

<https://doi.org/10.1038/s42003-024-07440-7>

# Devaluing memories of reward: a case for dopamine

Check for updates

Benjamin R. Fry<sup>1,8,9</sup>, Nicolette Russell<sup>1,9</sup>, Victoria Fex<sup>2,9</sup>, Bing Mo<sup>1</sup>, Nathan Pence<sup>1</sup>, Joseph A. Beatty<sup>3,4</sup>, Fredric P. Manfredsson<sup>5</sup>, Brandon A. Toth<sup>6</sup>, Christian R. Burgess<sup>6</sup>, Samuel Gershman<sup>7</sup> & Alexander W. Johnson<sup>1,4</sup> ✉

Midbrain dopamine cells encode differences in predictive and expected value to support learning through reward prediction error. Recent findings have questioned whether reward prediction error can fully account for dopamine function and suggest a more complex role for dopamine in encoding detailed features of the reward environment. In this series of studies, we describe a novel role for dopamine in devaluing sensory features of reward. Mesencephalic dopamine cells activated during a mediated devaluation phase were later chemogenetically reactivated. This retrieval of the devalued reward memory elicited a reduction in the hedonic evaluation of sucrose reward. Through optogenetic and chemogenetic manipulations, we confirm dopamine cells are both sufficient and necessary for mediated devaluation, and retrieval of these memories reflected dopamine release in the nucleus accumbens. Consistent with our computational modeling data, our findings indicate a critical role for dopamine in encoding predictive representations of the sensory features of reinforcement. Overall, we elucidate a novel role for dopamine function in mediated devaluation and illuminate a more elaborate framework through which dopamine encodes reinforcement signals.

Animals anticipate long-term future rewards to behave adaptively in the present. To achieve this, they can learn predictions of future rewards directly from their experiences and update these estimates using reward prediction errors (RPE)<sup>1,2</sup>. A venerable research history supports the idea that dopamine transients from midbrain cells encode these differences in predictive and expected value by relaying information throughout corticostriatal circuitry<sup>3–5</sup>. These studies have substantially increased the understanding of dopamine function and suggest that the nature of this encoding reflects a scalar-based model-free quantity that promotes the rewarding value of future events during learning.

Despite significant support across a range of model organisms and approaches<sup>6–11</sup>, more recent findings question the ubiquity of the RPE hypothesis in fully accounting for dopamine's role in reinforcement learning. This is due to the nature of dopamine RPE encoding, which reflects a value signal that does not take into consideration the detailed features of the encoded events experienced during learning. As such, model-free RPE struggles to account for findings that suggest dopamine is necessary for encoding the sensory features of predicted rewards<sup>12</sup>, sensory errors in

stimulus-stimulus learning<sup>13</sup>, and unexpected changes in the sensory features of rewards<sup>14</sup>. An additional limitation of model-free RPE learning is that state values are acquired through direct experience with the current state, this information is then leveraged to predict the likely occurrence of future rewarded states. However, many behaviors and decisions reflect the engagement of processes in which the organism can mentally navigate and adapt to its reward environment without directly experiencing the events themselves<sup>15–20</sup>. Accordingly, it has been suggested that dopamine can function in a model-based manner through which a detailed representation of the reward environment is internally navigated by the organism<sup>21–23</sup>, or that dopamine transients encode detailed sensory features of reward events that can be used to gauge a predictive feature map—the successor representation (SR)<sup>24,25</sup>.

Mediated devaluation is a particularly striking example of learning that is dependent both on encoding detailed sensory reinforcement features and the sequence of states through which rewarding events are acquired. This relatively understudied associative learning phenomenon occurs when a previously reward-paired conditioned stimulus (CS) can retrieve memories

<sup>1</sup>Department of Psychology, Michigan State University, East Lansing, MI, USA. <sup>2</sup>Lyman Briggs College, Michigan State University, East Lansing, MI, USA.

<sup>3</sup>Department of Physiology, Michigan State University, East Lansing, MI, USA. <sup>4</sup>Neuroscience Program, Michigan State University, East Lansing, MI, USA.

<sup>5</sup>Department of Translational Neuroscience, Barrow Neurological Institute, Phoenix, AZ, USA. <sup>6</sup>Michigan Neuroscience Institute, University of Michigan, Ann Arbor, MI, USA. <sup>7</sup>Department of Psychology, Harvard University, Cambridge, MA, USA. <sup>8</sup>Present address: Department of Biobehavioral Health, The Pennsylvania State University, University Park, PA, USA. <sup>9</sup>These authors contributed equally: Benjamin R. Fry, Nicolette Russell, Victoria Fex. ✉e-mail: [awj@msu.edu](mailto:awj@msu.edu)

of food rewards so detailed in nature that animals sensorially experience the absent reward<sup>26–29</sup>. As a result, when the auditory CS is paired with an injection of the gastric illness-evoking agent, LiCl, rodents will acquire a devaluation to the food reward. Despite never experiencing the food paired with LiCl, the CS-evoked representation is sufficiently strong to substitute for the food itself in the acquisition of the aversion<sup>30,31</sup>. If dopamine signaling plays a role in mediated devaluation, this would be particularly challenging for traditional accounts of dopamine function, given that in this setting, the acquisition of an aversion relies on encoding detailed reinforcement features<sup>32,33</sup> and flexible transitions between states under conditions when the reward is not experienced.

To determine whether dopamine influences mediated devaluation, mice were trained to acquire an association between an auditory CS and a liquid sucrose reinforcer. During the aversion phase, the CS alone was presented and paired with LiCl. We first examined whether the activity-dependent labeling of ventral tegmental area (VTA) cells following CS-LiCl aversion would lead to a disruption in the hedonic evaluation of the sucrose reward when these VTA cells were reactivated. Subsequently, using optogenetic and chemogenetic approaches we examined the sufficiency and necessity of VTA dopamine cells during aversion. In an additional series of experiments, we used *in vivo* fiber photometry to examine dopamine release in nucleus accumbens (NAc) targets and used computational modeling data to recapitulate core features of the dopaminergic manipulations on mediated devaluation. Overall, we show a novel function for dopamine in the devaluation of detailed sensory features of reward.

## Results

### Using representation-mediated learning to devalue the reward

To reveal the potential vast array of features underlying dopamine-dependent learning and memory, we adapted a mediated devaluation approach to use in mice<sup>26,30</sup>. One of the curious features of mediated devaluation is the transient nature of the phenomenon<sup>27,28</sup>, such that early in training CSs gain access to detailed reinforcement features that can substitute for the absent reward during aversion. However, this window of accessibility rapidly closes as training continues, and as a result, the capacity of a CS to guide mediated devaluation is lost. Mice received Pavlovian training in which an auditory CS was paired either 16 (minimal) (Supplementary Fig. 1A) or 64 (extensive) (Supplementary Fig. 1E) times with 0.2 M sucrose, followed by an aversion phase in which the CS alone was presented and preceded an injection of the gastric malaise inducing agent, LiCl (see Supplementary Methods). If the CS could retrieve detailed sensory features of the previously paired but absent sucrose solution (e.g., its taste), we would expect the devaluation produced by LiCl to diminish the perceived palatability of the sucrose when it was reintroduced<sup>34</sup>. To determine whether mediated devaluation was achieved, mice received a sucrose consumption test and we subsequently integrated rigorous quantitative analyses of rodent licking behavior, in which the temporal distribution of interlick intervals can be used to infer the perceived palatability of a consumed liquid reward<sup>35–39</sup>. Even though the sucrose solution was never paired with LiCl, the devaluation of the detailed sensory features of sucrose reward was sufficient in minimally trained mice to evoke a significant decrease in the perceived sweetness and palatability of the sucrose, compared to the saline-control condition (Supplementary Fig. 1D).

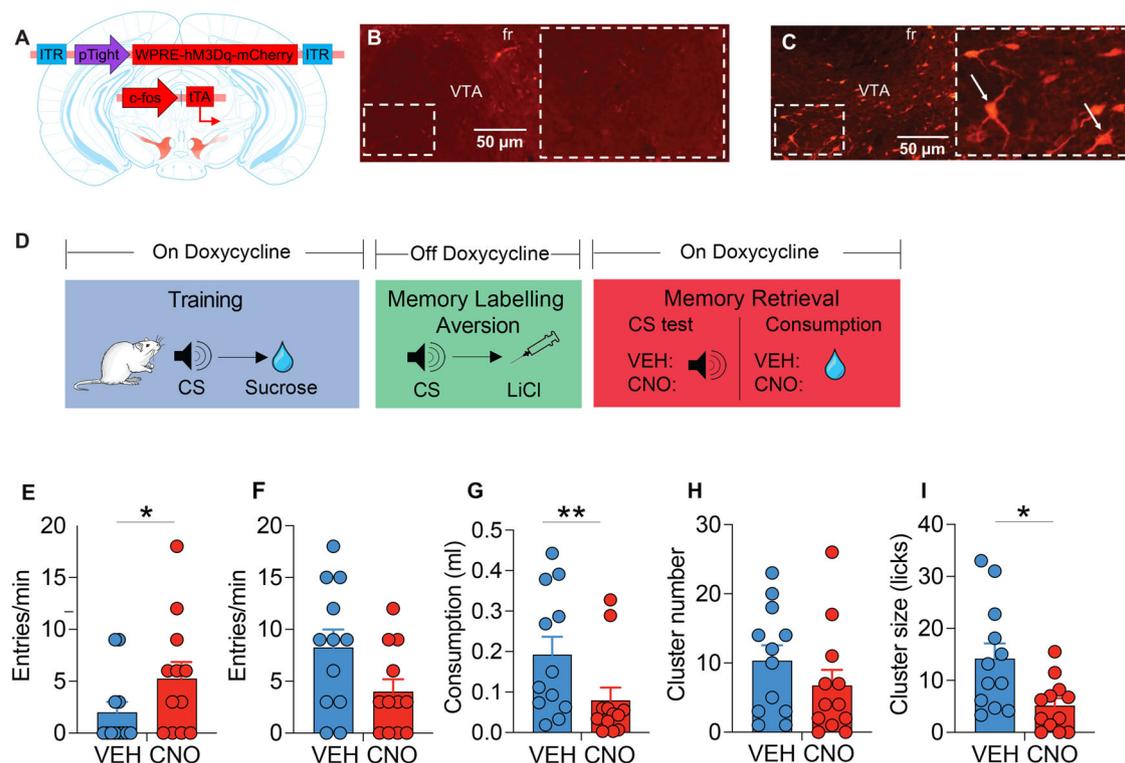
### Activity-dependent labeling of ventral tegmental area cells to devalue sucrose reward

To extend our examination of the neuronal basis of mediated devaluation of sucrose reward, we labeled VTA cells in an activity-dependent manner during CS-evoked mediated devaluation of sucrose reward and reactivated these cells using chemogenetics. Transgenic *cfos*-*htTA* mice—in which the expression of tetracycline transcriptional activator (TRE) is directed to

activated neurons by the *cfos* promoter—received bilateral infusions into VTA of an adeno-associated virus (AAV) encoding the excitatory DREADD, hM3Dq, via pAAV-PTRE-tight-hM3Dq-mCherry (Fig. 1A). hM3Dq expression was only evident in mice that received doxycycline (Dox) withdrawal (Fig. 1B, C; Supplementary Fig. 2). In mice that received activity-dependent labeling of DREADDs in VTA cells during aversion, reactivation of these cells via clozapine-N-oxide (CNO) during a cue test led to elevated pre-CS responses (Fig. 1E;  $p < 0.05$ ,  $d = 0.45$ ) but no significant effects over CS responding (Fig. 1F;  $p = 0.1$ ,  $d = 0.39$ ). Importantly, reactivation of these cells attenuated both sucrose consumption (Fig. 1G;  $p = 0.01$ ,  $d = 1.10$ ), and its perceived palatability (Fig. 1H;  $p < 0.05$ ,  $d = 2.23$ ) but not the motivation to consume sucrose (Fig. 1I;  $p = 0.1$ ,  $d < 0.5$ ). This effect appeared to be more prevalent in male rather than female *cfos*-*htTA* mice (Supplementary Fig. 3). Overall, these findings suggest that the VTA forms part of the neuronal circuitry underlying the encoding of detailed features of sucrose reward, such that they can be devalued and retrieved, leading to long-term changes in the evaluation of biologically meaningful rewarding events. When we examined the identity of cells labeled with hM3Dq, the majority (>85%) colocalized with tyrosine hydroxylase (Supplementary Fig. 2). Based on these findings, we turned our attention to the role of VTA dopamine cells in mediated devaluation.

### Ventral tegmental area dopamine cells are both sufficient and necessary for the mediated devaluation of sucrose reward

Mice expressing Cre-recombinase under the control of the tyrosine hydroxylase (TH) promoter-received unilateral injections of a Cre-dependent ChR2, (AAV5-Ef1 $\alpha$ -DIO-ChR2-eYFP) or control eYFP (AAV5-Ef1 $\alpha$ -DIO-eYFP) (Fig. 2A–D and Supplementary Fig. 4) along with an optical fiber cannula implanted slightly dorsal to the injection site. This approach allows temporally discrete and precise stimulation of ChR2-infected cells, resulting in depolarization in the presence of 473 nm optical stimuli<sup>40</sup> (Supplementary Fig. 5). All TH-Cre mice underwent mediated devaluation through CS-LiCl aversion. At this stage, VTA TH cell stimulation was timed to coincide with CS presentation during the aversion phase (Fig. 2E), with the prediction that this stimulation would promote further access to dopamine-dependent sensory features of reinforcement typically generated by the CS alone. Accordingly, this would enhance the strength of the retrieved reward memory such that mice might express more robust mediated devaluation when the sucrose was reintroduced. We also examined whether any potential augmentation in mediated devaluation required intact signaling of the dopamine D2 receptor<sup>33</sup>, by administering a subset of eYFP and ChR2 mice with haloperidol (a D2 receptor antagonist) prior to optogenetic stimulation. During the cue test, pre-CS entries did not differ as a function of virus or drug treatment (Fig. 2F); however, prior haloperidol treatment led to a decrease in CS-evoked food cup responses (Fig. 2G;  $p < 0.05$ ,  $d = 0.7$ ). During consumption testing, overall intake was comparable (Fig. 2H) as were the number of bursts initiated (Fig. 2I). Conversely, when we examined licking microstructure measures associated with reward palatability of the sucrose, ChR2 mice in the non-drug condition displayed a significant reduction relative to their eYFP counterparts (Fig. 2J;  $p < 0.01$ ,  $d = 1.93$ ). This augmented mediated devaluation effect that followed VTA TH cell stimulation was dependent on intact signaling of the dopamine D2 receptor, as ChR2 mice that received haloperidol treatment prior to optogenetic stimulation displayed increased hedonic evaluation relative to non-drug treated ChR2 mice ( $p < 0.0001$ ,  $d = 1.61$ ). Similar to *cfos*-*htTA* mice, stimulation of VTA cells more readily augmented mediated devaluation in male relative to female ChR2 mice (Supplementary Fig. 6). Furthermore, in a naïve group of ChR2 mice trained with 64 CS-US trials, which limits the expression of mediated devaluation (Supplementary Fig. 1G, H)<sup>27,28</sup>, optogenetic stimulation during aversion was nevertheless capable of eliciting a significant decrease in the overall intake (Supplementary Fig. 7C;  $p = 0.01$ ,  $d = 1.47$ ) and perceived



**Fig. 1 | Mediated devaluation of sucrose reward via activity-dependent labeling and chemogenetic activation of ventral tegmental area cells.** **A** *cfos*-*htTA* mice were bilaterally injected with pAAV-PTRE-tight-*hM3Dq*-*mCherry* into the VTA. **B** *cfos*-*htTA* mice maintained on a diet containing Dox do not express *hM3Dq*-*mCherry* in the ventral tegmental area, whereas **C** while off Dox mice displayed extensive bilateral infection of *hM3Dq*-*mCherry*. Arrows indicate somatic expression of *hM3Dq*. **D** Simplified schematic of behavioral training, testing, and activity-dependent labeling. Following Pavlovian training, mice were removed from the diet containing Dox to permit activity-dependent *hM3Dq*-*mCherry* labeling during CS-LiCl aversion. To target labeled cells that were active during aversion, *cfos*-*htTA* mice also received Dox injections and were subsequently maintained on a

diet containing Dox for the remainder of the study. Reactivation of labeled cells was achieved via CNO-evoked activation of *hM3Dq*-*mCherry* and assessed under CS and sucrose consumption test conditions. **E** During the cue test, CNO-evoked reactivation of *hM3Dq*-*mCherry* cells enhanced pre-CS responding (main effect of the drug,  $F(1,11) = 5.01, p < 0.05$ ) (**F**) but did not significantly influence CS responses. **G–I** Reactivation of *hM3Dq*-*mCherry* cells significantly attenuated both (**G**) overall intake ( $F(1,11) = 7.47, p = 0.01$ ) (**I**) and the palatability of sucrose reward ( $F(1,11) = 5.41, p < 0.05$ ), (**H**) but did not influence the motivation to consume it. Blue circles = vehicle; red circles = CNO. Error bars indicate the standard error of the mean (SEM). Abbreviations: fr = fasciculus retroflexus, VTA = ventral tegmental area. \* $p$ 's < 0.05; \*\* $p$  = 0.01.

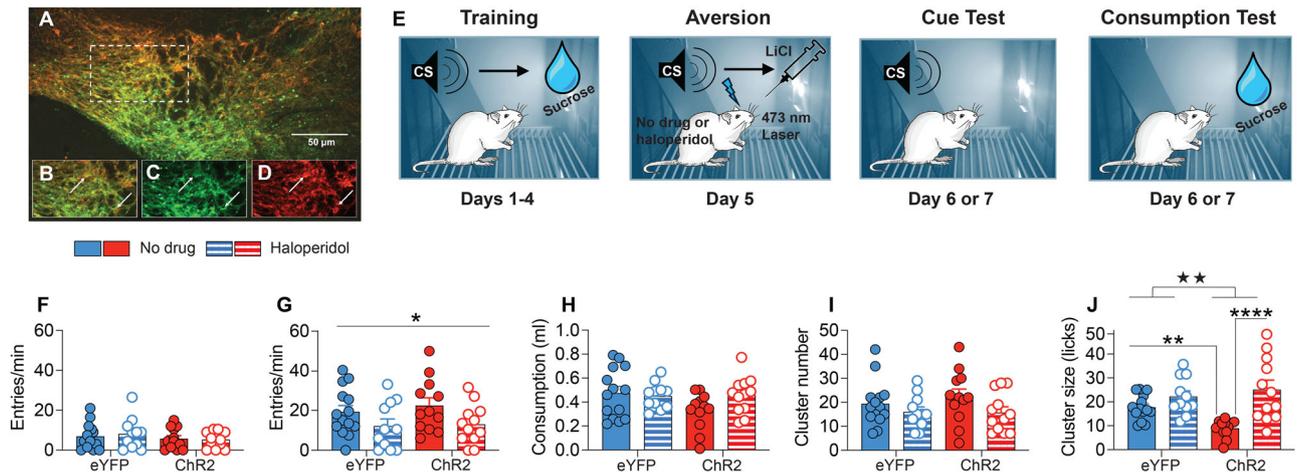
palatability of the sucrose (Supplementary Fig. 7E;  $p = 0.01, d = 1.45$ ), compared to eYFP mice trained with this more extensive Pavlovian conditioning design.

To examine whether dopamine cell activity is necessary for mediated devaluation, TH-Cre mice received bilateral injections of a Cre-dependent inhibitory DREADD virus (AAV8- *hSyn*-DIO-*hM4D*(Gi)-*mCherry*) or control eYFP into VTA (Fig. 3A–D and Supplementary Fig. 8). All *hM4Di*-treated mice underwent mediated devaluation via CS-LiCl pairing. During this stage, a subset of these mice received chemogenetic inactivation of VTA dopamine cells via CNO pretreatment (*hM4Di*-LiCl CNO group). The performance of these mice was compared to two other groups that were expected to show mediated devaluation—*hM4Di*-treated mice that received vehicle instead of CNO (*hM4Di*-LiCl VEH group), and an eYFP group that received CNO (eYFP-LiCl CNO group). An additional cohort of eYFP mice received saline rather than LiCl during aversion (eYFP-saline CNO group) and served as a control to compare to all other LiCl-paired mediated devaluation groups. To examine whether the CS became devalued following CS-LiCl pairings, all mice received a cue test. There was no evidence that the CS entered into an association with LiCl nor that inactivation of VTA dopamine cells during mediated devaluation influenced pre-CS or CS responding (Fig. 3F, G). Conversely, substantial group differences were revealed during the consumption test (Fig. 3H–J). Relative to the unperturbed eYFP-saline-control group, eYFP mice that received LiCl displayed a significant reduction in overall intake (Fig. 3H;  $p < 0.01, d = 1.68$ ), which did not impact the motivation to engage in sucrose intake (Fig. 3I) but reflected a

significant attenuation of reward palatability (Fig. 3J;  $p = 0.01, d = 1.28$ ). Thus, treatment with CNO did not prevent mediated devaluation in eYFP mice. By comparison, CNO-evoked chemogenetic inactivation in *hM4Di* mice disrupted mediated devaluation, such that mice in the *hM4Di*-LiCl CNO group displayed overall intake (Fig. 3H) and palatability responses (Fig. 3J) that did not differ significantly from the unperturbed eYFP-saline-control group. Furthermore, *hM4Di*-LiCl CNO mice displayed elevated overall intake compared to eYFP mice that received LiCl during aversion (Fig. 3H;  $p = 0.02, d = 1.31$ ). Finally, *hM4Di*-LiCl mice that received vehicle rather than CNO during aversion also displayed mediated devaluation, as indicated by a tendency to show reduced overall intake (Fig. 3H;  $p = 0.06, d = 1.12$ ) and a significant reduction in tastant palatability (Fig. 3J;  $p < 0.01, d = 1.38$ ) compared to eYFP mice that received saline during aversion.

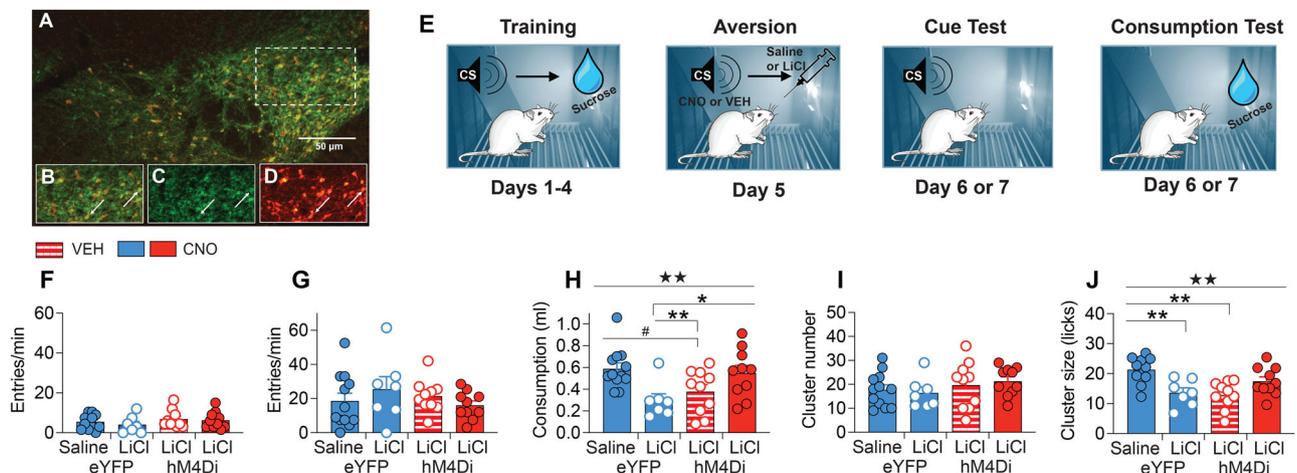
### Tracking physiological nucleus accumbens dopamine release dynamics that underlie the retrieval of mediated devaluation of reward

One of the caveats of optogenetic and chemogenetic approaches is that they are unlikely to mimic endogenous dopamine activity. A major output of ventral mesencephalic dopamine cells is the nucleus accumbens (NAc). This VTA→NAc circuit is critically implicated in a range of dopamine-dependent processes, including RPE<sup>41</sup>, incentive salience<sup>42</sup> and sensorimotor motivational processes<sup>43</sup>. Thus, to address whether VTA dopamine cell circuitry normally encodes detailed reinforcement signals consistent with a role in mediated devaluation, we used fiber photometry to dynamically track



**Fig. 2 | Optogenetic stimulation of ventral tegmental area dopamine cells enhances mediated devaluation of sucrose reward.** A, B Representative photomicrographs displaying colocalization of ChR2 expression (green) in tyrosine hydroxylase (TH) positive neurons (red). Separate photomicrographs for C ChR2 and D TH. Arrows indicate somatic expression of ChR2 and TH. E Simplified schematic of the behavioral testing and training procedures. Thirty minutes prior to aversion, half the mice from eYFP and ChR2 cohorts received 0.1 mg/kg intraperitoneal injection of haloperidol. During aversion mice received contemporaneous presentation of the CS and laser stimulation followed by LiCl injection to induce mediated devaluation. Subsequently, mice received testing with the sucrose (consumption test) and CS (cue test) alone. F Pre-CS food cup responses did not differ between viral group or drug conditions; however, G haloperidol treatment during

aversion led to a subsequent decrease in CS-evoked food cup responses. \* Main effect of drug, ( $F(1,45) = 5.98, p = 0.02$ ). H Overall intake and I cluster number were comparable across all groups. J Relative to the mediated devaluation expressed by non-drug treated eYFP controls, stimulation of VTA dopamine cells in the no drug ChR2 group during aversion subsequently enhanced devaluation of the palatability of sucrose reward ( $F(1,45) = 7.05, p = 0.01$ ). This augmented effect was dependent on intact D2R signaling as treatment with haloperidol in ChR2 mice led to an attenuation of mediated devaluation relative to non-drug exposed counterparts ( $F(1,45) = 22.25, p < 0.0001$ ). Filled blue circles = eYFP-no drug; open blue circles = eYFP-haloperidol; Filled red circles = ChR2-no drug; open red circles = ChR2-haloperidol. ★★Significant virus x drug interaction, ( $F(1,45) = 5.83, p = 0.01$ ), \*\* $p \leq 0.01$ , \*\*\*\* $p < 0.0001$ .

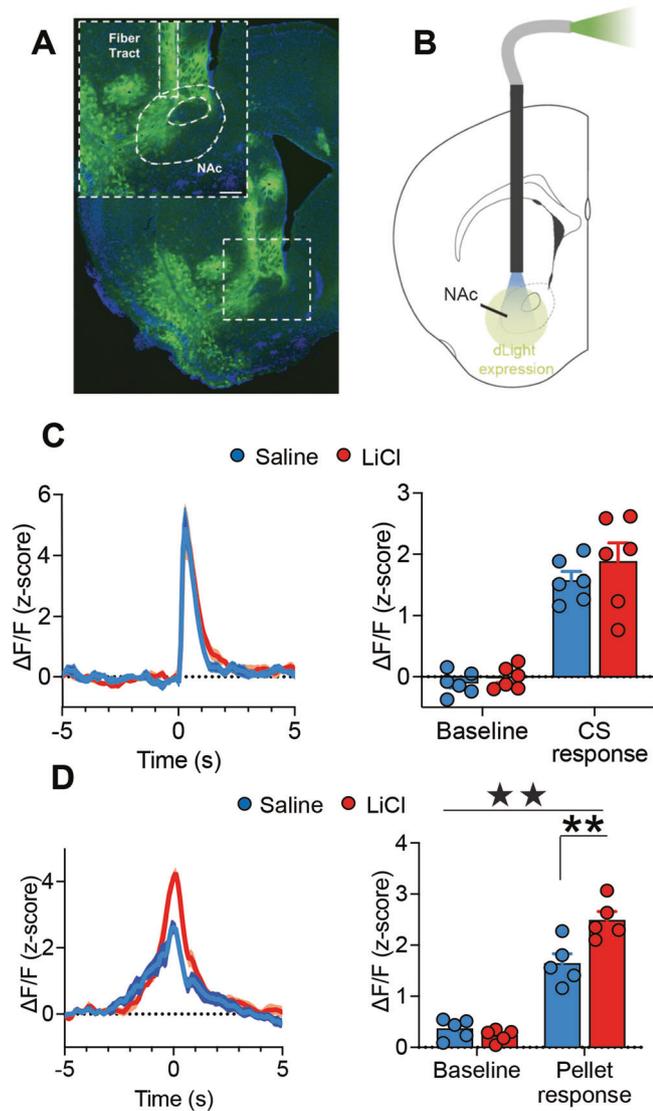


**Fig. 3 | Chemogenetic inhibition of ventral tegmental area dopamine cells disrupts mediated devaluation of sucrose reward.** A–D Representative photomicrographs of Cre-dependent hM4Di (red) in tyrosine hydroxylase (TH) positive neurons (green). E Simplified schematic of the behavioral training and testing procedures. Prior to aversion mice received injections of either vehicle (VEH) or clozapine-N-oxide (CNO) followed by presentation of the CS and injections of either saline, or LiCl to induce mediated devaluation. Subsequently, mice received testing with the CS (cue test) and sucrose (consumption test) alone. F pre-CS and G CS elicited food cup entries were comparable irrespective of chemogenetic manipulations or whether mice received saline (eYFP-saline-CNO) or LiCl (eYFP-LiCl-CNO, hM4Di-LiCl-VEH, hM4Di-LiCl-CNO) during aversion. H eYFP-LiCl-CNO and hM4Di-LiCl-VEH mice displayed mediated devaluation relative to eYFP mice that received saline during aversion. Chemogenetic inhibition of VTA dopamine cells

(hM4Di-LiCl-CNO) disrupted mediated devaluation relative to eYFP mice treated with LiCl (eYFP-LiCl-CNO). Main effect of group, ( $F(3,35) = 4.53, p < 0.01$ ). I Inactivation of VTA dopamine cells or mediated devaluation did not impact motivation to initiate consumption (cluster number). J eYFP-LiCl-CNO and hM4Di-LiCl-VEH mice displayed significant reduction in palatability of the sucrose reward relative to eYFP-saline-treated mice, whereas mice exposed to DREADD inhibition of VTA dopamine cells (hM4Di-LiCl-CNO) during CS-LiCl aversion did not differ significantly from eYFP mice exposed to CS-saline during aversion (eYFP-saline-CNO;  $p = 0.09$ ). Main effect of group ( $F(3,33) = 6.80, p = 0.01$ ). Filled blue circles = eYFP-saline-CNO; open blue circles = eYFP-LiCl-CNO; Filled red circles = hM4Di-LiCl-CNO; open red circles = hM4Di-LiCl-VEH. Overall group effect, ★★ $p < 0.01$ . Post-hoc group differences, \*\* $p \leq 0.02$ , \* $p < 0.05$ , # $p = 0.06$ .

local dopamine release in the NAc. Mice received an AAV for the engineered dopamine receptor (dLight1.1) into the NAc<sup>44</sup> (Fig. 4A). This fluorescent biosensor permits the recording of fluorescence upon receptor binding of dopamine, producing fast observable responses to local

dopamine release (Fig. 4B). Mice received Pavlovian training with two discriminable auditory CSs paired with the two distinct flavored pellets. This approach permitted a rigorous within-subject control for photometry to compare dLight activity. During the final stages of training, NAc dopamine



**Fig. 4 | Dopamine release in the nucleus accumbens is modulated by reward cue presentation and retrieval of mediated devaluation.** **A** Representative photomicrograph of dLight expression (green) in nucleus accumbens (NAc) with DAPI (blue). Fiber tract for optic fiber is demarcated by straight dashed line. **B** Injection and implant schematic, wild-type mice were injected with dLight1.1 in NAc and optic fibers at the same sites. **C** Left panel, mean z-scored NAc dopamine timecourse response during CS test. Following aversion, NAc dopamine release was comparable during responses to CSs paired with either saline or LiCl (0 s = CS response). Right panel, responses collapsed across (pre-CS) baseline and CS response. Main effect of cue presentation only, ( $F(1,10) = 61.76, p < 0.001$ ). **D** Left panel, timecourse responses during consumption test, NAc dopamine activity was elevated during retrieval of devalued reward memory as mice made contact with the food pellets (0 s = pellet response). Right panel, responses collapsed across baseline (5 s prior to pellet response) and pellet response. Blue circles = saline; red circles = LiCl ★★ Time X condition interaction ( $F(1,8) = 19.2, p < 0.01$ ). Post-hoc Bonferroni group differences \*\* $p < 0.001$ .

responses displayed an expected increase in activity in response to the CSs and food USs<sup>41,45</sup>, and importantly prior to aversion there were no differences in stimulus responses (Supplementary Fig. 9). Following aversion, when mice received the cue test to examine whether previously pairing the CS with LiCl altered its encoding, findings (Fig. 4C) revealed comparable dLight activity in the NAc for each CS, irrespective of whether it was paired with LiCl or saline. This is consistent with the previously reported behavioral findings (Figs. 2G and 3G) indicating that CS processing was unaffected by

mediated devaluation. Strikingly, however, during the consumption test, we observed significant increases in dLight in the NAc as mice consumed the food pellet whose memory had been devalued by pairing its CS associate with LiCl (Fig. 4D;  $p < 0.01, d = 2.12$ ).

### Modeling mediated devaluation and dopamine activity using the successor representation model

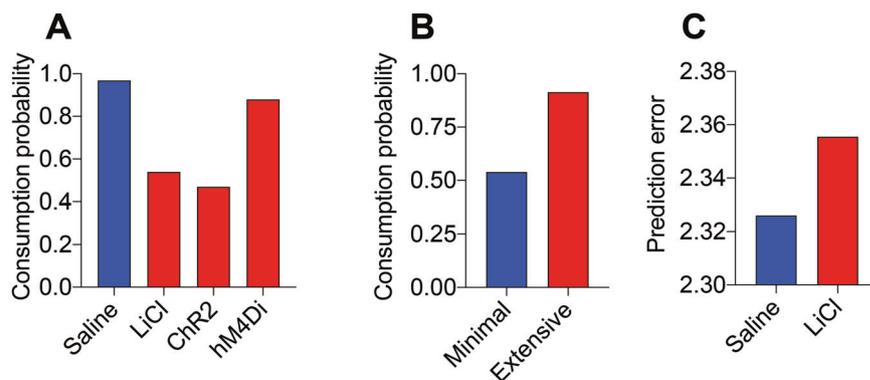
The observation that VTA dopamine cells appear to encode detailed reinforcement signals necessary for the devaluation of reward memories suggests a novel role of dopamine function that cannot be readily accounted for by standard RPE models of dopamine function<sup>45</sup>. These models have no mechanism by which the value of a reinforcer can be updated after CS devaluation. Dopamine neurons may receive “model-based” inputs that confer devaluation sensitivity<sup>23,46</sup>; however, it remains unclear what computational function RPEs play within a model-based reinforcement learning system. A related hypothesis is that dopamine neurons signal a vector-valued “generalized prediction error” over a collection of features, rather than just rewards<sup>25</sup>. This hypothesis can account for a wide range of non-classical dopamine responses, including the sensitivity of dopamine neurons to sensory prediction errors<sup>12,47,48</sup> and the causal role of dopamine in learning sensory predictions<sup>13,14</sup>.

We applied the successor representation (SR) model developed by Gardner and colleagues<sup>25</sup> to our mediated devaluation experiment (see Supplementary Materials for details). This model computes a prediction error for each sensory feature and treats the aggregated dopamine signal as a superposition of these errors. To capture optogenetic and chemogenetic perturbations, the model applies a modulation of the prediction errors. Sucrose consumption is modeled as a monotonic function of expected future reward. The model was able to recapitulate key empirical findings reported above, including the observed reduction in sucrose consumption after CS devaluation with LiCl (Fig. 5A), which was accounted for by the capacity of the CS to activate a predictive representation of the sucrose US during aversion, linking it to LiCl by error-driven learning. Moreover, consistent with our behavioral findings (Supplementary Fig. 1) and other studies<sup>27,28</sup>, the SR model predicts that sucrose consumption will be significantly lower in the minimal relative to the extensively trained condition (Fig. 5B). This reflects the acquisition of a stronger association between the US feature with reward after extensive training, partially counteracting the effects of devaluation during aversion. In addition, the model accurately predicted the dopamine manipulations, which include an enhanced devaluation effect after optogenetic stimulation (ChR2), and a reduced devaluation effect after chemogenetic inhibition (hM4Di) (Fig. 5A). These aspects of the model reflect its capacity to encode the predictive representation of sensory features, such that stimulating or inhibiting these errors produces directional changes in stimulus-stimulus learning<sup>49</sup>. Finally, consistent with the dLight findings, the model predicted higher dopamine transmission during the consumption test in the devaluation condition compared to the control condition (Fig. 5C), due to the surprising absence of gastric malaise that elicits a larger prediction error in the LiCl condition.

### Discussion

Our findings reveal a novel function for midbrain dopamine cells through which they gate access to a vast array of detailed features of reinforcement that are typically activated by biologically meaningful stimuli alone. Having confirmed the parameters underlying mediated devaluation, we demonstrated that mesencephalic cells encode CS-evoked mediated devaluation of sucrose reward. Accordingly, chemogenetic reactivation of these cells led to devaluation due to an associatively mediated reduction in the hedonic taste properties of the sucrose. We also showed that within the VTA, transient activation of dopamine cells during aversion was sufficient to enhance the encoding of detailed sensory reinforcement features, which promoted mediated devaluation when mice were subsequently tested in the absence of optogenetic manipulations. Conversely, chemogenetic inactivation of VTA dopamine neurons prevented associatively-evoked perceptual processing of the taste features of the sucrose reward, disrupting the capacity of the CS-

**Fig. 5 | Modeling mediated devaluation through sensory prediction errors.** **A** Successor Representation model simulation of sensory prediction errors revealed a reduction in sucrose consumption after memory devaluation (Saline vs. LiCl), an augmented devaluation effect following optogenetic stimulation (Chr2), and an attenuated devaluation effect as a result of chemogenetic inhibition (hM4Di). **B** The model predicts that mediated devaluation is weaker following extensive Pavlovian training due to the sucrose US feature acquiring a more robust association with reward, which counteracts the effects of CS devaluation. **C** The model predicts greater prediction errors during the consumption test in LiCl relative to saline conditions.



LiCl pairings to establish taste-illness associations and elicit a devaluation to the sucrose reward during consumption testing. The retrieval of these reinforcement features also reflected increased dopamine binding in the NAc. These features of dopamine encoding were accurately predicted by the SR model, in which a prediction error for each sensory feature was computed and revealed core features of the dopaminergic manipulations and activity profile of dopamine transients that we observed during mediated devaluation testing.

These findings stand in stark contrast to traditional accounts<sup>3,5,50,51</sup> that restrict dopamine function to encoding RPEs using model-free or cached value signals of future events. These algorithms limit dopamine's role to encoding value signals and do not provide access to specific details about the contents of learning. As such, VTA dopamine cell stimulation would be expected to enhance RPE leading to a greater influence over CS-LiCl learning; however, no evidence during either cue testing or when we separately examined dopamine binding in the NAc indicated such a role. Moreover, RPE cannot account for why acute dopamine stimulation during aversion would enhance the capacity of CS-LiCl pairings to subsequently augment mediated devaluation to the specific taste features of sucrose reward. Dopamine's role in this latter process suggests that in the presence of the CS, stimulation during aversion enhanced further access to detailed sensory (e.g., taste) processing of the associated absent sucrose US, whereas inactivating VTA dopamine neurons at this time prevented retrieval of detailed reinforcement features that were necessary for the generation of mediated devaluation. Moreover, model-free learning systems cannot account for why putative dopamine transients in the NAc would be greater for the food pellet whose memory had been devalued by pairing its CS associate with LiCl. In this situation, dopamine activity would be reliant on the transfer of value between cues retrospectively (e.g., food pellet—CS—LiCl), which is not factored into the computational architecture of model-free or cached value reinforcement agents. This heightened activity to the devalued food pellet is all the more challenging given the aversive nature of the phenomenon (i.e., via LiCl) and the prediction by RPE that this should serve to diminish dopamine responses<sup>25,28</sup>. In addition to RPE, dopamine also plays a critical role in modulating motivational salience toward both rewarding and aversive events<sup>42,52,53</sup>. As a result, following these events, animals display an increase in attention, orienting, and general motivation. However, our findings indicate that the salience of the CS was unchanged following aversive LiCl pairings, at least as measured by approach behavior. Salience effects would also struggle to account for how the CS mediated new learning to the absent associated reward and why these effects were specific to reductions in the palatability of sucrose, while at the same time leaving intact the animal's motivation to initiate sucrose consumption.

To account for dopamine's role in mediated devaluation, we modeled dopamine encoding using the SR model that assigns prediction errors to sensory signals and treats the overall dopamine signal as a superposition of these errors<sup>24,25</sup>. Through this approach, we were able to replicate key features of the mediated devaluation phenomenon. One of the notable features of mediated devaluation is its transient nature. Early on in training, CSs are

capable of retrieving detailed reinforcement features that are sufficiently salient that they can transfer the acquisition of the aversion to the absent food reward. As training proceeds, the window of CS-evoked access to detailed reward features narrows, meaning that the CS can no longer support a sufficiently detailed representation of the reward (e.g., its taste) to enable taste-illness associations following LiCl<sup>27,28</sup>. To account for the short-lived expression of mediated devaluation, studies have focused on the idea that the capacity for the CS to influence learning (its associability<sup>54</sup>) diminishes as training proceeds, and perhaps with it, its capacity to support mediated devaluation. However, approaches designed to increase associability after extensive training were subsequently ineffective in establishing a representation-mediated aversion to food<sup>28</sup>. This contrasts with our findings that acute VTA dopamine cell stimulation together with CS presentation during aversion was nevertheless effective in reestablishing mediated devaluation following prolonged training with the CS. To account for the weaker mediated devaluation after extensive training, the SR model predicts that the US feature has a stronger association with reward, and this counteracts the effects of CS devaluation. Moreover, the model makes the prediction that larger RPEs will result from consumption of the devalued food following minimal training (Supplementary Fig. 10). Importantly, the model also accurately predicted that dopamine stimulation would enhance, and inhibition reduce mediated devaluation, which is reflected in its capacity to encode predictive representation of sensory features to produce directional changes in stimulus-stimulus learning. As it related to the enhanced binding of dopamine in the NAc, the SR model also accounted for this through an enhanced error signal driven by the surprising absence of gastric malaise in the LiCl condition.

Our series of studies are the first to examine in detail the circuitry underlying mediated devaluation. These findings suggest that the encoding of detailed reinforcement signals by VTA dopamine cells is in part relayed to the NAc. The NAc is composed mainly of medium-size spiny neurons, which can be distinguished based on their expression of cells expressing D1Rs or D2Rs<sup>55,56</sup>. Interestingly, consistent with our past work<sup>33</sup> the capacity of acute stimulation of VTA dopamine neurons to augment mediated devaluation was dependent on D2R activation. Thus, it is tempting to speculate that the circuit underlying dopamine-based encoding of mediated devaluation includes VTA efferents to the NAc in a D2R-dependent manner. Of course, other target regions from VTA dopamine cells may play an important role in mediated devaluation, including the basolateral amygdala<sup>19</sup>, hippocampus<sup>57</sup>, and other cortical targets such as the insular<sup>58</sup> and orbitofrontal cortex<sup>59</sup>. Future studies targeting the nature of any underlying circuitry encoding detailed reinforcement signals via dopamine should be explored. Our studies also indicated potential biological sex differences, in that male mice appeared to be more vulnerable to mediated devaluation and the stimulatory effects that followed VTA dopamine cells. Given the small sample size, these latter findings should be interpreted with caution though they are nevertheless consistent with sex differences in dopaminergic signaling<sup>60,61</sup>.

Several limitations and unanswered questions remain. Many of our studies utilized a single-outcome design, which is not ideally suited to examine sensory-specific encoding<sup>15</sup>. However, we adopted rigorous analyses of ingestive behavior to confirm that our mediated devaluation effects typically attenuated the taste and sensory features of the sucrose reward, providing confidence that our manipulations disrupted sensory and not more diffuse features of reinforcement that would have been revealed through other licking (i.e., cluster number) and Pavlovian approach measures. Moreover, we implemented a multiple-outcome design with the fiber photometry experiments and revealed a pattern of dopamine binding consistent with sensory-specific encoding. Other concerns include the TH-Cre mouse line employed in the optogenetic and chemogenetic studies, which is known to suffer from ectopic expression of Cre-recombinase<sup>62</sup> and the possibility that CNO has the potential to reverse metabolize to clozapine, though the dose used in our studies is below that which typically leads to clozapine conversion<sup>63</sup>.

These limitations aside, our findings provide novel insight into the role of dopamine in reinforcement encoding and suggest a pivotal role in mediated devaluation. To achieve this, a subset of dopamine cells are required to undergo computational processes that are more elaborate than traditional models of dopamine function predict and may in part reflect dopamine's role in sensory prediction error<sup>25</sup>. Future studies targeting the nature of any underlying circuitry encoding these detailed reinforcement signals via dopamine should include the extent to which our findings relate to proposed roles in model-based encoding<sup>22</sup>, or signaling of surprise from afferent sensory systems<sup>64</sup> that are relayed to higher-order cortical sites<sup>23</sup>. Alternatively, given our approach intersects appetite and aversion, it is possible that we are engaging a heterogeneous population of dopamine cells that include those that respond to high-intensity sensory stimuli and are aversive when stimulated<sup>65</sup>. In addition to uncovering novel mechanisms of dopamine action in learning, our studies are relevant to neuropsychiatric endophenotypes of reality testing<sup>31</sup> and could be implemented as a tool to attenuate ingestive behavior<sup>66,67</sup> and substance abuse<sup>68</sup> through dampening, via mesencephalic DA manipulations, the memories associated with reward.

## Methods

### Animals

All mice were initially group housed prior to surgery and thereafter single housed for the duration of the study. The *cfos*-*htTA* mice were originally obtained from Jackson Labs (Strain # 018306) and contained two co-injected transgenes, *cfos*-*tTA* and *cfos*-*shEGFP*. The expression of tetracycline transcriptional activator (*tTA*) and green fluorescent protein (*shEGFP*) is directed to activated neurons by the *cfos* promoter. Mice were bred for a minimum of two generations with wild-type C57BLJ mice (Jackson Laboratory). *cfos*-*htTA* mice were 8–12 weeks at the time of surgery and maintained on a diet containing 40 mg/kg doxycycline. For the optogenetic and chemogenetic studies, TH-Cre mice expressing Cre-recombinase under the control of the tyrosine hydroxylase promoter (TH-Cre) (Jackson Laboratory, Strain #008601) were used. TH-Cre were bred up to four generations out with wild-type C57BLJ mice (Jackson Laboratory). TH-Cre mice were 8–12 weeks at the time of surgery. For the *dLight* studies wild-type C57BLJ mice (Jackson Laboratory) were 10–12 weeks at the time of surgery. Mice were maintained under a 12 hr light dark cycle (lights on at 7AM). All procedures were carried out between the hours of 11AM–4PM and approved by the Michigan State University Institutional Animal Care and Use Committee. We have complied with all relevant ethical regulations for animal use.

### Stereotaxic surgery

For the activity-dependent labeling study (Experiment 1), at 12 wks of age, *cfos*-*htTA* mice ( $n = 6\sigma, 7\varphi$ ) were anesthetized using 5% isoflurane and 1 mg/kg buprenorphine, placed in a stereotaxic apparatus and virally infused bilaterally with 0.25  $\mu$ l of a tet-responsive adenovirus-associated virus pAAV-PTRE-tight-hM3Dq-mCherry expressing hM3Dq at the level

of the VTA (AP  $-3.08$ , ML  $+/-0.6$ , DV  $-4.5$ ). See Supplementary Methods for additional details on virus constructs.

To examine the sufficiency of VTA dopamine cells in devaluing memories of food reward (Experiment 2), TH-Cre mice received 0.25  $\mu$ l of a Cre-dependent adenovirus-associated virus expressing channel-rhodopsin (AAV5-Ef1 $\alpha$ -DIO-ChR2-eYFP;  $n = 19\sigma, 16\varphi$ ) or control eYFP (AAV5-Ef1 $\alpha$ -DIO-eYFP;  $n = 23\sigma, 9\varphi$ ) (Vector Biolabs, Malvern, PA) unilaterally infused into the VTA in a manner counterbalanced for hemisphere. On infection and subsequent recombination in a Cre-expressing cell, the ChR2 version of the virus leads to the expression of modified Na<sup>+</sup> channels at the level of the cell membrane that elicit action potentials in the presence of 473 nm wavelength light. Viral infusions lacking the ChR2 sequence simply carried a generic eYFP reporter as a means of controlling for non-specific effects (e.g. behavior changes due to neural inflammation, tissue damage, etc). Following viral infusions, optic fiber cannulae (200  $\mu$ m core, 4.1 mm; Thorlabs, Newton, NJ) were implanted dorsal ( $\approx 0.3$  mm) to the injection site and affixed with dental acrylic (Lang Dental Manufacturing Co, Wheeling, IL).

To determine the necessity of VTA dopamine cells (Experiment 3), TH-Cre mice received 0.25  $\mu$ l of a Cre-dependent inhibitory DREADD virus (AAV8- hSyn-DIO-hM4D(Gi)-mCherry;  $n = 12\sigma, 12\varphi$ ) (Addgene) or control eYFP ( $n = 10\sigma, 9\varphi$ ) bilaterally injected into the VTA.

For fiber photometry experiments, 200  $\mu$ l of AAV5-CAG-*dLight1.1* (AddGene #111067-AAV5) was infused into the NAc (AP: 1.2 mm, ML:  $-1.3$  mm, DV:  $-4.1$ ,  $-4.5$  mm relative to bregma) of wild-type C57BLJ mice ( $n = 4\sigma, 2\varphi$ ). In addition, optical fibers were implanted during the same surgery. Thus, after viral injection, a metal ferrule optic fiber (400- $\mu$ m diameter core; BFH37-400 Multimode; NA 0.37; ThorLabs) was implanted unilaterally over NAc (A/P: 1.2 mm, M/L:  $-1.3$  mm, D/V:  $-4.1$  mm). Fibers were fixed to the skull using dental acrylic; after the completion of the experiments, mice were sacrificed, and the locations of optic fiber tips were identified based on the coordinates of Franklin and Paxinos<sup>69</sup>.

### Drug treatment

To activate the excitatory (Experiment 1) and inhibitory DREADD (Experiment 3), clozapine-N-oxide (CNO; NIDA Drug Supply Program) was diluted in 10% (2-Hydroxypropyl)- $\beta$ -cyclodextrin in 0.2 M sterile phosphate-buffered solution (PBS). Mice received intraperitoneal injections of CNO (0.3 mg/kg) 15 min prior to the memory retrieval and aversion phase. To examine whether the facilitatory effect of optogenetic stimulation of VTA dopamine cells (Experiment 2) during memory retrieval requires intact D2R signaling, mice ( $n = 13\sigma, 10\varphi$ ) received a 0.1 mg/kg intraperitoneal injection of haloperidol (MilliporeSigma, Burlington, MA) dissolved in a 10% Tween 80 (MilliporeSigma) and 90% sterile saline vehicle. This dose was chosen as it was previously shown to disrupt representation-mediated responding without influencing motoric actions<sup>33</sup>.

### Optogenetic stimulation

For optogenetic stimulation (Experiment 2), 473 nm blue light was delivered via a fiber-coupled laser source (ThorLabs, Newton, NJ) that was attached to a waveform generator (Agilent Technologies, Santa Clara, CA) integrated into the Med Associates apparatus. For the behavioral studies, prior to each session, the light intensity was tested and calibrated using a high-sensitivity power meter (ThorLabs) to emit  $\sim 20$  mW at the tip of the 200  $\mu$ m optical fiber, which was subsequently attached to the ferrule tip of the mouse. Laser stimulation occurred during cue-evoked memory retrieval during the latter 5 s of each CS presentation; the period of time that in training resulted in the delivery of a sucrose reward. Mice received 1 s of optogenetic stimulation (5 ms pulses at 20 Hz). For the in-vitro electrophysiology studies (see, supplementary methods), slices were exposed to light pulses of 10 Hz and 25 Hz.

### dLight recordings

Beginning 3 weeks after surgery, mice were connected to a fiber optic patch cable. Fiber optic patch cables (0.8 m long, 400  $\mu$ m diameter; Doric Lenses)

were firmly attached to the implanted fiber optic cannulae with zirconia sleeves (Doric Lenses). LEDs (Plexon; 473 nm) were set such that a light intensity of <0.1 mW entered the brain; light intensity was kept constant across sessions for each mouse. Emission light was passed through a filter cube (Doric) before being focused onto a sensitive photodetector (2151, Newport). Signals were digitized at 60 Hz using PyPhotometry, which allows for pulsed delivery of light, minimizing the amount of bleaching over the course of recordings.

To address photobleaching over the course of the recording period, the photometry signal was corrected by subtracting a double exponential fit and then adding back the mean of the trace. Signals were then smoothed with a 120 ms sliding window and the background was subtracted. The fluorescence signal was converted to  $\Delta F/F ((F - F_0)/F_0)$ ; where  $F_0$  was calculated as the 10th percentile of the entire fluorescence trace). These traces were then z-scored using the MATLAB (Mathworks, Natick, MA) *zscore* function to facilitate comparisons across days and mice.

### Immunohistochemistry

Mice were deeply anesthetized by way of intraperitoneal injection of sodium pentobarbital, then sacrificed via exsanguination during transcardial perfusion with 4% paraformaldehyde (Sigma-Aldrich, St. Louis, MO). Brains were extracted and placed in a 10% sucrose with 4% paraformaldehyde solution for 24 h at 4°C. Afterward, brains were sliced using a freezing microtome at 30  $\mu$ m and moved through six, 8 min washes in 0.1 M phosphate-buffered saline (PBS). Next, slices were placed in a solution consisting of 3% normal donkey serum (NDS; Catalog# 017-000-121; Jackson ImmunoResearch, West Grove, PA) and 10% Triton-x (MilliporeSigma, Burlington, MA) in PBS for 1 h, then 24 h in a solution of 3% NDS, 10% Triton-x, and rabbit-anti-TH primary (Catalog# P21962; MilliporeSigma) at 1:1000 concentration in PBS. The next day, slices were washed six times for 8 min each in PBS, then placed in a solution consisting of 3% NDS, 10% Triton-x; and for hM4Di-treated tissue, Alexa-Fluor donkey-anti-rabbit-488 secondary (Catalog# A21206; Invitrogen, Carlsbad, CA), or for ChR2 or eYFP-treated tissue, Alexa-Fluor donkey-anti-rabbit-568 secondary (Catalog# A10042; Invitrogen) at 1:1000 concentration in PBS for 24 h. Slices were then flow-mounted using 0.1 M phosphate buffer (PB) onto microscope slides and left to dry for 24 h. Finally, slides were treated with Prolong Gold Antifade Mountant with DAPI (Thermo Fisher Scientific, Waltham, MA) and left to cure for an additional 24 h prior to imaging and quantification.

### Behavioral studies

#### General behavioral procedures

**Food cup training.** Prior to behavioral testing, mice were food-restricted to 90% of their baseline weight by limiting food access to a single daily portion of lab chow, with the exception of cfos-htTA mice that received daily access to a diet containing 40 mg/kg Dox. Mice underwent one day of food cup training in Med Associates (Med Associates, St. Albans, VT) conditioning boxes (see Supplemental Methods for specific details of the training and testing apparatus). During these sessions, mice received 50  $\mu$ l of 0.2 M sucrose that was freely available at the start of the session. Once initial licking began, 16 sucrose deliveries occurred, with each delivery occurring under a random time 120 s schedule. At the start of each trial, a new 50  $\mu$ l bolus of 0.2 M sucrose was delivered congruently with magazine clicker presentation and made available for 10 s, upon which time it was vacuumed off. Mice were moved on to initial conditioning based on having met the criterion of at least 10 s spent licking while the reinforcer was available.

**Pavlovian conditioning.** To initially reveal the behavioral parameters for mediated devaluation, (group minimal) mice were split along the lines of either being paired with LiCl ( $n = 8$ ) or No-LiCl ( $n = 8$ ) and began four days of conditioning. Each training session lasted approximately 30 min, during which time they received 4 pseudo-randomly distributed presentations of the 10 s CS with a variable ITI of 450 s. 5 s into each CS, a delivery of a 50  $\mu$ l

0.2 M liquid sucrose reward (US) occurred. This limited training of 16 CS-US pairings was selected based on pilot studies aimed at setting the conditions for early-stage learning, a period in which mediated devaluation is thought to most readily occur. To confirm the transient nature of mediated devaluation, an additional group of mice (group extensive) was similarly split along whether they were to be paired with LiCl ( $n = 8$ ) or No-LiCl ( $n = 8$ ), however in each session mice received 16 pseudo-randomly distributed trials with the CS and US. This more extensive training exposure with 64 CS-US pairings was expected to prevent subsequent mediated devaluation. To examine whether optogenetic stimulation during aversion could reignite mediated devaluation a separate group of eYFP ( $n = 9$ ) and ChR2 ( $n = 7$ ) received 64 CS-US pairings during training.

**Aversion.** To enable activity-dependent labelling, the diet for cfos-htTA mice was switched to regular lab chow 24 hrs prior to the aversion stage. On day 5 mice were placed back in the conditioning chamber and allowed to habituate in the absence of cues for 6 min, this was done so as to minimize potential confounding associations as a result of handling stress and to ensure that mice were paying attention to the cue. Following habituation, the CS was played once every 30 s for 5 min in the absence of US delivery, with the goal of rapidly triggering a substitutive CS-evoked representation of the sucrose US. On completion of the session, in a manner counterbalanced for performance during the final two training sessions (i.e., entries/min in the magazine), mice from each group immediately received a 0.6 M intraperitoneal injection of LiCl at 0.15 mg/kg before being returned to their home cages. Mice were not fed for 3 h post-LiCl to avoid any association between chow and illness. At this stage, cfos-htTA mice ( $n = 12$ ) were also injected with Dox (66 mg/kg in 10 ml/kg) and placed back onto a diet containing 40 mg/kg Dox to further prevent labeling. Subsequently, approximately half the mice received cue followed by consumption testing, for the remaining mice testing order was reversed.

**Cue testing.** In order to determine whether the CS entered into direct associations with illness, mice were placed back in their original training context and given four presentations of the CS in the absence of US, entries in magazines during CS presentation were recorded. Mice received four 10 s CS presentations separated by a 450 s fixed ITI.

**Consumption testing.** Mice were placed back in the conditioning chamber for 5 min and given free access to the US in the absence of cues. Licks were recorded for analysis of microstructure with the expectation that mediated devaluation should occur in the form of decreased average cluster size (the amount of licks contained within a 500 ms pause criterion), a measure known to reflect stimulus palatability.

**dLight studies.** All experiments were conducted as within-subject tests, with mice tested with both flavors of pellets. Only one recording session was conducted on a given day. Animals were food-restricted to 85–90% of their ad lib fed body weight prior to starting behavioral experiments and maintained on a food-restricted diet for the duration of the recording period.

**Habituation to FED3.** To reduce neophobic responses, animals were allowed one 30 min session to acclimate to the recording chamber and freely feed from the in-house FED3 with standard pellets (Dustless Precision Pellets (20 mg); Bio-Serv).

**Habituation to reward.** Animals were allowed two 30 min sessions to freely feed from the in-house FED3 with either a banana-flavored pellet or berry-flavored pellet (Bio-Serv). Animals did not show an inherent preference for either pellet flavor.

**Pavlovian conditioning.** Animals were then trained across 8 alternating 40 min sessions to associate a 1 kHz tone with the delivery of the banana-flavored pellet and a 2.8 kHz tone with the delivery of the berry-flavored pellet. ~11 tones were played per session for 10 s, with pellet delivery occurring at 5 s

after tone onset, and a variable 3–4 min ITI (using custom Arduino code). After conditioning, mice were provided with an injection of saline.

**Aversion.** During a 10 min session, animals were presented with 10 trials of the 1 kHz tone, with a fixed 30 s ITI. At the end of this session, animals were administered an injection of LiCl. On the following day, this was repeated with the 2.8 kHz tone, after which animals received an injection of saline.

**CS test.** During a 40 min session, animals were presented with 11 trials of the 1 kHz tone, with a 3–4 min ITI. The following day, this was repeated with the 2.8 kHz tone.

**Reward test.** Animals were allowed one 30 min session to freely feed from the FED3 with a single flavor of pellet and a fixed 1 min ITI. The following day, this was repeated with the other flavor of the pellet. During testing with the banana-flavored pellet, the patch cable for one mouse became disconnected therefore their reward test data were excluded from the analyses. Mice received a total of 11 trials, the dopamine transients from these trials were averaged for the  $\Delta F/F$  z-scores.

### Statistics and Reproducibility

Data were subjected to analysis of variance (ANOVA). The initial parametric studies to confirm mediated devaluation (Supplementary Fig. 1), test data were analyzed separately for minimal and extensive conditions using a within-subject two-way condition (saline, LiCl) and time (1–5 min) ANOVA. Follow-up analyses were conducted using Bonferroni post-hoc. The activity-dependent labeling study (Fig. 1) was analyzed using a one-way ANOVA, with the within-subject variable of the drug (vehicle, CNO) for each response measure. Follow-up analyses included sex as a between-subject factor, and drug x sex interactions were followed up by tests of simple main effects. For the optogenetic study, a two-way ANOVA with between-subject variables of virus (eYFP, ChR2) and drug (no drug, haloperidol) on each response measure was incorporated (Fig. 2). Significant interactions were followed up by tests of simple main effects. The chemogenetic study was analyzed using a one-way ANOVA to confirm overall group differences between eYFP-saline-CNO, eYFP-LiCl-CNO, hM4Di-LiCl-VEH, hM4Di-LiCl-CNO. Group differences were analyzed subsequently with post-hoc Bonferroni comparisons (Fig. 3). The  $\Delta F/F$  z-scores from fiber photometry were subjected to a two-way ANOVA with within-subject variables of time (baseline, CS or pellet response) and condition (saline, LiCl) (Fig. 4). The  $\alpha$  level for significance was .05 and all analyses were conducted using Statistica (Statsoft, Tulsa, OK). In addition, Cohen’s  $d$  effect sizes measures were calculated (effect size interpretation: small,  $d = 0.20$ ; medium,  $d = 0.50$ ; large,  $d = 0.80$ ). Subjects were excluded if they were outliers according to the median absolute deviation. Using this method,  $n = 3$  and  $n = 2$  eYFP-saline-CNO mice were respectively excluded from the cluster number and cluster size analysis in the chemogenetic inhibition experiment. All data collected from these studies are available upon request.

### Computational modeling

We adapted the model developed by Gardner et al.<sup>25</sup>, which we summarize here. The mediated devaluation paradigm was analyzed into a set of states that are traversed sequentially. Each state  $s$  was represented by a feature vector  $f(s) = [f_1(s), \dots, f_b(s)]$ . We defined 4 features: a constant context feature, a CS feature, a US (sucrose) feature, and a “physiological state” feature (−1 for LiCl, 0 for saline/baseline).

The model formalizes sensory predictions using the *successor representation* [SR]; the expected discounted activation of each feature  $j$  conditional on state  $s_t$  is the state at time  $t$ :

$$M(s_t, j) = E \left[ \sum_{k=0}^{\infty} \gamma^k f_j(s_k) \right]$$

where  $E[\cdot]$  is the expectation operator (which averages its argument over randomness in future state transitions) and  $\gamma$  is a discount factor (controlling the effective time horizon of the SR). Because in general, the state space is large (or possibly infinite), we use a linear approximation of the SR parametrized by weight matrix  $W$ :

$$\widehat{M}(s_t, j) = \sum_i f_i(s) W_{ij}$$

The weights can be learned using a form of temporal difference learning, analogous to models of reinforcement learning but generalized to arbitrary features:

$$\Delta W_{ij} = \alpha_w \delta_t f_i(s_t)$$

where  $\alpha_w$  is a learning rate, and  $\delta_t$  is a vector-valued prediction error:

$$\delta_t(j) = f_j(s_t) + \gamma \widehat{M}(s_{t+1}, j) - \widehat{M}(s_t, j)$$

We model the total dopamine signal in our experiments as a superposition of these errors:

$$DA_t = \sum_j \delta_t(j)$$

Optogenetic and chemogenetic perturbations were modeled by the following equations:

$$\delta'_t(j) = \begin{cases} (1 + \eta) \delta_t(j), & \eta < 0 \\ f_j(s_t) \eta + \delta_t(j), & \eta > 0 \end{cases}$$

where  $\eta = 1.0$  for excitation and  $\eta = -0.8$  for inhibition (see ref. 25 for justification of this functional form). Obviously, it is a gross oversimplification to model optogenetic and chemogenetic perturbations in the same way, but for our purposes, this simplification was adequate to account for the experimental results.

Value computation was modeled by assuming a linear approximation:

$$\widehat{V}(s_t) = \sum_i f_i(s_t) \sum_j U_j W_{ij}$$

where  $U_j$  is a reward prediction weight for feature  $j$ , updated by an error-driven learning rule:

$$\Delta U_j = \alpha_U f_j(s_t) [r_t - \widehat{V}(s_t)]$$

with learning rate  $\alpha_U$ . To model consumption choice in the test phase, we transformed the values into choice probabilities using a sigmoidal transformation  $F(\widehat{V}(s_t) - \tau)$ , where  $F$  is the standard normal cumulative distribution function, and  $\tau$  is a response threshold.

To comply with the empirical observation that negative prediction errors have a smaller dynamic range than positive prediction errors, we rescaled the negative errors for both  $W$  and  $U$  by  $\frac{1}{4}$  (though our results don’t depend strongly on this assumption).

We used the same parameter values as in ref. 25:  $\gamma = 0.95$ ,  $\alpha_w = 0.06$ ,  $\alpha_U = 0.03$ . In addition, we set the response threshold  $\tau = 1$ , but the results are qualitatively unchanged for other choices of threshold.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Data availability

All source data for the data figures in this manuscript are provided as Supplementary Data (Excel files; Supplementary Data 1). All other data are available from the corresponding author on request and are currently being stored on a server at the Department of Psychology, Michigan State University.

## Code availability

The code that was used to generate the computational modeling with the Gardner model can be obtained at <https://github.com/mphgardner/TDSR>.

Received: 6 March 2024; Accepted: 23 December 2024;

Published online: 03 February 2025

## References

- Bowman, E. M., Aigner, T. G. & Richmond, B. J. Neural signals in the monkey ventral striatum related to motivation for juice and cocaine rewards. *J. Neurol.* **75**, 1061–1073 (1996).
- Sutton, R. S. Simulated evolution and learning, second Asia-Pacific conference on simulated evolution and learning, SEAL'98 Canberra, Australia, November 24–27, 1998 selected papers. *Lect Notes Comput Sc* 195–197 (1999).
- Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science* **275**, 1593–1599 (1997).
- Schultz, W. Predictive reward signal of dopamine neurons. *J. Neurophysiol.* **80**, 1–27 (1998).
- Schultz, W. & Dickinson, A. Neuronal coding of prediction errors. *Ann. Rev. Neurosci.* **23**, 473–500 (2000).
- Bayer, H. M. & Glimcher, P. W. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* **47**, 129–141 (2005).
- Eshel, N., Tian, J., Bukwich, M. & Uchida, N. Dopamine neurons share common response function for reward prediction error. *Nat. Neurosci.* **19**, 479–486 (2016).
- Steinberg, E. E. et al. A causal link between prediction errors, dopamine neurons and learning. *Nat. Neurosci.* **16**, 966–973 (2013).
- Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J. & Frith, C. D. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* **442**, 1042–1045 (2006).
- Bayer, J., Rusch, T., Zhang, L., Gläscher, J. & Sommer, T. Dose-dependent effects of estrogen on prediction error related neural activity in the nucleus accumbens of healthy young women. *Psychopharmacology* **237**, 745–755 (2020).
- Torigoe, M. et al. Zebrafish capable of generating future state prediction error show improved active avoidance behavior in virtual reality. *Nat. Commun.* **12**, 5712 (2021).
- Keiflin, R., Pribut, H. J., Shah, N. B. & Janak, P. H. Ventral tegmental dopamine neurons participate in reward identity predictions. *Curr. Biol.* **29**, 93–103.e3 (2019).
- Sharpe, M. J. et al. Dopamine transients are sufficient and necessary for acquisition of model-based associations. *Nat. Neurosci.* **20**, 735–742 (2017).
- Chang, C. Y., Gardner, M., Tillio, M. G. D. & Schoenbaum, G. Optogenetic blockade of dopamine transients prevents learning induced by changes in reward features. *Curr. Biol.* **27**, 3480–3486.e3 (2017).
- Balleine, B. W. & Dickinson, A. The role of incentive learning in instrumental outcome revaluation by sensory-specific satiety. *Anim. Learn. Behav.* **26**, 46–59 (1998).
- Holland, P. C. Relations between Pavlovian-instrumental transfer and reinforcer devaluation. *J. Exp. Anim. Behav. Process.* **30**, 104–117 (2004).
- Daw, N. D., Niv, Y. & Dayan, P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* **8**, 1704–1711 (2005).
- Johnson, A. W., Bannerman, D., Rawlins, N., Sprengel, R. & Good, M. A. Targeted deletion of the GluR-1 AMPA receptor in mice dissociates general and outcome-specific influences of appetitive rewards on learning. *Behav. Neurosci.* **121**, 1192–1202 (2007).
- Johnson, A. W., Gallagher, M. & Holland, P. C. The basolateral amygdala is critical to the expression of Pavlovian and instrumental outcome-specific reinforcer devaluation effects. *J. Neurosci.* **29**, 696–704 (2009).
- Johnson, A. W., Bannerman, D. M., Rawlins, N. R., Sprengel, R. & Good, M. A. Impaired Outcome-Specific Devaluation of Instrumental Responding in Mice with a Targeted Deletion of the AMPA Receptor Glutamate Receptor 1 Subunit. *J. Neurosci.* **25**, 2359–2365 (2005).
- Wunderlich, K., Smittenaar, P. & Dolan, R. J. Dopamine enhances model-based over model-free choice behavior. *Neuron* **75**, 418–424 (2012).
- Langdon, A. J., Sharpe, M. J., Schoenbaum, G. & Niv, Y. Model-based predictions for dopamine. *Curr. Opin. Neurobiol.* **49**, 1–7 (2018).
- Akam, T. & Walton, M. E. What is dopamine doing in model-based reinforcement learning? *Curr. Opin. Behav. Sci.* **38**, 74–82 (2021).
- Gershman, S. J. The successor representation: its computational logic and neural substrates. *J. Neurosci.* **38**, 7193–7200 (2018).
- Gardner, M. P. H., Schoenbaum, G. & Gershman, S. J. Rethinking dopamine as generalized prediction error. *Proc. Biol. Sci.* **285**, 20181645 (2018).
- Holland, P. C. Acquisition of representation-mediated conditioned food aversions. *Learn. Motiv.* **12**, 1–18 (1981).
- Holland, P. Amount of training affects associatively-activated event representation. *Neuropharmacol* **37**, 461–469 (1998).
- Holland, P. C. Amount of training effects in representation-mediated food aversion learning: no evidence of a role for associability changes. *Learn. Behav.* **33**, 464–478 (2005).
- Fry, B. R., Fex, V., Sawa, A., Niwa, M. & Johnson, A. W. The role of midbrain dopamine cells projecting to the insular cortex in mediated performance: implications for animal models of reality testing. *Behav. Neurosci.* **138**, 164–177 (2024).
- Holland, P. C. Event representation in Pavlovian conditioning: Image and action. *Cognition* **37**, 105–131 (1990).
- Fry, B. R., Roberts, D., Thakkar, K. N. & Johnson, A. W. Variables influencing conditioning-evoked hallucinations: overview and future applications. *Psychol. Med.* **52**, 2937–2949 (2022).
- Holland, P. C., Lasseter, H. & Agarwal, I. Amount of training and cue-evoked taste-reactivity responding in reinforcer devaluation. *J. Exp. Anim. Behav. Process.* **34**, 119–132 (2008).
- Fry, B. R. et al. Assessing reality testing in mice through dopamine-dependent associatively evoked processing of absent gustatory stimuli. *Schizophr. Bull.* **30**, 67–14 (2019).
- Garcia J. Food for Tolman: cognition and cathexis in concert. In *Aversion, Avoidance and Anxiety* (eds Archer T. & Nilsson L.-G.) 45–85. (Erlbaum, Hillsdale, NJ, 1989).
- Johnson, A. W. et al. An analysis of licking microstructure in three strains of mice. *Appetite* **54**, 320–330 (2010).
- Johnson, A. W. & Gallagher, M. Greater effort boosts the affective taste properties of food. *Proc. Biol. Sci.* **278**, 1450–1456 (2011).
- Johnson, A. W. Examining the influence of CS duration and US density on cue-potentiated feeding through analyses of licking microstructure. *Learn. Motiv.* **61**, 85–96 (2018).
- Johnson, A. W. Characterizing ingestive behavior through licking microstructure: Underlying neurobiology and its use in the study of obesity in animal models. *Int. J. Dev. Neurosci.* **64**, 38–47 (2017).
- Robles, C. F. & Johnson, A. W. Disruptions in effort-based decision-making and consummatory behavior following antagonism of the dopamine D2 receptor. *Behav. Brain Res.* **320**, 431–439 (2017).
- Zhang, F. et al. Optogenetic interrogation of neural circuits: technology for probing mammalian brain structures. *Nat. Protoc.* **5**, 439–456 (2010).
- Hart, A. S., Rutledge, R. B., Glimcher, P. W. & Phillips, P. E. M. Phasic dopamine release in the rat nucleus accumbens symmetrically encodes a reward prediction error term. *J. Neurosci.* **34**, 698–704 (2014).
- Wyvell, C. L. & Berridge, K. C. Intra-accumbens amphetamine increases the conditioned incentive salience of sucrose reward:

- enhancement of reward “wanting” without enhanced “liking” or response reinforcement. *J. Neurosci.* **20**, 8122–8130 (2000).
43. Salamone, J. D., Correa, M., Farrar, A. & Mingote, S. M. Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. *Psychopharmacology* **191**, 461–482 (2007).
44. Palissery, G. K. & Lerner, T. N. A dLight-ful new view of neuromodulation. *Trend Neurosci.* **41**, 566–568 (2018).
45. Schultz, W. Reward prediction error. *Curr. Biol.* **27**, R369–R371 (2017).
46. Dayan, P. & Berridge, K. C. Model-based and model-free Pavlovian reward learning: revaluation, revision, and revelation. *Cogn. Affect. Behav. Neurosci.* **14**, 573–492 (2014).
47. Takahashi, Y. K. et al. Dopamine neurons respond to errors in the prediction of sensory features of expected rewards. *Neuron* **95**, 1395–1405.e3 (2017).
48. Stalnaker, T. A. et al. Dopamine neuron ensembles signal the content of sensory prediction errors. *eLife* **8**, e49315 (2019).
49. Hall, G. Associative structures in Pavlovian and instrumental conditioning. In *Steven’s Handbook of Experimental Psychology: Learning, Motivation, and Emotion* (eds Pashler H. & Gallistel, R.) 1–45 (John Wiley & Sons Inc., 2002).
50. Schultz, W., Apicella, P., Scarnati, E. & Ljungberg, T. Neuronal activity in monkey ventral striatum related to the expectation of reward. *J. Neurosci.* **12**, 4595–4610 (1992).
51. Schultz, W. Behavioral dopamine signals. *Trend Neurosci.* **30**, 203–210 (2007).
52. Bromberg-Martin, E. S., Matsumoto, M. & Hikosaka, O. Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron* **68**, 815–834 (2010).
53. Howes, O. D. & Nour, M. M. Dopamine and the aberrant salience hypothesis of schizophrenia. *World Psychiatry* **15**, 3–4 (2016).
54. Pearce, J. M. & Hall, G. A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol. Rev.* **87**, 532–552 (1980).
55. Smith, Y., Bevan, M. D., Shink, E. & Bolam, J. P. Microcircuitry of the direct and indirect pathways of the basal ganglia. *Neuroscience* **86**, 353–387 (1998).
56. Humphries, M. D. & Prescott, T. J. The ventral basal ganglia, a selection mechanism at the crossroads of space, strategy, and reward. *Prog. Neurobiol.* **90**, 385–417 (2010).
57. Busquets-Garcia, A. et al. Hippocampal CB1 receptors control incidental associations. *Neuron* **99**, 1247–1259.e7 (2018).
58. Fry, et al. The role of midbrain dopamine cells projecting to the insular cortex in mediated performance: Implications for animal models of reality testing. *Behav. Neurosci.* **138**, 164–177 (2024).
59. McDannald, M. & Schoenbaum, G. Toward a model of impaired reality testing in rats. *Schizophr. Bull.* **35**, 664–667 (2009).
60. Walker, Q. D., Ray, R. & Kuhn, C. M. Sex differences in neurochemical effects of dopaminergic drugs in rat striatum. *Neuropsychopharmacology* **31**, 1193–1202 (2006).
61. Rivera-Garcia, M. T., McCane, A. M., Chowdhury, T. G., Wallin-Miller, K. G. & Moghaddam, B. Sex and strain differences in dynamic and static properties of the mesolimbic dopamine system. *Neuropsychopharmacology* **45**, 2079–2086 (2020).
62. Stuber, G. D., Stamatakis, A. M. & Kantak, P. A. Considerations when using Cre-driver rodent lines for studying ventral tegmental area circuitry. *Neuron* **85**, 439–445 (2015).
63. Manvich, D. F. et al. The DREADD agonist clozapine N-oxide (CNO) is reverse-metabolized to clozapine and produces clozapine-like interoceptive stimulus effects in rats and mice. *Sci. Rep.* **8**, 3840 (2018).
64. Redgrave, P. & Gurney, K. The short-latency dopamine signal: a role in discovering novel actions? *Nat. Rev. Neurosci.* **7**, 967–975 (2006).
65. Menegas, W., Akiti, K., Amo, R., Uchida, N. & Watabe-Uchida, M. Dopamine neurons projecting to the posterior striatum reinforce avoidance of threatening stimuli. *Nat. Neurosci.* **21**, 1421–1430 (2018).
66. Johnson, A. W. Eating beyond metabolic need: how environmental cues influence feeding behavior. *Trends Neurosci.* **36**, 101–109 (2013).
67. Sherwood, A., Holland, P. C., Adamantidis, A. & Johnson, A. W. Deletion of Melanin Concentrating Hormone Receptor-1 disrupts overeating in the presence of food cues. *Physiol. Behav.* **152**, 402–407 (2015).
68. Lovinger, D. M. & Gremel, C. M. A circuit-based information approach to substance abuse research. *Trends Neurosci.* **44**, 122–135 (2021).
69. Watson, C., Paxinos, G. & Puelles, L. (eds) *The Mouse Nervous System* (Academic Press, 2011).

## Acknowledgements

This research was funded in part by NIH grant R01-DK111575 and internal discretionary funds provided by the Department of Psychology and Neuroscience Program at Michigan State University, to A.W.J.

## Author contributions

B.R.F. played an equal role in the investigation and a supporting role in writing—review and editing. N.R. played an equal role in the investigation. V.F. and B.M. played an equal role in conducting tissue analyses. N.P. assisted with the investigation. J.A.B. provided slice recordings. F.P.M. designed and provided viral constructs for pAAV-PTRE-tight-hM3Dq-mCherry. B.A.T. and C.R.B. played an equal role in fiber photometry data collection and analysis. S.G. provided computational analyses. A.W.J. played a lead role in the conceptualization, investigation, methodology, and writing—original draft.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s42003-024-07440-7>.

**Correspondence** and requests for materials should be addressed to Alexander W. Johnson.

**Peer review information** *Communications Biology* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editors: Fereshteh Nugent and Benjamin Bessieres.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025