predicted positive or negative reinforcement were assessed during the first training session and again in the same mice after they had met acquisition criteria.

In response to the discriminative cue that predicted positive reinforcement, we observed a positive calcium transient in D1 MSNs; however, this response did not change over training (Figures 2G, 2H, and S2E). The same pattern was present in the negative reinforcement task, where the discriminative cue evoked a positive, time-locked calcium transient in D1 MSNs, and this cue-evoked response was not sensitive to training history (Figures 2I, 2J, and S2F). Conversely, D2 MSNs also showed a positive time-locked calcium transient in response to both the discriminative cue that signaled positive reinforcement



**Figure 3. D2 MSN responses to cues increase over learning, while D1 MSN responses do not change, in Pavlovian tasks**

(A) (i) Mice received a 5-s cue followed by a half-second shock in a Pavlovian fear-conditioning task. (ii) During extinction, the cue was presented for 5 s but the shock was not delivered.

(B) Freezing across trials during fear-conditioning session 1 (FC1), fear-conditioning session 4 (FC4), and extinction session 4 (EXT4) (repeated measures ANOVA, trial  $\times$  group interaction  $F_{(5,017,55,19)} = 2.88$ ,  $p = 0.022$ ).

(C) D1 MSN response to the cue in the first (FC1) and last (FC4) fear-conditioning session.

(D and E) D1 MSNs showed no difference in response to the cue (nested ANOVA,  $F_{(1,71)} = 0.02$ ,  $p = 0.897$ ,  $n = 6$  mice) or the shock (nested ANOVA  $F_{(1,71)} = 1.73$ ,  $p = 0.194$ ,  $n = 6$  mice) over sessions. Dark gray dots are individual trials across all animals, light gray dots are averaged responses for each animal.

(F) D2 MSN response to the cue in FC1 and FC4.

(G) D2 MSNs showed an increase in response to the cue over sessions (nested ANOVA,  $F_{(1,71)} = 11.28$ ,  $p = 0.0014$ ,  $n = 6$  mice).

(H) The peak responses to the shock in D2 MSNs normalized to the pre-trial baseline (nested ANOVA,  $F_{(1,71)} = 0.42$ ,  $p = 0.522$ ,  $n = 6$  mice).

(legend continued on next page)

(Figures 2K and 2L) and the discriminative cue that signaled negative reinforcement (Figures 2M and 2N). The response in both conditions was further increased as performance increased over training (also see Figures S2G and S2H). These data—combined with the data above—suggest that D1 MSNs respond to stimuli generally, while D2 MSNs dynamically respond to changes in predictive cues.

### MSN responses are consistent across tasks

While these results were consistent across different operant tasks, it was not clear whether these signals were specific to motivated action. D1/D2 MSN responses could signal information about cues that were not sensitive to contingencies/actions (e.g., if/how strongly the cue predicted an outcome). We hypothesized that D2 MSN responses dynamically change with the predictive value of cues, regardless of whether the task required an action or not. We tested this hypothesis using a Pavlovian fear-conditioning paradigm. Mice were presented with a 5-s cue followed by a brief, unavoidable footshock. Mice were trained over four fear-conditioning sessions (6 trials per session, Figure 3Ai), followed by four sessions of extinction where the cue, but no shock, was presented (Figure 3Aii). Freezing was analyzed during (1) fear-conditioning session 1 (FC1) when learning of the cue-shock association was forming, (2) in the final session (fear-conditioning session 4 [FC4]), when the association was established, and (3) on the final session of extinction, when the cue-shock association was extinguished (EXT4). Both groups of mice (D1-cre and A2A-Cre) increased freezing during acquisition and decreased freezing following extinction (Figures 3B, S3A, and S3B).

We observed a time-locked positive D1 MSN calcium transient in response to both the cue and the shock in early fear conditioning (FC1) and late fear conditioning (FC4; Figures 3C–3E, 3O, S3C, and S3D). The magnitude of this response did not change over learning. Conversely, D2 MSNs exhibited a positive, cue-evoked calcium transient that was further increased with learning (Figures 3F–3H, 3P, S3E, and S3F).

When mice went through extinction, freezing was reduced (Figure 3B). However, we observed no difference in the positive, cue-evoked D1 MSN calcium transient (Figures 3I, 3J, and 3O; see Figures S7G and S7H for additional analysis). Also, the positive, shock-evoked D1 MSN calcium transient disappeared during extinction (as the shock was not presented in extinction trials, Figures 3I and 3K). The positive, cue-evoked D2 MSN calcium transients were reduced with extinction (Figures 3L, 3M, 3P, S3I, and S3J). Further, there was no D2 MSN response in either direction at the time of the omitted shock during the final extinc-

tion session when shock was not presented and no longer expected (Figures 3L and 3N).

These data suggest that cue-evoked calcium transients in D2 MSNs track the strength of the cue-outcome association. This occurs in the same direction (a positive response) regardless of whether the outcome value is positive or negative (Figures 2 and 3) and does not depend on the action required, as this same pattern was present for both reinforcement (motivated action, Figure 2) and fear conditioning (freezing, Figure 3). Thus, D2 MSNs appear to track predictions, while D1 MSNs respond to stimuli generally. These data also further rule out that the D2 MSN response tracks general motor responses—freezing or otherwise—as the same signal was observed when the animals were freezing (Figure 3) and when mice were making a motivated operant response (Figures 2, S2G, and S2H).

### D2 MSNs track prediction errors

We hypothesized that D2 MSN responses tracked a valueless prediction error. In the case of valueless prediction error, responses to cues are initially small when they are neutral and novel (regardless of valence) and increase as those cues acquire associative value. Conversely, in response to the stimulus itself (what the cue predicts), initially the response is positive and large, and this response decreases as the stimulus becomes more predicted by the antecedent cue. Thus, one key feature of valueless and other prediction-error signals is that in addition to cue responses changing over learning, relative stimulus responses also change.<sup>15</sup>

While we observed in Figure 1C that D2 MSNs showed a negative calcium transient at the time of sucrose collection in well-trained animals, when unexpected sucrose was given to naive mice D2 MSN calcium responses were positive in response to sucrose consumption (time-locked to the first sucrose lick, Figure 4A). Thus, the D2 MSN calcium response to an unexpected reward was positive (Figure 4B; see Figure S4A), and this response to the same rewarding stimulus was reduced as the reward became more predicted (Figure 1C). We observed similar increases in D2 MSN calcium responses that occurred in response to unpredicted vs predicted aversive stimuli.

The same pattern was present with aversive stimuli. Shock-evoked transients were assessed during the first 2 trials of fear conditioning (FC1, when animals did not expect the footshock) and during the first 2 trials of the final fear-conditioning session (FC4, when the footshock was anticipated). First, we analyzed D1 and D2 MSN responses aligned to the onset of the shock (instead of the cue as in Figures 3D and 3G) to probe the shock response relative to baseline. Shock-evoked D1 MSN responses

(I) D1 MSN cue and shock responses in the last fear-conditioning session (FC4) compared with the last session of extinction (data from FC4, replotted from C).

(J) There was no change in the D1 MSN response to the cue following extinction (nested ANOVA,  $F_{(1,71)} = 0$ ,  $p = 0.99$ ,  $n = 6$  mice).

(K) D1 MSN responses to shock period. In FC4, the shock was presented, and in extinction (EXT4), it was not (nested ANOVA,  $F_{(1,71)} = 68.59$ ,  $p < 0.0001$ ,  $n = 6$  mice).

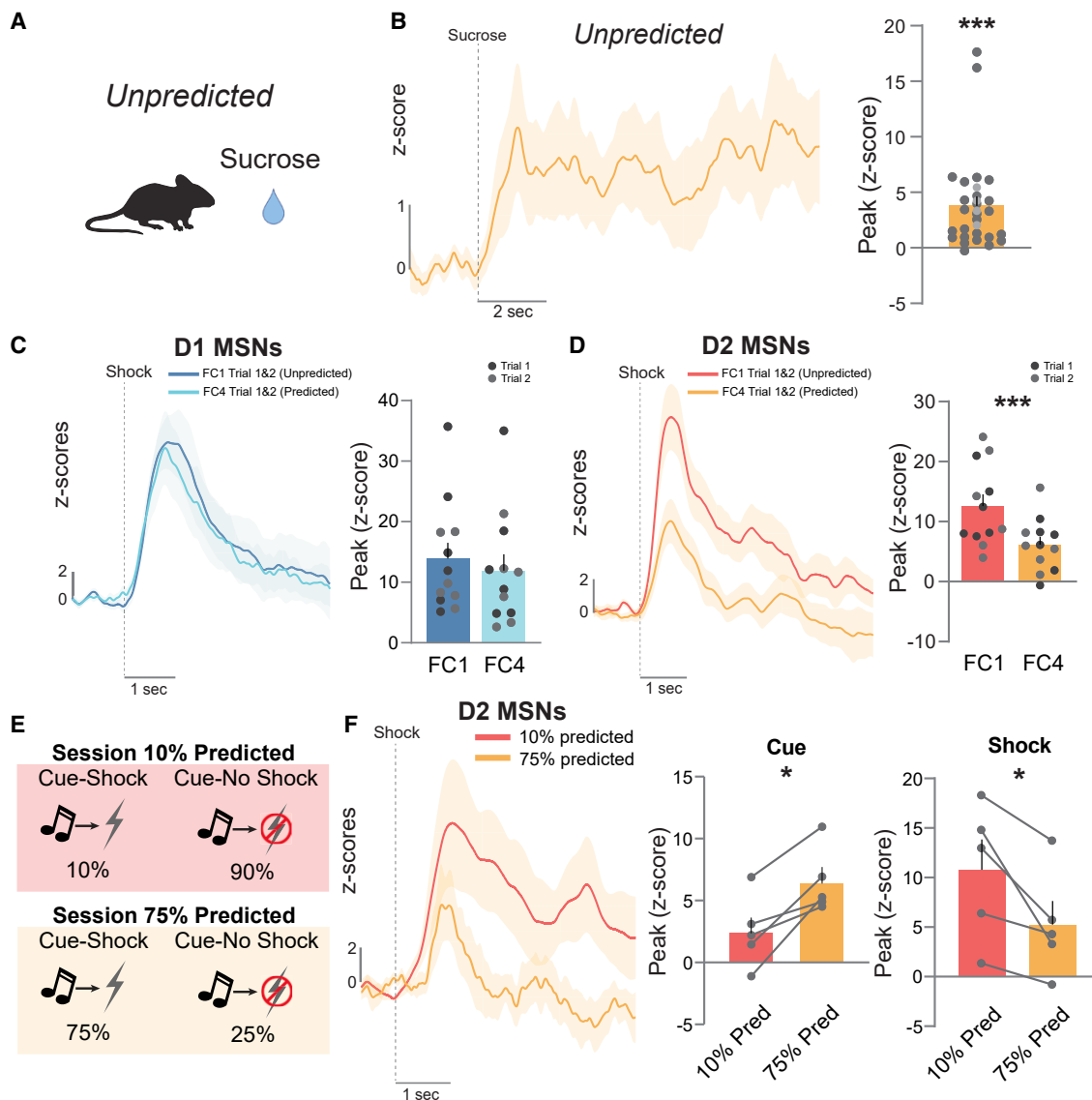
(L) D2 MSN response to the cue and shock in FC4 as compared with extinction (FC4 data plotted from F).

(M) There was a decrease in the cue response following extinction (nested ANOVA,  $F_{(1,71)} = 16.32$ ,  $p = 0.0002$ ,  $n = 6$  mice).

(N) The response during the shock period was also reduced (nested ANOVA,  $F_{(1,71)} = 56.31$ ,  $p < 0.0001$ ,  $n = 6$  mice). In FC4, the shock was presented, and in extinction (EXT4), it was not.

(O and P) Heatmap of D1 MSN (O) and D2 MSN (P) responses to task parameters ordered by response magnitude to cue (from largest at top to smallest at bottom) during early (FC1) and late (FC4) fear conditioning and late fear extinction (EXT4).

Data represented as mean  $\pm$  SEM. \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ . (fear-conditioning session 1 [FC1]; fear-conditioning session 4 [FC4]; extinction session 4 [EXT4].)



**Figure 4. D2 MSN responses to stimuli are modulated by prior predictions**

(A) Mice were given *ad libitum* access to sucrose in the delivery port and signals were analyzed around the first lick in the first lick bout.  
 (B) When sucrose was not predicted, D2 MSN response was increased (independent sample t test,  $t_{25} = 4.41$ ,  $p = 0.0002$ ,  $n = 4$  mice), rather than decreased (Figure 1C). We next determined whether D1 MSN or D2 MSN responses to footshocks were changed based on prediction.  
 (C) Signals were Z scored around the baseline preceding the onset of the shock in the first two trials of fear-conditioning session 1 (FC1) when the mouse experiences the cue and the shock for the first and second time, and fear-conditioning session 4 (FC4) when the mouse has extensive experience with the cue-shock association. Dark gray dots are trial 1 response, light gray dots are trial 2 response. There was no effect on D1 MSNs responses to the shock under these conditions (nested ANOVA  $F_{(1,23)} = 0.74$ ,  $p = 0.41$ ,  $n = 6$  mice).  
 (D) D2 MSN responses to the footshock became smaller as the prediction between cue and shock became stronger (nested ANOVA,  $F_{(1,23)} = 18.22$ ,  $p = 0.0011$ ,  $n = 6$  mice).  
 (E) The likelihood of a cue-shock pairing was manipulated (shock occurs 10% of the trials in the session or 75% of the trials in the session).  
 (F, left) Fiber photometry trace from D2 MSNs during the shock responses depending on the probability of the prediction.  
 (F, right) D2 MSNs showed an increase in the cue response (paired t test,  $t_4 = 4.54$ ,  $p = 0.0105$ ,  $n = 5$  mice) and decrease in the magnitude of the shock response (paired t test,  $t_4 = 3.12$ ,  $p = 0.0356$ ,  $n = 5$  mice) with greater predictability of the shock outcome.  
 Data represented as mean  $\pm$  SEM. \* $p < 0.05$ , \*\*\* $p < 0.001$ . (fear-conditioning session 1 [FC1]; fear-conditioning session 4 [FC4]).

were positive but did not change with expectation (Figures 4C and S4B). This was true whether data were assessed over days (FC1 vs. FC4) or within-session (trial 1 versus trial 6; Figures S4D and S4E). However, the positive D2 MSN calcium

transient evoked by the shock was smaller when it was expected (Figures 4D and S4C).

Finally, we directly manipulated the probability of shock presentation in two separate sessions: the first session when the shock



was predicted by the cue 10% of the time and the second session when the shock was predicted by the cue 75% of the time (Figure 4E). Accordingly, the size of the positive cue-evoked calcium transient in the D2 MSNs reflected how well the shock was predicted, where the higher probability cue resulted in a stronger positive D2 MSN calcium response to the cue (Figures 4F and S4F–S4H) and a reduced response to the shock, an effect that was consistent across repeated testing (Figures S4F–S4H). Thus, as the cue becomes more predictive, the cue response is largest and the outcome response is smallest, similar to what has been observed by other prediction-based signals.<sup>57,58</sup>

### Learning increases D2 MSN recruitment

To further understand how these signals developed over learning, we employed microendoscopic cellular resolution calcium imaging (see Figure S1C for the gradient index [GRIN] lens placements). GCaMP6m was expressed in either D1 or D2 MSNs as described above, and using a GRIN lens for optical access, we recorded single-cell calcium transients in these identified populations during fear conditioning (Figures 5A–5C; see Figures S5A–S5D).

Replicating our results, we found that D1 MSNs—when represented as a whole field trace or the average of all identified single-cell responses—responded to both the cue and the footshock during FC1 (Figures 5D–5F, S5C, and S5D, for whole field calcium traces; Figures S5E–S5H for area under the curve). While there was heterogeneity in the responses (Figure 5G, left), most cells responded positively (55.41% of all identified cells recorded during the session) to the shock (Figure 5G, right). Following extended training, when behavioral responses were asymptotic (FC4), the proportionality of D1 MSN responses did not change (Figure 5H). Consistent with the idea that D1 MSNs respond to the presence of stimuli, the number of D1 MSNs that were active at the time of the previous footshock was reduced when the shock was not present in extinction (to 23.81%; Figures S5I–S5M). Thus, D1 MSNs track stimulus presence and show that the population-level responses in the NAC core are representative of a majority of the population.

D2 MSNs showed the same population response patterns observed above—both when represented as field of view and when presented as the average of all identified single-cell responses (Figures 5I–5K and S5C–S5H). Cue- and shock-responsive cells were identified as any cell that had a significant increased response ( $>1.96$  Z scores from baseline) within 2 s of cue onset and 2 s of shock onset, respectively. The increase in population-level response to the cue over training was explained by an increase in the size of the D2 MSN ensemble that was evoked by the presentation of the predictive cue, which was largest on the final fear condition session (FC4) compared with the initial fear-conditioning session (FC1). Indeed, the percentage of cue-responsive D2 MSNs increased from 39.26% to 75.68% of identified cells between FC1 and FC4 (Figures 5L and 5M, left). Shock-responsive D2 MSNs were also increased and went from 54.21% to 81.99% of the population (Figures 5L and 5M, right). Critically, the opposite was observed when animals moved from fear conditioning to fear extinction where the percentages of responsive cells were reduced to 42.86% for the cue and to 34.92% for the footshock (Figures S5Q and S5R). These results

support the hypothesis that D2 MSNs signal predictions as this signal progressively develops with learning (over fear conditioning) and is updated when new information is encountered (during extinction). More specifically, the D2 MSN response to the cue becomes stronger as the animals learn the association between the cue and the footshock and weaken as this association is extinguished (Figures 5I–5K and S5N–S5R).

### D2 MSN dynamics change over learning

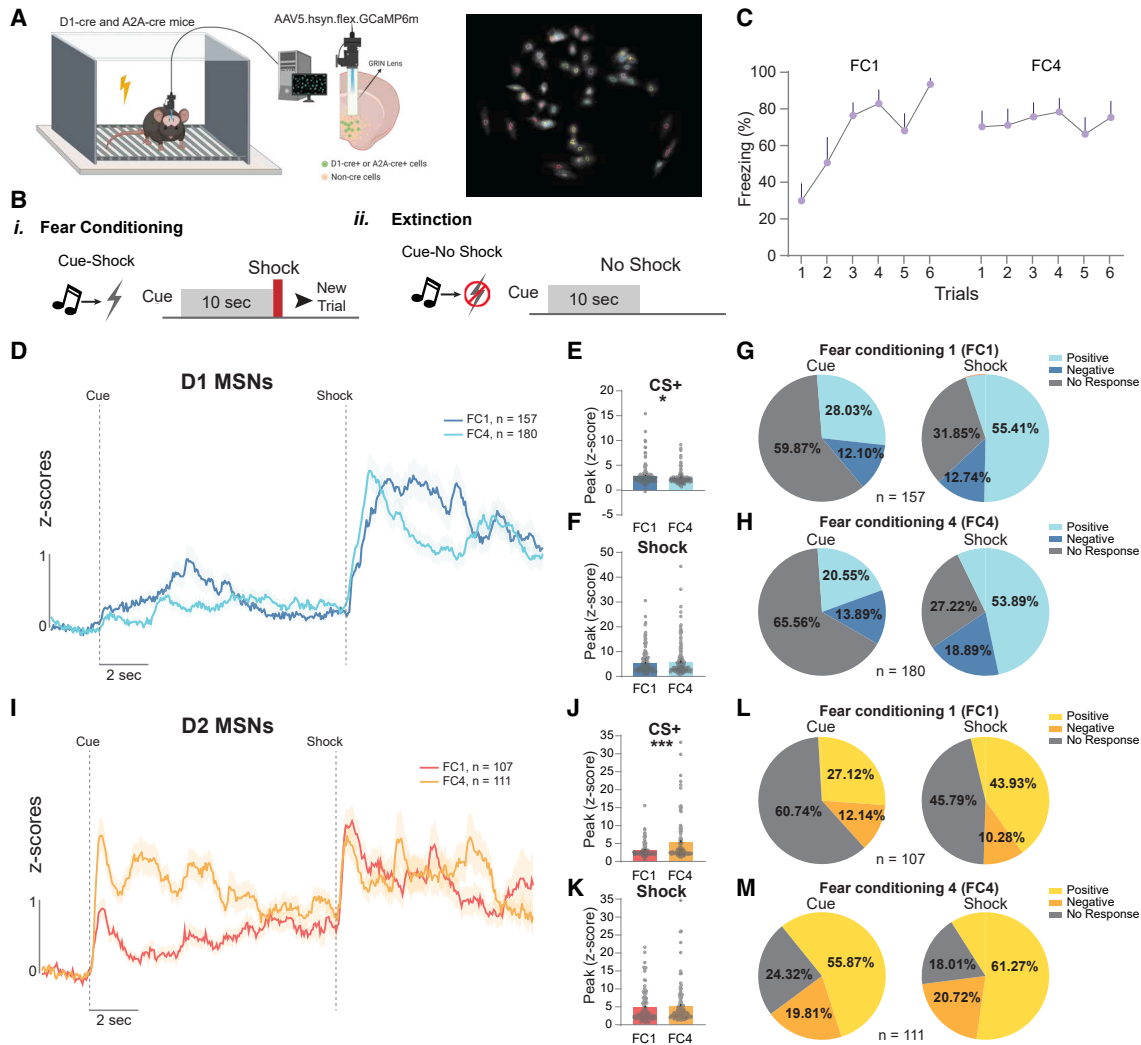
To better understand how D1 and D2 MSN populations were changing over experience, we used dimensionality reduction approaches to visualize how these populations were changing both within and across sessions (Figures 6Ai–6Aiv). We performed a principal-component analysis (PCA) and plotted the trajectory summarizing the activity of all recorded neurons through the space defined by the first and second principal components (PCs). The large variation of the trajectory during the cue period in the final fear-conditioning session (FC4, Figure 6Aiv) as compared with the first fear-conditioning session (FC1, Figure 6Aiii) is consistent with the finding that more neurons are engaged at this time in FC4 compared with FC1 (Figures 5L and 5M). This difference was not apparent when comparing FC1 and FC4 trajectories of D1-MSNs (Figures 6Ai and 6Aii), again supporting that only D2-MSNs undergo changes at the population level. While learning-dependent changes occur in D2 MSNs, we wanted to further understand the cell response patterns that characterized these changes and how the temporal patterns of these responses changed over time.

### D2 MSNs bridge cues and outcomes

For D2 MSNs to function as a prediction error signal, both the cue and the outcome response should be represented in the same cells, rather than two separate populations. Analysis of the single-cell responses during the first fear-conditioning session (FC1) revealed 3 groups of D2 MSNs: (1) D2 MSNs that showed a positive response only to the cue (cue only), (2) D2 MSNs that showed a positive response only to the shock (shock only), and (3) D2 MSNs that showed a positive response to both the cue and shock (both cue and shock; Figure 6B). Consistent with the data presented in Figure 6A showing little change in the D1 MSN responses to stimuli across sessions, we found that more than half of D1 MSNs (55.1%) responded to the shock only (Figure 6C), and this pattern was roughly the same between FC1 (Figure 6C) and FC 4 (Figure 6D).

However, D2 MSN response patterns changed over learning. Initially, only a small percentage of the D2 MSNs respond to both the cue and shock (14.28%) while the rest of the cells respond only to the cue (46.94%) or only to the shock (38.78%; Figure 6E). However, when animals learned the association between the cue and shock—in the fourth fear-conditioning session (FC4)—the majority of the D2 MSNs showed a positive response to both the cue and the shock (65.38%; Figure 6F). Next, we used activity-based co-registration to determine whether cells showing one of the three activity signatures were more or less likely to be active in subsequent sessions. Co-registered cells were identified based on activity during the first fear-conditioning session (FC1) and compared against activity signatures in the last fear-conditioning session (FC4). Thus, only cells that were active in both FC1 and FC4 were considered co-registered (Figure 6G; see Figures S6A and



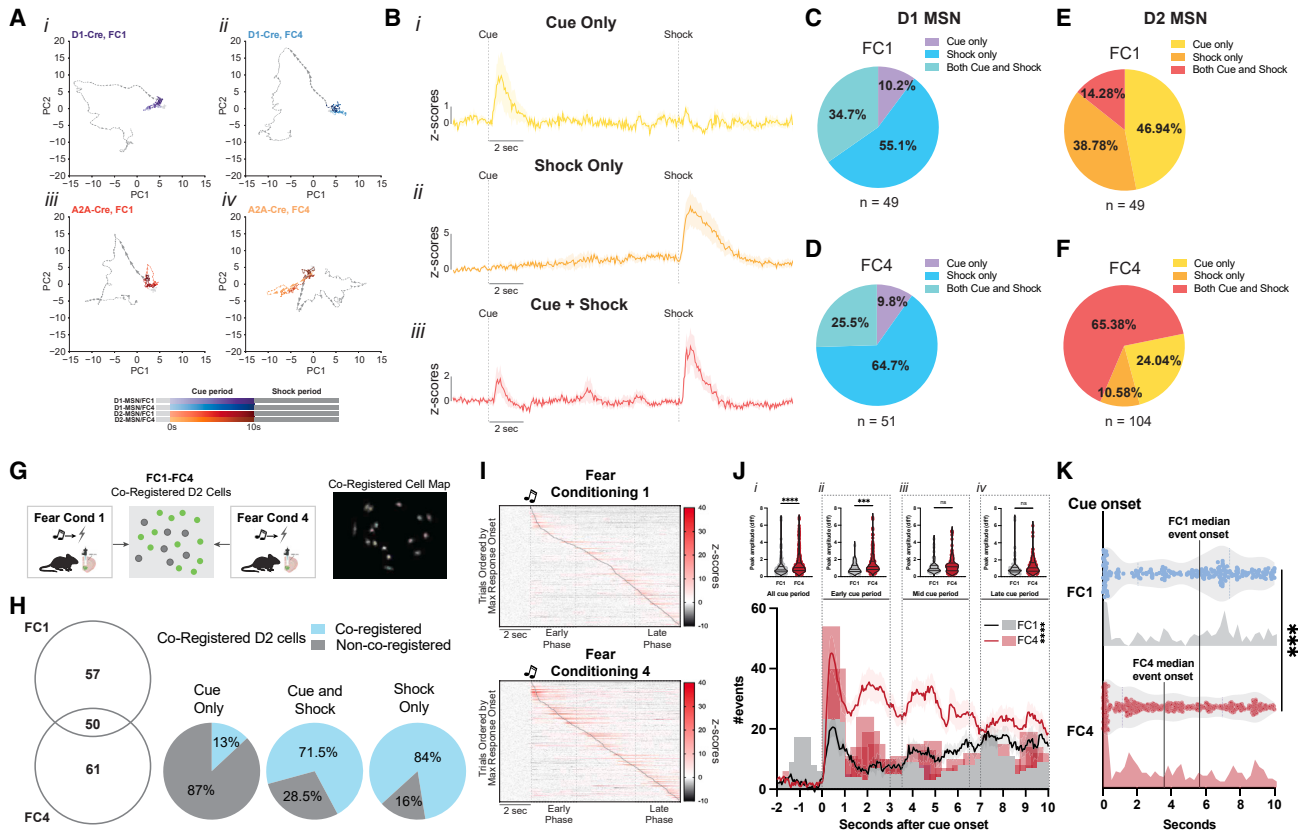


**Figure 5. D2 MSN recruitment is increased over learning and individual D2 MSNs respond to both the cue and the shock**

(A) D1 and D2 MSN responses were recorded via cell-type-specific expression of GCaMP6m as described. A GRIN lens was implanted above the NAC core for optical access.  
 (B) Fear conditioning. Mice received a 10-s cue followed by a 0.5-s shock.  
 (C) Freezing responses increased over training (repeated measures ANOVA, trial  $\times$  group interaction  $F_{(3,990,35,91)} = 4.21$ ,  $p = 0.0068$ ). Session 1 (FC1), session 4 (FC4).  
 (D) D1 MSNs responses across detected cells (157 cells in FC1, 180 cells in FC4).  
 (E) There was a moderate decrease in the peak response to the cue in the last session as compared with the first (independent sample t test,  $t_{335} = 2.51$ ,  $p = 0.013$ ,  $n = 5$  mice).  
 (F) The shock response did not change (independent sample t test,  $t_{335} = 0.6697$ ,  $p = 0.5035$ ,  $n = 5$  mice).  
 (G and H) Percentage of the total D1 MSNs detected in each session that increased (positive), decreased (negative), or showed no response (no response) to the cue or shock in the first (G, FC1) or last session (H, FC4). D1 MSN responses did not change over learning.  
 (I) D2 MSNs responses (107 cells in FC1, 111 cells in FC4).  
 (J) D2 MSN responses to the cue were increased over sessions (independent sample t test,  $t_{216} = 3.44$ ,  $p = 0.0007$ ,  $n = 5$  mice).  
 (K) No difference in the shock response (independent sample t test,  $t_{216} = 0.71$ ,  $p = 0.478$ ,  $n = 5$  mice).  
 (L and M) Percentage of D2 MSNs that increased (positive), decreased (negative), or did not respond (no response) to the cue and the shock in the first session (L, FC1) and the last (M, FC4). The number of D2 MSNs that responded to the cue changed over learning.  
 Data represented as mean  $\pm$  SEM. \* $p < 0.05$ , \*\*\* $p < 0.001$ . (fear-conditioning session 1 [FC1]; fear-conditioning session 4 [FC4].)

S6B for co-registration approach). Critically, cells that responded to the cue only were not likely to be identified in FC4 (13% co-registered), while a majority of the D2 MSNs that responded only to the shock (84% co-registered) or both to the cue and shock (71.5%

co-registered) were re-recruited and active in FC4 (Figure 6H). Additionally, of the cells that were shock responsive in FC1 and co-registered in FC4, >60% showed a response to the cue in FC4 (Figure S6C).



**Figure 6. D2 MSNs, but not D1 MSNs, change dynamically over learning in both the pattern and timing of responses to learned cues**

(A) Neural trajectories summarizing the activity of D1-MSNs in fear-conditioning session 1 (FC1, [i],  $n = 157$  neurons), D1-MSNs in fear-conditioning session 4 (FC4, [ii],  $n = 180$  neurons), D2-MSNs in FC1 ([iii],  $n = 107$  neurons), and D2-MSNs in FC4 ([iv],  $n = 111$  neurons.) Each time point is depicted as an arrow pointing in the direction of the next time point. The size of each arrow is proportional to the delay until the next time point (i.e., how fast the activity is moving along the trajectory with large arrows depicting more rapid changes). The pre-cue baseline period is colored light gray, the cue period is color coded (D2-MSN/FC1, red; D2-MSN/FC4, orange; D1-MSN/FC1, dark blue; D1-MSN/FC4, light blue), and the shock period is colored dark gray. As mice learn the cue-footshock contingency, D2 MSN cue responses, but not D1 MSN responses, become more variable.

(B) D2 MSNs were categorized based on observed activity patterns in the NAc during the initial fear-conditioning session (FC1) in the following categories: (i) response only to the cue, (ii) response only to the shock, and (iii) response both to the cue and shock.

(C) In D1 MSNs, most of the cells only responded to the shock during the initial fear-conditioning session (FC1).

(D) The D1 MSN cell recruitment to the cue and the shock was similar in the last fear-conditioning session (FC4), with a majority of cells responding only to the shock.

(E) In D2 MSNs, initially (on FC1) only a small percentage of cells responded to both the cue and shock.

(F) In FC4, a majority of D2 MSNs responded to both the cue and shock.

(G) D2 MSNs were recorded on the first session (FC1) and cells detected during this session were longitudinally co-registered with cells in the last session (FC4) based on activity during each session.

(H) Most of the cells that only responded to the cue in FC1 were not detected as active during the final fear-conditioning session (FC4, only 13% co-registered). The majority of D2 MSNs that responded to the shock (either shock alone, or both cue and shock) were re-recruited in FC4.

(I) Heatmaps showing cue responses for fear-conditioning session 1 and 4 ordered by the tune of response following cue presentation.

(J) Histogram of event numbers for each second of the cue period, superimposed on the Z scored averaged calcium responses. Event analysis showed that the number of D2 MSN events within the cue period increased with learning (chi square = 34.32,  $p < 0.0001$ ) and the amplitude of those events became larger as well ([ii] the whole cue period, unpaired t test,  $t_{718} = 4.26$ ,  $p < 0.0001$ ,  $n = 239$ –481 events). When clustered based on the timing of the response the peak event amplitude was larger in FC4 during the early segment ([ii] from the cue onset to 3 s; unpaired t test,  $t_{301} = 3.76$ ,  $p = 0.0002$ ,  $n = 78$ –225 events) but not during the middle ([iii] 3.5 s to 6.5 s, unpaired t test,  $t_{169} = 1.13$ ,  $p = 0.26$ ,  $n = 57$ –114 events) or the late ([iv] 7 s to the cue offset, unpaired t test,  $t_{191} = 1.33$ ,  $p = 0.18$ ,  $n = 73$ –120 events) segments of the cue period.

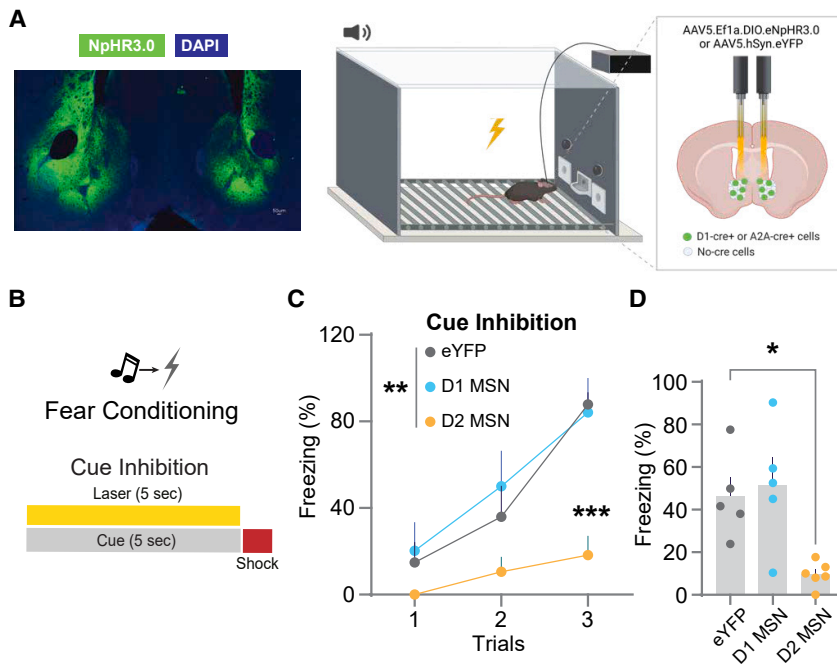
(K) The event onset was earlier in FC4 compared with FC1 (unpaired t test,  $t_{718} = 3.61$ ,  $p = 0.0003$ ,  $n = 239$ –481 events).

Data represented as mean  $\pm$  SEM, \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ , ns, not significant. (fear-conditioning session 1 [FC1]; fear-conditioning session 4 [FC4]).

### Temporal pattern of D2 MSN response changes with learning

We next used a hierarchical clustering approach to identify different patterns of responses that were present within the D2

MSN population in response to the cue during fear conditioning (Figures S6D and S6E). There were three major clusters of cells: cells that responded early in the cue period, cells that responded in the middle of the cue period, and cells that responded near the



**Figure 7. Optogenetic inhibition of D2 MSN responses during the cue slows associative learning**

(A) The inhibitory opsin halorhodopsin (AAV5Efl1a-DIO.eNpHR3.0) was selectively expressed in D1 or D2 MSNs. Representative histology.

(B) A laser (constant, 5 s, 8 mW, 590 nm) was illuminated at the time of cue onset.

(C) When D2, but not D1 MSNs, were inhibited mice developed a freezing response at a slower rate (RM ANOVA trial  $\times$  group interaction  $F_{(2,18)} = 8.17$ ,  $p = 0.0030$ ; multiple comparison D2 MSN versus eYFP session 3,  $p = 0.0005$ ).

(D) D2 MSN inhibition reduced freezing during the 3 training trials as compared with eYFP animals (one-way ANOVA  $F_{(2,13)} = 7.52$ ,  $p = 0.0067$ ; Bonferroni's post-hoc eYFP versus D1,  $p > 0.9999$ ; eYFP versus A2A,  $p = 0.018$ ).

Data represented as mean  $\pm$  SEM. \* $p < 0.05$ , \*\* $p < 0.01$ . Scale bar = 50  $\mu$ m.

end of the cue period. Between FC1 and FC4, the cluster of cells that responded early in the cue increased both in magnitude and number of cells that made up this population (Figure S6E), suggesting that with learning, D2 MSN responses to the cue become more time-locked to cue onset. Indeed, when D2 MSN cue responses were ordered based response time during the cue period for FC1 (top) and FC4 (bottom), there appeared to be a shift in the timing, where responses occurred closer to cue onset (Figure 6I). To statistically test this, D2 MSN calcium events that occurred any time within the cue window were identified and analyzed. Between FC1 and FC4 the amplitude of events increased (Figure 6J, all events), and the number of D2 MSN events within the cue period increased (Figure 6J, bottom histogram), without any change in the kinetics of individual events (Figure S6F).

When events were clustered and analyzed based on the timing of the response during the cue period, these effects were largest during the first 3 s of the 10-s cue period (Figures 6ii–6iv). In the last fear-conditioning session (FC4), over 55% of the events occurred in the first 3 s following cue presentation, and 10% occurred in the last 3 s (Figure S6G), suggesting a shift in the timing of the response toward earlier in the cue period over learning. Indeed, there was a significant decrease in the average response time of all events (i.e., earlier in the cue period) between FC1 and FC4 (Figure 6K).

Together these data show that the number of D2 MSNs that respond to a cue is increased over learning, the cells that are recruited evolve to respond to both the cue and the shock over learning, and the temporal profile of these responses shifts to be more aligned with the cue onset as the cue becomes more predictive.

### D2 MSN cue responses are necessary for learning

Finally, to test whether D2 MSN activity at the time of a cue was necessary for associative learning, we expressed halorhodopsin

(NpHR), an inhibitory opsin, selectively in D1 or D2 MSNs, using D1- and A2A-Cre mice, respectively (Figure 7A). Mice underwent fear conditioning, as described above, where a 5-s cue was paired with a brief shock. On each trial, D1 and D2 MSNs were inhibited for the duration of the cue presentation (590 nm, 8 mW, constant; Figure 7B). Inhibiting D2 MSNs during the cue presentation resulted in a deficit in learning the cue-outcome association as compared with the D1-cre and eYFP controls (Figures 7C and 7D). D1 MSN inhibition at the time of the predictive cue did not have any effect on the trajectory of associative learning (Figures 7C and 7D). These effects cannot be explained by changes in motor activity, as these stimulations had no effect on open-field locomotion (Figures S7A–S7C). Additionally, when D2 MSNs were inhibited during appetitive conditioning, this also slowed the trajectory of learning for sucrose in a reinforcement task (Figures S7D and S7F). In the larger context of our study, these results strongly suggest that the positive and cue-evoked D2 MSN signal in the NAc that occurs at the time of the predictive cue causally mediates cue-outcome associations, regardless of valence.

### DISCUSSION

Here, using a wide range of paradigms that result in behavioral responses to stimuli of both positive and negative valence, we characterized the precise role that NAc core D1 and D2 MSN activity plays in behavioral control. At the population and single-cell level, we showed that D1 MSNs in the NAc core responded to the presence of unconditioned stimuli—regardless of valence. Conversely, D2 MSNs in the NAc core responded in a prediction-based fashion, increasing with learning, scaling with prediction, and causally controlling the trajectory of associative learning. Overall, these data show that D1 and D2 MSNs in the NAc core do not have opposing roles in behavioral control. Rather, they work in tandem to provide information regarding specific valence-independent aspects of associative learning.

While the NAc has been referred to as a reward-associated brain region,<sup>23,32,59,60</sup> these studies are not the first to show that signaling within the NAc, and also the dorsal striatum,<sup>61</sup> is causally related to both aversive and appetitive stimulus processing.<sup>17,19,21–26,55,62,63</sup> While some of these studies explained these results based on the idea of bidirectional valence coding, the results from our experiments, which examine a wider range of behavioral responses and contingencies, suggest that signaling in the NAc—especially the core region<sup>63–67</sup>—may best be explained by valence-independent factors that allow for adaptive behavioral control across contexts and conditions. This is also consistent with data across a range of experiments and fields showing that D1 and D2 MSNs receive overlapping inputs of sensory and cortical systems<sup>68</sup> and stimulation of inputs into the NAc from the prefrontal cortex, hippocampus, or basolateral amygdala inputs are reinforcing.<sup>31,69,70</sup> Indeed, we and others show that the stimulation of both D1 and D2 MSNs supports reinforcement,<sup>47,49–51</sup> indicating that these populations do not track valence or opposing motivational drive. It is important to note, however, that there is regional heterogeneity when it comes to behavioral control in the NAc,<sup>67</sup> and it is possible that these effects are specific to the NAc core. Nevertheless, these data still show that the idea of opposing valence cannot explain the role of these populations ubiquitously throughout the ventral striatum.

The framework presented within this manuscript is consistent with a large amount of emerging data on the role of these populations in behavioral control. Recent work has shown that manipulating D2 MSNs alters behavior, especially under unpredictable conditions (risky choice) and when behavioral updating is necessary.<sup>53</sup> Zalocusky et al.<sup>53</sup> showed that in a model of risk preference in rats, stimulation of D2 MSNs during a period where animals were presented with a choice to emit an operant response or not resulted in fewer risky choices. Here, we show that D2 MSNs track the predictability of cues and are engaged during similar types of operant tasks in a fashion that tracks performance. Nishioka et al.<sup>48</sup> showed that optogenetic inhibition of D2 MSNs in the NAc during trials where there were errors in outcomes negatively impacted performance in the subsequent trial. Here, we show that D2 MSNs track predictions and signal when outcomes are unexpected, something that would be necessary for error-based updating.

These data are also consistent with previous work using optogenetics, where activation of D2 MSNs during a reward-predictive cue enhances motivation.<sup>50,71</sup> This would be predicted based on the idea that D2 MSNs transmit a prediction signal, and this would enhance the association between cues and outcomes, thus increasing operant performance. This same work also showed that triggering activity during the food pellet delivery reduces motivation,<sup>50,71</sup> which would also be expected if D2 MSNs were transmitting a prediction error signal as an increase in the D2 MSN response at the time of the outcome would signal an error in prediction—i.e., that food is no longer presented following that action—and reduce future responding. The data included within this manuscript are also consistent with other recent work showing that both D1 and D2 MSN activity is evoked in response to food rewards,<sup>72–74</sup> where we show here that unexpected sucrose delivery results in increased D1 MSN responses (signaling stimulus presence) and D2 MSN responses (signaling an unexpected outcome).

Together, the data presented within the current manuscript and data from others converge to show that D2 MSN activity is critical in situations where updating is necessary. Overall, the ability of D1 and D2 MSNs to signal precise types of information along with the fact that there is evidence of collateral transmission between MSNs suggests an ongoing relationship between D1 and D2 MSNs,<sup>75,76</sup> which likely allows for the ability to refine responses to relevant environmental stimuli.

Finally, these data are critical for conceptualizing how experience-dependent changes within these two cellular populations give rise to behavioral maladaptation associated with disease states. There has been extensive characterization of the effects of drugs of abuse on D1 and D2 MSNs at the molecular and cellular level. Acute drug exposure has been shown to enhance D1 MSN, while suppressing D2 MSN activity—likely through activation of dopamine receptors on these populations.<sup>38,77</sup> Additionally, repeated drug exposure has been shown to result in long-lasting enhancement of synaptic activity in D1 MSNs relative to D2 MSNs.<sup>38,39,78–81</sup> D2 MSNs have also been linked to drug-associated plasticity and seeking, especially as it relates to cue-driven behavior like drug seeking, or drug-taking under more variable reinforcement schedules,<sup>37,82</sup> a result that would be predicted based on these data as well. Given the critical role we have identified here for D1 and D2 MSNs in associative learning, it is plausible that drugs of abuse may disrupt the balance critical for forming associations between predictive cues and outcomes.<sup>33–35,37,43,83</sup>

By experimentation that integrates learning across contexts, behavioral action, and valence, we provide a framework for accumbal MSNs in valence-independent behavioral control. We conclude that NAc core MSNs help coordinate the learning of associations to drive adaptive behavior in all cases. The roles that D1 and D2 MSNs play in adaptive behavior thus provide new insights into many of the psychopathologies associated with disruptions in this brain region—including substance use disorders, gambling, and depression among others.<sup>3–5,7,8,84</sup> In effect, we can frame many of these disorders as dysregulation in adaptive associative learning processes, a perspective that can ultimately reshape how we treat them.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Subjects
- METHOD DETAILS
  - General Surgical Procedures
  - Histology
- FIBER PHOTOMETRY
  - Fiber photometry general approach



- Fiber photometry analysis
- **CELLULAR RESOLUTION CALCIUM IMAGING VIA MICROENDOSCOPES**
  - Single cell imaging general approach
  - Image processing and signal extraction
  - Calcium activity quantification
  - Longitudinal co-registration
  - Neural trajectory analysis
  - Hierarchical clustering
- **BEHAVIORAL EXPERIMENTS**
  - Apparatus
  - Positive Reinforcement
  - Negative Reinforcement
  - Unsignaled footshocks of varying intensities
  - Unsignaled sucrose delivery
  - Fear conditioning followed by extinction
  - Fear conditioning with changing cue-outcome probabilities
- **OPTOGENETICS**
  - Intracranial Self-Stimulation (ICSS)
  - Optical inhibition of D1- and D2-MSNs during fear conditioning
  - Optical inhibition of D1- and D2-MSNs during positive reinforcement
  - Optical inhibition of D1- and D2-MSNs during open field test
  - Optogenetic excitation approach validation via fiber photometry
- **QUANTIFICATION AND STATISTICAL ANALYSIS**

#### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.neuron.2023.11.023>.

#### ACKNOWLEDGMENTS

This work was supported by NIH grants DA048931, AA030931, and DA052317 to E.S.C., R21MH132052 and KL2TR002245 to M.G.K., and GM07628 to J.E.Z. as well as by funds from the VUMC Faculty Research Scholar Award to M.G.K., the Pfeil Foundation to M.G.K., Brain and Behavior Research Foundation to M.G.K. and E.S.C., the Whitehall Foundation to E.S.C., and the Edward Mallinckrodt, Jr. Foundation to E.S.C.

#### AUTHOR CONTRIBUTIONS

Concept, study coordination, and experimental design: J.E.Z., M.G.K., and E.S.C. Experimental execution (including surgeries, behavioral, and optogenetics experiments, histology): J.E.Z., M.G.K., E.S.C., D.D.P., A.G., P.R.M., A.U.I., H.J.Y., M.Z.L., M.C., V.K., K.C.T., R.B., and J.T. Data analysis and pipeline: J.E.Z., M.G.K., and M.C. Wrote the manuscript: J.E.Z., M.G.K., and E.S.C. Scientific guidance: E.S.C., M.E.J., B.A.G., P.J.C., and D.J.C. All authors edited and reviewed the manuscript.

#### DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: December 20, 2022

Revised: October 3, 2023

Accepted: November 22, 2023

Published: December 21, 2023

#### REFERENCES

1. Cardinal, R.N. (2006). Neural systems implicated in delayed and probabilistic reinforcement. *Neural Netw.* *19*, 1277–1301.
2. Green, L., and Myerson, J. (2004). A discounting framework for choice with delayed and probabilistic rewards. *Psychol. Bull.* *130*, 769–792.
3. Avanzi, M., Uber, E., and Bonfà, F. (2004). Pathological gambling in two patients on dopamine replacement therapy for Parkinson's disease. *Neurol. Sci.* *25*, 98–101.
4. Chang, P.C., Pollema-Mays, S.L., Centeno, M.V., Proccissi, D., Contini, M., Baria, A.T., Martina, M., and Apkarian, A.V. (2014). Role of nucleus accumbens in neuropathic pain: linked multi-scale evidence in the rat transitioning to neuropathic pain. *Pain.* *155*, 1128–1139.
5. Dodd, M.L., Klos, K.J., Bower, J.H., Geda, Y.E., Josephs, K.A., and Ahlskog, J.E. (2005). Pathological gambling caused by drugs used to treat Parkinson disease. *Arch. Neurol.* *62*, 1377–1381.
6. Keiflin, R., and Janak, P.H. (2015). Dopamine prediction errors in reward learning and addiction: from theory to neural circuitry. *Neuron* *88*, 247–263.
7. Redish, A.D. (2004). Addiction as a computational process gone awry. *Science* *306*, 1944–1947.
8. Schultz, W. (2011). Potential vulnerabilities of neuronal reward, risk, and decision mechanisms to addictive drugs. *Neuron* *69*, 603–617.
9. Wise, R.A., and Koob, G.F. (2014). The development and maintenance of drug addiction. *Neuropsychopharmacology* *39*, 254–262.
10. Bussey, T. (1999). Novelty in the brain. *Trends Cogn. Sci.* *3*, 126.
11. Lubow, R.E. (1973). Latent inhibition. *Psychol. Bull.* *79*, 398–407.
12. Lubow, R.E. (1973). Latent inhibition as a means of behavior prophylaxis. *Psychol. Rep.* *32*, 1247–1252.
13. Parkin, A.J. (1997). Human memory: novelty, association and the brain. *Curr. Biol.* *7*, R768–R769.
14. Quintero, E., Díaz, E., Vargas, J.P., Schmajuk, N., López, J.C., and De la Casa, L.G. (2011). Effects of context novelty vs. familiarity on latent inhibition with a conditioned taste aversion procedure. *Behav. Processes* *86*, 242–249.
15. Rescorla, R., and Wagner, A. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. *Class. Cond.: Curr. Res. Theor.* *2*, 64–99.
16. Russell, J.A. (2003). Core affect and the psychological construction of emotion. *Psychol. Rev.* *110*, 145–172.
17. Kutlu, M.G., Zachry, J.E., Brady, L.J., Melugin, P.R., Kelly, S.J., Sanders, C., Tat, J., Johnson, A.R., Thibeault, K., Lopez, A.J., et al. (2020). A novel multidimensional reinforcement task in mice elucidates sex-specific behavioral strategies. *Neuropsychopharmacology* *45*, 1463–1472.
18. Thibeault, K.C., Kutlu, M.G., Sanders, C., and Calipari, E.S. (2019). Cell-type and projection-specific dopaminergic encoding of aversive stimuli in addiction. *Brain Res.* *1713*, 1–15.
19. Kutlu, M.G., Zachry, J.E., Melugin, P.R., Cajigas, S.A., Chevee, M.F., Kelly, S.J., Kutlu, B., Tian, L., Siciliano, C.A., and Calipari, E.S. (2021). Dopamine release in the nucleus accumbens core signals perceived saliency. *Curr. Biol.* *31*, 4748–4761.e8.
20. Oleson, E.B., Gentry, R.N., Chioma, V.C., and Cheer, J.F. (2012). Subsecond dopamine release in the nucleus accumbens predicts conditioned punishment and its successful avoidance. *J. Neurosci.* *32*, 14804–14808.
21. Wenzel, J.M., Oleson, E.B., Gove, W.N., Cole, A.B., Gyawali, U., Dantrassy, H.M., Bluett, R.J., Dryanovski, D.I., Stuber, G.D., Deisseroth, K., et al. (2018). Phasic dopamine signals in the nucleus accumbens that cause active avoidance require endocannabinoid mobilization in the midbrain. *Curr. Biol.* *28*, 1392–1404.e5.
22. Bin Saifullah, M.A., Nagai, T., Kuroda, K., Wulaer, B., Nabeshima, T., Kaibuchi, K., and Yamada, K. (2018). Cell type-specific activation of mitogen-activated protein kinase in D1 receptor-expressing neurons of the nucleus accumbens potentiates stimulus-reward learning in mice. *Sci. Rep.* *8*, 14413.



23. Day, J.J., and Carelli, R.M. (2007). The nucleus accumbens and Pavlovian reward learning. *Neuroscientist* *13*, 148–159.
24. de Jong, J.W., Afjei, S.A., Pollak Dorocic, I., Peck, J.R., Liu, C., Kim, C.K., Tian, L., Deisseroth, K., and Lammel, S. (2019). A neural circuit mechanism for encoding aversive stimuli in the mesolimbic dopamine system. *Neuron* *101*, 133–151.e7.
25. Ray, M.H., Russ, A.N., Walker, R.A., and McDannald, M.A. (2020). The nucleus accumbens core is necessary to scale fear to degree of threat. *J. Neurosci.* *40*, 4750–4760.
26. Sugam, J.A., Saddoris, M.P., and Carelli, R.M. (2014). Nucleus accumbens neurons track behavioral preferences and reward outcomes during risky decision making. *Biol. Psychiatry* *75*, 807–816.
27. Albin, R.L., Young, A.B., and Penney, J.B. (1989). The functional anatomy of basal ganglia disorders. *Trends Neurosci.* *12*, 366–375.
28. Gerfen, C.R. (1992). The neostriatal mosaic: multiple levels of compartmental organization in the basal ganglia. *Annu. Rev. Neurosci.* *15*, 285–320.
29. Gerfen, C.R., Engber, T.M., Mahan, L.C., Susel, Z., Chase, T.N., Monsma, F.J., and Sibley, D.R. (1990). D1 and D2 dopamine receptor-regulated gene expression of striatonigral and striatopallidal neurons. *Science* *250*, 1429–1432.
30. Kupchik, Y.M., Brown, R.M., Heinsbroek, J.A., Lobo, M.K., Schwartz, D.J., and Kalivas, P.W. (2015). Coding the direct/indirect pathways by D1 and D2 receptors is not valid for accumbens projections. *Nat. Neurosci.* *18*, 1230–1232.
31. Britt, J.P., Benaliouad, F., McDevitt, R.A., Stuber, G.D., Wise, R.A., and Bonci, A. (2012). Synaptic and behavioral profile of multiple glutamatergic inputs to the nucleus accumbens. *Neuron* *76*, 790–803.
32. Russo, S.J., and Nestler, E.J. (2013). The brain reward circuitry in mood disorders. *Nat. Rev. Neurosci.* *14*, 609–625.
33. Bateup, H.S., Santini, E., Shen, W., Birnbaum, S., Valjent, E., Surmeier, D.J., Fisone, G., Nestler, E.J., and Greengard, P. (2010). Distinct subclasses of medium spiny neurons differentially regulate striatal motor behaviors. *Proc. Natl. Acad. Sci. USA* *107*, 14845–14850.
34. Kravitz, A.V., Freeze, B.S., Parker, P.R.L., Kay, K., Thwin, M.T., Deisseroth, K., and Kreitzer, A.C. (2010). Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry. *Nature* *466*, 622–626.
35. Kravitz, A.V., Tye, L.D., and Kreitzer, A.C. (2012). Distinct roles for direct and indirect pathway striatal neurons in reinforcement. *Nat. Neurosci.* *15*, 816–818.
36. Parker, J.G., Marshall, J.D., Ahanonu, B., Wu, Y.W., Kim, T.H., Grewe, B.F., Zhang, Y., Li, J.Z., Ding, J.B., Ehlers, M.D., et al. (2018). Diametric neural ensemble dynamics in parkinsonian and dyskinetic states. *Nature* *557*, 177–182.
37. Bock, R., Shin, J.H., Kaplan, A.R., Dobi, A., Markey, E., Kramer, P.F., Gremel, C.M., Christensen, C.H., Adrover, M.F., and Alvarez, V.A. (2013). Strengthening the accumbal indirect pathway promotes resilience to compulsive cocaine use. *Nat. Neurosci.* *16*, 632–638.
38. Calipari, E.S., Bagot, R.C., Purushothaman, I., Davidson, T.J., Yorgason, J.T., Peña, C.J., Walker, D.M., Pirpinias, S.T., Guise, K.G., Ramakrishnan, C., et al. (2016). In vivo imaging identifies temporal signature of D1 and D2 medium spiny neurons in cocaine reward. *Proc. Natl. Acad. Sci. USA* *113*, 2726–2731.
39. Creed, M., Ntamati, N.R., Chandra, R., Lobo, M.K., and Lüscher, C. (2016). Convergence of reinforcing and anhedonic cocaine effects in the ventral pallidum. *Neuron* *92*, 214–226.
40. Danjo, T., Yoshimi, K., Funabiki, K., Yawata, S., and Nakanishi, S. (2014). Aversive behavior induced by optogenetic inactivation of ventral tegmental area dopamine neurons is mediated by dopamine D2 receptors in the nucleus accumbens. *Proc. Natl. Acad. Sci. USA* *111*, 6455–6460.
41. Francis, T.C., Chandra, R., Friend, D.M., Finkel, E., Dayrit, G., Miranda, J., Brooks, J.M., Iñiguez, S.D., O'Donnell, P., Kravitz, A., et al. (2015). Nucleus accumbens medium spiny neuron subtypes mediate depression-related outcomes to social defeat stress. *Biol. Psychiatry* *77*, 212–222.
42. Hearing, M., Graziane, N., Dong, Y., and Thomas, M.J. (2018). Opioid and psychostimulant plasticity: targeting overlap in nucleus accumbens glutamate signaling. *Trends Pharmacol. Sci.* *39*, 276–294.
43. Lobo, M.K., Covington, H.E., Chaudhury, D., Friedman, A.K., Sun, H., Damez-Werno, D., Dietz, D.M., Zaman, S., Koo, J.W., Kennedy, P.J., et al. (2010). Cell type-specific loss of BDNF signaling mimics optogenetic control of cocaine reward. *Science* *330*, 385–390.
44. Lobo, M.K., and Nestler, E.J. (2011). The striatal balancing act in drug addiction: distinct roles of direct and indirect pathway medium spiny neurons. *Front. Neuroanat.* *5*, 41.
45. Jaskir, A., and Frank, M.J. (2023). On the normative advantages of dopamine and striatal opponency for learning and choice. *eLife* *12*, e85107.
46. Al-Hasani, R., McCall, J.G., Shin, G., Gomez, A.M., Schmitz, G.P., Bernardi, J.M., Pyo, C.O., Park, S.I., Marcinkiewicz, C.M., Crowley, N.A., et al. (2015). Distinct subpopulations of nucleus accumbens dynorphin neurons drive aversion and reward. *Neuron* *87*, 1063–1077.
47. Cole, S.L., Robinson, M.J.F., and Berridge, K.C. (2018). Optogenetic self-stimulation in the nucleus accumbens: D1 reward versus D2 ambivalence. *PLoS One* *13*, e0207694.
48. Nishioka, T., Macpherson, T., Hamaguchi, K., and Hikida, T. (2021). Distinct roles of dopamine D1 and D2 receptor-expressing neurons in the nucleus accumbens for a strategy dependent decision making. *bioRxiv*.
49. Soares-Cunha, C., Coimbra, B., Sousa, N., and Rodrigues, A.J. (2016). Reappraising striatal D1- and D2-neurons in reward and aversion. *Neurosci. Biobehav. Rev.* *68*, 370–386.
50. Soares-Cunha, C., Coimbra, B., Domingues, A.V., Vasconcelos, N., Sousa, N., and Rodrigues, A.J. (2018). Nucleus accumbens microcircuit underlying D2-MSN-driven increase in motivation. *eNeuro* *5*.
51. Soares-Cunha, C., de Vasconcelos, N.A.P., Coimbra, B., Domingues, A.V., Silva, J.M., Loureiro-Campos, E., Gaspar, R., Sotiropoulos, I., Sousa, N., and Rodrigues, A.J. (2020). Nucleus accumbens medium spiny neurons subtypes signal both reward and aversion. *Mol. Psychiatry* *25*, 3241–3255.
52. Soares-Cunha, C., Domingues, A.V., Correia, R., Coimbra, B., Veitas-Gaspar, N., de Vasconcelos, N.A.P., Pinto, L., Sousa, N., and Rodrigues, A.J. (2022). Distinct role of nucleus accumbens D2-MSN projections to ventral pallidum in different phases of motivated behavior. *Cell Rep.* *38*, 110380.
53. Zalocusky, K.A., Ramakrishnan, C., Lerner, T.N., Davidson, T.J., Knutson, B., and Deisseroth, K. (2016). Nucleus accumbens D2R cells signal prior outcomes and control risky decision-making. *Nature* *531*, 642–646.
54. Gallo, E.F., Meszaros, J., Sherman, J.D., Chohan, M.O., Teboul, E., Choi, C.S., Moore, H., Javitch, J.A., and Kellendonk, C. (2018). Accumbens dopamine D2 receptors increase motivation by decreasing inhibitory transmission to the ventral pallidum. *Nat. Commun.* *9*, 1086.
55. Roitman, M.F., Wheeler, R.A., and Carelli, R.M. (2005). Nucleus accumbens neurons are innately tuned for rewarding and aversive taste stimuli, encode their predictors, and are linked to motor output. *Neuron* *45*, 587–597.
56. Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. *Science* *275*, 1593–1599.
57. Fiorillo, C.D., Tobler, P.N., and Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* *299*, 1898–1902.
58. Tobler, P.N., Fiorillo, C.D., and Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science* *307*, 1642–1645.
59. Nicola, S.M. (2007). The nucleus accumbens as part of a basal ganglia action selection circuit. *Psychopharmacology* *191*, 521–550.
60. Willmore, L., Cameron, C., Yang, J., Witten, I.B., and Falkner, A.L. (2022). Behavioural and dopaminergic signatures of resilience. *Nature* *611*, 124–132.

61. Bloem, B., Huda, R., Amemori, K.I., Abate, A.S., Krishna, G., Wilson, A.L., Carter, C.W., Sur, M., and Graybiel, A.M. (2022). Multiplexed action-outcome representation by striatal striosome-matrix compartments detected with a mouse cost-benefit foraging task. *Nat. Commun.* *13*, 1541.
62. Iordanova, M.D., Yau, J.O.-Y., McDannald, M.A., and Corbit, L.H. (2021). Neural substrates of appetitive and aversive prediction error. *Neurosci. Biobehav. Rev.* *123*, 337–351.
63. Li, S.S.Y., and McNally, G.P. (2015). A role of nucleus accumbens dopamine receptors in the nucleus accumbens core, but not shell, in fear prediction error. *Behav. Neurosci.* *129*, 450–456.
64. Ambroggi, F., Ghazizadeh, A., Nicola, S.M., and Fields, H.L. (2011). Roles of nucleus accumbens core and shell in incentive-cue responding and behavioral inhibition. *J. Neurosci.* *31*, 6820–6830.
65. Corbit, L.H., Muir, J.L., and Balleine, B.W. (2001). The role of the nucleus accumbens in instrumental conditioning: evidence of a functional dissociation between accumbens core and shell. *J. Neurosci.* *21*, 3251–3260.
66. Floresco, S.B., McLaughlin, R.J., and Haluk, D.M. (2008). Opposing roles for the nucleus accumbens core and shell in cue-induced reinstatement of food-seeking behavior. *Neuroscience* *154*, 877–884.
67. West, E.A., and Carelli, R.M. (2016). Nucleus accumbens core and shell differentially encode reward-associated cues after reinforcer devaluation. *J. Neurosci.* *36*, 1128–1139.
68. Guo, Q., Wang, D., He, X., Feng, Q., Lin, R., Xu, F., Fu, L., and Luo, M. (2015). Whole-brain mapping of inputs to projection neurons and cholinergic interneurons in the dorsal striatum. *PLoS One* *10*, e0123381.
69. Otis, J.M., Nambodiri, V.M.K., Matan, A.M., Voets, E.S., Mohorn, E.P., Kosyk, O., McHenry, J.A., Robinson, J.E., Resendez, S.L., Rossi, M.A., et al. (2017). Prefrontal cortex output circuits guide reward seeking through divergent cue encoding. *Nature* *543*, 103–107.
70. Stuber, G.D., Sparta, D.R., Stamatakis, A.M., van Leeuwen, W.A., Hardjoprajitno, J.E., Cho, S., Tye, K.M., Kempadoo, K.A., Zhang, F., Deisseroth, K., et al. (2011). Excitatory transmission from the amygdala to nucleus accumbens facilitates reward seeking. *Nature* *475*, 377–380.
71. Soares-Cunha, C., Coimbra, B., David-Pereira, A., Borges, S., Pinto, L., Costa, P., Sousa, N., and Rodrigues, A.J. (2016). Activation of D2 dopamine receptor-expressing neurons in the nucleus accumbens increases motivation. *Nat Commun* *7*, 11829.
72. Baldo, B.A., and Kelley, A.E. (2007). Discrete neurochemical coding of distinguishable motivational processes: insights from nucleus accumbens control of feeding. *Psychopharmacology* *191*, 439–459.
73. Carlezon, W.A., and Thomas, M.J. (2009). Biological substrates of reward and aversion: a nucleus accumbens activity hypothesis. *Neuropharmacology* *56* (Supplement 1), 122–132.
74. Natsubori, A., Tsutsui-Kimura, I., Nishida, H., Boucekkioua, Y., Sekiya, H., Uchigashima, M., Watanabe, M., de Kerchove d'Exaerde, A.de K., Mimura, M., Takata, N., et al. (2017). Ventrolateral striatal medium spiny neurons positively regulate food-incentive, goal-directed behavior independently of D1 and D2 selectivity. *J. Neurosci.* *37*, 2723–2733.
75. Burke, D.A., Rotstein, H.G., and Alvarez, V.A. (2017). Striatal local circuitry: a new framework for lateral inhibition. *Neuron* *96*, 267–284.
76. Dobbs, L.K., Kaplan, A.R., Lemos, J.C., Matsui, A., Rubinstein, M., and Alvarez, V.A. (2016). Dopamine regulation of lateral inhibition between striatal neurons gates the stimulant actions of cocaine. *Neuron* *90*, 1100–1113.
77. Luo, Z., Volkow, N.D., Heintz, N., Pan, Y., and Du, C. (2011). Acute cocaine induces fast activation of D1 receptor and progressive deactivation of D2 receptor striatal neurons: in vivo optical microprobe [Ca<sup>2+</sup>] imaging. *J. Neurosci.* *31*, 13180–13190.
78. MacAskill, A.F., Cassel, J.M., and Carter, A.G. (2014). Cocaine exposure reorganizes cell-type and input-specific connectivity in the nucleus accumbens. *Nat. Neurosci.* *17*, 1198–1207.
79. Pascoli, V., Terrier, J., Espallargues, J., Valjent, E., O'Connor, E.C., and Lüscher, C. (2014). Contrasting forms of cocaine-evoked plasticity control components of relapse. *Nature* *509*, 459–464.
80. Joffe, M.E., and Grueter, B.A. (2016). Cocaine experience enhances thalamo-accumbens N-methyl-D-aspartate receptor function. *Biol. Psychiatry* *80*, 671–681.
81. Joffe, M.E., Turner, B.D., Delpire, E., and Grueter, B.A. (2018). Genetic loss of GluN2B in D1-expressing cell types enhances long-term cocaine reward and potentiation of thalamo-accumbens synapses. *Neuropsychopharmacology* *43*, 2383–2389.
82. Heinsbroek, J.A., Neuhofer, D.N., Griffin, W.C., Siegel, G.S., Bobadilla, A.C., Kupchik, Y.M., and Kalivas, P.W. (2017). Loss of plasticity in the D2-accumbens pallidal pathway promotes cocaine seeking. *J. Neurosci.* *37*, 757–767.
83. Durieux, P.F., Bearzatto, B., Guiducci, S., Buch, T., Waisman, A., Zoli, M., Schiffmann, S.N., and de Kerchove d'Exaerde, A. (2009). D2R striatopallidal neurons inhibit both locomotor and drug reward processes. *Nat. Neurosci.* *12*, 393–395.
84. Xu, L., Nan, J., and Lan, Y. (2020). The nucleus accumbens: A common target in the comorbidity of depression and addiction. *Front. Neural Circuits* *14*, 37.
85. Chen, T.W., Wardill, T.J., Sun, Y., Pulver, S.R., Renninger, S.L., Baohan, A., Schreiter, E.R., Kerr, R.A., Orger, M.B., Jayaraman, V., et al. (2013). Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* *499*, 295–300.
86. Gradinaru, V., Zhang, F., Ramakrishnan, C., Mattis, J., Prakash, R., Diester, I., Goshen, I., Thompson, K.R., and Deisseroth, K. (2010). Molecular and cellular approaches for diversifying and extending optogenetics. *Cell* *141*, 154–165.
87. Mattis, J., Tye, K.M., Ferenczi, E.A., Ramakrishnan, C., O'Shea, D.J., Prakash, R., Gunaydin, L.A., Hyun, M., Fenno, L.E., Gradinaru, V., et al. (2012). Principles for applying optogenetic tools derived from direct comparative analysis of microbial opsins. *Nature methods* *9*, 159–172.
88. Challis, R.C., Ravindra Kumar, S., Chan, K.Y., Challis, C., Beadle, K., Jang, M.J., Kim, H.M., Rajendran, P.S., Tompkins, J.D., Shivkumar, K., et al. (2019). Systemic AAV vectors for widespread and targeted gene delivery in rodents. *Nature protocols* *14*, 379–414.
89. Klapoetke, N.C., Murata, Y., Kim, S.S., Pulver, S.R., Birdsey-Benson, A., Cho, Y.K., Morimoto, T.K., Chuong, A.S., Carpenter, E.J., Tian, Z., et al. (2014). Independent optical excitation of distinct neural populations. *Nature methods* *11*, 338–346.
90. Zhou, P., Resendez, S.L., Rodriguez-Romaguera, J., Jimenez, J.C., Neufeld, S.Q., Giovannucci, A., Friedrich, J., Pnevmatikakis, E.A., Stuber, G.D., Hen, R., et al. (2018). Efficient and accurate extraction of in vivo calcium signals from microendoscopic video data. *eLife* *7*, e28728.
91. Dana, H., Sun, Y., Mohar, B., Hulse, B.K., Kerlin, A.M., Hasseman, J.P., Tsegaye, G., Tsang, A., Wong, A., Patel, R., et al. (2019). High-performance calcium sensors for imaging activity in neuronal populations and microcompartments. *Nat. Methods* *16*, 649–657.
92. Lerner, T.N., Shilyansky, C., Davidson, T.J., Evans, K.E., Beier, K.T., Zalocusky, K.A., Crow, A.K., Malenka, R.C., Luo, L., Tomer, R., et al. (2015). Intact-brain analyses reveal distinct information carried by SNc dopamine subcircuits. *Cell* *162*, 635–647.
93. Tian, L., Hires, S.A., Mao, T., Huber, D., Chiappe, M.E., Chalasani, S.H., Petreanu, L., Akerboom, J., McKinney, S.A., Schreiter, E.R., et al. (2009). Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators. *Nat. Methods* *6*, 875–881.
94. Kutlu, M.G., Zachry, J.E., Melugin, P.R., Tat, J., Cajigas, S., Isiktas, A.U., Patel, D.D., Siciliano, C.A., Schoenbaum, G., Sharpe, M.J., et al. (2022). Dopamine signaling in the nucleus accumbens core mediates latent inhibition. *Nat. Neurosci.* *25*, 1071–1081.
95. Pnevmatikakis, E.A., Soudry, D., Gao, Y., Machado, T.A., Merel, J., Pfau, D., Reardon, T., Mu, Y., Lacefield, C., Yang, W., et al. (2016). Simultaneous denoising, deconvolution, and demixing of calcium imaging data. *Neuron* *89*, 285–299.

## STAR★METHODS

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Antibodies</b>		
anti-GFP	Abcam	#AB13970, RRID:AB_300798
anti-chicken AlexaFluor 488	Life Technologies	#A-11039, RRID:AB_2534096
<b>Bacterial and virus strains</b>		
AAV5.Syn.Flex.GCaMP6f.WPRE.SV40	Chen et al. <sup>85</sup>	Addgene #100833-AAV5, RRID:Addgene_100833
AAV5.CAG.Flex.GCaMP6m.WPRE.SV40	Chen et al. <sup>85</sup>	Addgene #100839-AAV5, RRID:Addgene_100839
AAV5. Ef1a.DIO.NpHR3.0-YFP	Deisseroth Lab Gradinaru et al. <sup>86</sup>	Addgene # 26966-AAV5, RRID:Addgene_26966
AAV5.Ef1a.DIO.hChR2	Mattis et al. <sup>87</sup>	Addgene #35509-AAV5, RRID:Addgene_35509
AAV5.hSyn1.eYFP	Challis et al. <sup>88</sup>	Addgene #117382-AAV5, RRID:Addgene_117382
AAV5-Syn-FLEX-rc[ChrimsonR-tdTomato]	Klapoetke et al. <sup>89</sup>	Addgene #62723-AAV5, RRID:Addgene_62723
<b>Experimental models: Organisms/strains</b>		
C57BL/6J mice	Jackson Laboratories	SN: 000664
D1-Cre mice	Jackson Laboratories	SN: 030329
Adora2a (A2A)-Cre mice	Mutant Mouse Resource and Research Centers	SN: 036158-UCD
<b>Software and algorithms</b>		
Matlab R2019b	Mathworks	N/A
Prism9.0	Graphpad	N/A
Constrained non-negative matrix factorization algorithm (CNMF-E)	Zhou et al. <sup>90</sup>	N/A
Inscopix Data Process Software (IDPS)	Inscopix	N/A

### RESOURCE AVAILABILITY

#### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Erin Calipari ([erin.calipari@vanderbilt.edu](mailto:erin.calipari@vanderbilt.edu)).

#### Materials availability

This study did not generate new unique reagents.

#### Data and code availability

- Data reported in this paper are available from the [lead contact](#) upon reasonable request.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

### EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

#### Subjects

Male and female 6- to 8-week-old D1:Cre (Jax: #030329) and A2A:Cre (MMRRC RRID: *MMRRC\_036158-UCD*) mice were obtained from Jackson Laboratories and the Mutant Mouse Resource and Research Centers (MMRC) repository (University of Missouri, Columbia, MO), respectively. All animals were maintained on a 12h reverse light/dark cycle. Animals had ad-libitum access to food and water except for the studies that employed sucrose as the main reinforcer, where mice were food restricted to 90% of free-feeding weight for the duration of the studies. Mice were weighed every other day to ensure that weight was maintained. All experiments were conducted in accordance with the guidelines of the Institutional Animal Care and Use Committee (IACUC) at Vanderbilt University School of Medicine, which approved and supervised all animal protocols.

## METHOD DETAILS

### General Surgical Procedures

Ketoprofen (5mg/kg; subcutaneous injection) was administered at least 30 mins before any surgical procedure. Ophthalmic ointment was applied to the eyes before and throughout the duration of surgery. Under Isoflurane anesthesia (induced at 4%, maintained at 1.5% in  $O_2$ ), mice were positioned in a stereotaxic frame (Kopf Instruments, Tujunga, CA). Using aseptic techniques, a midline incision was made down the scalp and a craniotomy was made using a dental drill. A 10-mL Nanofil Hamilton syringe (WPI) with a 34-gauge beveled metal needle was used to infuse viral constructs into the NAc core as follows:

#### For fiber photometry and optogenetics

The NAc core was targeted for virus injection (from bregma, A/P = +1.4 mm; ML = +1.5 mm; D/V = -4.3 mm; 10° angle). Targeting was unilateral for fiber photometry and optogenetic excitation (channelrhodopsin) experiments and bilateral for optogenetic inhibition (halorhodopsin) experiments. Virus was infused at a rate of 50 nL/min for a total of 500 nL. All viruses were adeno-associated viruses (AAV), and which viruses were used for each experiment are outlined in the experimental details. Following AAV infusion, the needle was kept at the injection site for seven minutes and then slowly withdrawn. Permanent implantable 2.5 mm fiber optic canula were implanted in the NAc [(Doric Lenses, Quebec, Canada) (Fiber Photometry: 400 $\mu$ M fiber core, 0.48 NA; Optogenetics: 200 $\mu$ M fiber core, 0.22 NA)]. Fiberoptic cannula were positioned above the viral injection site (from bregma, A/P = +1.4 mm; ML = +1.5 mm; D/V = -4.2 mm; 10° angle) and were cemented to the skull using C&B Metabond adhesive cement system (C&B Metabond; Parkell).

#### For microendoscopic imaging

Viral injection targeting of the NAc core was done as described above (from bregma, A/P = +1.4 mm; M/L = +1.0 mm; DV = -3.8; 0° angle). These coordinates targeted the same area of the NAc core as above but without an angle. After viral infusion, the craniotomy was enlarged, and the dura mater was removed with a 26-gauge beveled needle. No tissue was aspirated. Proview integrated lenses (Inscopix) were used for all mice in these studies. These lenses are a combination of a traditional gradient index (GRIN) lens with a baseplate already attached, allowing for lens implantation and baseplate surgery to be done in the same surgical procedure. A 0.6 mm diameter, GRIN lens (7.3mm length, Inscopix) was lowered stereotaxically through the craniotomy -3.7 mm into the brain. The remaining portion of the lens was above the skull and attached to the baseplate. Lenses were implanted slightly above the track for virus infusions to avoid tissue damage in the imaging plane. the lens was positioned above the viral injection site (from bregma, A/P = +1.4 mm; M/L = +1.0 mm; D/V = -3.7 mm; 0° angle) and was cemented to the skull using adhesive cement (C&B Metabond; Parkell). The baseplate, which is attached to the lens, was cemented around the lens and to the skull to support the connection of the miniaturized microscope for freely moving imaging. Mice were allowed to recover for 6-8 weeks until testing for signal and beginning experimental testing.

### Histology

At the end of each of the behavioral/imaging experiments outlined below, subjects were deeply anesthetized with an intraperitoneal injection of Ketamine/Xylazine (100mg/kg/10mg/kg) and transcardially perfused with 10 mL of PBS solution followed by 10 mL of cold 4% PFA in 1x PBS. Animals were quickly decapitated, the brain was extracted and placed in 4% PFA solution and stored at 4 °C for at least 48-hours. Brains were then transferred to a 30% sucrose solution in 1x PBS and allowed to sit until brains sank to the bottom of the conical tube at 4 °C. After sinking, brains were sectioned at 35 $\mu$ m on a freezing sliding microtome (Leica SM2010R). Sections were stored in a cryoprotectant solution (7.5% sucrose + 15% ethylene glycol in 0.1 M PB) at -20 °C until immunohistochemical processing. We immunohistochemically stained all NAc slices with an anti-GFP antibody (chicken anti-GFP; Abcam #AB13970, 1:2000 in 5% BSA in PBS; room temperature overnight) for GCaMP6f, GCaMP6m, channelrhodopsin, and halorhodopsin for the validation of viral placement. Sections were then incubated with secondary antibodies [gfp: goat anti-chicken AlexaFluor 488 (Life Technologies #A-11039, 1:1000 in 5% BSA in PBS)] overnight at 4-degree temperature. After washing, sections were incubated for 5 min with DAPI (NucBlue, Invitrogen) to achieve counterstaining of nuclei before mounting in Prolong Gold (Invitrogen). Fluorescent images were taken using a Keyence BZ-X700 inverted fluorescence microscope (Keyence), under a dry 10x objective (Nikon). The injection site location and the fiber implant or GRIN lens placements were determined via serial imaging in all animals. We identified sections that displayed the NAc core, viral expression, and fiber optic tip.

## FIBER PHOTOMETRY

### Fiber photometry general approach

A Cre-recombinase-dependent virus carrying the fluorescent calcium indicator GCaMP6f (AAV5.hSyn.Flex.GCaMP6f.WPRE.SV40) was expressed in the NAc core of transgenic mice that express cre-recombinase in either D1 MSNs (D1-Cre mice) or D2 MSNs [Adora2a (A2A)-Cre mice]. Resulting fluorescent signals were recorded through a permanently implanted fiberoptic (400 $\mu$ m fiber diameter, 0.48 NA) affixed to each mouse's skull (described in surgical procedures). The fiber photometry recording system uses two light-emitting diodes (LED, Thorlabs) controlled by an LED driver (Thorlabs) at 490nm (run through a 470nm filter to produce 470nm excitation - the excitation peak of GCaMP) and 405nm (an isosbestic control channel, <sup>91-94</sup>). LED emissions pass through several filters and are reflected off of a series of dichroic mirrors (Fluorescence MiniCube, Doric) allowing for emission and the recording of resulting excitation through the same optical system. LEDs were controlled by a real-time signal processor (RZ5P; Tucker-Davis Technologies) and emission signals from each LED stimulation were determined via multiplexing. The fluorescent signals were collected via a photoreceiver



(Newport Visible Femtowatt Photoreceiver Module, Doric). Synapse software (Tucker-Davis Technologies) was used to control the timing and intensity of the LEDs and to record the emitted fluorescent signals. The LED intensity was set to 125 $\mu$ W for each LED and was measured daily to ensure that it was constant across trials and experiments. For each event of interest (e.g., predictive cue, head entries, licks, shock), transistor-transistor logic (TTL) signals were used to timestamp onset times from Med-PC V software (Med Associates Inc.) and were detected via the RZ5P in the synapse software (explained in more detail in the analysis section below).

### Fiber photometry analysis

The analysis of the fiber photometry data was conducted using a custom Matlab pipeline, as we have described previously.<sup>19,95</sup> Raw 470nm (F470 channel) and isosbestic 405nm (F405 channel) traces were collected at a rate of 1000 samples per second (1kHz). The raw data from each individual channel (450nm or 470nm) was then minimally filtered using a lowess filter before calculating  $\Delta f/f$  values via polynomial curve fitting. For the lowess filter, the number of data points for calculating the filtered value was set to 0.0004 (values closer to 1 indicates stronger smoothing).  $\Delta f/f$  for the entire trace was calculated as (F470nm-F405nm)/F405nm. This transformation uses the isosbestic F405nm channel, which is not responsive to fluctuations in calcium, to control for calcium-independent fluctuations in the signal and to control for photobleaching. Then, the data from the resulting  $\Delta f/f$  trace was cropped around behavioral events using TTL pulses. For each experiment 2s of pre-TTL and 18s of post-TTL  $\Delta f/f$  values were analyzed. Z-scores were calculated from the cropped trace by taking the pre-TTL  $\Delta f/f$  values as baseline ( $z\text{-score} = (\text{TTLsignal} - b\_mean)/b\_stdev$ , where the TTL signal is the  $\Delta f/f$  value for each post-TTL time point,  $b\_mean$  is the baseline mean, and  $b\_stdev$  is the baseline standard deviation). This allowed for the determination of calcium events that occurred at the precise moment of each significant behavioral event. For statistical analysis, peak height, and area under the curve (AUC) values were calculated for each individual trace around identified behavioral relevant events via trapezoidal numerical integration on each of the z-scores across a fixed timescale which varied based on experiment. The duration of the peak height and AUC data collection was determined by limiting the analysis to the z-scores between time 0 (TTL signal onset) and the time where the calcium signal returns to the pre-event baseline.

## CELLULAR RESOLUTION CALCIUM IMAGING VIA MICROENDOSCOPES

### Single cell imaging general approach

For calcium imaging at the single cell level, we used endoscopic miniature microscopes (nVista miniature microscope, Inscopix). Calcium signals were recorded via the calcium sensor GCaMP6m (GCaMP (AAV5.CAG.Flex.GCaMP6m.WPRE.SV40)). Signals were recorded on the miniature microscopes through the GRIN lens allowing for the resolution of single cells. Single cell activity in the NAc core was recorded in awake and behaving animals. During each behavioral session, the miniscope was attached to the baseplate that was implanted previously (see surgical section for description). The imaging parameters (gain, LED power, focus) were determined for each animal to ensure recording quality and kept constant throughout the study. At the end of the recording session, the miniscope was removed and the baseplate cover was replaced.

### Image processing and signal extraction

Data was acquired at 20 frames per second using nVista miniature microscopes (Inscopix). Image processing was accomplished using Mosaic software (v.1.3.1., Inscopix Data Processing Software v1.3.1 (Mountain View, CA)). Raw videos were pre-processed by applying 2x spatial downsampling to reduce file size and processing time, and isolated dropped frames were corrected. No temporal downsampling was applied. Lateral movement was corrected by using a portion of a single reference frame using Inscopix Data Processing Software (IDPS v1.3.1). Images were cropped to remove post-registration borders and sections in which cells were not observed. Videos were then exported as TIF stacks for analysis. After motion correction and cropping, a constrained non-negative matrix factorization algorithm optimized for micro-endoscopic imaging (CNMF-E) was utilized to extract fluorescence traces from neurons.<sup>90,95</sup> CNMF-E cell detection parameters were as follows:  $patch\_dims = 50, 50$ ;  $K = 20$ ;  $gSiz = 20$ ;  $gSig = 12$ ;  $min\_pnr = 20$ ;  $min\_corr = 0.8$ ;  $max\_tau = 0.400$ . Considering calcium fluctuations can exhibit negative transients, associated with a pause in firing,<sup>69</sup> we did not constrain temporal components to  $\geq 0$ . The  $\Delta f/f$  values were computed for the whole field of view as the output pixel value was represented as a relative percent change from the baseline. Raw CNMF-E traces were used for all analyses. The spatial mask and calcium time series of each cell were manually inspected using the IDPS interface. Cells found to be duplicated or misdetected due to neuropils or other artifacts were discarded.

### Calcium activity quantification

Stimulus-evoked activity was assessed by aligning calcium activity traces around cue or stimulus onset. Transistor-transistor logic (TTL) signals from MedPC were directly fed to the nVista system. The behavioral apparatus was interfaced with the miniature microscope via BNC cables and the onset of behavioral events - cues and shocks - were associated with a frame of the video using TTLs. The  $\Delta f/f$  values for each movie frame were calculated as  $M'(x,y,t) = (M(x,y,t) - F_{baseline}(x,y))/F_{baseline}(x,y)$  where  $M'$  is the output movie with  $\Delta f/f$  values,  $M(x,y,t)$  is the value for the pixel coordinate  $(x,y)$  at the  $t$  frame of the movie, and  $F_{baseline}(x,y)$  is the baseline value for the  $(x,y)$  coordinate. The raw  $\Delta f/f$  data from the CNMF-E program was exported and used for peri-event analysis. Data was cropped around each significant event (cue presentations; TTL) and z-scored to normalize for baseline differences. Z-scores were calculated by taking the pre-TTL  $\Delta f/f$  values as baseline ( $z\text{-score} = (\text{TTLsignal} - b\_mean)/b\_stdev$ , where TTL signal is the  $\Delta f/f$  value for each post-TTL time point,  $b\_mean$  is the



baseline mean, and  $b\_stdev$  is the baseline standard deviation). For traces aligned around a stimulus, a baseline window of 2 seconds prior to stimulus onset was used. Z-scored traces were then averaged across trials to create one trace per neuron for each stimulus type. To quantify the magnitude of the response, peak height (the maximum value of the calcium transient) and area under the curve (AUC) for each animal's averaged population trace was calculated. Values across a two second window beginning at stimulus onset were averaged, and this value was used to determine the response profile of each individual cell. We then calculated whether the cell response to the cues and shock outcomes were significant to determine responsive and non-responsive cells as well as the direction of the response (positive, negative, or no response). For this analysis, we calculated peak heights of the cue or shock evoked response from an individual cell averaged from 6 fear-conditioning trials as the maximum z-score achieved during a 2 second post-TTL window. We then ran two separate one-tailed independent t-tests to determine whether the response was significantly higher than +1.96 or lower than -1.96, the critical z-score for significance at  $p=0.05$ . For the AUCs, we used one-tailed independent t-tests to determine whether the mean AUC value was significantly different than 0. The cells that showed an averaged maximum response through 6 trials higher than the threshold were labeled as "Positive" cells. The cells that showed a significant negative response were labeled "Negative" cells. All other cells were determined as "No response" cells. We conducted this analysis for the cue and shock cell responses.

### Longitudinal co-registration

To identify cells and their responses across days within the same animal we used the longitudinal registration pipeline, defined in the Inscopix Data Process Software (IDPS) Guide (Section 4.9.1 Longitudinal Registration) identify the same cell across recording sessions in longitudinal series. Briefly, the algorithm uses the CNMF-E processed cell-map, which is an average of cellular activity responses across a recording session. The program then takes the number of cell segmentation outputs and creates an aligned "summary" cell map made of binarized spatial maps of all the ROIs that appear in the cell sets for each recording session. Each cell map was then aligned to the first cell map using a rigid transformation that accounts for translation and rotation. The aligned cell map is then designated global IDs and a normalized cross correlation (NCC) value is calculated. This NCC value is a measure of the ratio of overlap between a pair of ROIs from the "summary" cell map and the individual cell map. Matches are selected by identifying the maximal NCC value from the correlation matrix of ROI comparisons. Because cells may only be active in particular cell sets, a "match" means that an active cell in FC1 and FC4 were located in the same place in the field. For this set of experiments the images of the first cell set (the first day of fear conditioning, FC1) were defined as the global cell set against which the other cell sets are matched (this map was compared to the map that was detected on the final session of fear conditioning, FC4, and then independently to the cell map on the final session of extinction, EXT4). We then find the pair of cell images between the global cell sets and other aligned cell sets that maximize the normalized cross correlation (NCC). The program then generates an output that aligns the same cell from across sessions. We are thus able to determine if the same cell is active across sessions. To validate the accuracy of the co-registration pipeline, we also manually overlaid the co-registered cell maps onto the FC4 cell map (Figures S11 and S12). We then manually identified the co-registered cells to ensure there was no misalignment. We found that the co-registration maps are highly accurate in terms of identifying the cells active both during FC1 and FC4.

### Neural trajectory analysis

This analysis was performed using custom code written in python (v3.10.9). For each cell, we computed the mean activity across the six trials of each session. We applied a gaussian filter with  $\sigma=1$  bin (0.05s) and standardized the resulting smoothed average using the `scipy.stats.zscore()` method (scipy v1.10.0). All recorded neurons for each condition were combined into a timepoint x neuron matrix (one matrix for each of D1-MSN/FC1, D1-MSN/FC4, D2-MSN/FC1, D2-MSN/FC4). We then performed principal component analysis (PCA) by computing the covariance matrix using `numpy.cov()` followed by eigen value decomposition using `numpy.linalg.eigh()` (numpy v1.23.5). The resulting principal components were sorted and the top two were used for visualizing the trajectory through PCA space.

### Hierarchical clustering

Using the D2 MSN calcium traces from all the cells detected in the fear conditioning session 4, we employed a hierarchical clustering approach to group cells based on their cue responses in an unbiased way. We used the "clustergram" Matlab function and the "correlation" distance metric to group the cell activity.

## BEHAVIORAL EXPERIMENTS

### Apparatus

Mice were trained and tested daily in individual Med Associates (St. Albans, Vermont) operant conditioning chambers fitted with two illuminated nose pokes on either side of an illuminated sucrose delivery port, all of which featured an infrared beam break to assess head entries and nose pokes as well as a lickometer (Med Associates) to record tongue contacts on the sipper. One nose poke functioned as the active and the other as the inactive nose poke depending on the phase of the experiment (described below). Responses on both nose pokes were recorded throughout the duration of the experiments. Chambers were fitted with additional visual stimuli including a standard house light and two yellow LEDs located above each nose poke. Auditory stimuli included a white noise generator (which were used at 85 dB in these experiments) and a 16-channel tone generator capable of outputting frequencies between the

range of 1 and 20 kHz (also presented at 85 dB). Each box was outfitted with an infrared camera that recorded behavioral sessions and allowed for the analysis of freezing behavior (described below).

Each code that was written to run the MedPCV program for each of the behavioral tests below was coded to also send a TTL pulse at the onset of each behaviorally relevant stimulus. This was sent through a BNC connection that interfaced with either the TDT fiber photometry synapse software (for fiber photometry recordings) or with the Inscopix recording software (for microendoscopic imaging). A series of behavioral experiments were run throughout this study to link activity within D1 and D2 MSNs to behavioral responding in Pavlovian and reinforcement contexts. They are outlined in detail below:

### Positive Reinforcement

Mice were trained to nose poke on an active nose poke - denoted by its illumination - for delivery of sucrose in a trial-based fashion. Following a correct response, the sucrose delivery port was illuminated for 5 seconds, and sucrose was delivered (1s duration of delivery, 10% sucrose w/v, 10ul volume per delivery). To create a trial-based procedure, a discriminative stimulus ( $S^d_{sucrose}$ ) – an auditory cue – signaled that responses emitted during the presentation of the  $S^d$  would result in the delivery of sucrose. In these experiments, the  $S^d$  was white noise (85dB). Responses made during the  $S^d$  resulted in sucrose delivery and responses made at any other time in the session were recorded, but not reinforced. During the initial training, the  $S^d$  was presented throughout the entirety of each 1-hour session and animals could respond for sucrose without interruption. When animals reached  $\geq 60$  active responses in a single session, they were then moved to a discrete trial-based structure in subsequent 1-hour sessions, wherein the  $S^d$  was presented for 30 seconds at the beginning of each trial with a variable inter-trial interval (ITI) – on average the interval was 30 seconds, with interval times ranged from 20 to 40 seconds. Each trial ended following a correct response and associated sucrose delivery or at the end of a 30 second period with no active response. At the end of the trial both the  $S^d$ , and sucrose were terminated. Animals that exhibited active responses in  $\geq 80\%$  of trials during a session then proceeded to the final phase of training wherein the duration of  $S^d$  sucrose was reduced to 10 seconds. Upon reaching the 80% criterion during this phase (i.e., acquisition), post-training calcium responses in D1 and D2 MSNs were recorded over a 30 min session.

### Negative Reinforcement

Mice were trained to emit an operant response on the opposite, non-sucrose-paired nose poke for shock avoidance. A second auditory discriminative stimulus ( $S^d_{shock}$ ; 85dB at 2.5kHz) was presented at the beginning of each trial following a variable inter-trial interval (ITI) – on average the interval was 76 seconds, with interval times ranged from 30 to 150 seconds. In each trial, the discriminative stimulus was presented for 30 seconds after which a series of 20 footshocks (1mA, 0.5 second duration) was delivered with a 10 second inter-stimulus interval. Trials ended when animals responded on the correct nose poke or at the end of the shock period. The end of the shock period was denoted by the presentation of a house light cue that signaled the end of the trial and was illuminated for one second. During these trials, mice could respond during the initial 30 second  $S^d_{shock}$  period to avoid shocks completely, respond any time during the shock period to terminate the remaining shocks, or not respond at all. If mice did not respond both the trial and  $S^d_{shock}$  were terminated after all 20 shocks had been presented (230 seconds total). Acquisition during negative reinforcement training was defined as receiving fewer than 20% of total shocks in a single one-hour session.

### Unsignaled footshocks of varying intensities

Behaviorally naïve mice were placed into the same operant boxes used above, however, none of the ports or cues were active. A total of 8 footshocks were delivered in a non-contingent and inescapable fashion over a 12-minute period. Shocks were delivered at 0.3mA and 1mA intensities (4 presentations for each shock intensity). Shocks were delivered in a pseudo-random order with variable inter-stimulus intervals (mean ITI = 30 sec). All shock intensities were presented within the same test session. Shocks were delivered by the Standalone Aversive Simulator/Scrambler (Med Associates, Env 4145). Shock intensities were measured prior to the session via ammeter (Med Associates, ENV-421-B) with two clip leads attached to the grid rods in the Med Associates chamber. During the session, shock intensities were manually changed on the Standalone Aversive Simulator/Scrambler interface.

### Unsignaled sucrose delivery

Mice with no prior training history were placed into the operant box and were given ad libitum access to the port, which was manually filled with sucrose (10% sucrose w/v), prior to the session. Sessions were 30 minutes in length. Tongue contacts on the sucrose cup were recorded using an electrical lickometer (sampled at 1 kHz), and via continuous video is collected with an overhead camera. Each lick triggered a TTL output to the fiber photometry recording software allows for signal alignment around the first sucrose lick in each mouse. If mice did not lick, the session was run the next day until mice made a lick for sucrose.

### Fear conditioning followed by extinction

Mice received a single footshock (1mA, 0.5 second duration) immediately following a 5 second auditory cue (5kHz tone; 85dB) for 6 pairings. Mice underwent 4 fear conditioning sessions over four consecutive days, followed by 4 extinction sessions in which the cue was presented, but shocks were omitted entirely. For cellular resolution imaging experiments, the cue was presented for 10s, instead

of 5, during the fear conditioning and fear extinction sessions. This was due to the reduced movement of the mice in response to the sizeable scope. Every session was captured with an infrared camera. Freezing analysis was assessed during the cue presentation for each animal. Freezing was hand-scored by an observer who was blind to the experimental conditions. Time spent freezing was recorded (immobile – lack of any movement including sniffing) from the cue onset to cue offset.

### **Fear conditioning with changing cue-outcome probabilities**

Mice underwent two sessions of fear conditioning in which the probability of the cue being followed by a foot shock was changed to probe how neural responses changed to the cue and the shock under these conditions. In session 1, on 10% of the trials the auditory cue (5 second; 1kHz tone; 85dB) was followed by a single brief shock (1mA, 0.5 second duration) at the offset of the cue. On the remaining 90% of trials, the same cue was presented without the shock. Trials were interleaved in a pseudo-random order. In session 2, 75% of the trials were cue-shock pairings and 25% were cue-no shock pairings. We analyzed the cue and shock response on the trial in which mice had received 10% of the cue-shock pairings (session 1) or 75% (session 2). A separate experiment was run where two sessions were run on two consecutive days where on 10% of the trials the auditory cue (5 second; 1kHz tone; 85dB) was followed by a single brief shock (1mA, 0.5 second duration) at the offset of the cue. In the remaining 90% of trials, the same cue was presented without the shock.

## **OPTOGENETICS**

### **Intracranial Self-Stimulation (ICSS)**

In a group of D1 or A2A-Cre mice, channelrhodopsin [excitatory opsin (AAV5.Ef1a.DIO.hChR2)], or eYFP [control vector (AAV5.hSyn1.eYFP)] was unilaterally expressed in the NAc core as described about in the surgical procedures section. A 200um fiber optic implant (Doric, 0.22 NA; MFC\_200/240-0.22\_4.7mm\_MF2.5\_FLT) was placed into the NAc core above the viral injection site. Mice were placed in an operant chamber where responses on an active nose poke resulted in laser stimulation (470nm, 2s, 14Hz, 8mW), and responses on the inactive nose poke had no programmed consequences. No other stimuli were present during this task (house lights, cue lights, etc.). The number of active and inactive nose pokes was recorded during the two-hour session for four days. The active and inactive nose poke were counterbalanced across mice.

### **Optical inhibition of D1- and D2-MSNs during fear conditioning**

Halorhodopsin [inhibitory opsin (AAV5.Ef1a.DIO.eNpHR3.0.YFP)] or eYFP [control vector (AAV5.hSyn1.eYFP)] was bilaterally expressed in D1-Cre and A2A-Cre mice using the strategies previously described. In a fear conditioning session, mice received a single footshock (1mA, 0.5 second duration) immediately following a 5 second auditory cue for 3 pairings. The cue was a tone (5kHz tone; 85dB). During this initial training session, each presentation of the cue was optically inhibited (5 sec, 8 mw, continuous laser) for the duration of the cue presentation. Freezing was hand-scored for the 5 second pre-footshock cue period for each trial in a blind fashion. The freezing response was defined as the time (seconds) that mice were immobile (lack of any movement including sniffing) during the tone period and calculated as the percentage of total cue time.

### **Optical inhibition of D1- and D2-MSNs during positive reinforcement**

Halorhodopsin [inhibitory opsin (AAV5.Ef1a.DIO.eNpHR3.0.YFP)] or eYFP [control vector (AAV5.hSyn1.eYFP)] was bilaterally expressed in A2A-Cre mice using the strategies described for the fear conditioning experiments above. All mice received the same positive reinforcement training described for the fiber photometry experiments above. During the initial training, the S<sup>d</sup> was presented throughout the entirety of each 1-hour session and animals could respond for sucrose without interruption. When animals reached  $\geq 60$  active responses in a single session, they were then moved to a discrete trial-based structure in subsequent 1-hour sessions, wherein the S<sup>d</sup> was presented for 30 seconds at the beginning of each trial with a variable ITI. During this discrete cue session, each presentation of the cue was optically inhibited (590nm, 5 sec, 8 mw, continuous laser) for the initial 10 seconds of the cue presentation. Mice received 4 such training sessions and their nose poke responses during the cue presentations were recorded (active nose pokes). We converted the number of active nose pokes for each session to the percentage of the active nose poke response numbers during the first training session (%Baseline active response) in order to eliminate the initial baseline learning differences between mice.

### **Optical inhibition of D1- and D2-MSNs during open field test**

Halorhodopsin [inhibitory opsin (AAV5.Ef1a.DIO.eNpHR3.0.YFP)] or eYFP [control vector (AAV5.hSyn1.eYFP)] was bilaterally expressed in D1-Cre and A2A-Cre mice using the strategies previously described. For the locomotor activity testing, all mice were placed in an open field arena and following a 5-minute habituation phase, mice were given yellow laser stimulations on a random interval schedule (590nm, 5 sec, 8 mw, continuous laser). The behavioral session was recorded via a camera above the open field arena. Velocity (cm/second) and Distance traveled (cm) during the intervals between each laser stimulation were measured using the EthoVision XT - Video tracking software (Noldus). Data are reported as average velocity per minute (cm/s/min) and average distance traveled per minute (cm/min).

### Optogenetic excitation approach validation via fiber photometry

A group of D1-cre and A2A-cre mice were injected with Chrimson.FLEX: AAV5-Syn-FLEX-rc[ChrimsonR-tdTomato] (Chrimson; Addgene) into the NAc core as described above. A Cre-recombinase-dependent virus carrying the fluorescent calcium indicator GCaMP6f (AAV5.hSyn.Flex.GCaMP6f.WPRE.SV40) was also expressed in the NAc core of transgenic mice as described for the fiber photometry experiments. Control mice (No Chrimson) received only the GCaMP6f injections. A 400  $\mu\text{m}$  fiber optic was implanted into the NAc core. Using the same stimulation parameters described above for optogenetic stimulation, a laser stimulation (590nm, 1 second, 20Hz, 8mW) was delivered into the NAc core while recording fluorescent signals emitted from GCaMP6f in the same animal using fiber photometry. Chrimson and No Chrimson control mice received 6 laser stimulations during Session 1 and again 24 hours later during Session 2. The size of the calcium response elicited by the Chrimson stimulations was recorded.

### QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analyses were performed using GraphPad Prism (version 9; GraphPad Software, Inc, La Jolla, CA) and Matlab (Mathworks, Natick, MA). Z-scores were calculated as explained above (see fiber photometry analysis). We used nested ANOVAs as well as paired, unpaired and independent t-tests where appropriate for analyzing peak height and AUC values for fiber photometry experiments. Repeated measures ANOVAs were used for the behavioral data from reinforcement studies (positive and negative reinforcement) as well as for the optogenetic studies with multiple trials (ICSS and cue inhibition). For the positive and negative reinforcement tasks, the probability of responding during the discriminative cue was calculated by: number of nose pokes on the active port during the predictive cue only / number of total nose pokes on the active side during both the cue and ITI. We then ran independent sample tests to compare the probability to chance in each task. For the positive reinforcement, the chance level for making a response on the active port during the cue presentation was 50% (30 second cue period versus 30 second averaged ITI period). For the negative response, making a response on the active port during the predictive cue at chance level was 75.1% (230 second cue period versus 76 second averaged ITI period). We also used one-way ANOVAs for the optogenetic studies, which were followed by Dunnett or Bonferroni post-hoc analyses to compare groups. For the single cell imaging experiments maximum z-scores were identified and statistical difference from the critical z-scores at  $p=0.05$  level (+1.96 or -1.96) was computed using independent-t-tests. We further compared FC1 and FC4 cell responses (peak height or AUC) using paired t-tests. Alpha was 0.05 for all statistical analysis. All data were depicted as group mean  $\pm$  standard error of the mean (S.E.M.). We assumed normal distribution of sample means for all t and F statistics.