

Published in final edited form as:

Cell Rep. 2024 April 23; 43(4): 114080. doi:10.1016/j.celrep.2024.114080.

## Distributional coding of associative learning in discrete populations of midbrain dopamine neurons

Riccardo Avvisati<sup>#1,2</sup>, Anna-Kristin Kaufmann<sup>#2</sup>, Callum J. Young<sup>1,3</sup>, Gabriella E. Portlock<sup>1</sup>, Sophie Cancemi<sup>1</sup>, Rui Ponte Costa<sup>3</sup>, Peter J. Magill<sup>2</sup>, Paul D. Dodson<sup>1,2,5,\*</sup>

<sup>1</sup>School of Physiology, Pharmacology, and Neuroscience, University of Bristol, BS8 1TD, UK

<sup>2</sup>Medical Research Council Brain Network Dynamics Unit, Nuffield Department of Clinical Neurosciences, University of Oxford, OX1 3TH, UK

<sup>3</sup>Computational Neuroscience Unit, Department of Computer Science, SCEEM, Faculty of Engineering, University of Bristol, BS8 1UB, UK

# These authors contributed equally to this work.

### Summary

Midbrain dopamine neurons are thought to play key roles in learning by conveying the difference between expected and actual outcomes. Recent evidence suggests diversity in dopamine signaling, yet it remains poorly understood how heterogeneous signals might be organized to facilitate the role of downstream circuits mediating distinct aspects of behavior. Here we investigated the organizational logic of dopaminergic signaling by recording and labeling individual midbrain dopamine neurons during associative behavior. Our findings show that reward information and behavioral parameters are not only heterogeneously encoded, but also differentially distributed across populations of dopamine neurons. Retrograde tracing and fiber photometry suggest that populations of dopamine neurons projecting to different striatal regions convey distinct signals. These data, supported by computational modelling, indicate that such distributional coding can maximize dynamic range and tailor dopamine signals to facilitate specialized roles of different striatal regions.

### Introduction

Learning to anticipate positive or negative consequences from environmental cues is essential for survival. Midbrain dopamine neurons are thought to play key roles in

---

This work is licensed under a [BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

\*Correspondence: paul.dodson@bristol.ac.uk

<sup>5</sup>Lead contact

#### Author Contributions

Conceptualization, P.D.D; Methodology, P.D.D, A.K.K, R.A, and C.J.Y; Formal Analysis, R.A, A.K.K, C.J.Y, P.D.D, and G.E.P; Investigation, A.K.K, R.A, S.C, and G.E.P; Writing – Original Draft, P.D.D and R.A; Writing – Review – Editing, P.D.D, R.A, A.K.K, P.J.M, G.E.P, C.J.Y, R.P.C; Visualization, P.D.D; Supervision, P.D.D, P.J.M, and R.P.C; Funding Acquisition, P.D.D, P.J.M, and R.P.C.

#### Declaration Of Interests

The authors declare no competing interests.

#### Inclusion and Diversity

We support inclusive, diverse, and equitable conduct of research.

this process by signaling the difference between expected and actual outcomes (reward prediction error; RPE)<sup>1</sup>. Such a fundamental teaching signal has traditionally been thought to necessitate a uniform message<sup>2–4</sup>. However, recent work has revealed that dopamine neurons and the signals they convey might be heterogeneous. For example, differences in dopamine release and dopamine axon activity have been observed in different striatal regions<sup>5–11</sup> with temporally-distinct activity patterns recorded across dorsal striatum<sup>12</sup>. Furthermore, dopamine neurons seem to signal more than just reward; encoding movement onset, movement kinematics, and multiple variables involved in decision-making<sup>13–20</sup>. Dopamine neurons also seem to be molecularly, physiologically, and anatomically diverse; single-cell RNA sequencing has identified multiple groups of midbrain dopamine neurons that can be distinguished by the combinatorial expression of different molecular markers<sup>21–27</sup>, and there is evidence that dopamine neurons exhibit differences in ion channels, other protein expression, firing properties, and input-output connectivity<sup>21,22,26,28–35</sup>. Together, these findings suggest that there might be functionally-specialized midbrain populations that each convey different information to discrete brain areas. However, it is not yet clear how such heterogeneous signals might instruct different striatal regions, which themselves have diverse and complimentary functional/behavioral roles. For example, the dorsolateral striatum (DLS) is thought to subservise sensorimotor functions in stimulus-response associations and habits, whereas the dorsomedial striatum (DMS) plays roles in response-outcome associations for goal-directed actions<sup>36</sup>. In further contrast, the ventrolateral striatum (VLS) is thought to be important for motivation<sup>37</sup>, whereas the core of the nucleus accumbens (NAc) has been ascribed roles in outcome evaluation<sup>8,38</sup>.

To define the organizational principles underlying heterogeneous dopamine signaling, we recorded and molecularly-identified individual dopamine neurons in mice during Pavlovian conditioned behavior. We find that there is no broad spatial organization of encoding in midbrain, but instead we show that neurons likely projecting to the same target are more homogenous in their firing patterns and these patterns match the activity of dopamine axons in the striatal target region. Temporal difference modelling predicts that distributional coding within these populations can tailor them to support different aspects of associative learning.

## Results

We trained head-fixed mice in a Pavlovian conditioning paradigm in which an auditory cue (1 s, 4 kHz tone) signaled reward delivery after a fixed delay (2 s from cue onset). Mice rapidly learned to associate the cue with reward as indicated by anticipatory licking during cues (Figure 1A and Figure S1). To investigate the firing of different dopamine neurons once mice had learnt the association, we extracellularly recorded individual neurons and then juxtacellularly labeled them to precisely determine their location and confirm they were dopaminergic (Figure 1B).

### Dopamine neurons heterogeneously encode reward behavior

The majority of dopamine neurons altered their firing rate at the onset of the cue and/or reward (Figure 1B and C). However, while changes at reward were generally increases in firing, changes at cue onset were a mix of increases and decreases in rate. Firing at cue

and reward may reflect the encoding of reward prediction<sup>39</sup>, or alternatively, encoding of actions with which to obtain reward<sup>15,18,19,40</sup>. To investigate whether changes in firing rate correlated with cue, reward, licking, or other kinds of movement (e.g. walking, running, or postural adjustments), we used a general linear model (GLM) (Figure 1D, 1E, and Figure S2) to investigate behavior-related firing across the session. We found that many neurons encoded reward (16 of 52) and/or cue (9 of 52). However, we also found that a significant number of neurons encoded parameters which were not obvious from the peristimulus time histogram, including licking (17 of 52) or movement (8 of 52). A considerable proportion of neurons (20 of 52) did not significantly encode ( $p > 0.05$ ) any of the features we examined (Figure 1D and 1E), which suggests that there may be populations of dopamine neurons that encode other facets of behavior, or that these neurons show no response to the fully predicted reward. Interestingly, many neurons encoded more than one parameter (13 of the 32 neurons encoding behavioral parameters), concordant with the idea that dopamine neurons multiplex signals<sup>16,20</sup>. These data suggest that there is not a uniform signal across midbrain dopamine neurons, but instead support a framework of heterogeneity with some neurons encoding single parameters and others multiplexing signals to encode several distinct aspects of behavior (Figure 1E).

### Encoding is not clearly defined by anatomical location

What underlies the heterogeneity we observe in reward-signaling? Given the different roles ascribed to the ventral tegmental area (VTA) and the substantia nigra pars compacta (SNc), one might predict that neurons in these regions would encode the paradigm differently<sup>41–43</sup>. To investigate this possibility, we divided neurons into those located in the VTA ( $n = 27$ ) or SNc ( $n = 25$ ) and compared their responses during the session. Despite previous observations of differences in encoding of spontaneous body movements by VTA and SNc neurons<sup>14</sup>, the responses of these two groups during Pavlovian conditioned behavior were nearly identical (Figure 2A–C). The comparable proportions of neurons in each region encoding cue, reward, licking, and movement (Figure 2A and 2B) also suggest that neurons in the two regions encode the paradigm in a similar manner. However, it is possible that such a blunt subdivision may mask finer spatial organization. We therefore considered whether encoding was organized so that neurons in close proximity signaled similar parameters. To probe this possibility, we plotted the dominant parameter (defined by the GLM coefficient) encoded by each neuron in Cartesian space (Figure 3A and 3B). We found that all parameters were represented across the anteroposterior and mediolateral extent, arguing against a precise focal encoding of parameters in different regions of the dopaminergic midbrain.

### Neurons projecting to different striatal regions express alternate combinations of proteins

If encoding is not spatially organized in the midbrain, is there another structural principle underlying response heterogeneity? Emerging evidence suggests that dopamine neurons projecting to particular target regions may differently encode parameters<sup>6,7,30,38</sup>. We therefore considered whether distinct midbrain-striatal circuits may account for some of the heterogeneity we observe. We injected retrograde tracer (cholera toxin B subunit; CTB) into the dorsomedial striatum (DMS), dorsolateral striatum (DLS), ventrolateral striatum (VLS), or nucleus accumbens core (NAc core) and examined the expression of three proteins

(Aldh1a1, Sox6, and calbindin) known to be differentially expressed in the dopaminergic midbrain<sup>26</sup> (Figure 4A). We found that most dopamine neurons projecting to DLS expressed Sox6 and Aldh1a1, but not calbindin, and were located in SNc, whereas those projecting to NAc core were located in VTA and had the opposite expression pattern (i.e. calbindin, but not Aldh1a1 or Sox6; Figure 4B–E). The marker expression of neurons projecting to DMS was less binary, with moderately prevalent expression of Aldh1a1 and Sox6, and rare expression of calbindin (Figure 4E). Notably, we found that these neurons are tightly clustered in the medial part of SNc (SNCM; Figure 4A–B)<sup>34</sup>. VLS-projecting neurons were predominantly localized to the parabrachial pigmented area of the VTA (PBP) and SNc<sup>34</sup>, and expressed Sox6 (Figure 4A–B and 4E). We found a significant interaction between marker expression and region (2 way parametric ANOVA with region and marker as factors;  $P < 0.0001$ ) suggesting populations can be distinguished using a combination of location and marker expression.

### Dopamine neuron populations encode distinct aspects of behavior

We next tested whether populations defined by both protein expression and anatomical location showed differential encoding of associative behavior. The activity of recorded dopamine neurons which putatively project to DMS (classified as such by their location in medial SNc and their expression of Aldh1a1 and Sox6; Figure 5E;  $n = 10$  neurons) differed considerably from the population mean (Figure 5A and 1C), with no increase in firing at reward presentation (Figure 5A). The putative DLS-projecting population encoded the cue (Figure 5A and S3) but, in contrast to DMS-projecting neurons, exhibited a mean increase in firing to reward (Figure 5D). The VLS-projecting population did not to change their firing upon cue presentation (despite the anticipatory licking suggesting that the mice registered and learned the predictive value of the cue), but robustly increased their firing shortly after reward presentation (Figure 5A –5D). While VLS-projecting neurons generally had the largest reward-related firing increases, the population was not significantly different from the DLS-projecting population (Figure 5D). Putative NAc core-projecting neurons also showed a distinct response, increasing their firing at both cue and reward. However, rather than being time-locked to the onset of these events, these increases in firing were delayed by a few hundred milliseconds (Figure 5B), coinciding with periods when the mice were still licking. All populations exhibited some degree of multiplexing, but it was prevalent in the DLS-projecting population where two-thirds of the neurons encoded two or more parameters (Figure 5A). Putative DLS- and DMS-projecting populations decreased firing at movement onset, whereas the firing of VLS or NAc core populations did not change (Figure S4). It has recently been observed in anaesthetized mice that neurons projecting to different regions have different firing properties (Farassat et al., 2019). We therefore tested whether the populations we identified had distinct firing in awake mice at rest (i.e. during the ITI, outside of engagement with the paradigm). We found that DMS, DLS, and VLS populations shared similar properties, whereas the NAc core-projecting population had significantly slower firing rates ( $p < 0.05$ ) and exhibited longer pauses ( $p < 0.01$ ) (Figure S5). To examine whether these populations could account for the heterogeneity observed across dopamine neurons we performed hierarchical clustering. Clusters were enriched with neurons from a given population, but did not exclusively contain one population (Figure 5F).

For example, 75% of neurons in one cluster were putative VLS-projecting neurons but the other VLS-projecting neurons were ascribed to another cluster.

If the populations of neurons that we putatively linked to projection targets accurately encompass the neurons projecting to each striatal region, we would predict that not only would the activity of dopamine axons in each target region be distinct, but also that axonal signals will resemble the signals recorded at the soma. To test these predictions, we used acute fiber photometry to record calcium signals in dopaminergic axons in different parts of striatum (Figure 6A–C). We recorded from DMS, NAc core, DLS, and VLS in each mouse (N=4) in different sessions. We observed considerable differences in the activity of dopamine axons in each region, particularly just after reward delivery (Figure 6A). Importantly these signals largely mirrored the action potential firing patterns from our putative projection-defined populations (Figure 5A). Axons in DLS and VLS showed relatively large increases in fluorescence following the fully-predicted reward, whereas axons in DMS and NAc core exhibited negligible changes (Figure 6B); only the DMS/NAc core signals are consistent with models of reward prediction error (RPE)<sup>1</sup>. To probe this further, we examined the normalized firing at reward for each of our putative projection-defined neuronal populations. We found that putative DLS- and VLS-projecting populations had a broader range of responses than DMS- and NAc core-populations suggesting that some neurons were inaccurately estimating future reward (Figure 7A). Recent work has suggested that encoding optimism as a probability distribution across dopamine neurons may confer advantage to reward learning<sup>44</sup>. Our data suggest that distributional coding may differ in populations projecting to different parts of striatum in both the width of the distribution and the skew (Figure 7A).

To explore the potential effect that different distributions might have on reward learning, we used a distributional temporal difference (TD) model where an agent learns state-value associations (Figure 7B). We then tested whether populations putatively projecting to different striatal targets would perform differently compared with a unified population of midbrain dopamine neurons<sup>2–4,44</sup>. Positive and negative learning rates, for an array of neurons, were fit to the juxtacellular data to generate projection-defined agents (Figure 7C). This resulted in each agent having state-value and state-error distributions, with each neuron converging on a different estimate of reward value (Figure 7B). To probe whether different distributions could confer a general advantage, we tested the accuracy of value estimations made by each projection-defined agent (Figure 7D). The model suggests that DMS- and NAc core-projecting populations would make significantly more accurate state-value associations (i.e., smaller value-estimate errors) than a unified population of midbrain dopamine neurons (created from the overall distribution of all recorded neurons; Figure 7D). DLS- and VLS-projecting populations consistently underestimated reward across a range of reward magnitudes (Figure 7E), which would result in dopamine release to fully-predicted reward. These populations therefore performed significantly worse at state-value estimations than DMS and NAc core. Taken together, this suggests that populations projecting to medial regions of the striatum (DMS and NAc core) might convey dopamine signals that are tuned to support state-value learning, whereas populations projecting to lateral regions (VLS and DLS) might be less well suited to this role.

## Discussion

Here, we defined at millisecond resolution the behavior-related activity of individual, precisely localized, dopamine neurons. In doing so, we identified considerable heterogeneity in the encoding of reward-related signals by midbrain dopamine neurons. Heterogeneity could not be well explained by anatomical subdivisions nor spatial location, whereas grouping neurons according to the striatal regions they might innervate revealed populations with divergent properties. The differential encoding of reward we observed was also evident in dopamine axons in the corresponding target regions of striatum. We show that individual dopamine neurons not only multiplex signals by encoding different combinations of egocentric and allocentric parameters, but they also exhibit different magnitudes of encoding from the rest of the population. Our TD modelling predicts that such distributional coding not only maximizes the dynamic range of dopamine signals, but also tailors them to support specialized functions of different striatal regions.

The role of dopamine neurons in predicting reward forms an important foundation for the understanding of how animals learn<sup>39</sup>. The dopamine signal has traditionally been considered to be uniform, broadcasting a common teaching signal across many brain circuits<sup>3,4</sup>. Instead, we find that such signals are far from uniform and our data suggest that different striatal regions receive specialized dopamine signals. Previous studies have identified that heterogeneity in dopamine neuron signaling can be parsed according to spatial localization in the midbrain<sup>14,16,45</sup>. While we cannot rule out the possibility of spatial organization, our analyses did not identify clear homogenous responses segregated by location (Figure 3); however, when we considered neurons that putatively project to the same projection targets (Figure 5), we observed more homogeneous responses. This suggests that the combination of cell body location along with molecular profiles provides a better description of cell populations than either property by itself. The anatomical location of these projection-defined groups suggests that there are “hot spots” containing neurons projecting to the same region e.g. in parts of the medial substantia nigra pars compacta (SNCM) and the VTA (Figure 5E). Indeed these “hot spots” may explain why some studies observe more uniform responses when recording sites are more localized<sup>2,16</sup>.

Recent work has suggested that dopamine neurons not only encode RPE, but may also encode other parameters including movement onset and kinematics<sup>13–18,46</sup>. In addition to neurons encoding general body movements, we identified a number of neurons that encoded licking (Figure 1D, 1E, and Figure S2). It is not clear whether these signals represent a motor or perceptual response; in principle, firing at licking could signify the initiation of a tongue movement, the sensory properties of contact with the spout, or a reward-related signal<sup>15</sup>. This is further complicated by the possibility that there could simultaneously be a motor response in neurons projecting to DLS but an incentive response in those projecting to the NAc core; further work will be needed to disambiguate these possibilities. Many individual dopamine neurons encoded the cue; however, in contrast to previous studies<sup>1,47</sup>, we did not observe a significant net response to the cue across the whole population (Figure 1C and 6A). Previous work has suggested that the cue serves an alerting role<sup>48–53</sup>. In support of this idea, dopamine neurons do not respond when the offset of a sound is used as a cue, they show larger responses to strong sensory stimuli, and they exhibit diminished

responses to cues predicting rewards with 100% reliability<sup>49,52,54,55</sup>. The cue we used only had a modest volume (62 dB), which is considerably quieter than many commercial systems (which can be 75 – 86 dB). In primates, a 72 dB tone only elicited a small change in dopamine firing whereas a 90 dB tone caused a large phasic increase<sup>55</sup>. It is therefore possible that introducing louder cues or changing reward probabilities would unmask a larger increase in firing to the cue<sup>15</sup>. It has also been suggested that the cue response signals motivational salience, with higher value stimuli eliciting a larger response<sup>51,56</sup>. In our experiments, we used relatively mild motivation strategies<sup>57</sup>, and one might therefore predict that if the motivational drive of the mice were very high, there would be a larger dopamine signal to the cue<sup>58,59</sup>. Regardless of the explanation, one important observation from our data is that a positive dopamine response at cue presentation is not necessary for Pavlovian conditioning.

Our data suggest that dopamine neurons projecting to different regions have distinct firing patterns; we confirmed these observations by measuring distinct signals in dopamine axons in different regions of striatum. These results argue against a model<sup>60</sup> where the firing of dopamine neurons is distinct from activity in dopaminergic axons, but instead support ideas that there are distinct profiles of dopamine release in different parts of striatum<sup>5–13,30,61</sup>. Perhaps the most striking difference between responses we observed was that the putative DMS-projecting group did not respond to predicted reward; this finding is in agreement with some studies<sup>7,9</sup> but not others<sup>5,11,30</sup>. The fact that reward probability was deterministic in our experiments may help to reconcile these apparent discrepancies; a recent study compared dopaminergic axon terminal responses in DMS during fixed-vs variable-probability reward and observed a similar lack of response to fixed reward which was rescued as rewards became probabilistic<sup>9</sup>. In contrast to DMS, we observed that the VLS-projecting population responded strongly to predicted reward and that NAc core-projecting neurons responded during licking. Similar pronounced reward signals have previously been observed in dopamine neuron terminals within VLS, and delayed signals in medial regions which might be consistent with licking rather than reward<sup>6</sup>. Interestingly, aversive taste is reported to result in dopamine release preferentially in NAc core, suggesting a possible evaluative role for these licking-related signals<sup>10</sup>.

What are the implications of projection-selective-encoding? Because dopamine signals likely result in different outcomes depending on the target region (e.g. cue attraction vs movement invigoration)<sup>38,61</sup>, it follows that different striatal territories might receive distinct dopamine signals. Such specialized signals would permit flexibility and a wide dynamic range; for example, different regions might receive a common signal in one learning scenario for appetitive situations where approach is desired, but tailored signals in aversive scenarios where avoidance would be the appropriate behavior<sup>6,11,30</sup>. Our modelling suggests that responses may be tuned to different parts of the reward spectrum. For example, the positively skewed DLS- and VLS-projecting responses (Figure 7A) suggest populations of dopamine neurons which tend to underestimate reward. One might expect such patterns of dopamine release to reinforce actions which could support habit development (a role that has been previously ascribed to DLS)<sup>62,63</sup>. This segregation of signaling profiles could facilitate simultaneous accurate reward evaluation in medial regions and action reinforcement in lateral striatum. As such, distributional coding within discrete

projection-defined populations may impart additional benefit compared with coding by a single population<sup>9,44</sup>. The heterogeneity we observe may also be compounded by the possibility that dopamine neuron populations projecting to different regions may co-release glutamate, GABA, or neuropeptides<sup>29,64,65</sup>. Furthermore, we report dopamine signals at the soma and axon, but dopamine release dynamics may be shaped locally and the striatum itself is heterogeneous with differences in dopamine transporters, cholinergic signaling, and striosome and matrix compartments across the striatum<sup>66,67</sup>. Further investigation is required to understand how differences in dopamine signaling interact with this additional complexity.

In conclusion, we find that even in simple learning paradigms, dopamine neurons represent multiple behavioral parameters in a heterogeneous manner. However, our data reveal an organizational logic where different striatal regions receive dopamine signals that are specialized to support different aspects of learning.

### Limitations of the study

One of the challenges of studying dopamine neuron subtypes is to fully define all existing populations. Single-cell transcriptomics has been used to identify putative dopamine neuron subsets based on expression of common sets of genes<sup>21–26,46,68–70</sup> and has identified at least seven populations<sup>26</sup>, although it is likely that there are further subgroups<sup>46</sup>. In our study we attempted to identify populations based on the striatal regions they innervate. We identified combinations of marker expression and cell body location that could be used to delineate which striatal region a dopamine neuron is likely to target. While DLS- and NAc core-projecting populations exhibited “all or nothing” expression of three key markers, DMS- and VLS-projecting populations were less clear cut. This raises the possibility that there may be more than one population of dopamine neurons projecting to these regions. For example, in addition to the Sox6+ Aldh1a1-population projecting to VLS we identified, there could also be some Sox6-Aldh1a1+ dopamine neurons which are likely to target intermediate regions (i.e. between VLS and DLS) as Aldh1a1 expression decreases in ventral regions<sup>68–70</sup>. Similarly, some of the remaining heterogeneity within our four populations could be accounted for by the presence of additional subgroups within these populations; for example, the recently identified Anxa1-expressing subtype of dopaminergic neuron that projects to DLS<sup>46</sup>. We also cannot rule out the possibility that a proportion of neurons with a marker and localization profile ascribed to a striatal target region might project to another brain region<sup>69</sup> (e.g. a proportion of neurons in medial SNc expressing Aldh1a1 and Sox6 could project to a region other than DMS). However, the concordance between neuronal firing and the activity of dopamine axons (recorded with photometry) in the ascribed striatal regions provides confidence that this is either rare, or these populations respond similarly during behavior.



## Star★Methods

### Resource Availability

#### Lead contact

Information and requests for resources and reagents should be directed to the and will be fulfilled by the Lead Contact, Dr Paul Dodson (paul.dodson@bristol.ac.uk)

#### Materials availability

This study did not generate new unique reagents.

### Experimental Model And Subject Details

#### Experimental animals

All experimental procedures on animals were conducted in accordance with the Animals (Scientific Procedures) Act, 1986 (United Kingdom) and approved by the animal welfare and ethical review boards at the University of Bristol and the University of Oxford. N=47 C57Bl6/J 3 – 4 month-old male mice (Charles River Laboratories) were used for recording and tracing and N=4 DAT<sup>IRESc<sup>re</sup></sup> (JAX:006660) 2 – 3 month-old male mice (heterozygous for the transgene) were used for fiber photometry experiments. Mice were group housed (except when isolated to prevent fighting or for experimental needs) in open-top (Bristol) or individually-ventilated (Oxford) cages. Cages were enriched with a house, cardboard tube and wooden chew block. Mice were kept in temperature-controlled conditions (21°C) and on a 12:12-h light–dark cycle (lights OFF at 08:15, lights ON at 20:15); experimental procedures were performed during the light phase of the cycle. Standard laboratory chow (Purina, UK) and water was provided *ad libitum* (except during food or water restriction).

### Method Details

#### Surgeries

Mice were anesthetized using 1 – 2% (vol/vol) isoflurane and placed in a stereotaxic frame, on a homeothermic heating mat (Harvard Apparatus) to ensure stable body temperature. Corneal dehydration was prevented using carbomer liquid gel (Viscotears, Alcon) and mice were perioperatively injected with the analgesic buprenorphine (0.03 mg/kg s.c., Vetergesic, Bayern).

For electrophysiological recordings, a custom L-shaped headpost (0.7 – 0.8 g, stainless steel or aluminum) was attached to the skull using cyanoacrylate glue<sup>14</sup>. The 3 mm diameter window in the headpost-base was positioned above the substantia nigra of the right hemisphere (centred at AP -3 mm and ML +1.5 mm from bregma). A craniotomy for single-unit recordings was made within the window of the headpost either on the day of headpost implantation or 1 – 7 days prior to recording. Two stainless steel screws (0.8 mm diameter; Precision Technology Supplies) were implanted in the skull, one above the frontal cortex and a reference above the cerebellum of the left hemisphere. A coiled 0.23 mm diameter stainless-steel wire (AM Systems) was implanted between the layers of cervical muscle to record EMG activity (filtered at 0.3 – 0.5 kHz). Exposed skull, screws and

EMG wire were covered with dental acrylic resin (Jet Denture Repair; Lang Dental). The craniotomy was sealed with fast set removable silicone rubber (Body Double; Smooth-On).

For retrograde tracer injections, a craniotomy was performed above the target region and a calibrated glass micropipette (708707; Blaubrand IntraMark) with a tip diameter of ~25µm was lowered to the appropriate target; NAc core (AP +1.0, ML +1.0, DV -4.3), DLS (AP +1.1, ML +1.8, DV -3), DMS (AP +1.0, ML +1.2, DV -2.8), VLS (AP +1.0, ML +1.8, DV -4.2). 30 – 150 nl cholera toxin subunit b (CTB; 0.5% w/v; C9903; Sigma-Aldrich) was manually injected at a rate of ~50 nl/min and pipettes were left in place for 5 – 10 minutes after injection. 9 – 13 days after tracer injection, mice were given a lethal dose of anesthetic and transcardially perfused. In a minority of experiments (N = 8), we injected CTB into dorsomedial striatum prior to electrophysiological recording, to verify that recorded neurons projected to the putatively assigned target; in these experiments we recorded and juxtacellularly labeled two SNCM neurons, both of which were CTB positive.

### Behavioral training

Animals were head-fixed using a custom headpost holder connected to a stereotaxic frame and positioned upon a custom-made treadmill where they could run, walk, or rest at will. Mice (N=34) were trained to associate an auditory cue with the delivery of a reward in a Pavlovian conditioning paradigm using a custom Arduino-based apparatus. Trials consisted of cue presentation (1 second, 4 kHz, 62 dB) delivered by a piezo speaker (535-8253, RS components), 1 second delay, followed by reward delivery (5 µl of 10% sucrose). Inter-trial interval (ITI) durations were randomly drawn from an exponential distribution with a flat hazard function to ensure equal distribution of expectation (4 – 10 s, median 5.4). Mice were either food (to > 85% of baseline weight) or water restricted (4 hours of *ad libitum* water per day after training/recording sessions using an automated water delivery system <https://doi.org/10.5287/bodleian:Vj4YaGAOY>); no differences in electrophysiological responses were observed between these motivators. Animals were trained in daily sessions consisting of 100 rewards and all mice showed robust anticipatory licking to cue before recording, (licking rate > 2 standard deviation from baseline during cue; median 5 days training prior to recording, IQR 3). Licking was monitored using a piezoelectric sensor (285-784, RS components). Movement periods and licking bouts for single-unit recordings were determined for the whole recording session off-line using cervical EMG and video recordings (30 frames/s). Movement typically involved walking or running on the treadmill as well as postural adjustments. Lick-onset was defined as the first video frame with visually detectable lower jaw movement, lick-offset was defined as the first of a series of at least three subsequent video frames with no visually detectable jaw movement. Movement onset and offset were defined in the same way using body and limb movements.

### Electrophysiological recording

Extracellular single-unit recordings were made with borosilicate glass electrodes (tip diameter 1.0 – 1.5 µm, in situ resistance 10 – 25 MΩ; GC120F-10, Harvard Apparatus) filled with saline solution (0.5 M NaCl) containing Neurobiotin (1.5% w/v, Vector Laboratories). Sterile saline (0.9% w/v NaCl) was frequently applied around the craniotomy to prevent dehydration of the exposed cortex. Electrode signals were filtered at 0.3 – 5 kHz and

amplified 1000 times (ELX-01MX and DPA-2FS, NPI Electronic Instruments). A Humbug (Quest Scientific) was used to eliminate mains noise at 50 Hz. All biopotentials were digitized online at 20 kHz using a Power 1401 mk3 analog-digital converter (Cambridge Electronic Design) and acquired using Spike2 software (version 7 or 10; Cambridge Electronic Design). For the recording, electrodes were lowered into the brain using a micromanipulator (IVM-1000; Scientifica). To avoid possible sampling bias, on-line criteria were applied to guide recordings of dopamine neurons (spike duration threshold-to-trough for bandpass-filtered spikes > 0.8 ms and firing rates < 20 spikes/s)<sup>71</sup>. Following recording, single neurons were juxtacellularly labeled with Neurobiotin<sup>14</sup> to allow for their unambiguous identification and localization. At the end of the experiment, mice were given a lethal dose of pentobarbital and transcardially perfused with PBS followed by 4% w/v paraformaldehyde in 0.1 M phosphate buffer (PFA). Brains were placed in PFA overnight at 4°C and then stored in PBS containing 0.05% w/v sodium-azide.

### Immunohistochemistry

50 µm coronal sections were cut from the midbrain on a vibrating-blade microtome (VT1000S; Leica Microsystems or DTK-1000, DSK). To confirm the location and neurochemical identity of recorded and juxtacellularly-labeled neurons, sections were incubated for 4 h at room temperature in PBS with 0.3% (vol/vol) Triton X-100 (Sigma) containing Cy3-conjugated streptavidin (1:1000) (GE Healthcare). To probe expression of different molecular markers in labeled neurons, a two-step procedure was applied, sections were incubated overnight in PBS-Triton with mouse anti-Tyrosine Hydroxylase (TH, 1:1000, T2928, Sigma-Aldrich) or chicken anti-TH (1:500, ab76442, Abcam); guinea pig anti-Sox6 (1:1000, gift from M. Wegner, Friedrich-Alexander University Erlangen-Nuremberg; (Stolt et al., 2006)) or rabbit anti-Sox6 (1:500, ab30455, Abcam). Sections were washed in PBS, and then incubated in PBS-Triton for > 4 hours with AMCA-conjugated secondary antibodies (donkey anti-mouse IgG, 1:500; 715-155-150 or donkey anti-chicken IgG, 1:500, 703-155-155; Jackson ImmunoResearch) or Brilliant Violet 421-conjugated secondaries (donkey anti-chicken IgG 1:500, 703-675-155, Jackson ImmunoResearch) to visualize immunoreactivity for TH, and AlexaFluor 647- or Cy5-conjugated secondary antibody to visualize immunoreactivity for Sox6 (A647: donkey anti-guinea pig IgG, 1:500, 706-605-148, Jackson ImmunoResearch; Cy5: donkey anti-rabbit IgG, 1:500, 711-175-152, Jackson ImmunoResearch). After imaging, the second step consisted of incubating overnight in PBS-Triton with rabbit anti-Aldh1a1 (1:500, HPA002123, Sigma-Aldrich) and goat anti-calbindin (1:500, sc7691; Santa Cruz) or mouse anti-calbindin (1:500, CB300, Swant), washing in PBS and then incubating overnight at room temperature in PBS-Triton with AlexaFluor 647- or Cy5-conjugated secondary antibodies (the fluorophore used in the previous step to visualize Sox6) to visualize immunoreactivity for Aldh1a1 (AF647: donkey anti-rabbit IgG, 1:500, 711-605-152, Jackson ImmunoResearch; Cy5: donkey anti-rabbit IgG, 1:1000, 711-175-152, Jackson ImmunoResearch) and AlexaFluor 488-conjugated secondary antibodies for Calbindin (donkey anti-goat IgG, 1:500, A11055, Life Technologies; donkey anti-mouse IgG, 1:500, 715-545-150, Jackson ImmunoResearch). This way, we were able to clearly visualize immunoreactivity for Sox6 (nuclear) and Aldh1a1 (cytoplasmic) using the same fluorescence channel. Borders of VTA and SNc were delineated using Aldh1a1 and calbindin immunofluorescence<sup>72</sup>.

For retrograde tracing, a combinatorial approach with partial overlap was used for immunohistochemistry, so that TH and CTB immunoreactivity was tested in all samples, but different series from the same animal could be tested for three additional markers. A number of markers have been identified as being selectively expressed in populations of dopamine neurons<sup>21–26,46,68–70</sup>; we therefore selected the three proteins (Sox6, Aldh1a1, and calbindin) that show good population discrimination and can be reliably detected using immunohistochemistry. Sections were incubated overnight at room temperature in PBS-Triton with chicken anti-TH (1:250/500, ab76442, Abcam), mouse anti-CTB (1:500, ab35988, Abcam) or goat anti-CTB (1:5000), #703, List Biological Labs), rabbit anti-calbindin (1:1000, CB38, Swant) or goat anti-calbindin (1:500, sc7691, Santa Cruz) or mouse anti-calbindin (1:500, CB300, Swant), rabbit anti-Sox6 (1:4000, ab30455, Abcam), rabbit anti-Aldh1a1 (1:500, HPA002123, Sigma Merck). Sections were washed in PBS and then incubated for > 4 h at room temperature in PBS-Triton and secondary antibodies. To visualize immunoreactivity for TH, AMCA- or Brilliant Violet 421-conjugated secondary antibodies were used (AMCA: donkey anti-chicken IgG, 1:500, 703-155-155, Jackson ImmunoResearch; BV421: donkey anti-chicken IgG 1:500, 703-675-155, Jackson ImmunoResearch). CTB was visualized using Cy3- or AlexaFluor 488-conjugated secondaries (Cy3: donkey anti-mouse IgG, 1:500, 715-165-151, Jackson ImmunoResearch; AF488: donkey anti-goat IgG, 1:500, 705-545-147, Jackson ImmunoResearch). Aldh1a1 or Sox6 immunoreactivity was visualized using Cy5- or AlexaFluor 647-conjugated secondaries (Cy5: donkey anti-rabbit IgG, 1:1000, 711-175-152, Jackson ImmunoResearch; AF647: donkey anti-rabbit IgG, 1:500, 711-605-152, Jackson ImmunoResearch) and Calbindin with Cy3- or AlexaFluor 488-conjugated secondaries (Cy3: donkey anti-mouse IgG, 1:500, 715-165-150, Jackson ImmunoResearch; AF488: donkey anti-mouse IgG, 1:500, A-21202, Life Technologies).

### Microscopy and cell counting

Example images were acquired using a confocal laser-scanning microscope (20x objective, LSM710; Carl Zeiss, or SP8; Leica). Images for cell counting were acquired on an epifluorescence microscope (DMI6000; Leica, or AxioImage.M2; Carl Zeiss) equipped with a 20x objective. Images of the dopaminergic midbrain were acquired as a series of 21 tiles (7x,3y). Sections containing CTB positive SNc and VTA neurons with a clearly defined nucleus were counted using the ‘cell counter’ plugin on ImageJ, Fiji version 1.53q<sup>73</sup> or Stereo investigator software 9.0 (MBF Bioscience). During counting, the experimenter was blind to the region targeted with CTB. To obtain percentages of midbrain dopamine neurons expressing a particular marker, counts were collapsed across sections, then divided by the number of neurons positive for both CTB and TH in each sample. Every marker-combination was counted in a minimum of three animals per striatal region. CTB injection sites in striatum were represented as honeycomb plots; a tessellated hexagonal structure was superimposed onto each image, then hexagons that were >80% by CTB immunoreactivity were coloured red at 100% opacity, opacity of hexagons that included 50 – 80% CTB immunopositivity was set at 50%. Images from each animal were superimposed and opacity was normalized.

## Fiber Photometry

DAT<sup>IRES<sup>cre</sup></sup> mice (JAX:006660) (N=4, 2 – 3 month-old male mice heterozygous for the transgene) were injected with AAV1-CAG-flex-GCaMP6f and AAV1-CAG-flex-tdTomato (Addgene: final titers  $4.45 \times 10^{12}$  and  $1.475 \times 10^{12}$  vg/ml respectively) into the midbrain at AP - 3.2, ML +0.5, DV -4.0 and AP -3.2, ML + 1.5, DV -4.0 relative to Bregma (~250 nl total per site). 3 weeks later, mice were implanted with a headpost (as described above), and a craniotomy was made above striatum. Mice were trained for 5 – 7 days in the Pavlovian conditioning paradigm. At the beginning of the photometry recording session, a bare fiber-terminated patch cable (200  $\mu$ m diameter, 0.48 NA, Thorlabs) was lowered into the brain using a micromanipulator (IVM-1000; Scientifica: AP +1.0, ML +1.0 to +1.2 for DMS and NAc, and AP + 1.0, ML +1.8 to +2.0 for DLS and VLS; -2.3 to -2.7 from brain surface for DMS and DLS; -3.3 to -3.8 for VLS and NAc). Data are comprised of a single recording session for each striatal site in each of the four mice (i.e. one DLS, one VLS, one DMS, and one NAc core recording per mouse); only one striatal region was recorded from during each session. Sessions were conducted in a pseudorandom order (with dorsal sites recorded prior to ventral). Fiber positions were confirmed post-hoc in fixed brains by visualizing GFAP immunoreactivity (1:1000 rabbit anti GFAP; 16825-1-AP, Proteintech) surrounding the fiber track using a Brilliant Violet 421-conjugated secondary antibody (donkey anti-rabbit IgG 1:500, 711-675-152, Jackson ImmunoResearch). Photometry data were acquired at 130 Hz using a pyPhotometry<sup>74</sup> board (Open Ephys). Both signals were median (5 point kernel) and low pass filtered (second order Butterworth filter with a 20Hz cut-off) and a 0.001 Hz second order high pass filter was applied to correct for photobleaching. Motion correction was performed by subtracting the best linear fit of the tdTomato signal from the GCaMP signal. Baseline was obtained by filtering the GCaMP signal with a low pass 0.001 Hz, second order Butterworth filter. The motion-corrected signal was then divided by this baseline to obtain a dF/F and each sweep normalized to 1 second before the auditory cue.

## Data analysis

Single-unit activity was isolated using template matching, principal component analysis and supervised clustering within Spike2 (Cambridge Electronic Design) and data were exported to MATLAB (Mathworks). Firing activity of labeled neurons was normalized as z-scores and used to construct peri-stimulus time histograms (PSTH; bin size 40 ms, smoothed with a 5-point Gaussian filter, half-width 70 ms) using a baseline of 1 second preceding cue onset. The first 2 principal components of the PSTHs (singular value decomposition) were used for hierarchical clustering; dendrograms were computed using an average method linkage function with Euclidean distances. To analyze which factors accounted best for changes in firing of individual midbrain dopamine neurons, a Poisson generalized linear regression model (GLM; fitglm function, MATLAB) was used to obtain a least-squares fit of the selected predictors to the recording data across the whole session. The recording session (including ITIs) was broken down into 200 ms bins of spike counts aligned to cue and reward delivery for every trial. Predictors were defined as cue, reward, licking, and movement and coded as either present or absent for every bin. Bins 0 to 400 ms from the onset of reward and cue were coded as reward and cue positive, respectively. Bins were coded as licking or movement positive if they overlapped at least 75% with licking and movement bouts, respectively. The model used a log link function and was set to

predict spike counts in every bin based on the binary regressors. Deviance goodness of fit tests confirmed that firing of 78% of neurons were well fit by the GLM ( $P < 0.05$ ). To determine the impact of different features, irrespective of whether they resulted in increases or decreases in firing, the absolute values of coefficients were considered; an individual cell was considered responsive to one of the four parameters if the corresponding p-value was  $< 0.05$ . Because periods of licking often occurred soon after reward delivery, we confirmed that the firing of neurons classified as ‘licking’ was time locked to lick episodes but not reward delivery (Figure S2). To analyze dopamine neuron firing properties, we used a coefficient of variability of interspike intervals (CV2) to examine firing regularity<sup>71,75</sup> and robust Gaussian surprise<sup>76,77</sup> to detect bursts of at least three spikes with significantly shorter ISI’s than the population of spike trains.

### Computational modelling

For each observed state, a Temporal Difference (TD) algorithm<sup>1,78</sup> produced a series of value predictions ( $V_{t,i}$ ). Each of these Value predictions represents a single neuron ( $i$ , at time  $t$ ). TD error ( $\delta_{t,i}$ ) was calculated by comparing a neuron’s existing state-value prediction, with a bootstrapped (predicted) estimate of the state’s value ( $V_{t+1,i}$ ):

$$\delta_{t,i} = r_{t,i} + \gamma V_{t+1,i} - V_{t,i} \quad (1)$$

where ( $r_{t,i}$ ) is the reward and  $\gamma$  the discount factor. Value predictions were updated according to the following update rule:

$$\begin{aligned} V_{t,i} &\leftarrow (V_{t,i} + \alpha_i^+ \delta_{t,i}) && \text{if } \delta_{t,i} > 0 \\ V_{t,i} &\leftarrow (V_{t,i} + \alpha_i^- \delta_{t,i}) && \text{if } \delta_{t,i} < 0 \end{aligned} \quad (2)$$

Where  $\alpha_i^+$  and  $\alpha_i^-$  are the unique positive and negative learning rates applied to each neuron ( $\alpha_j \sim U(0,1)$ ). Distributional coding occurs due to each neuron converging on distinct state-value estimates – according to the balance of their positive and negative learning rates. These learning rates were randomly initialized and then fit to each neuron using a grid search. Learning rates were tailored to the projection-defined agents by minimizing the difference between  $\delta_{t,i}$  at rewarded states, and a sample drawn from ‘activity distributions’ for each subpopulation. Subpopulations were approximated from neural data using a kernel density estimation.

To test these algorithms, we created an environment in which agents deterministically transition between states. At one such state, the agent receives a numerical reward randomly selected from a gaussian distribution ( $r \sim \mathcal{N}(5,5)$ ). Trained agents have a ‘value distribution’ associated with each state in its environment; calculated from the distribution of state-error associations across simulated neurons.

After training, agents were tested with different rewards. Agents were tested on a wider range ( $r \sim U(0,20)$ ) of familiar (i.e.  $\sim 5$ ) and larger rewards. To determine how accurate each

agent was at estimating value, we calculated the mean squared error (MSE) between actual ( $Y$ ) and predicted ( $\hat{Y}$ ) value, for each cell at the rewarded state:

$$MSE = \frac{1}{n} \sum (Y - \hat{Y})^2 \quad (3)$$

Note that predicted value ( $\hat{Y}$ ) should approximate the reward received during testing.

## Quantification And Statistical Analysis

Continuous data are presented as means with SEM, boxplots display first quartile, median and third quartile. The Shapiro-Wilk test and the Levene test were used to judge whether data sets were normally distributed with homogeneous variances ( $p < 0.05$  to reject). For normally distributed data, a one-way ANOVA was used. If data failed normality tests, Mann-Whitney rank sum or Kruskal-Wallis one-way ANOVA on ranks with Tukey's post-hoc method for multiple comparisons were used (MATLAB, Mathworks). Significance for statistical tests was set at  $p < 0.05$ .

## Key Resources Table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse anti-Tyrosine Hydroxylase	Sigma-Aldrich	Cat# T2928, RRID:AB_477569
Chicken anti-Tyrosine Hydroxylase	Abcam	Cat# ab76442, RRID:AB_1524535
Guinea pig anti-Sox6	Michael Wegner; University of Erlangen-Nuremberg; Germany	Cat# Wegner_Sox6 gp, RRID:AB_2891329
Rabbit anti-Sox6	Abcam	Cat# ab30455, RRID:AB_1143033
Rabbit anti-Aldh1a1	Sigma-Aldrich	Cat# HPA002123, RRID:AB_1844722
Goat anti-calbindin D28K (C-20)	Santa Cruz Biotechnology	Cat# sc-7691, RRID:AB_634520
Mouse anti-calbindin	Swant	Cat# 300, RRID:AB_10000347
Rabbit anti GFAP	Proteintech	Cat# 16825-1-AP, AB_2109646
Bacterial and Virus Strains		
AAV1-CAG-flex-GCaMP6f	Addgene	Addgene #100835, RRID:Addgene_100835
AAV1-CAG-flex-tdTomato	Addgene	Addgene # 28306-AAV1, RRID:Addgene_28306
Experimental Models: Organisms/Strains		
Mouse: DATIRESc <sup>re(Slc6a3tm1.1(cre)B km)</sup>	The Jackson Laboratory	JAX006660, RRID:IMSR_JAX:006660
Mouse: C57BL/6	Charles River	Strain code: 632
Software and Algorithms		
MATLAB 2022b	MathWorks	<a href="https://www.mathworks.com/">https://www.mathworks.com/</a> , RRID: SCR_001622

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Spike2 (version 7 and 10)	Cambridge Electronic Design	<a href="https://ced.co.uk/">https://ced.co.uk/</a> , RRID:SCR_000903
PyPhotometry	<a href="https://pyphotometry.readthedocs.io/en/latest/">https://pyphotometry.readthedocs.io/en/latest/</a>	RRID:SCR_022940
ImageJ, Fiji	<a href="http://fiji.sc/">http://fiji.sc/</a>	RRID:SCR_002285
Stereo Investigator	MBF Bioscience	<a href="http://www.mbfbioscience.com/stereo-investigator">http://www.mbfbioscience.com/stereo-investigator</a> , RRID:SCR_002526

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

We thank Mark Walton for helpful advice and Joe Pemberton and Dabal Pedamonti for feedback on an early version of this manuscript. We also thank Bristol and MRC BNDU animal services staff for expert technical assistance and the Wolfson Bioimaging Facility (University of Bristol) and Jennifer Blackmore (University of Oxford) for microscopy support. This work was supported by BBSRC (BB/P006957/2 and BB/T013907/1 Awards to P.D.D), the MRC (Awards MC\_UU\_12024/2 and MC\_UU\_00003/5 to P.J.M), and a Monument Trust Discovery Award from Parkinson's UK (J-1403 to P.D.D and P.J.M). A.K.K was supported by an MRC Doctoral Training Award (MC\_ST\_U13067) and C.J.Y and G.E.P were supported by the BBSRC South West Biosciences Doctoral Training Programme (BB/T008741/1 and BB/M009122/1).

## Data and code availability

- All data reported in this paper will be shared by the lead contact upon request.
- This paper does not report original code.
- Any additional information required to reanalyse the data reported in this paper is available from the lead contact upon reasonable request.

## References

1. Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. *Science*. 1997; 275: 1593–1599. DOI: 10.1126/science.275.5306.1593 [PubMed: 9054347]
2. Eshel N, Tian J, Bukwich M, Uchida N. Dopamine neurons share common response function for reward prediction error. *Nat Neurosci*. 2016; 19: 479–486. DOI: 10.1038/nn.4239 [PubMed: 26854803]
3. Schultz W. Reward functions of the basal ganglia. *J Neural Transm*. 2016; 123: 679–693. DOI: 10.1007/s00702-016-1510-0 [PubMed: 26838982]
4. Schultz W. Dopamine reward prediction-error signalling: a two-component response. *Nat Rev Neurosci*. 2016; 17: 183–195. DOI: 10.1038/nrn.2015.26 [PubMed: 26865020]
5. Brown HD, McCutcheon JE, Cone JJ, Ragozzino ME, Roitman MF. Primary food reward and reward-predictive stimuli evoke different patterns of phasic dopamine signaling throughout the striatum. *European Journal of Neuroscience*. 2011; 34: 1997–2006. DOI: 10.1111/j.1460-9568.2011.07914.x [PubMed: 22122410]
6. de Jong JW, Afjei SA, Pollak Dorocic I, Peck JR, Liu C, Kim CK, Tian L, Deisseroth K, Lammel S. A Neural Circuit Mechanism for Encoding Aversive Stimuli in the Mesolimbic Dopamine System. *Neuron*. 2019; 101: 133–151. e7 doi: 10.1016/j.neuron.2018.11.005 [PubMed: 30503173]



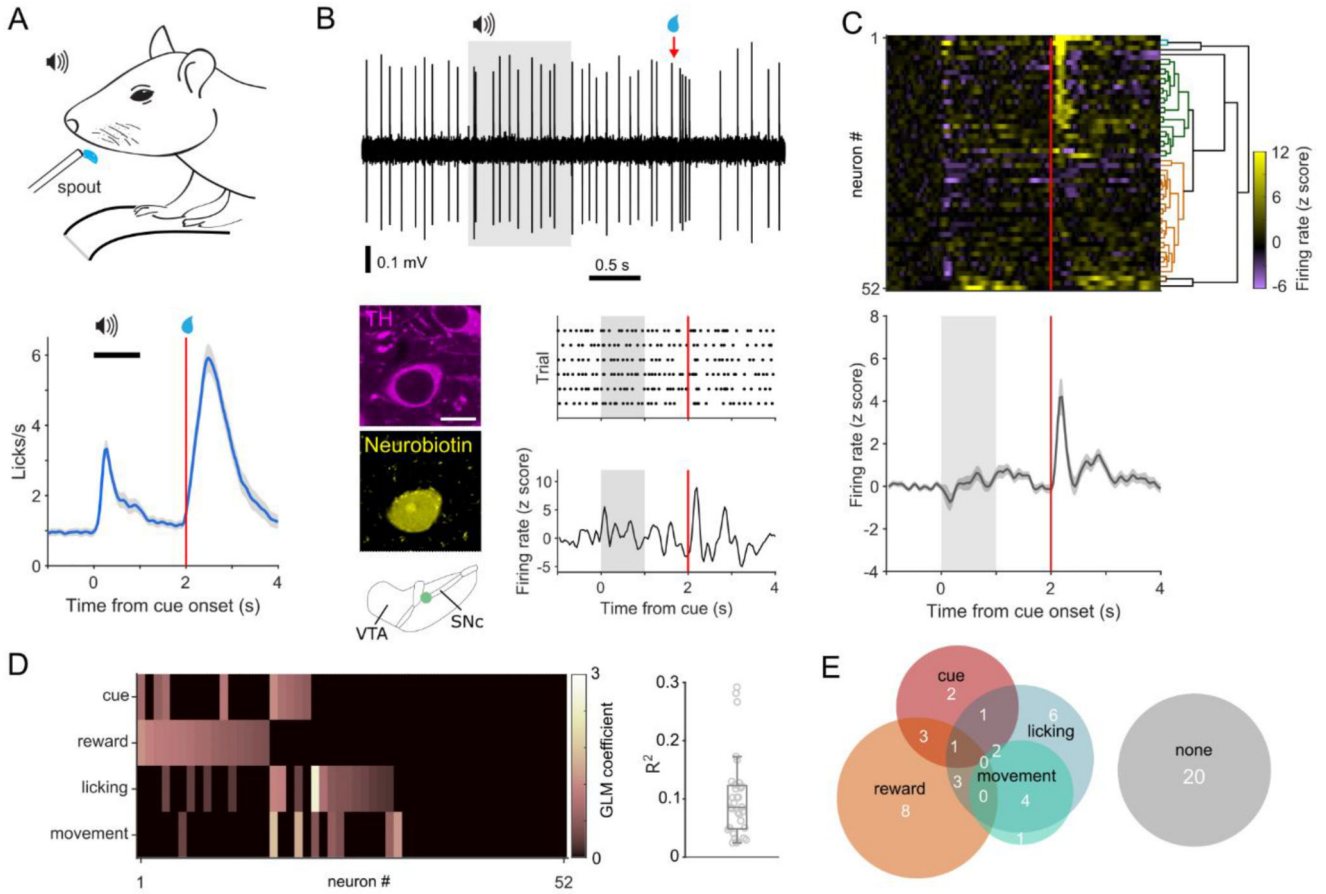
7. Parker NF, Cameron CM, Taliaferro JP, Lee J, Choi JY, Davidson TJ, Daw ND, Witten IB. Reward and choice encoding in terminals of midbrain dopamine neurons depends on striatal target. *Nat Neurosci.* 2016; 19: 845–854. DOI: 10.1038/nn.4287 [PubMed: 27110917]
8. Saddoris MP, Cacciapaglia F, Wightman RM, Carelli RM. Differential Dopamine Release Dynamics in the Nucleus Accumbens Core and Shell Reveal Complementary Signals for Error Prediction and Incentive Motivation. *J Neurosci.* 2015; 35: 11572–11582. DOI: 10.1523/JNEUROSCI.2344-15.2015 [PubMed: 26290234]
9. Tsutsui-Kimura I, Matsumoto H, Akiti K, Yamada MM, Uchida N, Watabe-Uchida M. Distinct temporal difference error signals in dopamine axons in three regions of the striatum in a decision-making task. *Elife.* 2020; 9: 1–39. DOI: 10.7554/eLife.62390
10. Yuan L, Dou YN, Sun YG. Topography of Reward and Aversion Encoding in the Mesolimbic Dopaminergic System. *J Neurosci.* 2019; 39: 6472–6481. DOI: 10.1523/JNEUROSCI.0271-19.2019 [PubMed: 31217328]
11. van Elzelingen W, Goedhoop J, Warnaar P, Denys D, Arbab T, Willuhn I. A unidirectional but not uniform striatal landscape of dopamine signaling for motivational stimuli. *Proceedings of the National Academy of Sciences.* 2022; 119: 1–12. DOI: 10.1073/pnas.2117270119
12. Hamid AA, Frank MJ, Moore CI. Wave-like dopamine dynamics as a mechanism for spatiotemporal credit assignment. *Cell.* 2021; 184: 2733–2749. e16 doi: 10.1016/j.cell.2021.03.046 [PubMed: 33861952]
13. Howe MW, Dombeck DA. Rapid signalling in distinct dopaminergic axons during locomotion and reward. *Nature.* 2016; 27: 1–22. DOI: 10.1038/nature18942
14. Dodson PD, Dreyer JK, Jennings KA, Syed ECJ, Wade-Martins R, Cragg SJ, Bolam JP, Magill PJ. Representation of spontaneous movement by dopaminergic neurons is cell-type selective and disrupted in parkinsonism. *Proceedings of the National Academy of Sciences.* 2016; 113: E2180–E2188. DOI: 10.1073/pnas.1515941113
15. Coddington LT, Dudman JT. The timing of action determines reward prediction signals in identified midbrain dopamine neurons. *Nat Neurosci.* 2018; 21: 1563–1573. DOI: 10.1038/s41593-018-0245-7 [PubMed: 30323275]
16. Engelhard B, Finkelstein J, Cox J, Fleming W, Jang HJ, Ornelas S, Koay SA, Thiberge SY, Daw ND, Tank DW, et al. Specialized coding of sensory, motor and cognitive variables in VTA dopamine neurons. *Nature.* 2019; 570: 509–513. DOI: 10.1038/s41586-019-1261-9 [PubMed: 31142844]
17. da Silva JA, Tecuapetla F, Paixão V, Costa RM. Dopamine neuron activity before action initiation gates and invigorates future movements. *Nature.* 2018; 554: 244–248. DOI: 10.1038/nature25457 [PubMed: 29420469]
18. Hughes RN, Bakhurin KI, Petter EA, Watson GDR, Kim N, Friedman AD, Yin HH. Ventral Tegmental Dopamine Neurons Control the Impulse Vector during Motivated Behavior. *Current Biology.* 2020; 30: 2681–2694. e5 doi: 10.1016/j.cub.2020.05.003 [PubMed: 32470362]
19. Barter JW, Li S, Lu D, Bartholomew Ra, Rossi Ma, Shoemaker CT, Salas-Meza D, Gaidis E, Yin HH. Beyond reward prediction errors: the role of dopamine in movement kinematics. *Front Integr Neurosci.* 2015; 9: 39. doi: 10.3389/fnint.2015.00039 [PubMed: 26074791]
20. Kremer Y, Flakowski J, Rohner C, Lüscher C. Context-Dependent Multiplexing by Individual VTA Dopamine Neurons. *J Neurosci.* 2020; 40: 7489–7509. DOI: 10.1523/JNEUROSCI.0502-20.2020 [PubMed: 32859713]
21. Poulin J-F, Zou J, Drouin-Ouellet J, Kim K-YA, Cicchetti F, Awatramani RB. Defining Midbrain Dopaminergic Neuron Diversity by Single-Cell Gene Expression Profiling. *Cell Rep.* 2014; 9: 930–943. DOI: 10.1016/j.celrep.2014.10.008 [PubMed: 25437550]
22. La Manno G, Gyllborg D, Codeluppi S, Nishimura K, Salto C, Zeisel A, Borm LE, Stott SRW, Toledo EM, Villaescusa JC, et al. Molecular Diversity of Midbrain Development in Mouse, Human, and Stem Cells. *Cell.* 2016; 167: 566–580. e19 doi: 10.1016/j.cell.2016.09.027 [PubMed: 27716510]
23. Tiklová K, Björklund ÅK, Lahti L, Fiorenzano A, Nolbrant S, Gillberg L, Volakakis N, Yokota C, Hilscher MM, Hauling T, et al. Single-cell RNA sequencing reveals midbrain dopamine neuron

- diversity emerging during mouse brain development. *Nat Commun.* 2019; 10: 581. doi: 10.1038/s41467-019-08453-1 [PubMed: 30718509]
24. Saunders A, Macosko EZ, Wysoker A, Goldman M, Krienen FM, de Rivera H, Bien E, Baum M, Bortolin L, Wang S, et al. Molecular Diversity and Specializations among the Cells of the Adult Mouse Brain. *Cell.* 2018; 174: 1015–1030. e16 doi: 10.1016/j.cell.2018.07.028 [PubMed: 30096299]
  25. Hook PW, McClymont SA, Cannon GH, Law WD, Morton AJ, Goff LA, McCallion AS. Single-Cell RNA-Seq of Mouse Dopaminergic Neurons Informs Candidate Gene Selection for Sporadic Parkinson Disease. *Am J Hum Genet.* 2018; 102: 427–446. DOI: 10.1016/j.ajhg.2018.02.001 [PubMed: 29499164]
  26. Poulin JF, Gaertner Z, Moreno-Ramos OA, Awatramani R. Classification of Midbrain Dopamine Neurons Using Single-Cell Gene Expression Profiling Approaches. *Trends Neurosci.* 2020; 43: 155–169. DOI: 10.1016/j.tins.2020.01.004 [PubMed: 32101709]
  27. Garritsen O, van Battum EY, Grossouw LM, Pasterkamp RJ. Development, wiring and function of dopamine neuron subtypes. *Nat Rev Neurosci.* 2023; doi: 10.1038/s41583-022-00669-3
  28. Schiemann J, Schlaudraff F, Klose V, Bingmer M, Seino S, Magill PJ, Zaghoul Ka, Schneider G, Liss B, Roeper J. K-ATP channels in dopamine substantia nigra neurons control bursting and novelty-induced exploration. *Nat Neurosci.* 2012; 15: 1272–1280. DOI: 10.1038/nn.3185 [PubMed: 22902720]
  29. Roeper J. Dissecting the diversity of midbrain dopamine neurons. *Trends Neurosci.* 2013; 36: 336–342. DOI: 10.1016/j.tins.2013.03.003 [PubMed: 23582338]
  30. Lerner TN, Shilyansky C, Davidson TJ, Evans KE, Beier KT, Zalocusky KA, Crow AK, Malenka RC, Luo L, Tomer R, et al. Intact-Brain Analyses Reveal Distinct Information Carried by SNc Dopamine Subcircuits. *Cell.* 2015; 162: 635–647. DOI: 10.1016/j.cell.2015.07.014 [PubMed: 26232229]
  31. Lammel S, Hetzel A, Häckel O, Jones I, Liss B, Roeper J. Unique properties of mesoprefrontal neurons within a dual mesocorticolimbic dopamine system. *Neuron.* 2008; 57: 760–773. DOI: 10.1016/j.neuron.2008.01.022 [PubMed: 18341995]
  32. Lammel S, Ion DI, Roeper J, Malenka RC. Projection-specific modulation of dopamine neuron synapses by aversive and rewarding stimuli. *Neuron.* 2011; 70: 855–862. DOI: 10.1016/j.neuron.2011.03.025 [PubMed: 21658580]
  33. Watabe-Uchida M, Zhu L, Ogawa SK, Vamanrao A, Uchida N. Whole-brain mapping of direct inputs to midbrain dopamine neurons. *Neuron.* 2012; 74: 858–873. DOI: 10.1016/j.neuron.2012.03.017 [PubMed: 22681690]
  34. Farassat N, Kauê Machado Costa, Stovanovic S, Albert S, Kovacheva L, Shin J, Egger R, Somayaji M, Duvarci S, Schneider G, et al. In vivo functional diversity of midbrain dopamine neurons within identified axonal projections. *Elife.* 2019; 8: 1–27. DOI: 10.7554/elife.48408
  35. Morales M, Margolis EB. Ventral tegmental area: cellular heterogeneity, connectivity and behaviour. *Nat Rev Neurosci.* 2017; 18: 73–85. DOI: 10.1038/nrn.2016.165 [PubMed: 28053327]
  36. Balleine BW, O’Doherty JP. Human and rodent homologues in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology.* 2010; 35: 48–69. DOI: 10.1038/npp.2009.131 [PubMed: 19776734]
  37. Tsutsui-Kimura I, Natsubori A, Mori M, Kobayashi K, Drew MR, de Kerchove d’Exaerde A, Mimura M, Tanaka KF. Distinct Roles of Ventromedial versus Ventrolateral Striatal Medium Spiny Neurons in Reward-Oriented Behavior. *Current Biology.* 2017; 27: 3042–3048. e4 doi: 10.1016/j.cub.2017.08.061 [PubMed: 28966085]
  38. Saunders BT, Richard JM, Margolis EB, Janak PH. Dopamine neurons create Pavlovian conditioned stimuli with circuit-defined motivational properties. *Nat Neurosci.* 2018; 21: 1072–1083. DOI: 10.1038/s41593-018-0191-4 [PubMed: 30038277]
  39. Schultz W. Behavioral dopamine signals. *Trends Neurosci.* 2007; 30: 203–210. DOI: 10.1016/j.tins.2007.03.007 [PubMed: 17400301]
  40. Coddington LT, Dudman JT. Learning from Action: Reconsidering Movement Signaling in Midbrain Dopamine Neuron Activity. *Neuron.* 2019; 104: 63–77. DOI: 10.1016/j.neuron.2019.08.036 [PubMed: 31600516]

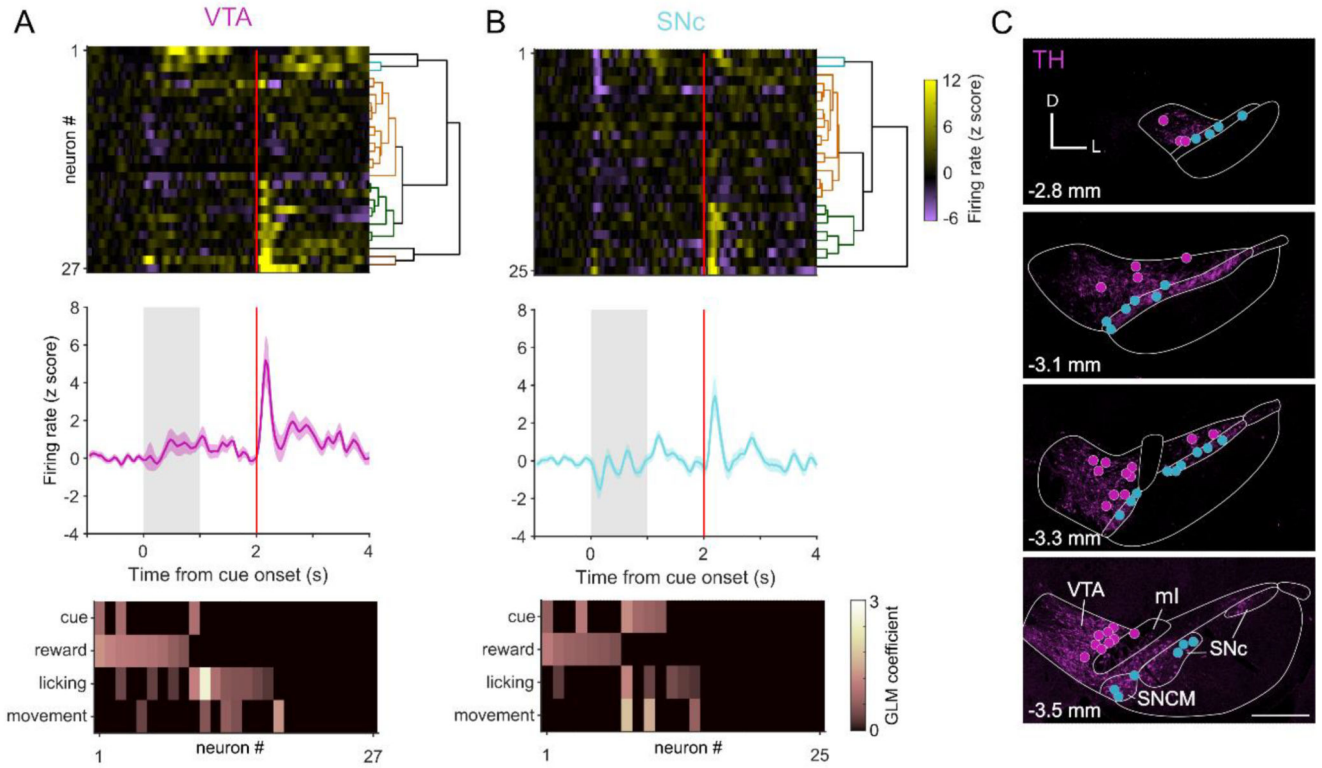
41. Cox J, Witten IB. Striatal circuits for reward learning and decision-making. *Nat Rev Neurosci*. 2019; 20 doi: 10.1038/s41583-019-0189-2
42. Collins AL, Saunders BT. Heterogeneity in striatal dopamine circuits: Form and function in dynamic reward seeking. *J Neurosci Res*. 2020; 98: 1046–1069. DOI: 10.1002/jnr.24587 [PubMed: 32056298]
43. Lerner TN, Holloway AL, Seiler JL. Dopamine, Updated: Reward Prediction Error and Beyond. *Curr Opin Neurobiol*. 2021; 67: 123–130. DOI: 10.1016/j.conb.2020.10.012 [PubMed: 33197709]
44. Dabney W, Kurth-Nelson Z, Uchida N, Starkweather CK, Hassabis D, Munos R, Botvinick M. A distributional code for value in dopamine-based reinforcement learning. *Nature*. 2020; 577: 671–675. DOI: 10.1038/s41586-019-1924-6 [PubMed: 31942076]
45. Matsumoto M, Hikosaka O. Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature*. 2009; 459: 837–841. DOI: 10.1038/nature08028 [PubMed: 19448610]
46. Azcorra M, Gaertner Z, Davidson C, He Q, Kim H, Nagappan S, Hayes CK, Ramakrishnan C, Fenno L, Kim YS, et al. Unique functional responses differentially map onto genetic subtypes of dopamine neurons. *Nature Neuroscience*. 2023; 26 (10) 1762–1774. DOI: 10.1038/s41593-023-01401-9 [PubMed: 37537242]
47. Flagel SB, Clark JJ, Robinson TE, Mayo L, Czuj A, Willuhn I, Akers Ca, Clinton SM, Phillips PEM, Akil H. A selective role for dopamine in stimulus-reward learning. *Nature*. 2011; 469: 53–57. DOI: 10.1038/nature09588 [PubMed: 21150898]
48. Schultz W. Responses of midbrain dopamine neurons to behavioral trigger stimuli in the monkey. *J Neurophysiol*. 1986. 1439–1461. [PubMed: 3794777]
49. Schultz W, Romo R. Dopamine neurons of the monkey midbrain: contingencies of responses to stimuli eliciting immediate behavioral reactions. *J Neurophysiol*. 1990; 63: 607–624. [PubMed: 2329364]
50. Horvitz J. Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. *Neuroscience*. 2000; 96: 651–656. [PubMed: 10727783]
51. Bromberg-Martin ES, Matsumoto M, Hikosaka O. Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron*. 2010; 68: 815–834. DOI: 10.1016/j.neuron.2010.11.022 [PubMed: 21144997]
52. Kutlu MG, Zachry JE, Melugin PR, Cajigas SA, Chevee MF, Kelly SJ, Kutlu B, Tian L, Siciliano CA, Calipari ES. Dopamine release in the nucleus accumbens core signals perceived saliency. *Current Biology*. 2021; 1–14. DOI: 10.1016/j.cub.2021.08.052
53. Strecker RE, Jacobs BL. Substantia nigra dopaminergic unit activity in behaving cats: Effect of arousal on spontaneous discharge and sensory evoked activity. *Brain Res*. 1985; 361: 339–350. DOI: 10.1016/0006-8993(85)91304-6 [PubMed: 4084803]
54. de Lafuente V, Romo R. Dopamine neurons code subjective sensory experience and uncertainty of perceptual decisions. *Proceedings of the National Academy of Sciences*. 2011; 108: 19767–19771. DOI: 10.1073/pnas.1117636108
55. Fiorillo C, Song M, Yun S. Multiphasic Temporal Dynamics in Responses of Midbrain Dopamine Neurons to Appetitive and Aversive Stimuli. *The Journal of Neuroscience*. 2013; 33: 4710–4725. DOI: 10.1523/JNEUROSCI.3883-12.2013 [PubMed: 23486944]
56. Tobler PN, Fiorillo CD, Schultz W. Adaptive Coding of Reward Value by Dopamine Neurons. *Science (1979)*. 2005; 307: 1642–1645. DOI: 10.1126/science.1105370
57. Barkus C, Bergmann C, Branco T, Carandini M, Chadderton PT, Galiñanes GL, Gilmour G, Huber D, Huxter JR, Khan AG, et al. Refinements to rodent head fixation and fluid/food control for neuroscience. *J Neurosci Methods*. 2022; 381 doi: 10.1016/J.JNEUMETH.2022.109705
58. Dayan P, Balleine BW. Reward, Motivation, and Reinforcement Learning. *Neuron*. 2002; 36: 285–298. DOI: 10.1016/S0896-6273(02)00963-7 [PubMed: 12383782]
59. Satoh T, Nakai S, Sato T, Kimura M. Correlated coding of motivation and outcome of decision by dopamine neurons. *J Neurosci*. 2003; 23: 9913–9923. DOI: 10.1523/JNEUROSCI.23-30-09913.2003 [PubMed: 14586021]

60. Mohebi A, Pettibone JR, Hamid AA, Wong JMT, Vinson LT, Patriarchi T, Tian L, Kennedy RT, Berke JD. Dissociable dopamine dynamics for learning and motivation. *Nature*. 2019; 570: 65–70. DOI: 10.1038/s41586-019-1235-y [PubMed: 31118513]
61. Heymann G, Jo YS, Reichard KL, McFarland N, Chavkin C, Palmiter RD, Soden ME, Zweifel LS. Synergy of Distinct Dopamine Projection Populations in Behavioral Reinforcement. *Neuron*. 2019; 1–12. DOI: 10.1016/j.neuron.2019.11.024
62. Yin HH, Knowlton BJ, Balleine BW. Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *European Journal of Neuroscience*. 2004; 19: 181–189. DOI: 10.1111/J.1460-9568.2004.03095.X [PubMed: 14750976]
63. Thorn CA, Atallah H, Howe M, Graybiel AM. Differential Dynamics of Activity Changes in Dorsolateral and Dorsomedial Striatal Loops during Learning. *Neuron*. 2010; 66: 781–795. DOI: 10.1016/j.neuron.2010.04.036 [PubMed: 20547134]
64. Tritsch NX, Granger AJ, Sabatini BL. Mechanisms and functions of GABA co-release. *Nature Reviews Neuroscience*. 2016; 17 (3) 139–145. DOI: 10.1038/nrn.2015.21 [PubMed: 26865019]
65. Eskenazi D, Malave L, Mingote S, Yetnikoff L, Ztaou S, Velicu V, Rayport S, Chuhma N. Dopamine Neurons That Cotransmit Glutamate, From Synapses to Circuits to Behavior. *Front Neural Circuits*. 2021; 15 665386 doi: 10.3389/FNCIR.2021.665386 [PubMed: 34093138]
66. Cragg SJ, Rice ME. DANCING past the DAT at a DA synapse. *Trends Neurosci*. 2004; 27: 270–277. DOI: 10.1016/j.tins.2004.03.011 [PubMed: 15111009]
67. Graybiel AM. Correspondence between the dopamine islands and striosomes of the mammalian striatum. *Neuroscience*. 1984; 13: 1157–1187. DOI: 10.1016/0306-4522(84)90293-8 [PubMed: 6152035]
68. Wu J, Kung J, Dong J, Chang L, Xie C, Habib A, Hawes S, Yang N, Chen V, Liu Z, et al. Distinct Connectivity and Functionality of Aldehyde Dehydrogenase 1a1-Positive Nigrostriatal Dopaminergic Neurons in Motor Learning. *Cell Rep*. 2019; 28: 1167–1181. e7 doi: 10.1016/j.celrep.2019.06.095 [PubMed: 31365862]
69. Poulin J, Caronia G, Hofer C, Cui Q, Helm B, Ramakrishnan C, Chan CS, Dombeck DA, Deisseroth K, Awatramani R. Mapping projections of molecularly defined dopamine neuron subtypes using intersectional genetic approaches. *Nat Neurosci*. 2018; 21: 1260–1271. DOI: 10.1038/s41593-018-0203-4 [PubMed: 30104732]
70. Pereira Luppi M, Azcorra M, Caronia-Brown G, Poulin J-F, Gaertner Z, Gatica S, Moreno-Ramos OA, Nouri N, Dubois M, Ma YC, et al. Sox6 expression distinguishes dorsally and ventrally biased dopamine neurons in the substantia nigra with distinctive properties and embryonic origins. *Cell Rep*. 2021; 37 109975 doi: 10.1016/j.celrep.2021.109975 [PubMed: 34758317]
71. Janezic S, Threlfell S, Dodson PD, Dowie MJ, Taylor TN, Potgieter D, Parkkinen L, Senior SL, Anwar S, Ryan B, et al. Deficits in dopaminergic transmission precede neuron loss and dysfunction in a new Parkinson model. *Proc Natl Acad Sci U S A*. 2013; 110: E4016–25. DOI: 10.1073/pnas.1309143110 [PubMed: 24082145]
72. Fu Y, Yuan Y, Halliday G, Rusznák Z, Watson C, Paxinos G. A cytoarchitectonic and chemoarchitectonic analysis of the dopamine cell groups in the substantia nigra, ventral tegmental area, and retrorubral field in the mouse. *Brain Struct Funct*. 2012; 217: 591–612. DOI: 10.1007/s00429-011-0349-2 [PubMed: 21935672]
73. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, et al. Fiji: an open-source platform for biological-image analysis. *Nat Methods*. 2012; 9: 676–682. DOI: 10.1038/nmeth.2019 [PubMed: 22743772]
74. Akam T, Walton ME. pyPhotometry: Open source Python based hardware and software for fiber photometry data acquisition. *Scientific Reports*. 2019; 9 (1) 1–11. DOI: 10.1038/s41598-019-39724-y [PubMed: 30626917]
75. Holt G, Softky W, Koch C, Douglas R. Comparison of discharge variability in vitro and in vivo in cat visual cortex neurons. *J Neurophysiol*. 1996; 75: 1806–1814. [PubMed: 8734581]
76. Ko D, Wilson CJ, Lobb CJ, Paladini CA. Detection of bursts and pauses in spike trains. *J Neurosci Methods*. 2012; 211: 145–158. DOI: 10.1016/j.jneumeth.2012.08.013 [PubMed: 22939922]
77. Sloan M, Alegre-Abarrategui J, Potgieter D, Kaufmann A-KK, Exley R, Deltheil T, Threlfell S, Connor-Robson N, Brimblecombe K, Wallings R, et al. LRRK2 BAC transgenic rats develop

- progressive, L-DOPA-responsive motor impairment, and deficits in dopamine circuit function. *Hum Mol Genet.* 2016; 25: 951–963. DOI: 10.1093/hmg/ddv628 [PubMed: 26744332]
78. Sutton RS. Learning to predict by the methods of temporal differences. *Mach Learn.* 1988; 3: 9–44. DOI: 10.1007/BF00115009

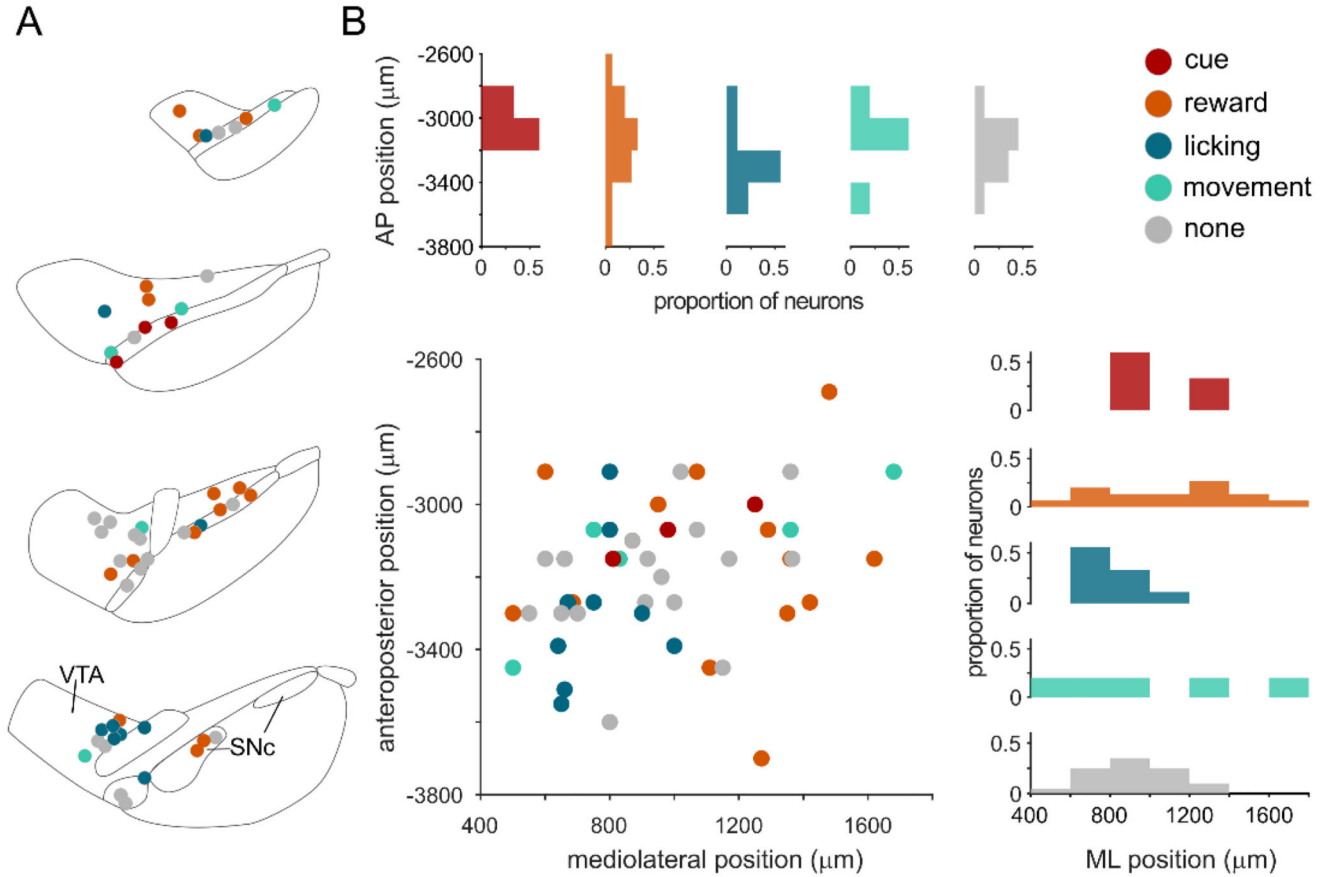


**Figure 1. Dopamine neurons heterogeneously encode reward and predictive cues.** (A) Head-fixed mice, positioned on a treadmill, were trained to associate a 4 kHz auditory tone (cue) with delivery of reward. After training, mice show robust anticipatory licking after the cue and licking to receive reward. (B) Extracellular recording of action potential firing (*top*) and corresponding peri-stimulus time histogram (PSTH; lower right) from an individual dopamine neuron during Pavlovian conditioned behavior. Grey shading indicates cue duration, red line indicates reward delivery. After recording, individual neurons were juxtacellularly labeled with neurobiotin and tested for immunoreactivity to tyrosine hydroxylase (TH: lower left; scale = 10  $\mu$ m) to confirm their dopaminergic identity and localization in the midbrain. The schematic depicts the location of the neuron in the dopaminergic midbrain. (C) Z-scored PSTH of individual responses from identified dopamine neurons (*rows*) grouped by hierarchical clustering (*top*), and mean response (*bottom*). (D) Features which correlate with changes in firing rate for each neuron, determined by a general linear model (GLM). (E) GLM output suggests that some neurons multiplex signals by encoding multiple aspects of behavior (numbers indicate the proportion of neurons; note that some combinations are unable to be displayed). N = 52 neurons from 30 mice.



**Figure 2. Anatomical subgroups do not account for response heterogeneity.**

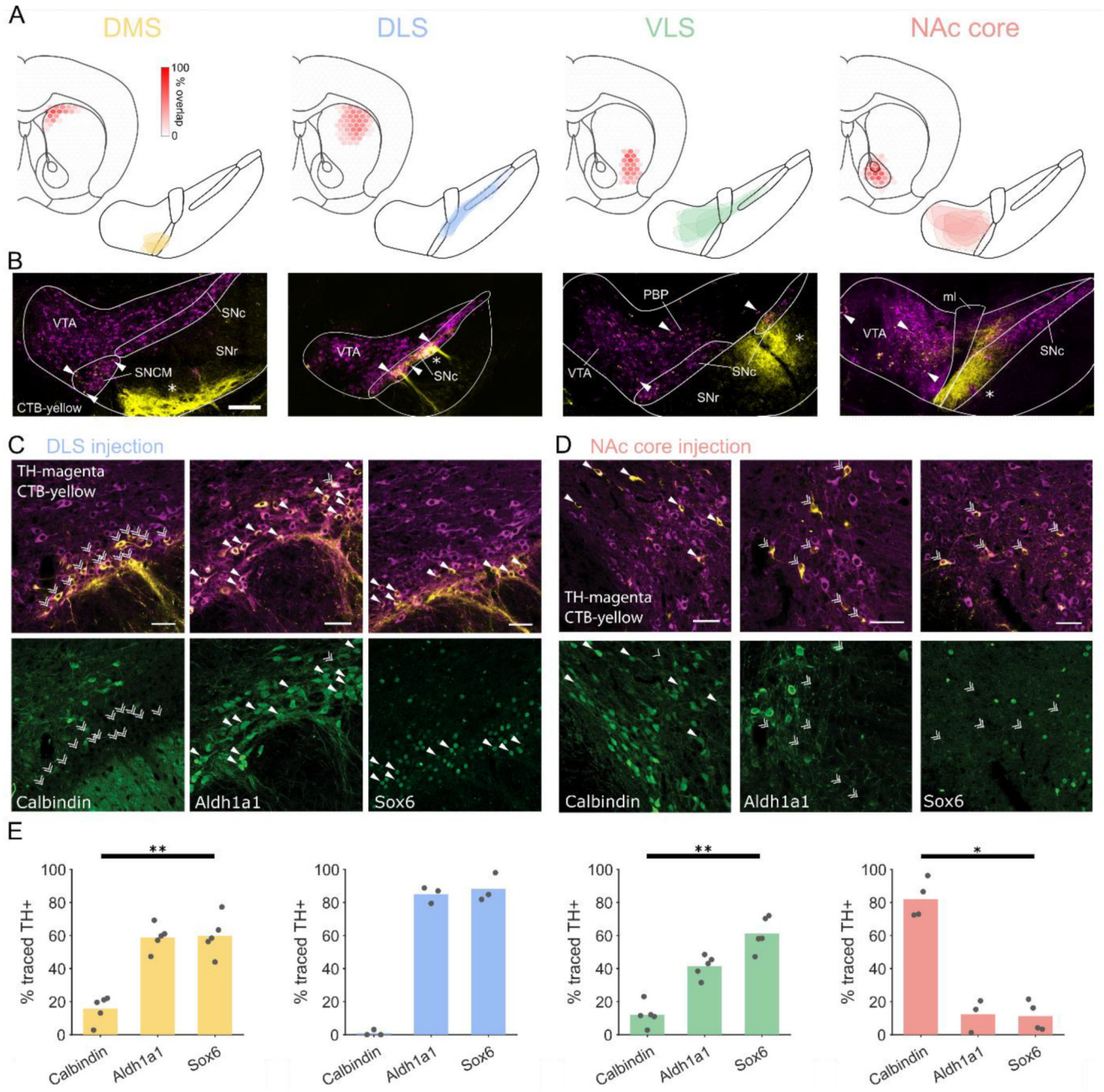
(A) VTA neurons heterogeneously encoded aspects of Pavlovian conditioned behavior (top) and firing could be classified as encoding multiple parameters by GLM (bottom). (B) SNc neurons exhibit a similar degree of heterogeneity in firing responses. (C) Schematic of the locations of recorded and juxtacellularly labeled neurons (VTA neurons magenta, SNc cyan) overlaid on example images of tyrosine hydroxylase (TH) immunofluorescence at four anteroposterior positions 2.8 to 3.5 mm from bregma. D = dorsal; L = lateral; VTA = ventral tegmental area; SNc = substantia nigra pars compacta; ml=medial lemniscus; SNCM=medial part of SNc. Scale = 500  $\mu$ m. N = 52 neurons from 30 mice.



**Figure 3. Encoding of behavioral parameters is not focally organized.**

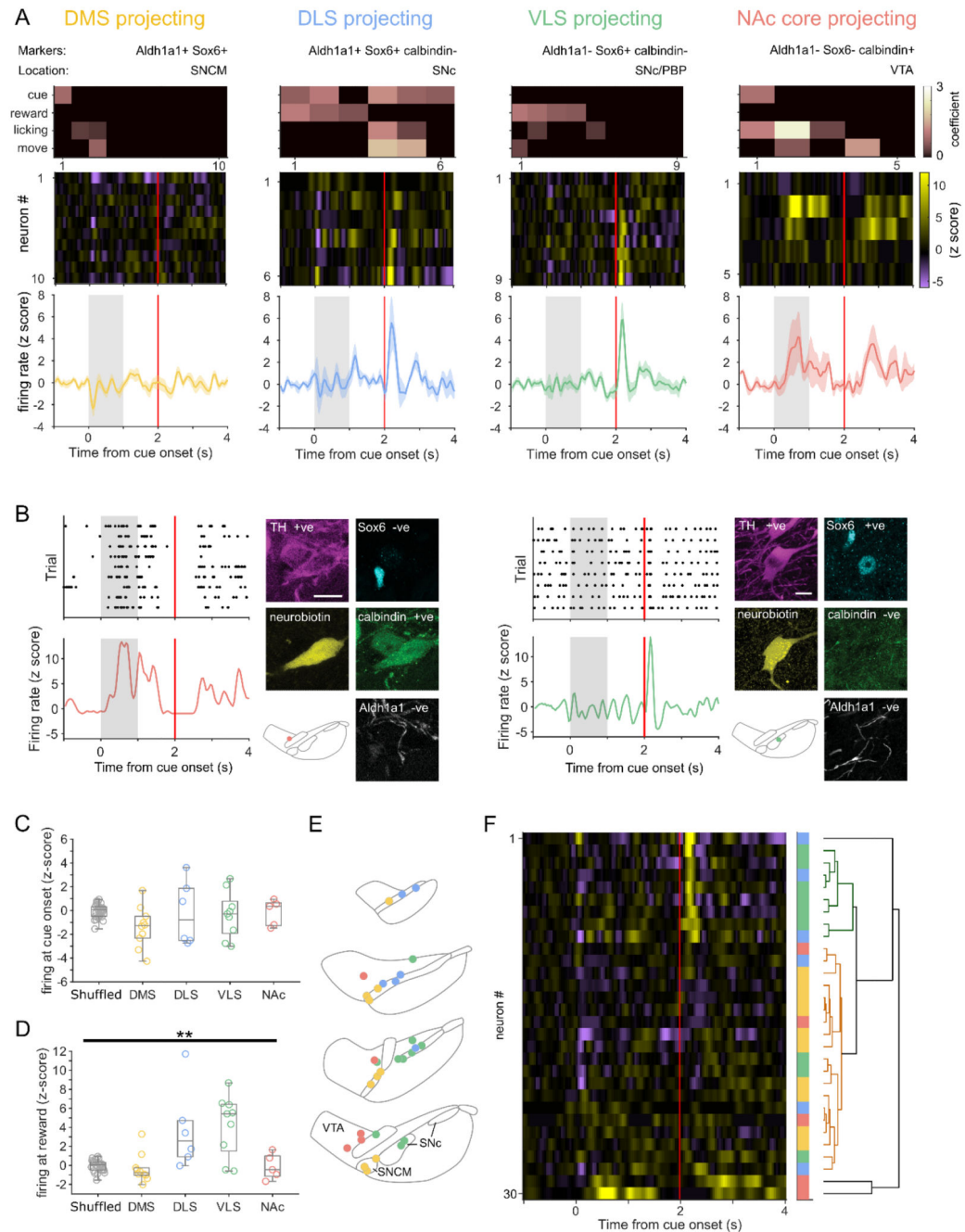
(A) Schematic of the dominant parameter encoded by each dopamine neuron according to their location in the midbrain. Note some of these neurons will also encode other parameters (e.g. cue in addition to reward). (B) The dominant parameter plotted according to the neurons anteroposterior (AP) or mediolateral (ML) position with respect to bregma. The proportion of neurons encoding each dominant parameter is shown in the AP (top) and ML plane (right). N=52 neurons from 30 mice.





**Figure 4. Differential marker expression in populations projecting to different parts of striatum.** (A) Schematics showing the extent of retrograde tracer injections into different striatal regions (top left: DMS, dorsomedial striatum; DLS, dorsolateral striatum; VLS, ventrolateral striatum; NAc core, core of the nucleus accumbens) and corresponding boundaries of CTB-labelled dopamine neurons in the midbrain (bottom right). Colored hexagons represent overlap of cholera toxin B (CTB) injections between animals. (B) Representative midbrain images following CTB injection into different striatal regions. Selected retrogradely traced dopamine neurons (filled arrowheads) illustrate the region of the dopaminergic midbrain which projects to each striatal target. Anterogradely traced fibers (asterisks) from striatal

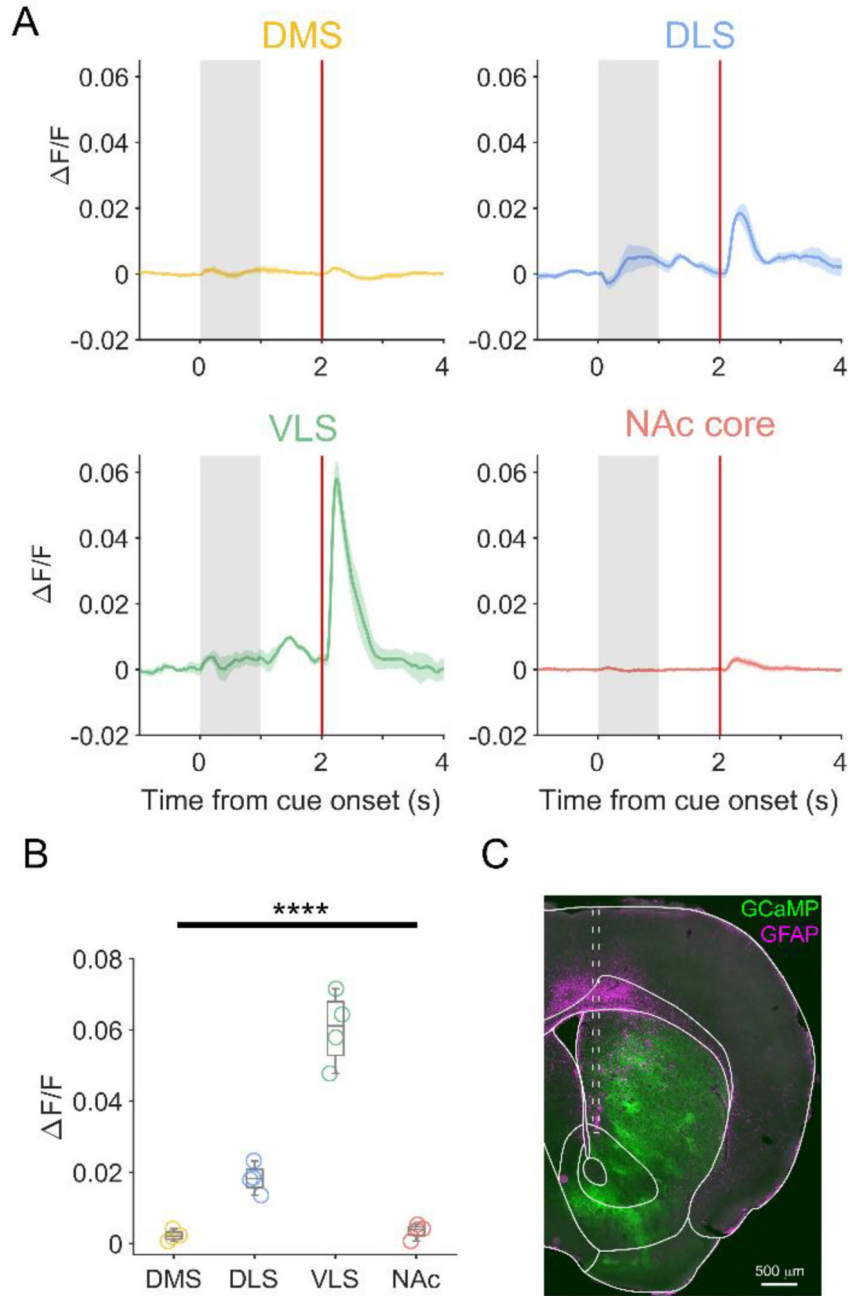
projection neurons are also visible in the SNc and SNr. VTA = ventral tegmental area; SNc = substantia nigra pars compacta; SNr = substantia nigra pars reticulata; ml = medial lemniscus; SNCM = medial part of SNc; PBP = parabrachial pigmented area of the VTA. Scale = 100  $\mu$ m. (C) Retrograde tracing with CTB injected into the DLS identified SNc dopamine neurons (positive for tyrosine hydroxylase, TH; top panels) which innervate this region. The majority of these neurons expressed Aldh1a1 and Sox6, but not calbindin. Filled arrowheads represent marker expression, double arrowheads, lack of expression. Scale = 50  $\mu$ m. (D) Retrograde tracing in NAc core identified VTA dopamine neurons expressing calbindin, but not Aldh1a1 or Sox6. Scale = 50  $\mu$ m. (E) Cell counting of CTB positive dopaminergic neurons revealed the percentage of neurons which express each marker. Each data point represents the counts from a single tracer injection. Kruskal-Wallis one-way ANOVA on ranks with Tukey's post-hoc test; \*  $p < 0.05$ , \*\*  $p < 0.01$ . N = 17 counts for calbindin and Sox6, 16 for Aldh1a1 from 17 animals.



**Figure 5. Neurons projecting to different striatal regions exhibit distinct responses.**

(A) mean PSTH, and individual heatmap plots for neurons putatively projecting to different striatal regions according to classifying marker combinations and location (bottom). Corresponding GLM plots show responses of each neuron to different behavioral parameters (top). (B) PSTH and marker expression of an example putative NAc core-projecting VTA neuron (left) and VLS-projecting SNc neuron (right). Scale = 10  $\mu$ m. (C) Firing at cue onset (0 to 240 ms after cue onset) for neurons projecting to each striatal region. There were no significant differences between regions or compared to shuffled baseline periods.

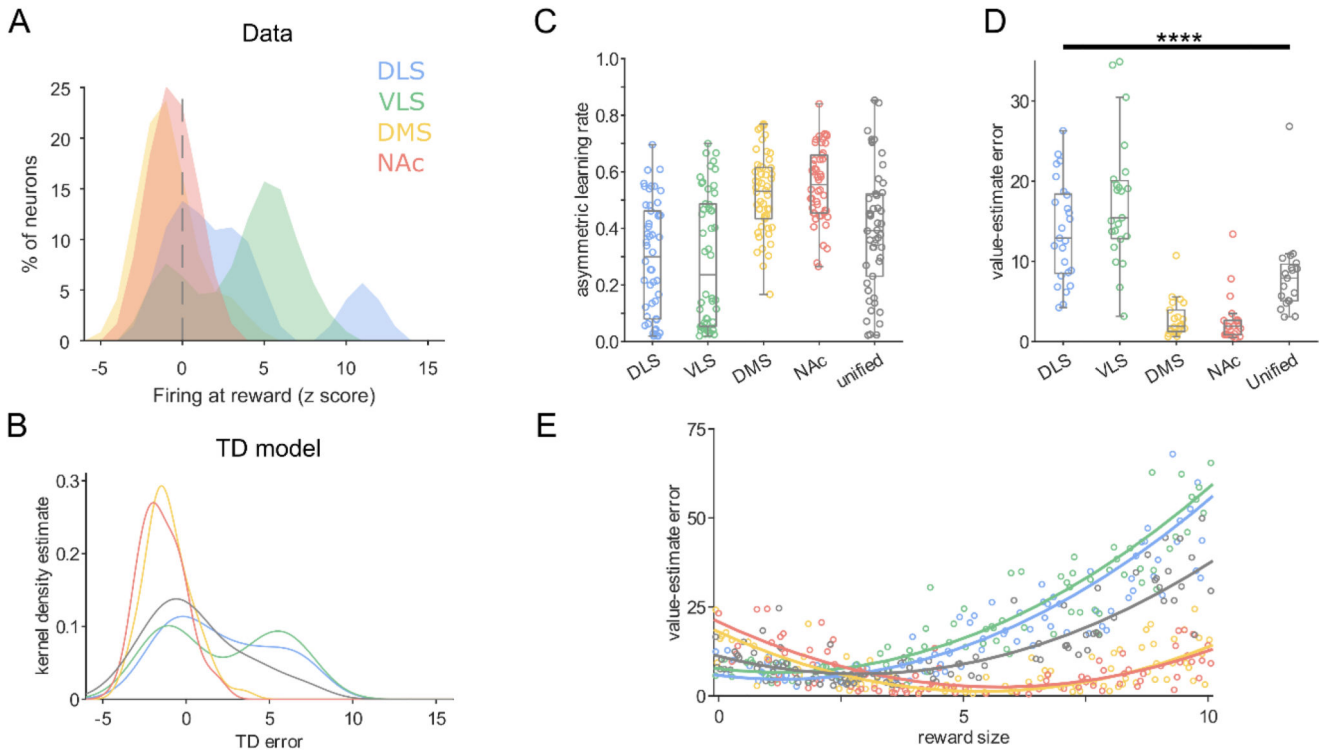
(D) Mean firing at reward (40 to 240 ms after reward delivery) for neurons projecting to each striatal region. Putative DLS- and VLS-projecting neurons showed significantly higher firing rates than DMS-projecting neurons ( $P < 0.05$  and  $P < 0.005$  respectively) and compared to shuffled baseline periods ( $P < 0.05$  and  $P < 0.05$ ), Kruskal-Wallis one-way ANOVA on ranks with Tukey's post-hoc test; \*\*  $p < 0.01$ . (E) Schematic of the location of recorded neurons, color coded by their putative projection-targets. (F) Hierarchically clustered individual responses (colored blocks denote putative projection target).  $N = 30$  neurons from 19 mice.



**Figure 6. Distinct reward-related signaling in different striatal regions.**

(A) Mean PSTH of the change in GCaMP6f fluorescence ( $\Delta F/F$ ) in dopaminergic axons in DMS (yellow), DLS (blue), VLS (green), and NAc core (red) recorded with fiber photometry. (B) Mean peak change in fluorescence at reward in each striatal region. DLS and VLS exhibited distinct signals at reward compared to each other population; there was no significant difference between signals in DMS and NAc ( $P > 0.05$  one way ANOVA with Tukey's post-hoc test; \*\*\*\*  $p < 0.0001$ ). (C) Representative image showing the expression

of GCaMP6f and the track (dashed line) of the optic fiber targeted to NAc core. N = 4 recording sessions in 4 mice, (1 recording session per region, per mouse).



**Figure 7. Different distributional coding of reward prediction error supports specialized roles in behavior.**

(A) Histogram of firing rate at reward for neurons in each putative projection-defined population. (B) Kernel density estimates from the temporal difference (TD) model represent the distribution of TD errors made by the ‘neurons’ comprising each projection-defined agent. The grey line represents a ‘unified population’ generated by sampling from all dopamine neurons in (a). (C) Asymmetric learning rates ( $\alpha_+ / (\alpha_+ + \alpha_-)$ ; balanced learning rates = 0.5) of ‘neurons’ fit from the data in (a) that comprise each agent (N = 50 per agent). (D) The difference between predicted and actual state-values (the mean squared error) to rewards near the center point (reward size of 5) of the training distribution, for each projection-defined agent. DLS- and VLS-projecting populations made significantly worse value-estimates than DMS- or NAc-projecting populations. DMS- and NAc-projecting populations were more accurate than the unified population whereas VLS made significantly larger errors. \*\*\*\* $p < 0.0001$  Kruskal-Wallis one-way ANOVA on ranks with Tukey’s post-hoc test. (E) Value-estimate errors of each agent for different numerical reward sizes.