nature neuroscience

Article

https://doi.org/10.1038/s41593-023-01547-6

State and rate-of-change encoding in parallel mesoaccumbal dopamine pathways

Received: 24 September 2022

Accepted: 7 December 2023

Published online: 11 January 2024

Check for updates

Johannes W. de Jong, Yilan Liang, Jeroen P. H. Verharen, Kurt M. Fraser & Stephan Lammel

The nervous system uses fast- and slow-adapting sensory detectors in parallel to enable neuronal representations of external states and their temporal dynamics. It is unknown whether this dichotomy also applies to internal representations that have no direct correlation in the physical world. Here we find that two distinct dopamine (DA) neuron subtypes encode either a state or its rate-of-change. In mice performing a reward-seeking task, we found that the animal's behavioral state and rate-of-change were encoded by the sustained activity of DA neurons in medial ventral tegmental area (VTA) DA neurons and transient activity in lateral VTA DA neurons, respectively. The neural activity patterns of VTA DA cell bodies matched DA release patterns within anatomically defined mesoaccumbal pathways. Based on these results, we propose a model in which the DA system uses two parallel lines for proportional–differential encoding of a state variable and its temporal dynamics.

Parallel state and rate-of-change encoding is ubiquitous throughout the peripheral and central nervous system[1–4]. Seminal work discussed in ref. 4 demonstrated that the firing rate of pressure sensors in the cat's toe simultaneously encodes the total amount of pressure applied to the toe (that is, a state quantity) as well as the rate by which this quantity is changed[4] (Extended Data Fig. 1a). Simultaneous state and rate-of-change encoding has been observed across various sensory systems including visual[1], olfactory[2] and nociceptive[3].

Rate-of-change encoding also exists in the midbrain dopamine (DA) system in the form of temporal difference (TD) reward prediction error (RPE) encoding[5–9]. In this case, TD RPE encoding is thought to reflect a learning signal that is generated by DA neurons in the ventral tegmental area (VTA) to track the change of a state value that represents the expected cumulative reward on a moment-by-moment basis (Extended Data Fig. 1b). DA release patterns in the nucleus accumbens (NAc), the prominent projection target of these neurons, however, have diverse characteristics and appear to correlate with behavioral output more closely and do not necessarily reflect a derivative-like signal. Importantly, 'motivation', 'wanting' or 'value of work', which all have been associated with DA release in the NAc, refer to value- or state-like qualities[10–13]. Whether VTA DA neurons homogenously encode TD RPE or also encode state-like signals has been controversial[6,7,10,14–20].

A recent study suggested that the slow, gradual release of DA (ramping) in the NAc is controlled locally in the distal axon, independent from the firing of VTA DA cell bodies, and encodes information about the value of the anticipated reward[10]. In contrast, another recent study used behavioral manipulations to disambiguate TD RPE encoding from value demonstrating that ramping of DA is consistent with TD RPE when it is assumed that the underlying value function is convex[17]. Although a convex value function indeed could explain the gradual increase in DA release that is observed when an animal approaches a reward (for example, as in ref. 21), other aspects of value encoding in DA release are more challenging to explain in this way. For instance, it is unclear how a convex value function can explain the correlation between minute-by-minute DA release and reward rate[10,11].

A major limitation of neurophysiological studies on this topic has been that recordings of DA cell body activity and release have been largely constrained to the lateral VTA (lVTA) and NAc core, respectively. DA neurons in the medial VTA (mVTA), however, greatly differ from lVTA DA neurons in their connectivity and molecular and electrophysiological characteristics suggesting divergent functions[17,22–26]. Whether mVTA DA neurons, like lVTA DA neurons, encode a derivative-like signal—both at the level of individual cells and at the axon terminal level—during motivated behaviors remains uncertain.

Department of Molecular and Cell Biology and Helen Wills Neuroscience Institute, University of California, Berkeley, CA, USA. e-mail: lammel@berkeley.edu

Here we study the neural activity of mVTA and lVTA DA neurons as well as downstream DA release in the context of the precise topographical organization of the mesoaccumbal DA system. In mice trained on a reward-seeking task, we observed heterogeneous patterns of DA neuron activity across a dorso-lateral to medio-ventral VTA gradient. Our findings indicate that VTA DA neurons encode in parallel both an abstract variable that is directly related to an animal's current 'state' and the rate of change of value over time. Based on these findings, we propose a conceptual model suggesting that, analogous to sensory systems, the midbrain DA system uses two parallel lines for proportional–differential encoding.

Results

Opto-tagging of VTA DA neurons using Neuropixels probes

The bias toward the lVTA (see Methods for details on anatomical nomenclature) may in part arise from many challenges to recording and identifying DA neurons with standard tetrode- or single-wire-based in vivo electrophysiology[27]. For example, prominent blood vessels along the midline of rodents and nonhuman primates impede surgical access to more mVTA subregions. Additionally, the smaller size of mVTA DA neurons[22] may decrease the probability of detecting these cells with standard electrophysiological approaches. To reliably record DA neurons in the mVTA from behaving mice, we used an in vivo electrophysiological approach that combines optogenetics with Neuropixels probes[28]. We reasoned that the small size and the ability to simultaneously record across the entire ventro-dorsal extent of the VTA makes Neuropixels probes an ideal tool to gain reliable access to mVTA DA neurons.

To achieve cell-type-specific recordings with Neuropixels, we implanted an optical fiber (angled by 15° toward the VTA) in DAT-Cre mice crossed to a channelrhodopsin-2 (ChR2) reporter mouse line (DAT-Cre × Ai32). As a result, we could simultaneously record VTA cells and provide light stimulation for optogenetic identification of VTA DA neurons. Neuropixels probes were acutely lowered into the mVTA or lVTA of head-fixed mice, and light pulses were applied to identify ChR2-expressing DA neurons (Fig. 1a–c). We defined mVTA and lVTA based on the anatomical location of projection-defined VTA DA neurons[22]. It is not strictly based only on the medio-lateral axis, but also incorporates the dorso-ventral axis. Because Neuropixels probes are light sensitive, we used 1 ms light pulses, which is shorter than the mean response latency of a tagged cell (Extended Data Fig. 2a,b). Light artifacts were identified and removed by fitting and subtracting the characteristic artifact shape from each channel without the risk of altering the action potential (AP) shape as they were well separated in time. We recorded 186 VTA cells in six mice, 67 of which met the criteria to be identified as dopaminergic (Extended Data Fig. 2c–i; mVTA: $n = 40$ DA cells and lVTA: $n = 27$ DA cells). Consistent with previous reports[10,29], VTA DA neurons had significantly lower baseline firing frequencies and a larger variation in waveform width compared to putative non-DA neurons (Extended Data Fig. 2j,k). However, we did not detect significant differences in baseline firing frequency or waveform width between mVTA and lVTA DA neurons (Extended Data Fig. 2l).

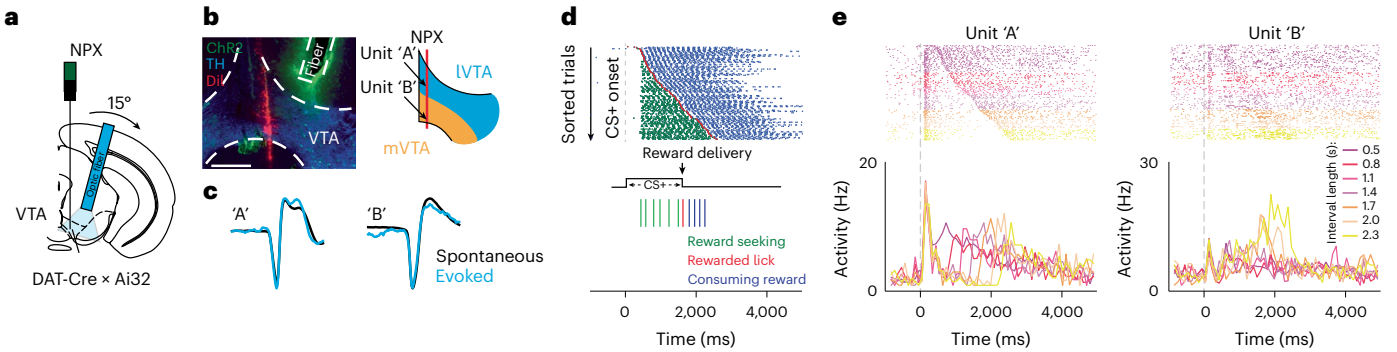
Sustained activity of VTA DA cells during motivated behavior

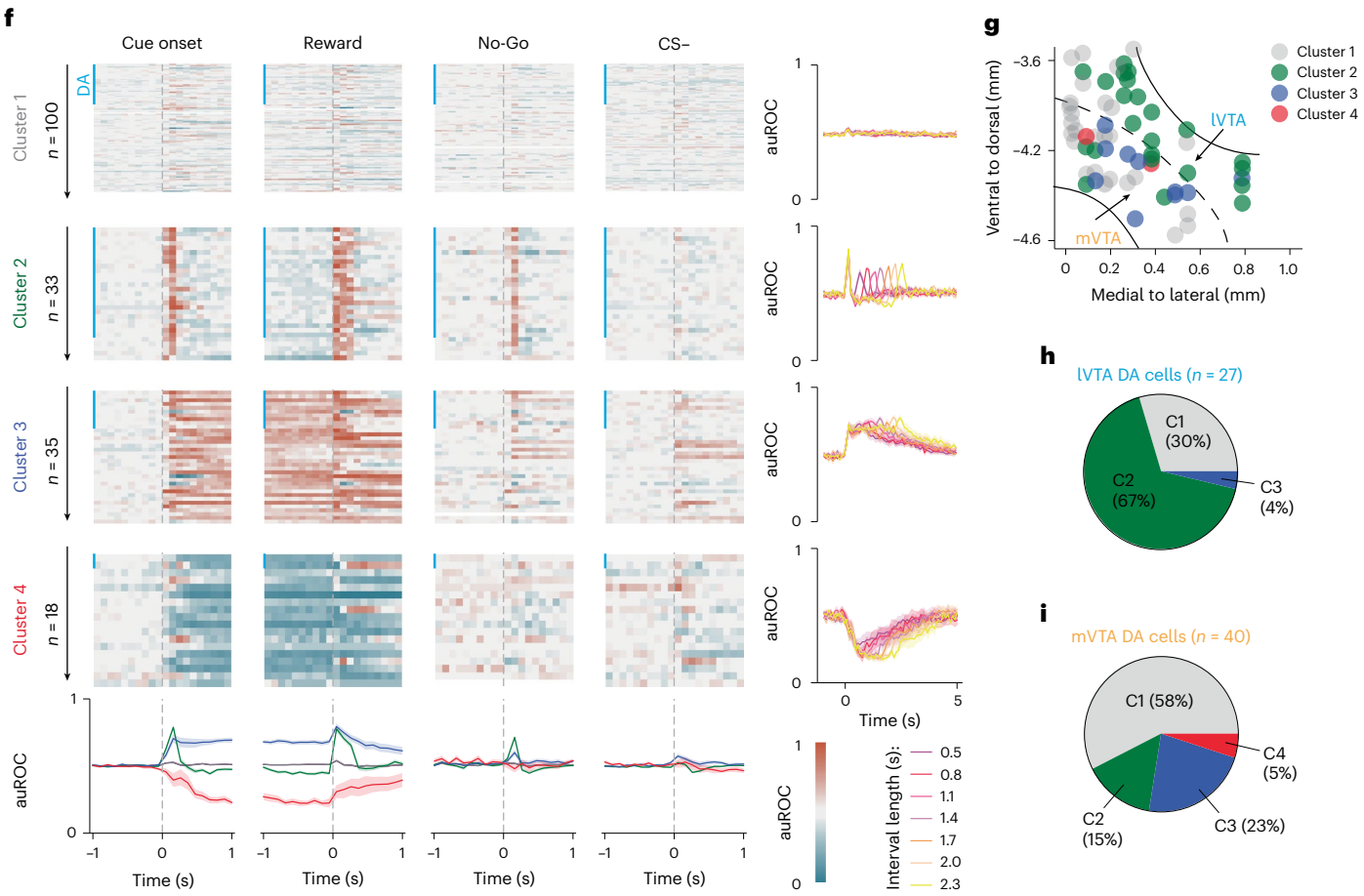
To assess neural activity patterns of DA neurons in different VTA subregions during motivated behavior, we performed Neuropixels recordings of genetically identified VTA DA neurons in mice that were trained on a cued reward-seeking task. In this task, the presence of one cue (CS+, conditioned stimulus) signaled the availability of a reward (1% sucrose solution) delivered following the first lick after the expiration of a random interval (range 0–2,500 ms drawn from a uniform distribution), while another cue (CS−) was presented without consequence. Thus, in the absence of information about the length of the random interval, the most successful strategy for obtaining a reward was continuous licking behavior during the CS+. Although mice had the possibility to not obtain a reward (that is, 'No-Go trials'), they typically displayed reward-seeking behavior during the presentation of the CS+ (Fig. 1d), which began ~500 ms after cue onset and lasted until the end of the random interval, suggesting both reward anticipation and motivation to work for the reward. Thus, the presentation of a CS+ is expected to lead to an immediate increase in reward expectation (that is, animals will obtain a reward if willing to work), but the exact time point of the reward delivery is unpredictable.

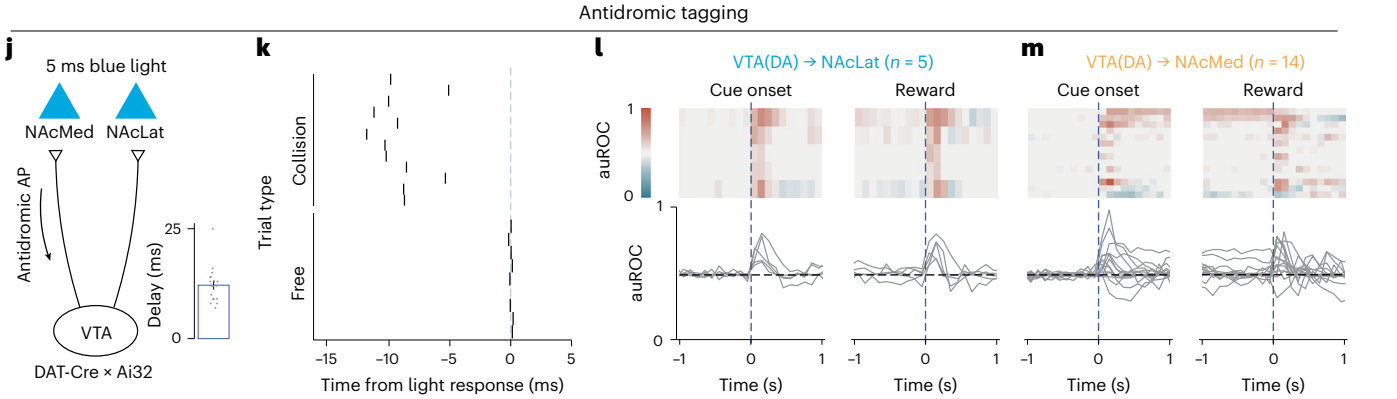
Following the recordings, we normalized neural activity patterns of all recorded VTA cells (area under the receiver operating characteristic, auROC; as in ref. 29). Sparse principal component analysis (PCA) in combination with k-means clustering ($k = 4$; see Methods for rationale) was applied to identify four clusters with distinct activity patterns (Fig. 1e,f and Extended Data Fig. 3a–d). The four clusters were not homogeneously represented throughout the VTA, nor among DA and non-DA neurons (Fig. 1g–i and Extended Data Fig. 3d,e). Yet, we found that DA neurons in the lVTA showed largely homogeneous activity patterns that were characterized by transient responses at CS+ onset and reward delivery, reminiscent of RPE-like responses[29,30] (cluster 2). These cells also responded to cue onset during both 'Go' trials (that is, reward-seeking behavior after CS+) and 'No-Go' trials (that is, no reward-seeking behavior after CS+). In contrast, mVTA DA neurons were more heterogeneous and included a subset of neurons that evinced sustained firing (that is, increased activity during motivated behavior in the interval between trial onset and reward delivery; cluster 3). Cells in cluster 1 showed small responses, which did not seem to encode any specific task-related events. Cells in cluster 4 showed sustained inhibition that occurred during reward-seeking behavior.

Fig. 1 | Four distinct activity patterns describe the neural activity of VTA DA and non-DA neurons during behavior. a, Experimental design. **b**, Left, sample recording site (scale bar 0.5 mm). Right, recording location of two DA cells in different VTA subregions (unit 'A' in lVTA (blue), unit 'B' in mVTA (orange)). **c**, Spontaneous and evoked waveforms of the two simultaneously recorded cells shown in **b**. **d**, Reward-seeking task. Top, all CS+ trials (sorted by interval length) of a sample session (green, licks during reward seeking; red, licks during reward delivery; blue, licks during reward consumption). Bottom, task structure. **e**, Peri-event histograms of the neural activity pattern of units 'A' and 'B' shown in **b** during behavior. Interval lengths are color-coded and grouped together in bins of 300 ms. **f**, Normalized (auROC) activity patterns of each VTA cell ($n = 186$ cells, $n = 6$ mice) are sorted based on cluster identity and task component. Right, mean normalized activity patterns of the four clusters during behavior color-coded by interval length. Vertical blue bar denotes DA neurons. Bottom, mean normalized activity responses during different task components for the four clusters. **g**, Anatomical location of opto-tagged VTA DA neurons (color-coded based on cluster identity; black dashed line indicates putative border between mVTA and lVTA). **h**,**i**, Percentage of opto-tagged DA neurons in the four clusters (C1–C4) based on location in lVTA (**h**) or mVTA (**i**). **j**, Schematic of antidromic opto-tagging approach. DA neurons projecting to NAcMed or NAcLat can be identified through antidromic activation of axons in these regions. This involves applying 5 ms blue light pulses to initiate an antidromic action potential (AP, denoted by an arrow), combined with simultaneous recordings in the VTA. Inset, bar graph showing mean delay of light response ($n = 19$ cells and $n = 7$ mice). **k**, Sample of collision test. Top, collision trials in which a spontaneous AP occurred in the 15 ms leading up to the time point of the mean light response. Note how the presence of spontaneous (that is, presumably orthodromic) APs blocks the light-evoked AP. Bottom, trials in which no spontaneous AP occurred in the 15 ms leading up to the acquisition of the light response (that is, 'free' trials). **l**, Normalized (auROC) responses of all NAcLat-projecting DA neurons ($n = 5$ cells) during cue onset (left) and reward delivery (right). Note the homogeneous response patterns across different cells. **m**, Same as in **l** but for NAcMed-projecting DA neurons ($n = 14$ cells). All data are represented as mean ± s.e.m. (shading). NPX, Neuropixels.

https://doi.org/10.1038/s41593-023-01547-6







Previous studies have shown that mVTA DA neurons project to NAc medial shell (NAcMed), whereas lVTA DA neurons project to NAc lateral shell (NAcLat)[22,27,31]. To examine activity patterns in projection-defined DA neurons, we leveraged antidromic-based opto-tagging to selectively record from NAcLat- or NAcMed-projecting VTA DA neurons (Fig. 1j,k). We found that NAcLat-projecting neurons show large transients at CS+ onset and reward delivery, whereas NAcMed-projecting DA neurons are more heterogeneous and include both nonresponding neurons as well as neurons that show sustained activation between CS+ onset and reward delivery (Fig. 1l,m; NAcMed: $n = 14$ cells, $n = 5$ mice and NAcLat: $n = 5$ cells, $n = 3$ mice). Together, these results suggest that DA neurons in the lVTA, including a subtype of DA cells projecting to NAcLat, show predominantly, but not exclusively, transient (that is, RPE-like) response patterns. In contrast, the activity pattern of DA neurons located toward more ventro-medial VTA subregions, which includes a subtype of DA neurons projecting to NAcMed, is more heterogeneous in that these cells are less likely to exhibit RPE-like response patterns but more frequently evince sustained activity patterns during reward-seeking behavior.

VTA DA cell body activity resembles DA release patterns

Next, we tested whether the differences in cue-evoked electrical activity—transient or sustained—of single mVTA or lVTA DA neurons could also be captured at the population level. We infused Cre-dependent GCaMP6m in the VTA of DAT-mice ($n = 5$ mice) and implanted an optical fiber in different VTA subregions (Fig. 2a,b). Four weeks later, mice were subjected to the same reward-seeking task as described above. We found that cue-evoked responses between mVTA and lVTA DA populations displayed sustained or transient patterns, respectively (Fig. 2c–f), reminiscent of the activity patterns we had observed in these VTA subregions in our electrophysiological recordings (Fig. 1). Similarly, transient patterns were most prominent in the most lateral aspect of the VTA and became more sustained when recording locations moved to more medio-ventral VTA subregions suggesting the presence of a gradient rather than categorical differences in neural activity patterns between mVTA and lVTA DA neurons (Fig. 2b,c).

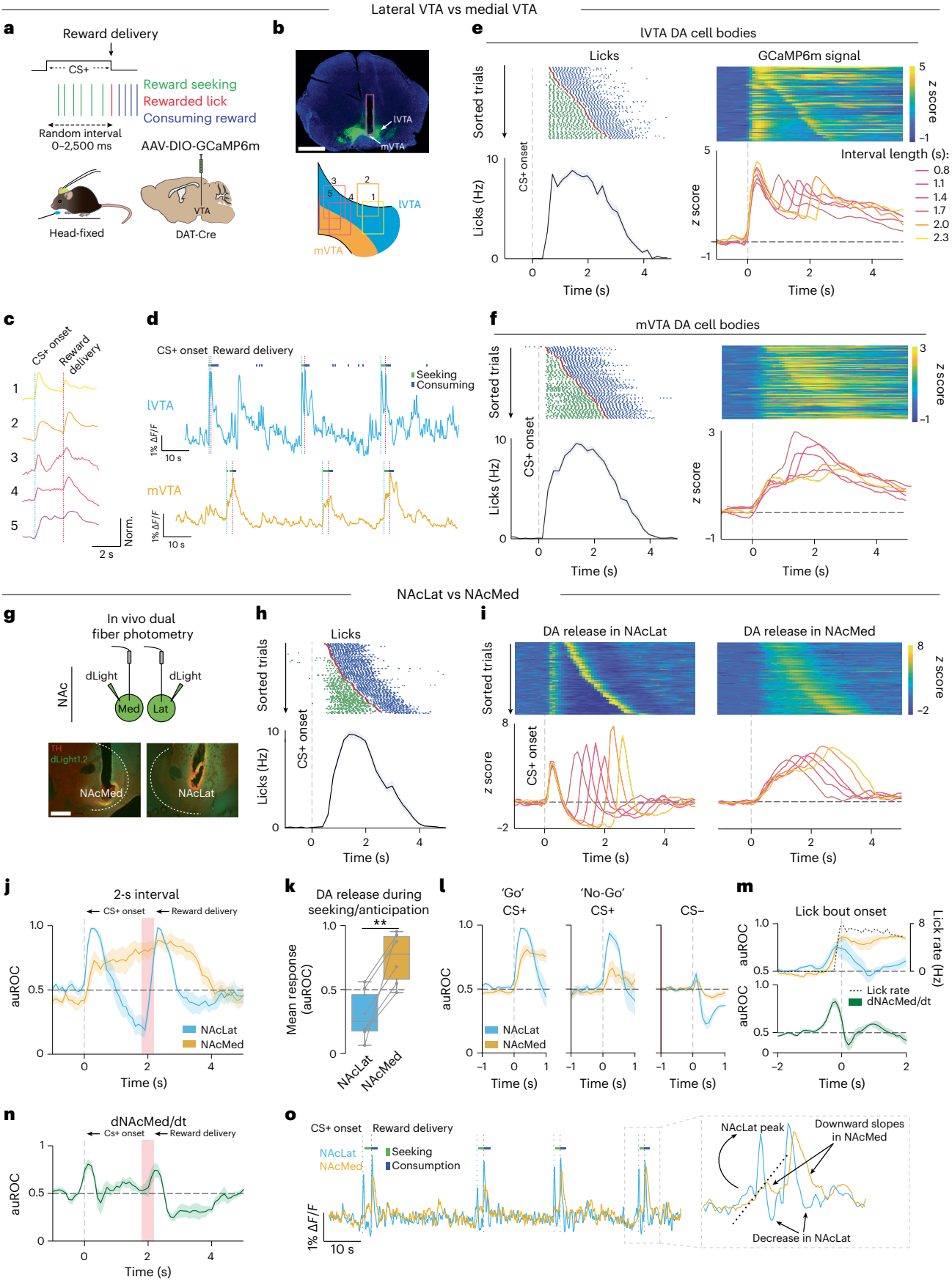
We then asked whether the temporal differences in cue-evoked calcium signals—transient versus sustained—in mVTA and lVTA DA neurons translate into similar differences of DA release in the corresponding axonal target regions in the NAc. To do this, we infused the DA sensor dLight1.2 into the NAcMed and NAcLat of C57BL/6 mice ($n = 6$ mice) and implanted optical fibers in these regions (Fig. 2g). Four weeks later, we performed simultaneous dLight fiber photometry recordings in the NAcMed and NAcLat during reward-seeking behavior (Fig. 2h). Similar to the somatic calcium responses, we observed transient DA release events in the NAcLat and sustained increase in DA release in the NAcMed (Fig. 2i–k), reminiscent of the 'ramping' responses observed by others[10,11,17,21,32–34]. Taken together, separate mesoaccumbal DA pathways display transient and sustained activity patterns in parallel during reward seeking, but we did not find compelling evidence for a dissociation of DA cell body activity and DA release as reported previously[10]. Thus, DA cell activity may directly translate to DA release within anatomically defined mesoaccumbal subsystems.

In our fiber photometry recordings, we observed sustained activity in the interval between cue onset and reward delivery in the mVTA and NAcMed (Fig. 2), but only a subset of mVTA DA cells showed these activity patterns in our electrophysiological recordings (Fig. 1). To test the hypothesis that this subset of VTA DA cell body activity is sufficient to mediate the DA release patterns that we observed in the NAcMed, we performed an experiment in which we optogenetically stimulated VTA DA neurons while measuring DA release in different NAc subregions. Specifically, we infused an adeno-associated virus (AAV)-carrying Cre-dependent ChR2-mCherry into the VTA of DAT-Cre mice ($n = 6$ mice) and dLight into the NAcMed and NAcLat of the same animals. Optical fibers were implanted in the VTA, NAcMed and NAcLat (Extended Data Fig. 4a). Four weeks later, we optogenetically stimulated VTA DA cell bodies using 1 ms light pulses and simultaneously recorded DA release in the NAcMed and NAcLat (Extended Data Fig. 4b). A 1 ms light pulse evoked a single AP (Extended Data Fig. 2). We then used the NAcLat and NAcMed DA release kernels that we obtained in this way to convolve the single-unit recordings of the largest clusters or responding cells in the lVTA (cluster 2) and mVTA (cluster 3; Extended Data Fig. 4c). As a result, we could predict DA release patterns in the NAcLat and NAcMed, which largely resembled the actual in vivo DA responses that were measured in these regions (Extended Data Fig. 4d–i). Thus, neural activity of VTA DA cell body activity, as observed during a reward-seeking task (Fig. 1), is sufficient to explain NAc DA release during that same task (Fig. 2), which includes sustained and transient DA release patterns in the NAcMed and NAcLat, respectively.

A previous study suggested that DA neural activity is dissociated from DA release[10]. In that study, DA release was recorded in the NAc core and DA cell body activity was recorded in the lVTA using a different behavioral assay and different species. To study DA release in the NAc core in our reward-seeking task, we obtained additional fiber photometry recordings targeting the NAc core. We found mixed activity patterns in the NAc core, which included sustained and transient activity patterns that largely depended on the precise recording location within the NAc core (that is, medial or lateral to the anterior commissure, respectively; Extended Data Fig. 5). Gaussian interpolation revealed that DA transients in the NAc during reward-seeking behavior can be best described by gradient with DA release in the most ventro-medial NAc showing sustained activity patterns and DA release in the NAcLat showing transient activity patterns (Extended Data Fig. 5d). We wondered if this discrepancy could be due to species differences (that is, rat versus mouse), but retrograde tracing experiments confirmed that the topographic organization of VTA DA neurons is conserved with

Fig. 2 | VTA DA cell body activity resembles DA release within defined mesoaccumbal subsystems. a, Experimental design; same task as in Fig. 1. **b**, Top, GCaMP6m (green) expression and recording site in mVTA (pink outline indicates recording location 4; scale bar 1 mm). Bottom, recording locations (labeled 1–5) in different VTA subregions for all recorded animals ($n = 5$ mice). **c**, Normalized photometry signals obtained during the 2.2–2.4 s interval between CS+ onset and reward delivery at five different VTA recording sites (numbers and colors correspond to **b**). Norm., normalized. **d**, lVTA DA cells, transient 'peaks' at CS+ onset and reward delivery; mVTA DA cells, sustained (ramping-like) activity patterns. **e**, Left, behavioral responses. Right, calcium transients of lVTA DA neurons. Top, individual trials (sorted by interval length). Bottom, mean plots grouped by interval length; different colors represent different interval lengths (bin width: 0.3 s). **f**, Same as in **e**, but for mVTA DA cells. **g**, Top, simultaneous recordings in NAcMed and NAcLat. Bottom, dLight expression and recording sites in NAcLat and NAcMed (scale bar 200 µm, representative example from $n = 6$ mice). **h**,**i**, As in **c** and **f**, but for DA release measurements in NAcLat and NAcMed. **j**, Group level analysis of the auROC-normalized 1.9–2.2 s (highlighted in red) interval for NAcLat and NAcMed ($n = 6$ mice, auROC calculated over one session (200 trials), one auROC trace for each mouse was included). **k**, Box plot (median and quantiles) showing significant differences in mean DA response between NAcMed and NAcLat during 1-s interval before reward delivery. Significance was calculated by means of two-sided paired t test; $^{**}P = 0.004$, $n = 6$ mice. **l**, dLight response in NAcLat and NAcMed to CS onset in 'Go' trials (left, $n = 6$ mice), 'No-Go' trials (middle, $n = 5$ mice) and following CS onset ($n = 6$ mice). **m**, Top, peri-event response to lick-bout onset (-500 ms after CS+ onset). Dashed line indicates lick rate. Bottom, rate-of-change of the NAcMed response during lick-bout onset. **n**, Same as in **j**, but for the rate-of-change response in NAcMed ($n = 6$ mice). **o**, Sample simultaneous recordings in NAcLat and NAcMed (green, reward-seeking licks and blue, reward consumption licks). Inset, single-trial response highlighting how the slope of the NAcMed trace is reflected in the NAcLat trace (black dotted line indicates slope in NAcMed DA release). All data are represented as mean ± s.e.m. (error bars or shading).

https://doi.org/10.1038/s41593-023-01547-6



https://doi.org/10.1038/s41593-023-01547-6

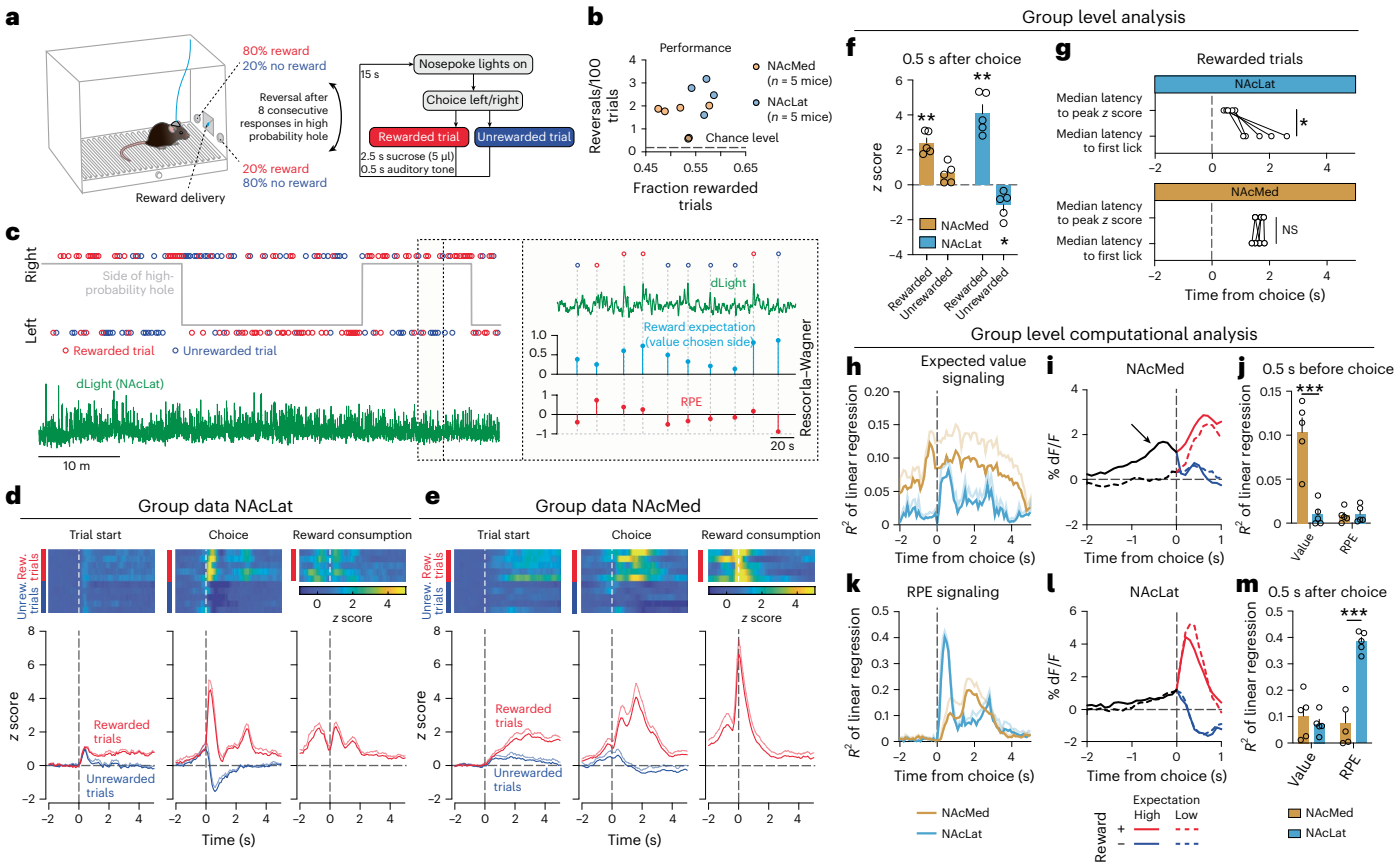


Fig. 3 | Separate mesoaccumbal DA pathways encode value and RPEs. **a**, Two-armed bandit task. **b**, Performance of individual mice (after training). **c**, Sample dLight (green) recording in NAcLat. Reward expectation (blue) and RPE (red) were extracted from trial-by-trial data using a Q learning model fitting procedure (see Extended Data Fig. 7 for model selection). **d**, Top, peri-stimulus time histograms of individual *z*-score-normalized DA release dynamics during the three main phases of the task. Bottom, mean response of all mice showing differential DA dynamics in NAcLat during rewarded (Rew.) and unrewarded (Unrew.) trials. **e**, Same as **d**, but for DA release in NAcMed. **f**, Mean dLight *z* score 0.5 s after choice (that is, after nosepoke). Circles indicate individual animals ($n = 5$ mice per group). Significance was calculated by means of two-tailed Student's *t* test different from 0; *P* values for the four comparisons in order of appearance: 0.0011 (**), 0.0582, 0.0014 (**), 0.0276 (*). **g**, Median latency until peak dLight *z* score, and median latency until first lick in rewarded trials. NAcLat, DA peak is observed before the animals made first lick; NAcMed, DA peak coincides with first lick. Circles indicate individual animals ($n = 5$ mice per group). Significance was calculated by means of Student's paired *t* test; $^{*}P = 0.0296$ for NAcLat, $P = 0.5527$ for NAcMed. **h**, Regression plots around the time of choice showing the fraction of trial-to-trial variance of dLight signal that could be explained by trial-to-trial modeled reward expectation. **i**, Mean NAcMed DA release during trials with high or low reward expectation highlighting (arrow) the correlation depicted in **h**. In trials with high reward expectations, NAcMed DA release is increased in the interval preceding the side-in response. **j**, Quantification of the data shown in **i** and **h**. Trial-by-trial variation in NAcMed DA release tracks reward expectation in the interval before choice ($n = 5$ mice per group). Asterisks indicate significant Sidak's post hoc test after a significant value component × brain area interaction effect in a two-way repeated measures (RM) ANOVA; $^{***}P < 0.0001$ for value, $P = 0.9931$ for RPE. **k**, Regression plots around the time of choice showing the fraction of trial-to-trial variance of the dLight signal that could be explained by the trial-to-trial modeled RPE. **l**, Mean NAcLat DA release highlighting the correlation depicted in panel **k**. NAcLat DA release tracks both positive and negative RPE immediately after the choice outcome is made known to the animal. **m**, Quantification of data shown in **k** ($n = 5$ mice per group). Asterisks indicate significant Sidak's post hoc test after a significant value component × brain area interaction effect in a two-way RM ANOVA; $P = 0.7318$ for value, $^{***}P < 0.0001$ for RPE. All data are represented as mean ± s.e.m. (error bars or shading). NS, not significant.

mVTA DA neurons mainly projecting to NAcMed and NAc core and lVTA DA neurons projecting to NAcLat in both rodent species (Extended Data Fig. 6). Thus, when taking the anatomical organization of VTA DA neurons into consideration, we did not find compelling evidence for a dissociation of DA cell body activity and release during behavior. Taken together, DA neural activity and release are topographically organized along a dorso-lateral to ventro-medial gradient, both at the cell body and terminal level.

Separate DA pathways encode reward expectation and RPE

Previous studies proposed that sustained (that is, ramping-like) DA release tracks between-trial variations in reward expectation[10,11]. To address this, we recorded DA release in the NAcLat and NAcMed in freely behaving mice during a two-armed bandit task[35,36]. Specifically, mice could make responses into two holes (a 'side-in' response), one of which delivered a sucrose reward with a high probability (80%), and the other one with a low probability (20%; Fig. 3a,b). Each time the animal made eight consecutive responses in the high-probability hole, a reversal in reinforcement contingencies occurred, such that the previously low-probability hole became the high-probability hole and vice versa. We estimated the trial-by-trial reward expectation and RPE values using the Rescorla–Wagner model (Fig. 3c and Extended Data Fig. 7a)[37]. In the NAcLat, we observed sharp peaks in DA release during reward delivery and a reduction in DA release during reward omission, which was significantly reduced in the NAcMed (Fig. 3d–g and Extended Data Fig. 8). Rather, NAcMed DA peaked during the first lick of reward consumption (Fig. 3g). A correlational analysis between the latent parameters of the Rescorla–Wagner model (that is, reward

expectation and RPE) and trial-by-trial DA release revealed a modest correlation between NAcMed DA release and the expected value of a choice before the outcome was known to the mouse (that is, in the 0.5 s before the choice; Fig. 3h–j). This correlation was highly specific to NAcMed DA release and not observed in the NAcLat (Fig. 3j). In contrast, NAcLat DA release did not differ between trials in which the mouse had a high or a low reward expectation but instead encoded both positive and negative RPEs robustly immediately after the choice (Fig. 3k–m). To further validate the apparent lack of negative RPE signaling in NAcMed DA release, we performed an additional experiment. Specifically, we trained head-fixed mice on a reward-conditioning task, in which a CS+ indicated the availability of a reward (that is, 1% sucrose). The reward was delivered after a single lick response (Extended Data Fig. 8). Trained mice rarely failed to make this single response and thus had a high reward expectation (93 ± 2% successful trials). During forced omission trials, in which no reward was delivered following the presentation of the CS+, we observed strong encoding of negative RPE in the NAcLat. In the NAcMed, DA release returned to the baseline following omission but did not dip below baseline. These findings collectively indicate a clear distinction in the encoding of DA signaling in the two brain regions, in which NAcMed DA release predominantly encodes reward expectation. Conversely, NAcLat DA release primarily signals the temporal derivative of reward expectation (that is, RPE), representing the 'error' or 'differential' signal over the underlying reward expectation.

Modeling DA activity as state and rate-of-change encoding

So far, we identified a subset of mVTA DA neurons that show sustained activity patterns both at the level of individual cells and at the axon terminal level in the NAcMed during motivated behavior (Figs. 1 and 2). We found that sustained NAcMed DA release tracks a state value (primarily related to reward expectation) as it changes between trials and NAcLat DA release tracks the error (or differential) of state value (Fig. 3). Consistent with this is the observation that NAcLat DA release patterns share similarities with the rate-of-change of NAcMed DA release (Fig. 2l–o and Extended Data Fig. 9). However, it has been proposed that the role of NAc DA release is not limited to encoding scalar state value[38,39]. For instance, DA release is increased in the ventro-medial NAc in response to aversive events and cues that predict them[23,40]. Additionally, DA release in the NAc seems to be involved in the encoding of novelty or general salience[41]. Based on the role of DA in such a wide range of domains and essential behaviors, we next aimed to develop a generalized model that incorporates both state value and TD RPE encoding in the DA system. In this model, DA activity across distinct mesoaccumbal pathways can be conceptualized as state and rate-of-change encoding (also known as proportional–differential encoding), analogous to what is observed in sensory systems[1–4].

Specifically, we developed a simple two-step computational model that involves two latent variables (Fig. 4a,b). The first latent variable refers to the 'state', which represents a generalized form of state value. In our model, this 'state' is linearly linked to the task-event kernels (CS+ presentation, reward seeking, reward delivery and reward consumption) through a set of 21 parameters (five for each input and one bias parameter). Of the four input parameters to the state variable, two relate to motivated behavior (reward seeking and consumption), whereas the other two relate to experimenter-controlled events (CS+ presentation and reward delivery). Notably, unlike the typical meaning of 'state value', this 'state' variable goes beyond representing reward expectation.

The second latent variable is the temporal derivative of the 'state', which we calculated at intervals of 100 ms. We fit these two latent variables to *z*-score-normalized neural activity patterns using a linear equation, with a bias parameter '*B*' indicating whether a cell encodes the state or its temporal derivative. Parameter '*A*' controls the amplitude of the best fit (Fig. 4a, inset).

Our model produced state and $\Delta state/\Delta t$ curves (Fig. 4c), which captured neural activity patterns during different aspects of the behavior (Fig. 4d). We found that the four different clusters encoded the two latent variables to a different extent. In line with our intuition from the actual activity of these neurons (Fig. 1), the model identified cluster 2 cells as those encoding $\Delta state/\Delta t$, as indicated by a low value of the bias parameter B. In contrast, clusters 3 and 4 cells were more likely to encode the state itself. The low amplitude A and low R^2 values of cluster 1 cells suggest that these cells did not robustly encode any specific aspects of task behavior (Fig. 4e). Notably, mVTA DA neurons were significantly more likely to encode state than lVTA DA neurons, but this did not depend on the smaller amplitude parameter of mVTA DA neurons. In contrast, lVTA DA neurons were mainly (and relatively homogeneously) encoding $\Delta state/\Delta t$ (Fig. 4f–i). Thus, our model demonstrates that sustained and transient activity patterns of mVTA and lVTA DA neurons during reward seeking can be conceptualized as parallel encoding of the latent variable 'state' and an associated rate-of-change, respectively.

Discussion

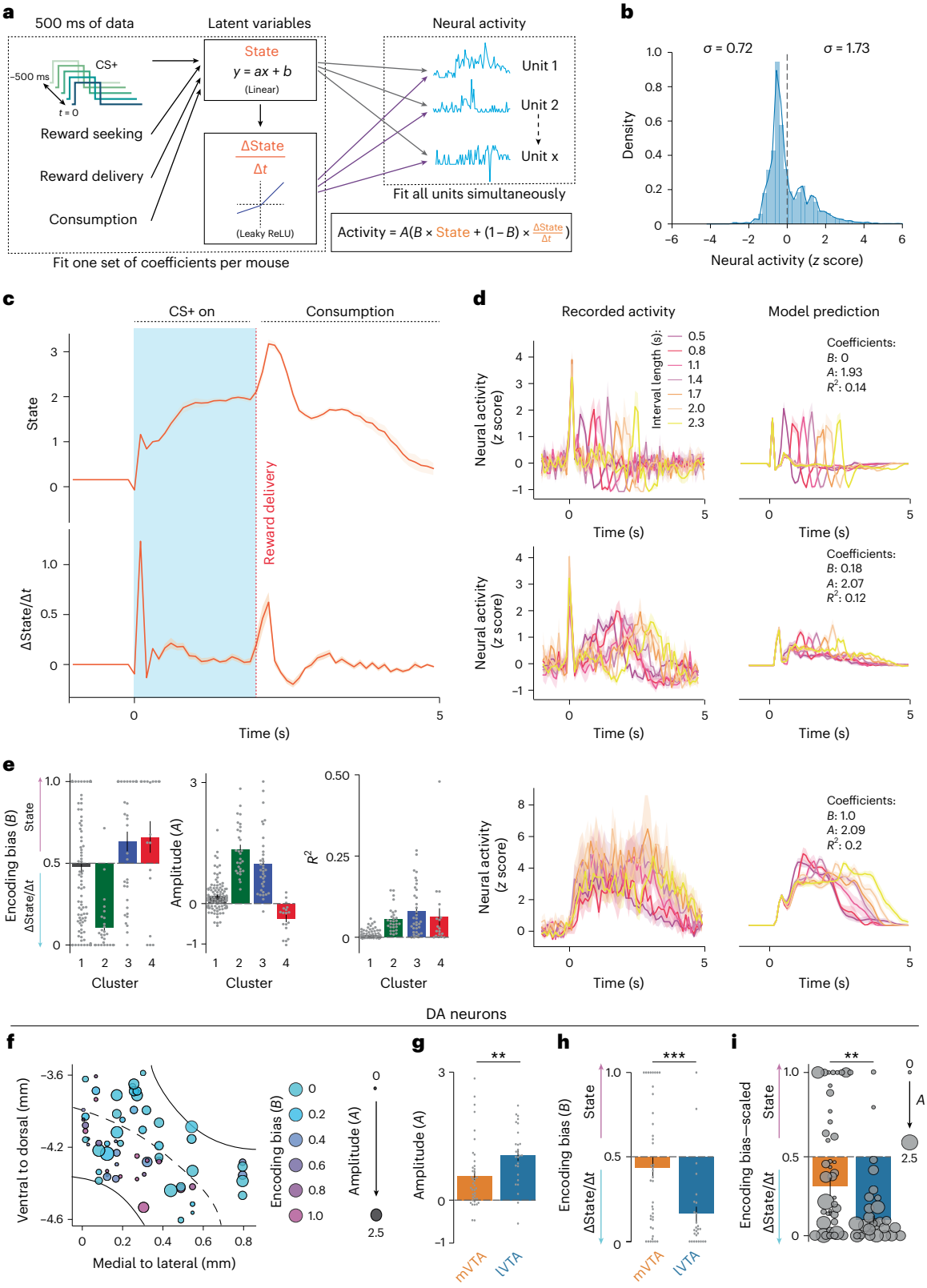
In this study, we did not find evidence that motivation-related DA release is dissociated from the activity of VTA DA neurons when the exact anatomy of the mesoaccumbal DA system is taken into consideration. Thus, our results provide an alternative framework that differs from previous models, which have posited that spike-independent modulation of DA release drives motivated behaviors[10]. Although we did not find evidence for behaviorally relevant DA release independent of the neural activity of VTA DA cell bodies within anatomically defined DA projection systems, there is evidence that local regulation of DA release in the striatum can occur independently of the activity of DA neurons[42,43]. Whether these local DA release mechanisms directly contribute to reward learning and motivated behaviors will be an important future research topic. Furthermore, we found that VTA DA neurons exhibit parallel encoding of a state and a temporal

Fig. 4 | State and rate-of-change are encoded by different VTA DA subpopulations. a, Schematic representation of the model. The task events of the last 500 ms were fitted to a one-dimensional state vector using linear regression. This state vector and its temporal derivative were fitted to a set of simultaneously recorded neurons using a formula with two free parameters: '*A*' controlled the amplitude of the predicted activity pattern and '*B*' controlled the bias of the best fit toward the state vector or its temporal derivative (the encoding bias). **b**, Differential distribution of *z*-score-normalized neural activity above and below the pretrial baseline suggests implementation of Leaky ReLU in our model. **c**, Sample showing state and $\Delta state/\Delta t$ traces averaged over all trials with an interval of 2 ± 0.15 s. **d**, Recorded activity (left) and model predictions (right) of three representative neurons with varying *B* coefficients and similar values for the *A* coefficient, as well as the coefficient of determination (R^2). **e**, Individual cells (gray dots) and means for each cluster. Left, bias (*B*) toward encoding either state or $\Delta state/\Delta t$ value. Middle, model amplitude (*A*). Right, R^2 values as a measure of the quality of fit. Cluster 1, $n = 100$; cluster 2, $n = 33$; cluster 3, $n = 35$ and cluster 4, $n = 18$ cells. **f**, Anatomical location of DA neurons with different color shades denoting '*B*' values; marker size refers to '*A*' values. Black dashed line indicates putative border between mVTA and lVTA. **g**, The model fits significantly lower amplitude parameters to mVTA neurons. Significance was calculated by means of two-sided *t* test; $^{**}P = 0.0095$. **h**, mVTA DA neurons were significantly more likely to encode the state compared to lVTA DA neurons. mVTA DA neurons include cells that are (i) biased in either direction or (ii) show mixed responses. Significance was calculated by means of two-sided Welch's *t* test; $^{***}P = 0.0008$. **i**, Same as in **h**, but individual data points were linearly weighted for amplitude (that is, their contribution to the mean, s.e.m. and statistical test scaled with the *A* parameter, which prevents cells with a small amplitude from confounding the analysis). Significance was calculated by means of (weighted) two-sided Welch's *t* test; $^{**}P = 0.0098$. All data are represented as mean ± s.e.m. (shading or error bars).

https://doi.org/10.1038/s41593-023-01547-6

derivative of value (for example, RPE) during motivated behavior. The utilization of proportional–differential encoding, involving the representation of a state and its derivative, has an extensive prevalence in sensory processing across both the central and peripheral nervous systems[1–4]. Hence, it is reasonable to postulate that the brain would adopt a comparable mechanism for encoding higher-order variables, such as state value.

Our finding that a subset of DA neurons track the animal's state, whereas other DA neurons encode both the derivative and the state itself is reminiscent of the 'fractional derivative' approach to understanding the computations within DA neurons[17]. However, our data suggest that there are a surprisingly larger number of VTA DA cells that do not exclusively encode a derivative signal. Additionally, our finding that NAcMed, but not NAcLat, DA release lacks a negative RPE



makes it unlikely that NAcMed DA release tracks a derivative-like signal (Fig. 3 and Extended Data Fig. 8). Although DA cells that do not appear to encode derivative-like signals are predominantly located in the mVTA, it is unlikely that DA neurons can be classified in categorical, mutually exclusive groups that can be distinguished solely based on their anatomical location in VTA. Rather, there seems to be a dorso-lateral to medio-ventral gradient of responses among VTA DA neurons. Further teasing apart VTA DA neuron heterogeneity along this gradient will be an important future research direction.

Previous studies, including our own work, have reported increased NAcMed DA release in response to salient environmental or aversive stimuli[23,40], which seems inconsistent with the idea that VTA DA neurons use proportional–differential encoding to simultaneously track reward expectation and its temporal derivative (Fig. 3). However, it is possible that DA signaling, particularly in the NAcMed, contains additional information, such as general salience, aversion, novelty or movement[38,39]. We hypothesize that the state variable encoded by mVTA DA neurons projecting to NAcMed is high-dimensional and encodes various aspects of the animal's motivational state, possibly including movement or action dynamics[18] as well as the increase in DA release following aversive stimuli[23,40]. To further test this hypothesis, we performed an experiment in which we recorded DA release in the NAcLat and NAcMed using dLight1.3 in head-fixed mice in response to unpredictable electrical tail shocks (Extended Data Fig. 10). We found that tail shocks caused a continuous increase in NAcMed DA release, whereas NAcLat DA release exhibited a sharp increase at the onset of the shock, which followed by a sustained decrease throughout the rest of the shock. Although there was some variability (for example, in the delay in the NAcMed DA response) between the hypothesized and experimental data, we argue that it is conceivable that DA release dynamics encode aversive events in terms of state and temporal derivative. While additional work is required to conclusively answer this question, our results open new research avenues about the nature of the state vector encoded in mVTA DA neurons and in NAcMed DA release and how it contributes to behavior in higher-dimensional scenarios.

Finally, the generalized conceptualization of state and rate-of-change encoding is not in conflict with current theories of the DA function in reinforcement learning[38,44–47]. There is mounting evidence that prediction errors reflect model-based reinforcement learning instead of a change in a scalar (one-dimensional) value signal[48,49], and it has been speculated that DA release may track such a high-dimensional signal[38,39]. Refining our understanding of the specific state variable encoded by mVTA DA neurons and describing variations in activity patterns and DA release across different mesoaccumbal DA pathways in the context of state and rate-of-change encoding could inspire a more unifying framework governing DA's role in reward learning and motivated behavior. This notion extends well beyond the DA system as it reflects a fundamental principle of the function of neural circuits across the peripheral and central nervous systems.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41593-023-01547-6.

References

1. Młynarski, W. F. & Hermundstad, A. M. Efficient and adaptive sensory codes. *Nat. Neurosci.* **24**, 998–1009 (2021).
2. Kim, A. J., Lazar, A. A. & Slutskiy, Y. B. Projection neurons in *Drosophila* antennal lobes signal the acceleration of odor concentrations. *eLife* **4**, e06651 (2015).
3. Luo, J., Shen, W. L. & Montell, C. TRPA1 mediates sensation of the rate of temperature change in *Drosophila* larvae. *Nat. Neurosci.* **20**, 34–41 (2017).
4. Adrian, E. D. & Zotterman, Y. The impulses produced by sensory nerve endings: part 3. Impulses set up by touch and pressure. *J. Physiol.* **61**, 465–483 (1926).
5. Schultz, W. Dopamine reward prediction-error signalling: a two-component response. *Nat. Rev. Neurosci.* **17**, 183–195 (2016).
6. Gershman, S. J. & Uchida, N. Believing in dopamine. *Nat. Rev. Neurosci.* **20**, 703–714 (2019).
7. Watabe-Uchida, M., Eshel, N. & Uchida, N. Neural circuitry of reward prediction error. *Annu. Rev. Neurosci.* **40**, 373–394 (2017).
8. Schultz, W. & Dickinson, A. Neuronal coding of prediction errors. *Annu. Rev. Neurosci.* **23**, 473–500 (2000).
9. Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science* **275**, 1593–1599 (1997).
10. Mohebi, A. et al. Dissociable dopamine dynamics for learning and motivation. *Nature* **570**, 65–70 (2019).
11. Hamid, A. A. et al. Mesolimbic dopamine signals the value of work. *Nat. Neurosci.* **19**, 117–126 (2016).
12. Wise, R. A. Dopamine, learning and motivation. *Nat. Rev. Neurosci.* **5**, 483–494 (2004).
13. Salamone, J. D. & Correa, M. The mysterious motivational functions of mesolimbic dopamine. *Neuron* **76**, 470–485 (2012).
14. Sun, F. et al. A genetically encoded fluorescent sensor enables rapid and specific detection of dopamine in flies, fish, and mice. *Cell* **174**, 481–496 (2018).
15. Tsutsui-Kimura, I. et al. Distinct temporal difference error signals in dopamine axons in three regions of the striatum in a decision-making task. *eLife* **9**, e62390 (2020).
16. Berke, J. D. What does dopamine mean? *Nat. Neurosci.* **21**, 787–793 (2018).
17. Kim, H. R. et al. A unified framework for dopamine signals across timescales. *Cell* **183**, 1600–1616 (2020).
18. Coddington, L. T. & Dudman, J. T. Learning from action: reconsidering movement signaling in midbrain dopamine neuron activity. *Neuron* **104**, 63–77 (2019).
19. Dayan, P. & Berridge, K. C. Model-based and model-free Pavlovian reward learning: revaluation, revision, and revelation. *Cogn. Affect. Behav. Neurosci.* **14**, 473–492 (2014).
20. Gardner, M. P. H., Schoenbaum, G. & Gershman, S. J. Rethinking dopamine as generalized prediction error. *Proc. Biol. Sci.* **285**, 20181645 (2018).
21. Howe, M. W., Tierney, P. L., Sandberg, S. G., Phillips, P. E. M. & Graybiel, A. M. Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. *Nature* **500**, 575–579 (2013).
22. Lammel, S. et al. Unique properties of mesoprefrontal neurons within a dual mesocorticolimbic dopamine system. *Neuron* **57**, 760–773 (2008).
23. De Jong, J. W. et al. A neural circuit mechanism for encoding aversive stimuli in the mesolimbic dopamine system. *Neuron* **101**, 133–151 (2019).
24. Engelhard, B. et al. Specialized coding of sensory, motor and cognitive variables in VTA dopamine neurons. *Nature* **570**, 509–513 (2019).
25. Poulin, J.-F., Gaertner, Z., Moreno-Ramos, O. A. & Awatramani, R. Classification of midbrain dopamine neurons using single-cell gene expression profiling approaches. *Trends Neurosci.* **43**, 155–169 (2020).
26. Ekstrand, M. I. et al. Molecular profiling of neurons based on connectivity. *Cell* **157**, 1230–1242 (2014).
27. De Jong, J. W., Fraser, K. M. & Lammel, S. Mesoaccumbal dopamine heterogeneity: what do dopamine firing and release have to do with it?. *Annu. Rev. Neurosci.* **45**, 109–129 (2022).

28. Steinmetz, N. A., Zatka-Haas, P., Carandini, M. & Harris, K. D. Distributed coding of choice, action and engagement across the mouse brain. *Nature* **576**, 266–273 (2019).
29. Cohen, J. Y., Haesler, S., Vong, L., Lowell, B. B. & Uchida, N. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature* **482**, 85–88 (2012).
30. Starkweather, C. K., Babayan, B. M., Uchida, N. & Gershman, S. J. Dopamine reward prediction errors reflect hidden-state inference across time. *Nat. Neurosci.* **20**, 581–589 (2017).
31. Beier, K. T. et al. Circuit architecture of VTA dopamine neurons revealed by systematic input-output mapping. *Cell* **162**, 622–634 (2015).
32. Wassum, K. M., Ostlund, S. B. & Maidment, N. T. Phasic mesolimbic dopamine signaling precedes and predicts performance of a self-initiated action sequence task. *Biol. Psychiatry* **71**, 846–854 (2012).
33. Phillips, P. E. M., Stuber, G. D., Heien, M. L. A. V., Wightman, R. M. & Carelli, R. M. Subsecond dopamine release promotes cocaine seeking. *Nature* **422**, 614–618 (2003).
34. Roitman, M. F. Dopamine operates as a subsecond modulator of food seeking. *J. Neurosci.* **24**, 1265–1271 (2004).
35. Verharen, J. P. H., Kentrop, J., Vanderschuren, L. J. M. J. & Adan, R. A. H. Reinforcement learning across the rat estrous cycle. *Psychoneuroendocrinology* **100**, 27–31 (2019).
36. Verharen, J. P. H., Adan, R. A. H. & Vanderschuren, L. J. M. J. Differential contributions of striatal dopamine D1 and D2 receptors to component processes of value-based decision making. *Neuropsychopharmacology* **44**, 2195–2204 (2019).
37. Rescorla, R. A. 'Configural' conditioning in discrete-trial bar pressing. *J. Comp. Physiol. Psychol.* **79**, 307–317 (1972).
38. Langdon, A. J., Sharpe, M. J., Schoenbaum, G. & Niv, Y. Model-based predictions for dopamine. *Curr. Opin. Neurobiol.* **49**, 1–7 (2018).
39. Cox, J. & Witten, I. B. Striatal circuits for reward learning and decision-making. *Nat. Rev. Neurosci.* **20**, 482–494 (2019).
40. Badrinarayan, A. et al. Aversive stimuli differentially modulate real-time dopamine transmission dynamics within the nucleus accumbens core and shell. *J. Neurosci.* **32**, 15779–15790 (2012).
41. Kutlu, M. G. et al. Dopamine signaling in the nucleus accumbens core mediates latent inhibition. *Nat. Neurosci.* **25**, 1071–1081 (2022).
42. Threlfell, S. et al. Striatal dopamine release is triggered by synchronized activity in cholinergic interneurons. *Neuron* **75**, 58–64 (2012).
43. Liu, C. et al. An action potential initiation mechanism in distal axons for the control of dopamine release. *Science* **375**, 1378–1385 (2022).
44. Syed, E. C. J. et al. Action initiation shapes mesolimbic dopamine encoding of future rewards. *Nat. Neurosci.* **19**, 34–36 (2016).
45. Soares, S., Atallah, B. V. & Paton, J. J. Midbrain dopamine neurons control judgment of time. *Science* **354**, 1273–1277 (2016).
46. Sadacca, B. F., Jones, J. L. & Schoenbaum, G. Midbrain dopamine neurons compute inferred and cached value prediction errors in a common framework. *eLife* **5**, e13665 (2016).
47. Niv, Y., Daw, N. D., Joel, D. & Dayan, P. Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl.)* **191**, 507–520 (2007).
48. Niv, Y. & Schoenbaum, G. Dialogues on prediction errors. *Trends Cogn. Sci.* **12**, 265–272 (2008).
49. Takahashi, Y. K. et al. Dopamine neurons respond to errors in the prediction of sensory features of expected rewards. *Neuron* **95**, 1395–1405 (2017).

Methods

Subjects

The following mouse lines (20–35 g, 8–20 weeks old, males and females) were used for experiments: C57BL/6J mice (Jackson Laboratory, stock: 000664), DAT-Cre (Jackson Laboratory, stock: 006660; strain code: B6.SJL-Slc6a3tm1.1(cre)Bkmn/J), DAT-Flp (Jackson Laboratory, stock: 035426; strain code: STOCK Slc6a3em1(flpo)Hbat/J) and Ai32 (Jackson Laboratory, stock: 012569; crossed to DAT-Cre mice). Furthermore, we used Long-Evans rats (Charles River, 500–800 g, 8–17 months old, male, strain code: 0006). Mice and rats were maintained on a 12-h light/12-h dark cycle (lights on at 07:00) with food ad libitum and room temperature of 22–25 °C and 55% humidity. All procedures complied with the animal care standards set forth by the National Institutes of Health and were approved by the University of California Berkeley's Administrative Panel on Laboratory Animal Care.

Stereotaxic surgeries

Surgeries were performed under general ketamine–dexmedetomidine anesthesia using a stereotaxic instrument (Kopf Instruments, model 1900). For retrograde tracing of mice (Extended Data Fig. 6a–d), animals were injected unilaterally with red fluorescent retrobeads (100 nl, LumaFluor) into the NAcMed (bregma, 1.0; lateral, 0.75 mm and ventral, −4.8 mm), NAc core (bregma, 1.0 mm; lateral, 1.0 mm and ventral, −4.5 mm) or NAcLat (bregma, 1.0 mm; lateral, 1.9 mm and ventral, −4.3 mm) using a 1 µl Hamilton syringe (Hamilton). For retrograde tracing of rats (Extended Data Figs. 6e–g), 4% FluoroGold (50 nl, Santa Cruz Biotechnology) and red fluorescent retrobeads (LumaFluor) were injected into the NAcLat (bregma, 2.0 mm; lateral, 2.8 mm and ventral, −7.1 mm; 200 nl retrobeads) and NAc core (bregma, 2.0 mm; lateral, 1.5 mm and ventral, −7.1 mm; 100 nl FluoroGold) of the same animals. The AAVs were from Addgene (AAV5-hSyn-dLight1.2, AAV5-hSyn-dLight1.3b, AAV1-CAG-FLEXFRT-ChR2(H134R)-mCherry), the UNC Vector Core (AAV5-DJ-EF1a-DIO-GCaMP6m, AAV5-hSyn-hChR2(H134R)-mCherry, AAV5-flex.taCasP3-TEVp). Canine adenovirus type 2 (CAV2-Cre) was from Plateforme de Vectorologie de Montpellier. In total, 400–500 nl of concentrated (~10^{12} infectious units per ml) AAV solution was injected into the NAc (NAcMed—bregma, 1.0 mm; lateral, ±0.75 mm and ventral, −4.8 mm; NAcLat—bregma, 1.0 mm; lateral, ±1.9 mm and ventral, −4.3 mm; NAc core—bregma, 1.0 mm; lateral, ±1.0 mm and ventral, −4.5 mm) or VTA (bregma, −3.4 mm; lateral, ±0.3 mm and ventral, −4.3 mm) using a syringe pump (Harvard Apparatus) at 150 nl min^{-1}. The injection needle was withdrawn 10 min after the end of the infusion.

For in vivo fiber photometry experiments (Figs. 2 and 3 and Extended Data Figs. 4, 5 and 7–10), mice were implanted unilaterally with an optical fiber (400 µm, numerical aperture (NA) = 0.48; Doric Lenses) targeted to the mVTA (bregma, −3.4 mm; lateral, 0.3 mm and ventral, −4.4 mm), lVTA (bregma, −3.4 mm; lateral, 0.8 mm and ventral, −4.1 mm) or NAc core (bregma, 1.0 mm; lateral, ±1.0 mm and ventral, −4.2 mm) or dual optical fibers in the NAcMed (bregma, 1.0 mm; lateral, ±0.75 mm and ventral, −4.6 mm) and NAcLat (bregma, 1.0 mm; lateral, ±1.9 mm and ventral, −4.1 mm; left and right side counterbalanced between animals) of the same animal.

When optogenetics in combination with fiber photometry was performed (Extended Data Fig. 4a–i), mice were implanted with bilateral optical fibers (200 µm, NA = 0.22; Newdoon) at 15°, which were directed toward the VTA. The tip of the fiber was ~0.5 mm above the dorsal extend of the VTA (bregma, −3.4 mm; lateral, ±0.8 mm and ventral, −3.8 mm). For in vivo electrophysiology experiments (Fig. 1), unilateral implantations at 15° toward the VTA were performed (coordinates same as above). A small craniotomy was prepared to allow access for the Neuropixels probes.

During all surgeries for in vivo electrophysiological or fiber photometry experiments, a headplate was attached to the skull to allow head fixation. One layer of adhesive cement (C&B Metabond; Parkell) was followed by acrylic (Jet Denture Repair; Lang Dental) to secure the fiber and the headplate to the skull. The incision was closed with a suture and tissue adhesive (Vetbond; 3M). The animals were kept on a heating pad until they recovered from anesthesia. Experiments were performed for 4–8 weeks (for AAVs) or 7 d (for retrograde tracing) after stereotactic injection. Injection sites and optical fiber placements were confirmed in all animals by preparing coronal sections (50 or 100 µm) of injection and implantation sites. We routinely carried out complete serial reconstruction of the injection sites and optical fiber placements.

Anatomical nomenclature

NAc. The NAc is an integral part of the ventral striatum and has been traditionally subdivided into shell and core substructures. The NAc shell is shaped like a crescent and is located medial and ventral to the core. The differences between shell and core have been defined by various histochemical (for example, calbindin, calretinin and substance P), electrophysiological, connectivity and functional criteria[50–52]. In this study and in previous studies[22,23,31], we describe an additional ventral striatal subregion, termed 'NAcLat', which is located laterally to the NAc core. We realize that the terminology 'NAcLat' may be misleading because it may suggest that the 'NAcLat' is part of the NAc shell, although these regions may be anatomically and functionally different. Nevertheless, we use the terminology 'NAc lateral shell' (termed NAcLat) as it refers to an anatomical region that is defined in the 'The Mouse Brain in Stereotaxic Coordinates' of ref. 53 and 'The Rat Brain in Stereotaxic Coordinates' of ref. 54 (mouse, bregma (1.34 mm to 0.74 mm); rat, bregma (2.16 mm to 0.84 mm)). We did not detect obvious differences in the topographic organization of the mesoaccumbal DA system between mice and rats (Extended Data Fig. 6). Using nomenclature based on widely established anatomical atlases allows other researchers to precisely target the same brain region to reproduce our results. Nevertheless, whether this region is more appropriately labeled NAcLat, ventrolateral striatum or lateral NAc remains a semantic question. More importantly, the NAcLat is the most prominent projection target of DA neurons that are located in the lateral part of the parabrachial pigmented nucleus, which is typically termed 'lVTA'[22,23,31,51].

VTA. We defined the mVTA as the region comprising the paranigral nucleus and interfascicular nucleus, whereas the lVTA was defined as the medio-dorsal and lateral parabrachial pigmented nucleus and the medial lemniscus region adjacent to the substantia nigra. Please note that the definition of mVTA versus lVTA is largely based on the anatomical location of projection-defined VTA DA neurons[22]. It is not strictly based only on the medio-lateral axis, but also incorporates the dorso-ventral axis (see schematic in Fig. 1b and Extended Data Fig. 6).

Behavioral assays

Reward-seeking assay. This refers to the assay shown in Figs. 1 and 2, and Extended Data Figs. 1, 2, 3, 5 and 9. Before the experiment, mice were water restricted (1 ml per day) and habituated (1 h) to a head-fixed setup that was located in a sound-attenuating chamber (55.9 × 38.1 × 40.6 cm, Med Associates). The experimental session consisted of 200 trials (100 CS+ and 100 CS− trials) with a randomly generated intertrial interval of 15–45 s. The conditioned stimulus (CS+) and a control stimulus that was without scheduled consequences (CS−) consisted of either a blue LED light, 11 kHz tone or white noise, counterbalanced between the CS+ and CS− and between different animals. Trial onset was signaled to the mice by the onset of the CS+ followed by a random interval between 0 and 2,500 ms. Licking after the expiration of the random interval resulted in reward delivery (3 µl of 1% sucrose in water). No reward was delivered if the animals did not lick within 500 ms following expiration of the random interval.

Reward omission assay. This refers to the assay shown in Extended Data Fig. 8. The experimental setup was identical to the reward-seeking

task. Water-restricted (1 ml per day) mice (n = 9 mice) were trained on a head-fixed reward-seeking task in which the presence of a 1-s CS+ indicated the availability of a reward (3 µl of a 1% sucrose solution). The reward was delivered immediately following the first lick during the CS+ presentation to promote high reward expectations. After training, mice were subjected to a session in which reward delivery was omitted in 10% of the trials. During this session, DA release in NAcMed and NAcLat was recorded simultaneously.

Two-armed bandit task. This refers to the paradigm shown in Fig. 3 and Extended Data Fig. 7. Mice were placed in an operant chamber (17.8 × 12.7 × 21.6 cm, Med Associates) within a sound-attenuating chamber (55.9 × 38.1 × 40.6 cm, Med Associates). One side of the chamber contained two illuminated nose-poke holes on the left and right side, a reward port in between the two holes (delivering 5 µl of a 10% sucrose solution) and on the other side a house light and a tone generator. The behavioral protocol we used was adapted from refs. 35,36. Animals received ten training sessions before the experimental fiber photometry recordings. Water-restricted animals could choose between two nose-poke holes that differed in the probability of reward delivery. At the start of the session, one nose-poke hole was randomly assigned as high probability for reward delivery (80%), while the other one was designated as low probability for reward delivery (20%). If the animal made eight consecutive responses at the high-probability nose-poke hole, reward contingencies reversed, such that the high-probability hole became the low-probability hole and vice versa. The reversal was not signaled to the animal, so it had to infer this change in reward contingencies by tracking the value of both choice options. A session started with the illumination of the house light and the illumination of the lights in the two nose-poke holes. A response in one of the nose-poke holes (side-in) turned off the lights in the two holes, and, in rewarded trials, led to reward delivery without a delay. During a rewarded trial, an auditory tone was played for 0.5 s, and a reward was delivered over the course of 2.5 s. During an unrewarded trial, both nose-poke lights turned off, but this was not accompanied by a tone or by reward delivery. Regardless of the trial outcome, a new trial was initiated 15 s after a side-in response from the animal, which was signaled to the animal by the illumination of both nose-poke holes. The house light remained illuminated during the entire length of the session, which lasted for 1 h. During the experimental recording session, animals were attached to the fiber photometry setup via a fiber-optic rotary joint (pigtailed with 400 µm, NA = 0.57 fiber; Doric Lenses), which allowed the animals to behave freely in the operant chamber.

Tail shock assay. This refers to the assay shown in Extended Data Fig. 10. Mice (n = 12 mice) were habituated to the same head-fixed setup that was used for the reward-seeking assay. Shocks (0.4 mA, 5 ms pulse, 20 Hz for 2 s) were delivered to the tail of the mouse using pregelled electrodes (Sonic Technology's adhesive pads for transcutaneous electrical nerve stimulation). Shocks were generated using a stimulator (SYS-A320, WPI). Ten shocks were delivered at a random interval between 15 and 35 s. Shock delivery occurred unpredicted and was not signaled to the mice.

Electrophysiology

Recording sessions were performed in mice (orthodromic n = 6 DAT-Cre × Ai32 mice and antidromic n = 8 DAT-Cre × Ai32 mice), which were trained in the reward-seeking assay for approximately 2 weeks. A Neuropixels probe was inserted into the brain, which was followed by a 20-min waiting period to allow settlement of the probe. Data were acquired using Open Ephys and the Neuropixels plugin (https://open-ephys.org/neuropixels) or SpikeGLX (https://billkarsh.github.io/SpikeGLX) and processed using Kilosort 3.0 (https://github.com/cortex-lab/kilosort) followed by manual cleaning using Phy2 (https://github.com/cortex-lab/phy) and/or post-kilosort (PKS) (https://github.com/handejong/post-kilosort).

The orthodromic opto-tagging sequence consisted of a ten-pulse train of 1 ms (470 nm) light pulses delivered every 3 s for 2 min. One millisecond light pulses were used because they caused a clear separation in time between the light pulse and the evoked APs, which allowed us to remove light artifacts from the dataset (Extended Data Fig. 2b–f). This is important because Neuropixels probes are light sensitive, and the light artifact at the pulse onset and offset would severely complicate clustering. The antidromic opto-tagging sequence consisted of one 5 ms pulse delivered at 1 Hz. There were no light artifacts during antidromic opto-tagging experiments.

Clean units (as automatically identified by Kilosort and manually verified using Phy2 and/or PKS) were considered opto-tagged if they showed an increase in firing frequency in the 6 ms (orthodromic) or 30 ms (antidromic) interval following light stimulation that was higher than the 99.9th percentile level of the highest peak obtained when we shuffled the unit response in the 0–100 ms interval 10,000 times[55]. In addition, we compared the waveform of the units in the 0–5 ms (orthodromic) or 5–30 ms (antidromic) interval to the spontaneous waveforms in the −90 to 0 ms interval (that is, before light pulse onset) and only included units when the correlation (Pearson's R) between spontaneous and evoked waveforms was >0.9. Finally, all units that fired less than one AP every ten light pulses (that is, firing fidelity of <0.1) were excluded even if they met all other tagging criteria. There was an additional inclusion criterion for the antidromic opto-tagging procedure. Units had to pass a 'collision test', which means that they showed a significant decrease in firing fidelity in response to a light pulse on trials where there was a spontaneous AP in the 15 ms interval before the light response pulse. This assumes that spontaneous APs travel in orthodromic direction and will collide with antidromic APs if they occur before the light-evoked antidromic AP has reached the cell body. Each mouse was recorded seven times along a grid pattern through the craniotomy. Before the last recording session, the Neuropixels probe was dipped in CM-DiI (Thermo Fischer Scientific, V22888) to allow for histological verification of the probe implantation coordinates. The recording sites of the first six recordings were calculated in relationship to the relative implantation coordinates to the last recording location. We never found opto-tagged units outside of the VTA, and the channel layout of the Neuropixels probes allowed us to determine the ventro-dorsal location of the recorded units (Figs. 1g and 4f). Recorded units were assigned to mVTA or lVTA solely based on anatomical criteria, and investigators were blind to the activity patterns of the recorded cells. Nevertheless, there may be some ambiguity in determining the precise recording location, in particular for units that were recorded at the border between mVTA and lVTA. Although Neuropixels probes allowed us to record far more than n = 186 VTA cells, we only included cells that were simultaneously recorded on the same or neighboring electrodes as the opto-tagged cells. Thus, only VTA cells and not cells from adjacent brain regions were included in our analysis.

Fiber photometry

Calcium and dLight transients were recorded using a custom-built fiber photometry rig as described previously[23,56]. Fluorescence signals were obtained by stimulating cells expressing GCaMP6m (Fig. 2a–f), dLight1.2 (Figs. 2g–o and 3 and Extended Data Fig. 4, 5, 8 and 9) or dLight (Extended Data Fig. 10), with a 470 nm LED (20 µW at fiber tip). In freely behaving mice (Fig. 3), isobestic signals were obtained by stimulating these cells with a 405 nm LED (20 µW at fiber tip). LED light (470 nm and 405 nm) was alternated at 20 Hz, and light emission was recorded using a complementary metal oxide semiconductor camera camera (Photometrics Prime or Point Grey Blackfly), which acquired video frames containing the fiber bundle (two fibers, 1 m in length, NA = 0.48, 400 µm core; Doric Lenses) at the same frequency. Video

frames were analyzed online, and fluorescent signals were acquired using custom acquisition code written in MATLAB. Experimental time stamps were acquired using transistor–transistor logic pulses generated by an Arduino or the Med Associates chamber. The fluorescent signal obtained after stimulation with 405 nm light was used to correct movement artifacts as follows. First, the 405 nm signal was fitted to the 470 nm signal using the first and second coefficients of the polynomial that was the best fit (least squares) to the 470 nm signal. The fitted 405 nm signal was then subtracted from the 470 nm signal to obtain the movement and bleaching-corrected signal.

For combined optogenetics and fiber photometry experiments (Extended Data Fig. 4), 100 (1 ms and 470 nm) laser pulses were delivered to the VTA at 15-s intervals. Laser light power was set to 5 mW at the tip of the optical fiber.

Histology, immunofluorescence and microscopy

Immunofluorescence and microscopy were performed as described previously[23]. After intracardial perfusion with 4% paraformaldehyde in PBS, pH 7.4, brains were postfixed overnight and coronal brain sections (50 or 100 µm) were prepared. We used the following primary antibodies: mouse anti-TH (1:1,000, Millipore) and chicken anti-GFP (1:1,000, Abcam). Brain sections were incubated overnight in a primary antibody solution. We used the following secondary antibodies: Alexa Fluor 647 goat anti-mouse (1:750, Thermo Fisher Scientific) and Alexa Fluor 477 goat anti-chicken (1:750, Abcam). Brain sections were incubated for 2 h in a secondary antibody solution. Image acquisition was performed on a Zeiss LSM710 laser scanning confocal microscope using ×20 or ×40 objectives and on a Zeiss AxioImager M2 upright widefield fluorescence/differential interference contrast microscope with charge-coupled device camera using ×5 and ×10 objectives. Zen Software 2.3 (Zeiss) was used for acquiring confocal and epifluorescence images. Images were analyzed using ImageJ. Sections were labeled relative to bregma using landmarks and neuroanatomical nomenclature as described in 'The Mouse Brain in Stereotaxic Coordinates' in ref. 53.

Computational modeling and statistical analysis

Clustering analysis. This refers to data shown in Fig. 1. Neural activity was binned into 100 ms bins and normalized using the auROC approach[29]. auROC has the advantage that it normalizes neural activity in that the resulting trace is not just a measure of the signal amplitude compared to baseline variation (for example, as in *z* score) but also takes intertrial variation into consideration. An additional advantage is that neural activity is normalized to the range of 0 to 1, which facilitates group analysis. As a baseline, we used the interval from −10 to −2 s before trial onset. We used sparce PCA ($\alpha = 1$) and *k*-means clustering to identify four clusters. The rationale for clustering with $k = 4$ is based on sparce PCA embedding (Extended Data Fig. 3a). Clustering with $k = 3$ will generally group cluster 1 and cluster 2 neurons together. However, ignoring the existence of small or nonresponding neurons (cluster 1) may not reflect the dataset accurately. Clustering with $k = 5$ results in dropping one of the clusters by the optimizer. Hierarchical clustering (in PCA space or on the complete dataset) results in largely similar clusters to those obtained using *k*-means clustering (data not shown). Notably, the main conclusion of our clustering analysis is not dependent on the selection of the clustering method or distance metric.

Computation modeling based on data obtained from Neuropixels recordings. This refers to data shown in Fig. 4. Discrete task events as well as neural activity were binned to 100 ms bins. Neural activity was *z*-score normalized. Task events were coded as '0' or '1' depending on whether the event was occurring at any given time point. Task events were shifted in time for five bins preceding $t = 0$ (that is, we used the last 500 ms of task-related activity). In total, we fitted (linear regression) 21 coefficients (4 × 5 + intercept) to the data to produce the state vector. Notably, when neurons were recorded in the same mouse, the same 21 coefficients were used to produce the state vector (note that mice were recorded multiple times). This is because we assume that the latent variables 'state' and 'Δstate/Δ*t*' are mouse-specific. By forcing the model to explain all neural activity patterns (36 ± 7.8 neurons per mouse), it increases the chance that the model will learn meaningful internal representations based on the behavioral task structure for individual mice. The derivative of the state vector was passed through a leaky rectified linear unit (leaky ReLU), which corrects negative values by a factor of 0.41. This was done because *z*-score-normalized neural activity is not equally distributed below and above the baseline (that is, a neuron with a baseline firing frequency of 5 Hz can fire at 25 Hz, but not at −20 Hz; as such, the dynamic range below the baseline is compacted). In the second layer of the model, the 'state' and 'Δstate/Δ*t*' traces were fitted to each individual neuron using two coefficients (no bias/intercept). The model was fitted using the Python package SciPy[57] based on the Broyden–Fletcher–Goldfarb–Shanno method, which minimizes the mean squared error between the predicted and observed neural activity. All coefficients for one mouse (21 + 2 × number of neurons) were fitted simultaneously. To address overfitting and local minima, the model was fitted using tenfold cross-validation and random initialization parameters, but the model converged on the same global minimum in each case.

Computational modeling based on data obtained from the two-armed bandit task. This refers to data shown in Fig. 3. Reward rate was calculated using a fixed time decay constant for all animals of 3% per second. In addition, for each individual trial, nose-poke values and RPE were calculated according to the Rescorla–Wagner model[37]. At task initiation, values for both nose-pokes holes, $Q_{\text{left},t=1}$ and $Q_{\text{right},t=1}$ were set at 0.5 and updated according to a Q learning rule on each individual trial:

$$Q_{s,t} = Q_{s,t-1} + \alpha \times \delta t$$

Here α is the learning rate, which was set to 0.33 to maximize the trial-by-trial variability in nose-poke values and RPE values, and δ_t is the RPE in trial *t*, given by:

$$\delta_t = \begin{cases} 1 - Q_{s,t-1} & \text{for rewarded trials} \\ 0 - Q_{s,t-1} & \text{for unrewarded trials} \end{cases}$$

Only the value of the chosen nose-poke hole *s* was updated on each individual trial so that the unchosen option retained its previous value. Reward expectation in trial *t*, as shown in Fig. 3b and used in Fig. 3f,g, was defined as the value that the chosen nose-poke hole had before choice (that is, $Q_{s,t-1}$).

Statistical analysis. R-squared and normalized cross-correlations were manually calculated in Python. Student's *t* tests (paired and unpaired) or Welch's *t* test for unequal variances were used to compare two groups. Data distribution was verified to be normal by plotting the Q–Q plot. No statistical methods were used to predetermine sample sizes, but our sample sizes are similar to those reported in previous work[10,17,23]. One-way or two-way repeated measures analysis of variance (ANOVA) was used to determine statistical differences in data with more than two groups. Tukey's post hoc test was applied when ANOVA showed a significant main effect. All statistics were performed using the SciPy[57] package in Python or using JASP[58] (version 0.15). Statistical significance was $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. All data are presented as means ± s.e.m. unless otherwise noted in the figure legends.

Reporting summary

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

Data availability

Source data are provided with this paper. All other data are available upon request from the corresponding author.

Code availability

Custom code used for the processing of raw in vivo electrophysiology or fiber photometry data is available on GitHub at https://github.com/handejong/post-kilosort (electrophysiology) and https://github.com/handejong/Fipster (fiber photometry). Code used for model fitting is available at https://github.com/lammellab/model_fitting.

References

50. Castro, D. C. & Bruchas, M. R. A motivational and neuropeptidergic hub: anatomical and functional diversity within the nucleus accumbens shell. *Neuron* **102**, 529–552 (2019).
51. Ikemoto, S. Dopamine reward circuitry: two projection systems from the ventral midbrain to the nucleus accumbens–olfactory tubercle complex. *Brain Res. Rev.* **56**, 27–78 (2007).
52. Zahm, D. S. & Brog, J. S. On the significance of subterritories in the 'accumbens' part of the rat ventral striatum. *Neuroscience* **50**, 751–767 (1992).
53. Franklin, K. & Paxinos, G. *Paxinos and Franklin's the Mouse Brain in Stereotaxic Coordinates, Compact* 5th edn (Academic Press, 2019).
54. Paxinos, G. & Watson, C. *The Rat Brain in Stereotaxic Coordinates. Sixth Edition* (Academic Press, 2007).
55. Cerniauskas, I. et al. Chronic stress induces activity, synaptic, and transcriptional remodeling of the lateral habenula associated with deficits in motivated behaviors. *Neuron* **104**, 899–915 (2019).
56. Kim, C. K. et al. Simultaneous fast measurement of circuit dynamics at multiple sites across the mammalian brain. *Nat. Methods* **13**, 325–328 (2016).
57. Virtanen, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020)
58. JASP Team. JASP (version 0.16.1). *JASP* https://jasp-stats.org (2022).

Acknowledgements

We thank everyone who provided advice throughout the study and on this manuscript—J. Roeper, R. Malenka, R. Awatramani, D. Dombeck and M. Yartsev. We thank the UNC vector core for AAVs, M. Kleinman and the Foster Lab at UC Berkeley for providing rats and assistance with tracing experiments. We thank J. Mankel for technical assistance with immunohistochemistry experiments. S.L. is a John P. Stock Faculty Fellow, Weill Neurohub Investigator and Rita Allen Scholar. This work was supported by the National Institutes of Health (grants R01-DA042889 and R01-MH123246 to S.L.), the Tobacco-Related Disease Research Program (T32IR5075 to S.L.), the One Mind Foundation (047483 to S.L.), the Weill Neurohub (to S.L.), the Rita Allen Foundation (to S.L.), the McKnight Foundation (to S.L.), the Wayne and Gladys Valley Foundation (to S.L.) and a NARSAD Young Investigator Award (27936 to J.W.d.J.).

Author contributions

Stereotaxic injections were performed by Y.L. and J.W.d.J. Neuropixels recordings were performed by J.W.d.J. Immunohistochemistry was performed by Y.L. Computational modeling was performed by J.W.d.J. and J.P.H.V. Fiber photometry was performed by J.W.d.J., Y.L. and J.P.H.V. Behavior experiments were performed by J.W.d.J., Y.L. and J.P.H.V. Data were analyzed by J.W.d.J. and J.P.H.V. The study was designed by J.W.d.J. and S.L. The manuscript was written by J.W.d.J., K.M.F. and S.L. and edited by all authors.

Competing interests

The authors declare no competing interests.

Additional information

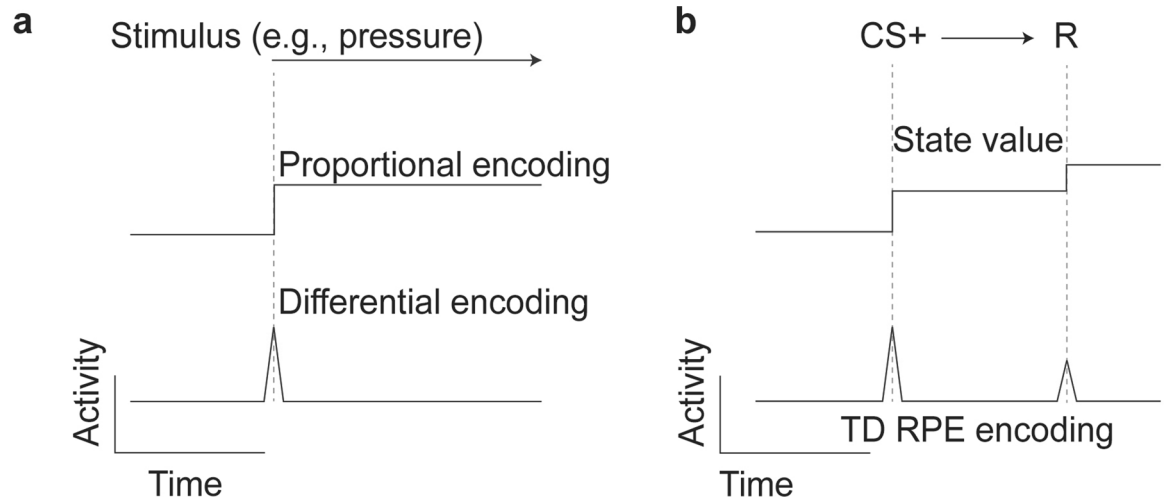
Extended data is available for this paper at https://doi.org/10.1038/s41593-023-01547-6.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41593-023-01547-6.

Correspondence and requests for materials should be addressed to Stephan Lammel.

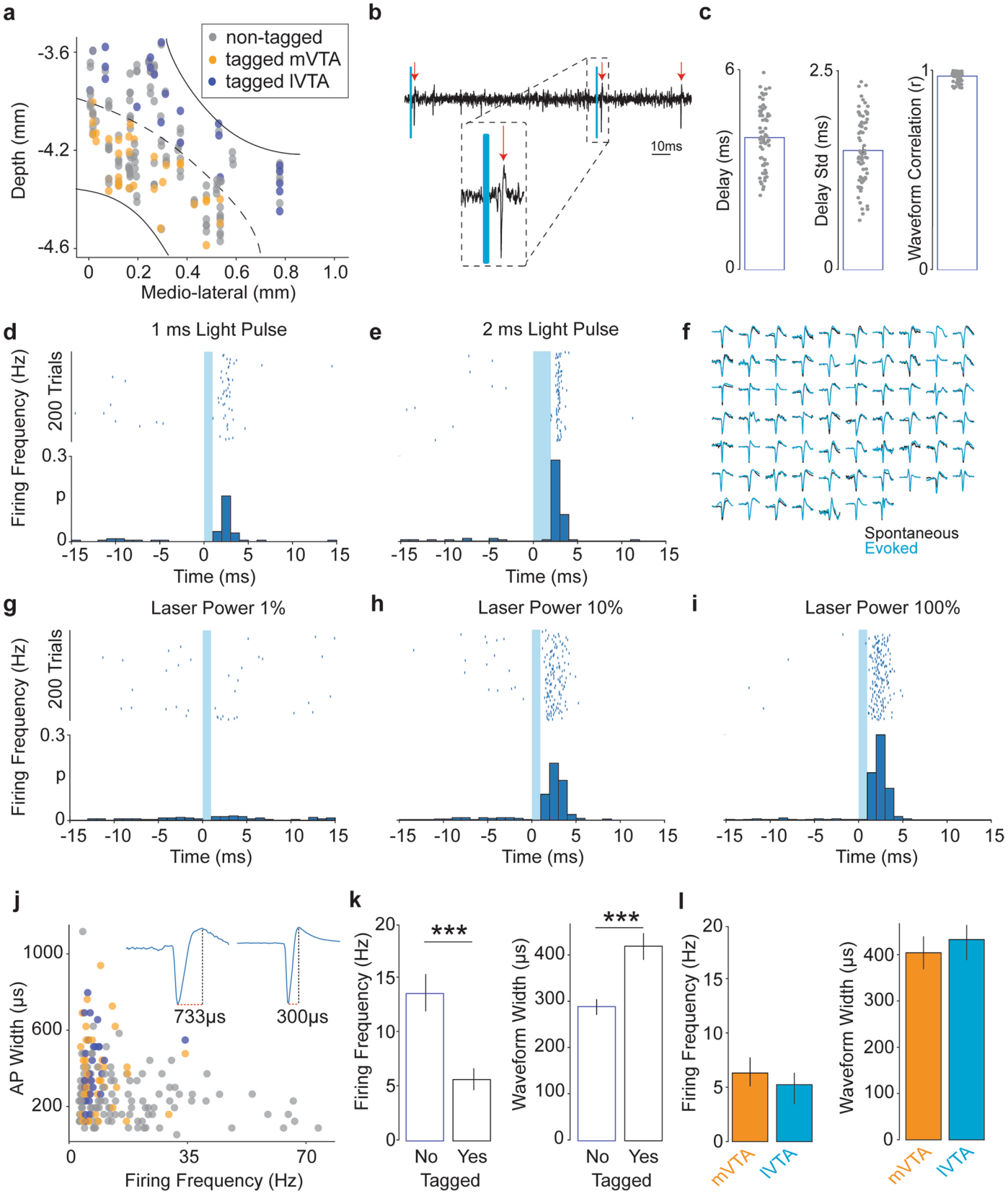
Peer review information *Nature Neuroscience* thanks the anonymous reviewers for their contribution to the peer review of this work.

https://doi.org/10.1038/s41593-023-01547-6



Extended Data Fig. 1 | Proportional-differential encoding in the peripheral and central nervous systems. (a) Schematic showing proportional and rate-of-change encoding of a stimulus (for example, pressure on skin) in the peripheral nervous system. (**b**) Schematic showing state value and rate-of-change encoding in the central nervous system. The state value and its rate-of-change during a Pavlovian conditioning task in which a reward is partially predicted (for example, as in the behavioral task used for experiments in Figs. 1 and 2). Note that in this schematic the 'state value' refers to both expected reward as well as the reward itself.

https://doi.org/10.1038/s41593-023-01547-6

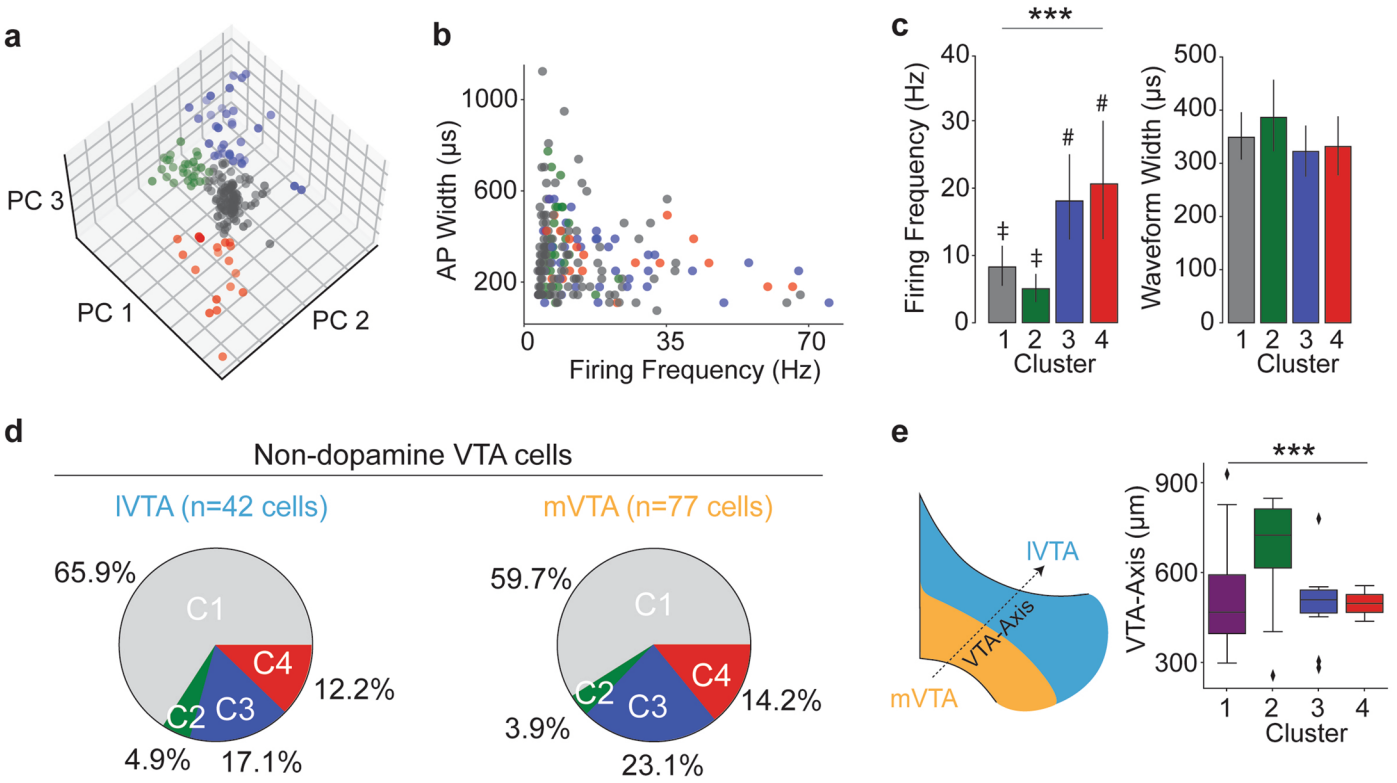


Extended Data Fig. 2 | See next page for caption.

https://doi.org/10.1038/s41593-023-01547-6

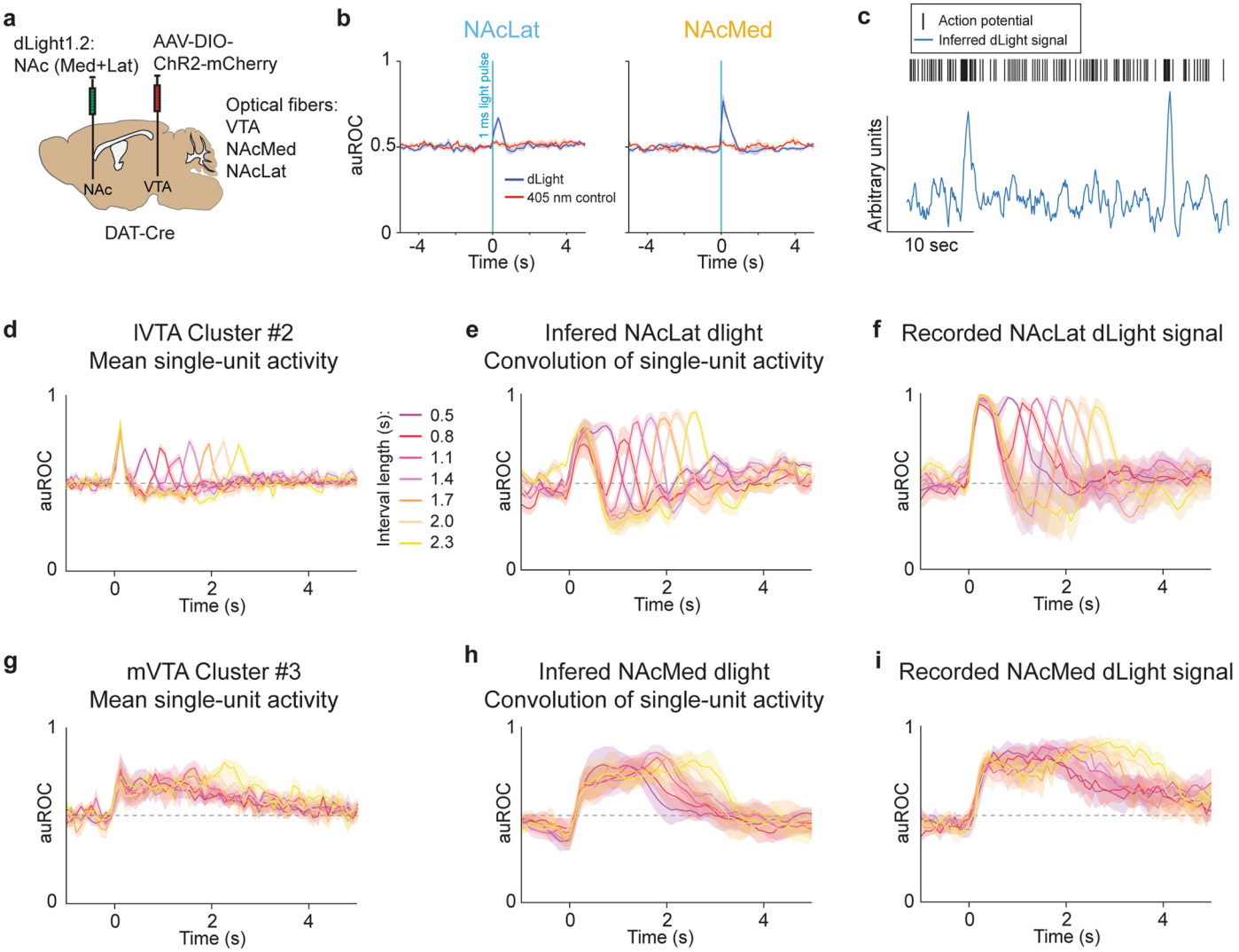
Extended Data Fig. 2 | Neuropixels-based opto-tagging of dopamine neurons. **(a)** Anatomical locations of the recorded VTA cells. Although Neuropixels probes allowed to record more than n = 186 cells (n = 6 mice), we only included cells that were simultaneously recorded on the same or neighboring electrodes as the opto-tagged cells. Thus, only VTA cells and not cells from adjacent brain regions were included. Different colors indicate locations of recorded cells in mVTA or lVTA and whether the cells were identified as dopaminergic ('tagged') or non-dopaminergic ('non-tagged'). **(b)** Representative single-channel recording showing 3 action potentials (APs) of the same unit (red arrows). Two of these APs are induced by a 1 ms light pulse (blue). Inset shows how the AP, and the light pulse are separated in time (scale bar 10 ms). **(c)** Mean delay, mean delay-variation to the first evoked AP, and waveform correlation between evoked and spontaneous APs for n = 67 opto-tagged cells (n = 6 mice). **(d)** Peri-event histogram of an opto-tagged cell following a 1 ms light pulse. 'P' indicates firing fidelity. Total firing fidelity is the sum of the bins between 0 and 5 ms. **(e)** Peri-event histogram of the same cell as in (**d**) but following a 2 ms light pulse, which increased firing fidelity, but at the risk of adulterating the waveforms with a light-offset artifact. **(f)** Evoked (blue) and spontaneous (black) waveforms for all opto-tagged cells. The waveform on the channel with the highest amplitude was chosen. **(g)** Peri-event histogram of an opto-tagged cell following a light pulse at 1% of the max laser power (max laser power: 25 mW at tip of optical fiber at a distance of ~1 mm from the recording site). **(h)** Same unit as in (**g**), but laser power was set to 10% of max. **(i)** Same unit as in (**g**, **h**), but following a 1 ms light pulse with laser at max intensity. Note, increase in firing fidelity when comparing (**g**, **h**, **i**). Firing fidelity was always < 1 (that is, one light pulse (1 ms) did not evoke more than a single AP). **(j)** AP width and baseline firing frequency (FF) for all recorded cells (n = 186 cells). Inset: Sample APs showing calculation of waveform width (that is, difference between AP peak and hyperpolarization peak). **(k)** Mean FF (left) and mean waveform widths for dopamine (tagged; n = 67 cells) and non-dopamine cells (non-tagged; n = 119 cells). Individual data points are presented in (**j**). Significance was calculated by means of two-sided Welch t-test; frequency ***P < 0.001. Opto-tagged cells had a significantly lower baseline FF and a longer waveform width compared to non-opto-tagged cells. **(l)** Mean FF (left) and mean waveform widths for dopamine neurons located in mVTA (n = 40 cells; orange) and lVTA (n = 27 cells; blue). Individual data points are shown in (**j**). Significance was calculated by means of two-sided Welch t-test. All data are represented as mean ± SEM.

https://doi.org/10.1038/s41593-023-01547-6



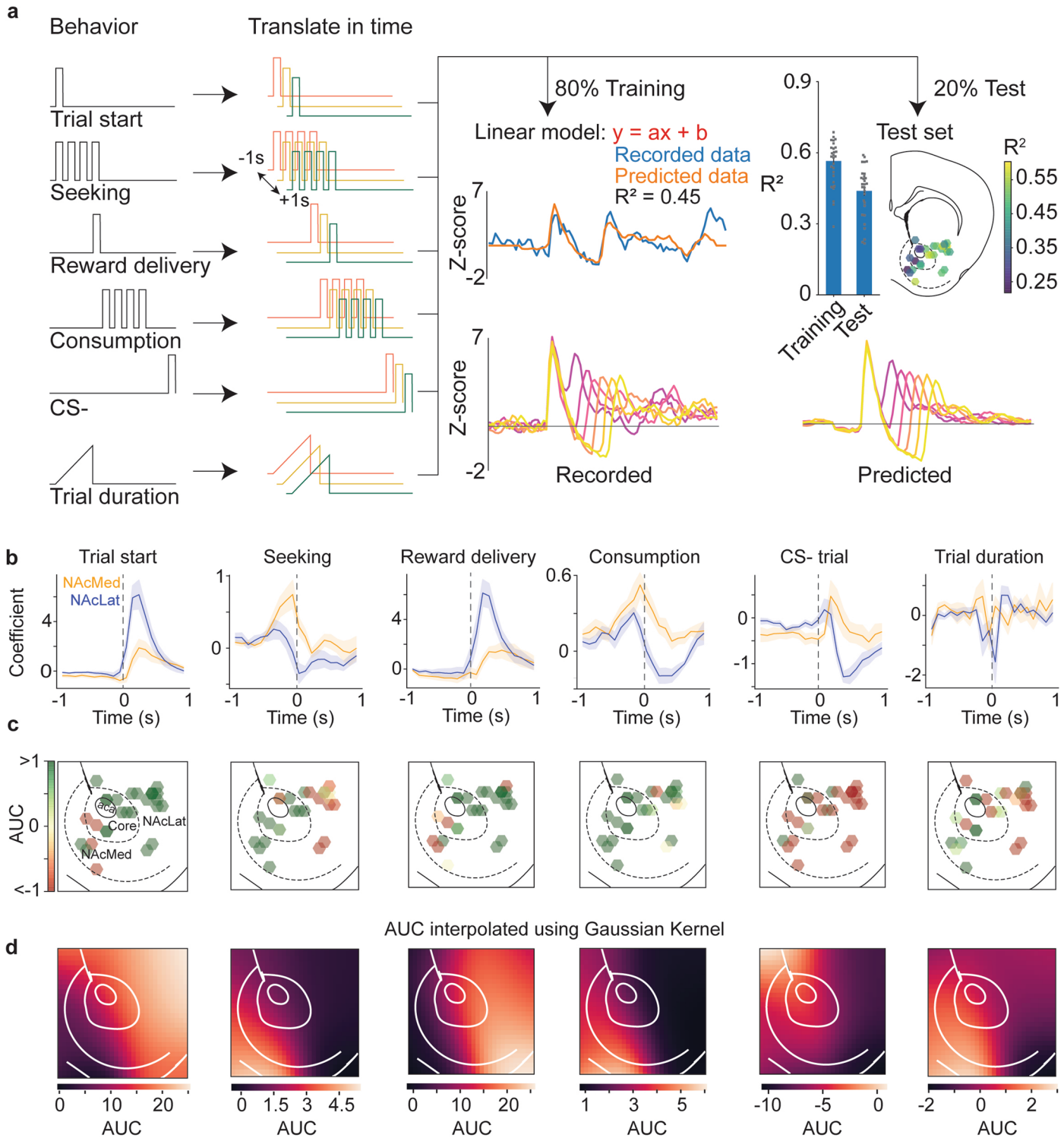
Extended Data Fig. 3 | Four distinct activity patterns describe neural activity of VTA cells during behavior. (a) Sparce PCA embedding of all recorded VTA cells (n = 186 cells; n = 6 mice); color coding indicates that four different activity patterns were obtained using k-means clustering. **(b)** Graph showing action potential (AP) width and baseline firing frequency of all recorded VTA cells (n = 186 cells); color coding indicates cluster identity (cluster 1: gray, cluster 2: green, cluster 3: blue, cluster 4: red). **(c)** Bar graphs showing mean baseline firing frequency (left) and action potential width (right) for all recorded VTA cells (n = 186 cells) based on cluster identity. Significance was calculated by means of 1 way ANOVA. The annotations ‡ and # indicate significant difference in Tukey's *post hoc* testing; ***P < 0.001. Data represented as mean ± SEM. **(d)** Percentage of non-opto-tagged (that is, non-dopaminergic neurons) neurons in lVTA (left) and mVTA (right) based on cluster identity. **(e)** Box plot (median, quantiles and outliers) depicting the anatomical location of all opto-tagged (that is, DA) neurons along a 45° (ventro-medial to dorso-lateral) axis through the VTA. Cluster 2 DA neurons were significantly more likely to be located in dorso-lateral part of the VTA compared to DA neurons in the other clusters. Cluster 1: n = 100 cells, Cluster 2: n = 33 cells, Cluster 3: n = 35 cells, Cluster 4: n = 18 cells. Significance was calculated by means of 1 way ANOVA; ***P < 0.001. Data are represented as median, interquartile range (shading) and total range (error bars).

https://doi.org/10.1038/s41593-023-01547-6



Extended Data Fig. 4 | VTA DA cell body activity is sufficient to explain DA release patterns in distinct NAc subregions. **(a)** AAV-DIO-ChR2-mCherry was injected bilaterally into the VTA of DAT-Cre mice. The same mice were injected with dLight into NAcLat and NAcMed. Optical fibers were implanted dorsal to the VTA, in the NALat and NAcMed (optical fibers in NAcMed and NAcLat were counterbalanced across all animals, n = 6 mice). **(b)** Hundred trials of 1 ms 470 nm light stimulation of VTA DA neurons were used to calculate DA release (auROC) in NAcLat (left) and NAcMed (right). Blue line indicates 1 ms light pulse in VTA. **(c)** Response kernels in (**b**) were used to convolute single-unit recordings. Top: Sample unit, each vertical line denotes a single action potential. Bottom: Corresponding inferred dLight trace when the sequence of action potentials is convoluted with the NAcLat kernel in (**b**). **(d)** Mean normalized (auROC) single-unit recordings of all cluster 2 neurons in the lVTA; color code indicates random interval length. CS+ onset was at t = 0 sec. **(e)** Activity traces of the neurons in (**d**) were convoluted as in (**c**). Trials were normalized (auROC) and averaged over all cells as in (**d**). **(f)** Recorded NAcLat DA release patterns normalized (auROC) and averaged across all mice (n = 6 mice). The 2 sec interval (orange) is the same as in Fig. 2j. Note the qualitative and quantitative similarities with (**e**), including the below-baseline reduction in the interval between CS+ onset and reward delivery. **(g–i)** Same as in (**d–f**), but for NAcMed. Note that this experiment does not unequivocally prove that mVTA cluster 3 single-unit activity is responsible for NAcMed DA release patterns. However, it does suggest that mVTA cluster 3 single-unit activity is sufficient to explain NAcMed DA release patterns. All data are represented as mean ± SEM (shading).

https://doi.org/10.1038/s41593-023-01547-6

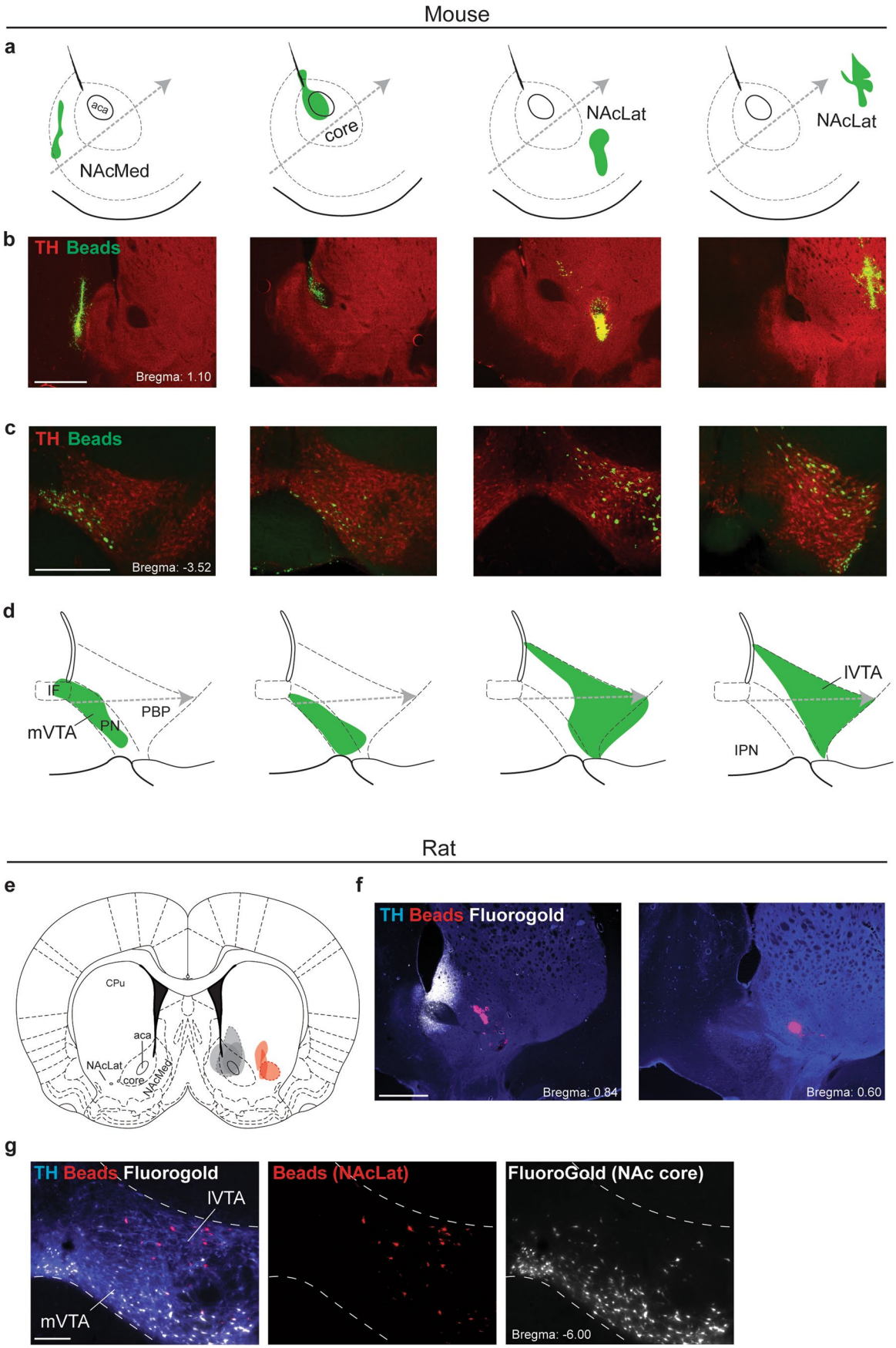


Extended Data Fig. 5 | See next page for caption.

https://doi.org/10.1038/s41593-023-01547-6

Extended Data Fig. 5 | A linear model to predict dopamine transients across all NAc subregions. (a) Each NAc recording was normalized, and a linear regression model was fitted to the discrete task events to predict NAc DA release. Left: Animal behavior was binned to 100 ms and shifted in time (1 sec in both directions, 20 coefficients were fitted for each task event). Task events were encoded as '1' or '0' depending on occurrence in any given time bin except for the trial duration, which was coded from 0 to 1 over the 0-2500 ms interval. Right: Data were split into a training (80%) and test (20%) set. The task events in the training set were fitted to the dLight data using linear regression. The sample trace shows a high overlap between recorded response (blue) and model prediction (orange). Bar graph shows model performance on training and test set (mean ± SEM, n = 29 recordings). Note that the model performed notably better in NAcLat, probably because it performed well at fitting the large transients at CS+ onset and reward delivery. Lower right: Recorded data (left) and model prediction (right); color coded based on trial duration (averaged over all trials). **(b)** Task-event kernels used to fit the behavior to the fiber photometry signal. **(c)** AUCs of the kernels in (**b**), hexagons denote individual recording locations. The color map was selected to highlight the direction (up or down) of the kernels (aca: anterior commissure). **(d)** Because the anatomical location of individual recordings was histologically verified, we used Gaussian interpolation to calculate the task-event kernels for each pixel in a coronal section of the NAc. Note the gradient from NAcMed to NAcLat with NAcMed showing positive AUC for kernels related to reward seeking, reward consumption and trial duration. Conversely, NAcLat kernels were specifically tuned to trial start (CS+ onset) and reward delivery. The NAc core represents an intermediate structure with signals tuned either toward NAcMed or NAcLat based on the position more medial or more lateral to the aca, respectively. All data are represented as mean ± SEM (error bars or shading).

https://doi.org/10.1038/s41593-023-01547-6

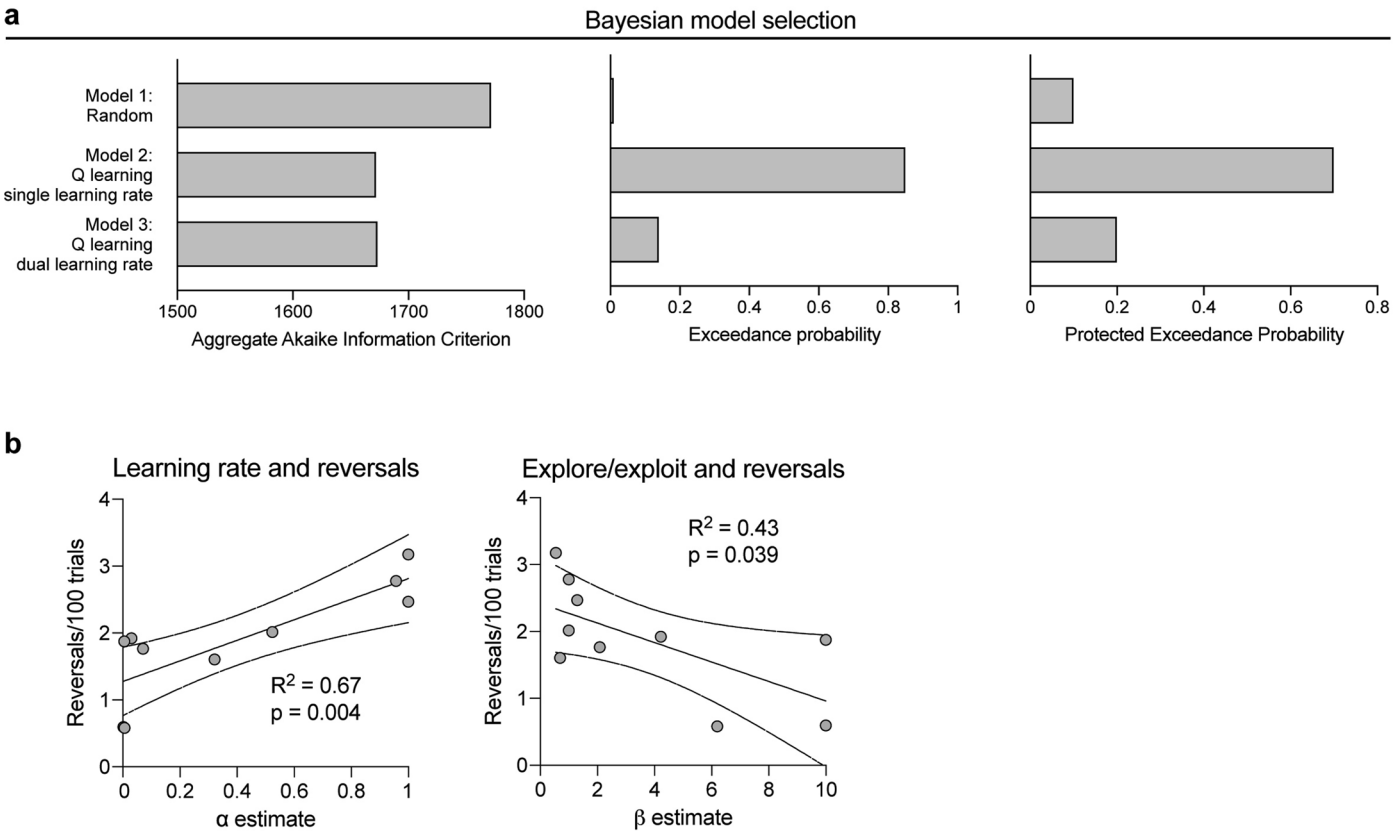


Extended Data Fig. 6 | See next page for caption.

https://doi.org/10.1038/s41593-023-01547-6

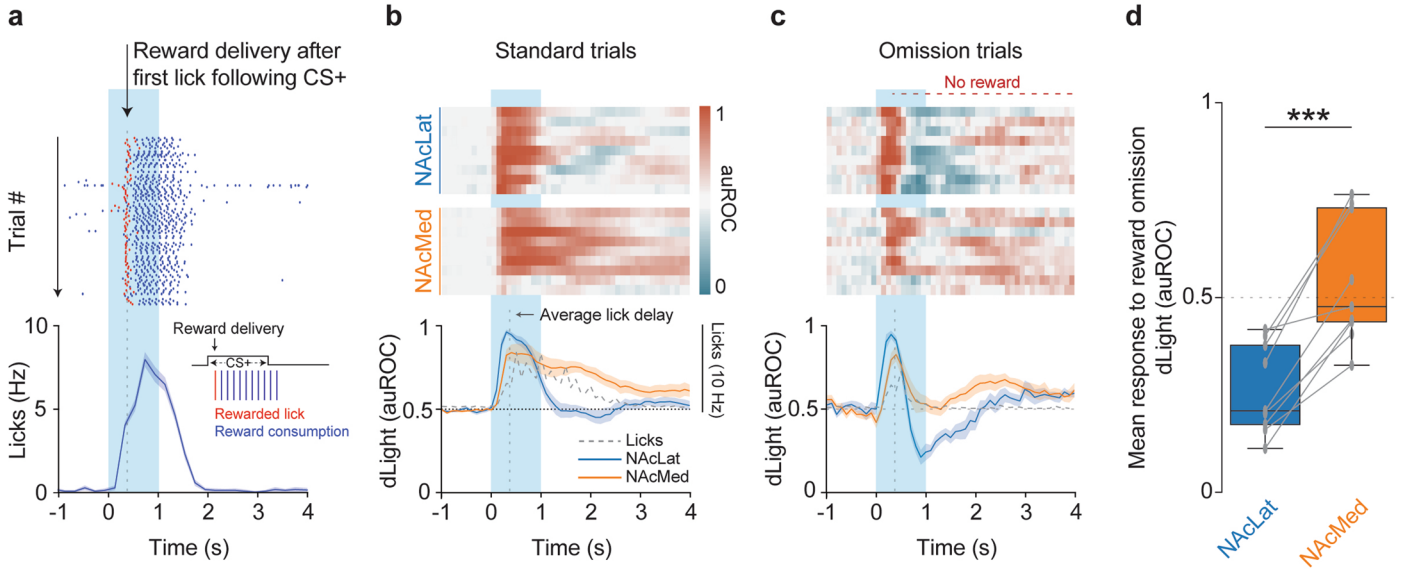
Extended Data Fig. 6 | Topography of the mesoaccumbal dopamine system in mice and rats. (a) Schematics showing representative fluorescent retrobeads locations (green) in different NAc subregions (NAcMed, NAc core, NAcLat) along the ventro-medial to dorso-lateral axis for different mice (n = 3 mice for each injection location; aca: anterior commissure). **(b)** Corresponding sample fluorescent images of coronal brain sections from the NAc (tyrosine hydroxylase (TH): red). **(c)** Sample fluorescent images of coronal midbrain sections from the same mice shown in (**a**,**b**). **(d)** Corresponding schematics highlighting locations of retrogradely labeled (that is, beads-positive, green) neurons in the medial VTA (mVTA) and lateral VTA (lVTA) (IF: interfascicular nucleus, PN: paranigral nucleus, PBP: parabrachial pigmented nucleus, IPN: interpeduncular nucleus; scale bars (**b**,**c**) 500 µm). **(e)** Schematic showing location of injection sites for fluorescent retrobeads (red) in the NAcLat and FluoroGold (gray) in the NAc core (n = 3 rats; CPu: Caudate Putamen). **(f)** Sample fluorescent images showing coronal brain sections of the NAcLat and NAc core at two different anterior-posterior coordinates (TH: blue, beads: red, FluoroGold: white; Scale bar 1 mm). **(g)** Fluorescent images showing coronal brain sections of the VTA from a rat that was injected with fluorescent retrobeads (red) into the NAcLat and FluoroGold (white) into the NAc core (scale bar 500 µm). Note, that mice and rats show a similar anatomical topography of NAcLat- and NAc core-projecting DA neurons (compare panels (**c**) and (**g**)). In both cases, NAcLat-projecting DA neurons are located in the lVTA, while NAc core-projecting DA neurons are located in the mVTA).

https://doi.org/10.1038/s41593-023-01547-6



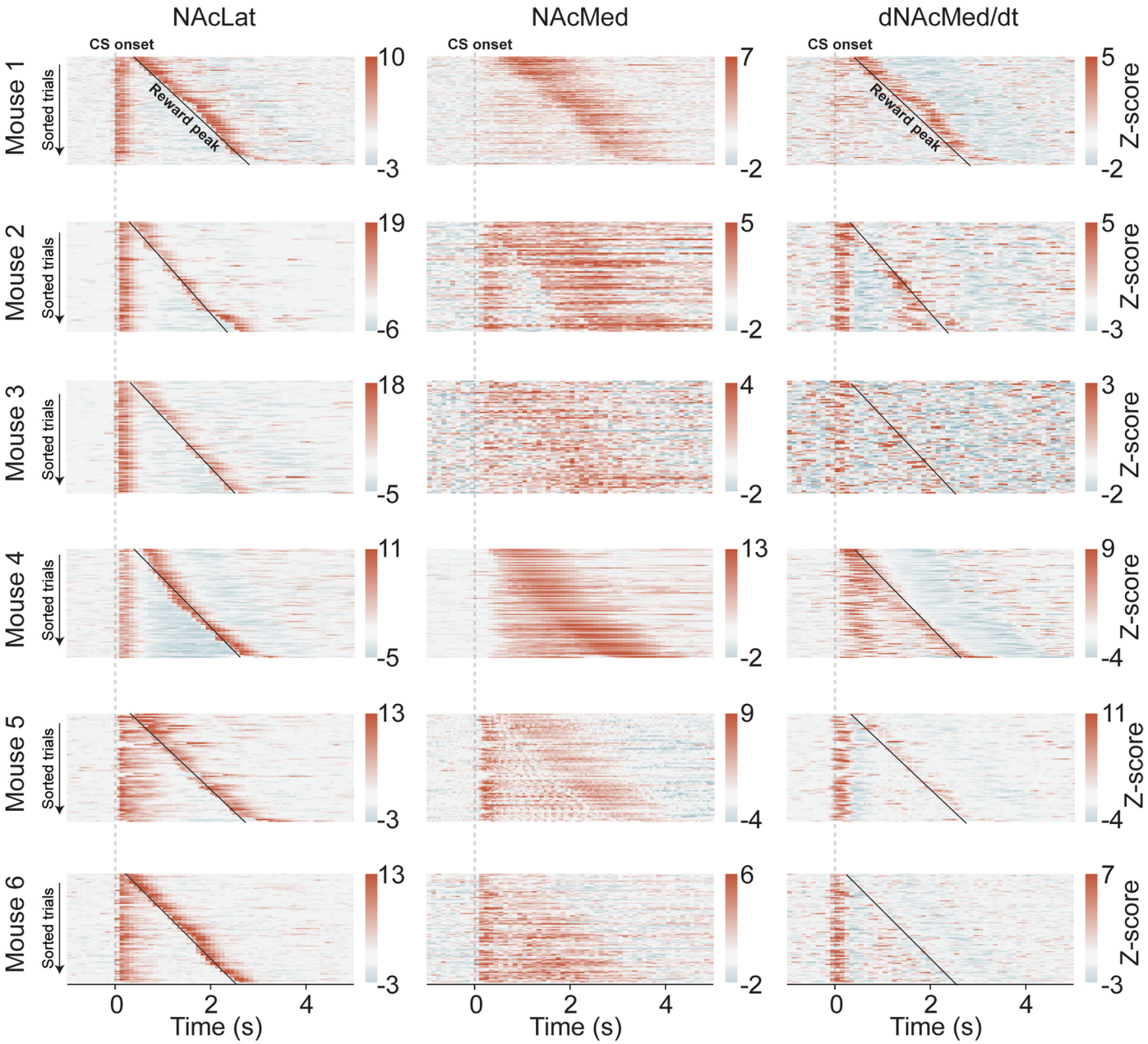
Extended Data Fig. 7 | Bayesian model selection showing that the behavior of the animals can be best described by a Q learning model with a single learning rate. (a) Results of Bayesian model selection. Model 2, a Q learning model with a single learning rate, has the highest protected exceedance probability and is, therefore, the best descriptor of the data of the three given models. Model 1 is a random choice model (in which every choice has a probability of 0.5); model 3 is a model with separate learning rates for learning from reward versus reward omission (n = 10 mice). **(b)** Best-fit model parameters for the experimental mice are shown in Fig. 3; each dot indicates an individual mouse/session. The two panels show the relation between learning rate α (left) and explore/exploit parameter β (right) versus the number of achieved reversal per 100 trials; lines indicate least-squares linear fit and its 95% confidence intervals; significance was tested by means of two-sided t-test.

https://doi.org/10.1038/s41593-023-01547-6



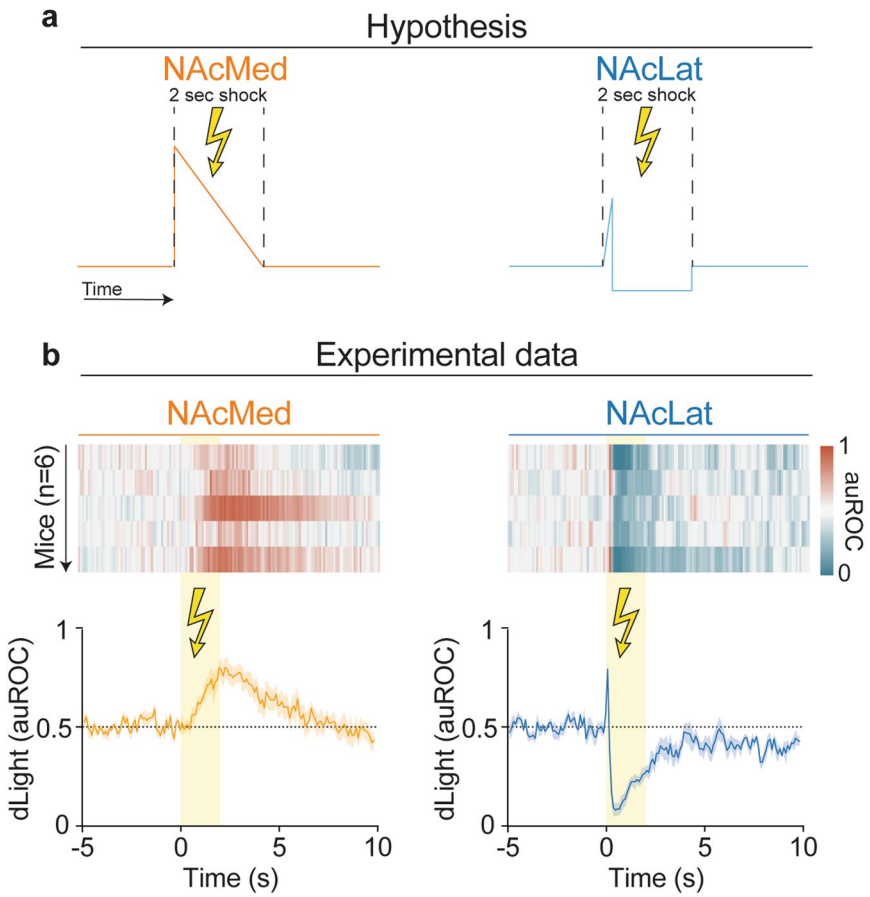
Extended Data Fig. 8 | Dopamine release in NAcLat, but not NAcMed, encodes negative RPE. (a) Graphs showing licking behavior of a representative sample animal during the task. Note that licking generally started ~500 ms after CS+ onset with only minimal variability. Inset shows schematic of experimental design: Head fixed mice were trained on a behavioral assay, in which the presence of a CS+ predicted the availability of a reward (3 µl of a 1% sucrose solution). The reward was delivered to the animal after one single lick (that is, it assumes minimal effort to the animal to obtain a reward). After training, mice were subjected to a session in which 10% of the trials were not rewarded (omission trials). DA release was measured simultaneously in the NAcLat and NAcMed using dLight1.2. **(b)** Top: Heatmaps showing auROC normalized dLight responses in the NAcLat and NAcMed for all animals (n = 9 mice) in standard trials in which reward delivery occurred after the first lick during CS+ presentation. Bottom: mean DA response in NAcLat (blue) and NAcMed (orange) as well as the mean lick rate (dashed line) for all mice. Note that NAcLat DA release tracks CS+ onset and reward delivery, while NAcMed DA tracks licking behavior, including reward consumption. **(c)** Same as in panel (**b**) but for omission trials in which no reward was delivered. Note that NAcMed DA release returns to baseline concurrent with the termination of licking behavior. Conversely, NAcLat DA release showed a negative deflection below baseline level during reward omission, which reflects a negative RPE. **(d)** Quantification (box plot, median and quantiles) of panels (**b**) and (**c**). Significance was calculated by means of a two-sided Students t-test; ***P < 0.001, n = 9 mice. All data, except panel (**d**), are represented as mean ± SEM (shading).

https://doi.org/10.1038/s41593-023-01547-6



Extended Data Fig. 9 | Dopamine release patterns for all mice used in experiment shown in Fig. 2g–o. NAcLat (left), NAcMed (middle) DA release was z-score normalized and trials were sorted by interval length. The derivative of NAcMed DA release (dNAcMed/dt, right) was calculated by determining the slope of a line that was fitted (linear regression) over a 500 ms interval at each time point. Mouse 4 is the sample animal shown in Fig. 2i; mouse 1 is the sample animal shown in Fig. 2o. Although there is substantive similarity in the pattern of DA release in the NAcLat with the temporal derivative of DA release in the NAcMed, there are two important limitations with such a comparison. First, the release of DA in the NAc reflects a complex response to activity in DA cell bodies in the VTA. As shown in Extended Data Fig. 4, NAc DA release is a convolution of VTA cell body activity. Second, there is significant between animal variation. For example, mouse 6 lacks a transient increase in the dNAcMed/dt trace in response to reward delivery. Additionally, while most mice show a dip in NAcLat DA release below the baseline between the onset of CS+ and reward delivery, only mouse 2 and mouse 5 exhibit a downward slope in NAcMed DA release during the same period (see also Fig. 2j,n).

https://doi.org/10.1038/s41593-023-01547-6



Extended Data Fig. 10 | Dopamine release patterns in NAcMed and NAcLat in response to aversive stimuli can be conceptualized in the context of state and rate-of-change encoding. (a) Hypothesis: the decrease in NAcLat DA release in response to a mild electric tail shock reflects the derivative of the downward slope in NAcMed DA release. If this is true, a sustained decrease in NAcLat DA release should be preceded by a brief increase that reflects the derivative of the onset of DA release in the NAcMed. **(b)** Experimental data: Head-fixed mice (n = 12 mice) were subjected to 10 unpredictable tail shocks while DA release in NAcMed and NAcLat was recorded using dLight 1.3. Left: Sustained increase in DA release in the NAcMed. Right: Sharp increase at shock onset that is followed by sustained below-baseline reduction of NAcLat DA release. While the gradual downward slope in the NAcMed and the transient increase in the NAcLat at shock onset suggest that our hypothesis is correct, there are two limitations that may be due to variability between experimental animals. First, the NAcMed DA response appears to be delayed compared to the NAcLat DA increase. Second, there is a sharp dip in NAcLat DA release that did not directly correspond to a sharp decrease in NAcMed DA release. All data are represented as mean ± SEM (shading).

nature portfolio

Corresponding author(s): Stephan Lammel

Last updated by author(s): Nov 30, 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Zen Software 2.3 (Zeiss) was used for acquiring confocal and epifluorescence images. Dopamine transients were measured using a custom built fiber photometry rig as described previously (de Jong et al., 2019; Neuron). Data acquisition and analys code is available at: https://github.com/handejong/Fipster. In vivo electrophysiology data were obtained using Open Ephys (open-ephys.org) or SpikeGLX (https://billkarsh.github.io/SpikeGLX). Individual units were sorted using Kilosort3.0 (https://github.com/MouseLand/Kilosort) and manually cleaned using Phy2 (https://github.com/cortex-lab/phy) or PKS (https://github.com/handejong/post-kilosort).
Data analysis	ImageJ (NIH, 64-bit Java 1.8.0_172, NIH) was used for analysis of fluorescence and confocal images. For fiber photometry experiments, video frames were analyzed online and fluorescent signals were acquired using custom acquisition code written in Matlab (R2019b) (https://github.com/handejong/Fipster). Time series data were analyzed using custom Python code (https://github.com/handejong/post-kilosort). Statistical tests were performed using JASP 0.15 (University of Amsterdam) or using the Python package Scipy v1.10.0.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Source data files are provided with the manuscript. All other data that support the findings of this study are available from the corresponding author upon request.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	*Use the terms sex (biological attribute) and gender (shaped by social and cultural circumstances) carefully in order to avoid confusing both terms. Indicate if findings apply to only one sex or gender; describe whether sex and gender were considered in study design; whether sex and/or gender was determined based on self-reporting or assigned and methods used.* *Provide in the source data disaggregated sex and gender data, where this information has been collected, and if consent has been obtained for sharing of individual-level data; provide overall numbers in this Reporting Summary. Please state if this information has not been collected.* *Report sex- and gender-based analyses where performed, justify reasons for lack of sex- and gender-based analysis.*
Reporting on race, ethnicity, or other socially relevant groupings	*Please specify the socially constructed or socially relevant categorization variable(s) used in your manuscript and explain why they were used. Please note that such variables should not be used as proxies for other socially constructed/relevant variables (for example, race or ethnicity should not be used as a proxy for socioeconomic status).* *Provide clear definitions of the relevant terms used, how they were provided (by the participants/respondents, the researchers, or third parties), and the method(s) used to classify people into the different categories (e.g. self-report, census or administrative data, social media data, etc.)* *Please provide details about how you controlled for confounding variables in your analyses.*
Population characteristics	*Describe the covariate-relevant population characteristics of the human research participants (e.g. age, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."*
Recruitment	*Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.*
Ethics oversight	*Identify the organization(s) that approved the study protocol.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to predetermine sample size, which were based on work in previous publications (e.g., de Jong et al., 2019; Neuron).
Data exclusions	No data exclusions.
Replication	All datasets were compiled using data obtained in multiple mice (n listed in figure legends). The same behavioral task was used for photometry and electrophysiology experiments and behavior appeared consistent between mice and cohorts.
Randomization	The identity of the CS+ and CS- (white noise, 11 kHz tone, blue light) was randomized over the experimental subjects. For simultaneous NAcLat and NAcMed recordings, fiber placements (left or right side of the brain) were randomized. Randomized allocation to experimental groups does not apply since there are no experimental groups in the manuscripts (i.e., all experiments are within-subject).
Blinding	Investigators were blind to recording location (e.g., "mVTA" or "lVTA" during data analysis.). No other blinding was applicable to the study as it does not contain between-subject comparisons.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used: Primary Antibodies: mouse anti-TH (1:1000, Millipore, MAB318), chicken anti-GFP (1:1000 Abcam, ab13970). Secondary antibodies: Alexa Fluor 647 goat anti-mouse (1:750, Thermo Fisher Scientific, A-21235), Alexa Fluor 477 goat anti-chicken (1:750, Abcam, ab150169).

Validation: All antibodies mentioned above have been extensively validated in mice in each of these previous studies (Lammel et al., 2008; Neuron; Lammel et al., 2012; Nature; Cerniauskas et al., 2019; Neuron, Yang et al., 2021; Nature Neuroscience).

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals: Laboratory mice (20-35g, 8-20 weeks old, male and female, group housed) from the following lines were used for the experiments: C57BL/6J (Jackson Laboratory; stock number: 000664), DAT-Cre mice (Jackson Laboratory, stock number: 006660, strain code: B6.SJL-Slc6a3tm1.1(cre)Bkmn/J), DAT-Flp (Jackson Laboratory, stock number: 035426, strain code: STOCK Slc6a3em1(flpo)Hbat/J), Ai32 (Jackson Laboratory, stock number: 012569; crossed to DAT-Cre mice). Additionally, we used Long-Evans rats (Charles River, 500-800 g, 8-17 months old, male, strain code: 0006). Mice were maintained on a 12:12 light cycle (lights on at 07:00) with food ad libitum and room temperature of 22-25°C and 55% humidity.

Wild animals: No wild animals were used.

Reporting on sex: Sex was not considered as a co-variable. All experiments were performed in mix cohorts of mice. No between-mice comparisons are made in the study.

Field-collected samples: No field-collected samples were used.

Ethics oversight: All procedures complied with the animal care standards set forth by the National Institutes of Health and were approved by University of California, Berkeley's Administrative Panel on Laboratory Animal Care.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Plants

Seed stocks: *Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.*

Novel plant genotypes: *Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.*

Authentication: *Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.*