UNIVERSITY OF CALIFORNIA

Los Angeles

Physiological constraints on the rapid dopaminergic modulation of striatal reward activity

A dissertation submitted in partial satisfaction of the

requirements for the degree

Doctor of Philosophy in Neuroscience

by

Charltien Long

2024

ABSTRACT OF THE DISSERTATION

Physiological constraints on the rapid dopaminergic modulation of striatal reward activity

by

Charltien Long

Doctor of Philosophy in Neuroscience

University of California, Los Angeles, 2024

Professor Sotirios Masmanidis, Chair

While the crucial role of dopaminergic (DA) neurons in associative learning is firmly established, there is less consensus about whether they also play a major regulatory function in behavioral control on short, subsecond timescales. Mechanistically, it is thought that DA neurons drive such behaviors by rapidly modulating striatal spiking activity – an effect whose magnitude is based on incomplete evidence. However, a view has emerged that only artificially high (i.e., supra-physiological) DA signals can robustly alter behavioral performance on fast timescales. This raises the intriguing possibility that moment-to-moment striatal spiking activity is actually not strongly shaped by DA signals in the physiological regime. To test this, we monitored spiking responses in the ventral striatum of behaving mice, while transiently raising or lowering DA levels. Surprisingly, most of these manipulations led to only weak changes in striatal activity, with the

notable exception of when DA release exceeded reward-matched levels. These findings suggest that in contrast to the widely held assumption, DA neurons normally play only a minor function in the rapid control of striatal dynamics in relation to other inputs. This constrained role demonstrates the importance of discerning DA neuron contributions to brain function under physiological and potentially non-physiological conditions.

The dissertation of Charltien Long is approved.

Christopher J. Evans

Carlos Portera-Cailliau

Kate Wassum

Sotirios Masmanidis, Committee Chair

University of California, Los Angeles

2024

This work is dedicated to my mother, father, and big sister.

I wouldn't be where I am or have become who I am today without them.

TABLE OF CONTENTS

LIST OF FIGURES

ABBREVIATIONS

AAV – adeno-associated virus

auROC – area under the receiver operating characteristic curve

cAMP – cyclic adenosine monophosphate

D1 – D1-type dopamine receptor

D2 – D2-type dopamine receptor

DA – dopamine

FSI – fast spiking interneuron

GABA – gamma-aminobutyric acid

GPCR – G protein-coupled receptor

GRIN – gradient-index

ITI – intertrial interval

MSN – medium spiny projection neuron

PKA – protein kinase A

ROC – receiver operating characteristic

RPE – reward prediction error

SNC – substantia nigra pars compacta

STDP – spike timing dependent plasticity

SVM – support vector machine

TAN – tonically active neuron

VTA – ventral tegmental area

ACKNOWLEDGEMENTS

This dissertation is a revised and expanded version of “Physiological constraints on the rapid dopaminergic modulation of striatal reward activity” which is a preprint that is not yet published, but has been available on bioRxiv since 09/2022 (doi:10.1101/2022.09.16.508310) [1]. The preprint was co-authored by Charltien Long (graduate student researcher, co-first author), Kwang Lee (project scientist, co-first author), Long Yang (project scientist), Theresia Dafalias (laboratory technician), Alexander K. Wu (undergraduate student researcher), and Sotiris C. Masmanidis (principal investigator, advisor, committee chair). Charltien Long, Kwang Lee, and Sotiris Masmanidis designed the study and experiments. Data collection was conducted by Charltien Long, Kwang Lee, Alexander K. Wu, and Theresia Dafalias. Data analysis and interpretation of results was done by Charltien Long, Kwang Lee, Long Yang, and Sotiris Masmanidis. Charltien Long, Kwang Lee, and Sotiris Masmanidis wrote the manuscript. All authors approved the current version of the manuscript. This dissertation contains additional background, commentary, theoretical proposals, and scientific views that do not necessarily reflect the opinions of the other authors and were not expressly endorsed by them.

I want to thank Sotiris Masmanidis for his tireless mentorship over the years. It was tough starting grad school during the pandemic, but he was always there to help and lend an ear. It means a lot to me that he entrusted me with this project and has faith in me as a graduate student. I also thank Kwang Lee for training me well during his final months as a postdoc before he started his own lab, and Long Yang for his support over the years during his time in our lab. I want to thank Dr. Carlos Portera-Cailliau for his continued mentorship since my first day as an MSTP student. I thank Dr. Kate Wassum for her mentorship and guidance both as a member of my committee as well as in her leadership role in the TNDA T32 program. I also want to thank Dr. Christopher Evans for his support and guidance. Lastly, I thank the leadership of the UCLA MSTP, currently Dr. Olujimi Ajijola and Dr. David Dawson, for their support.

VITA

EDUCATION

B.S., Neuroscience, with Honors 2017

Johns Hopkins University, Baltimore, MD

SELECTED PREVIOUS RESEARCH POSITIONS

Vinogradov Lab, UCSF, Dept of Psychiatry	05/2013 – 08/2013, 05/2014 – 08/2014
Huganir Lab, Johns Hopkins Medical Institute, Dept of Neuroscience	08/2013 – 05/2014
Hillis Lab, Johns Hopkins Medical Institute, Dept of Neurology	08/2014 – 05/2017
Senior Honors Thesis, Johns Hopkins Humanities Center*	01/2017 – 05/2017
Masmanidis Lab, UCLA, Dept of Neurobiology	08/2020 - Present

*Renamed to Department of Comparative Thought and Literature.

PUBLICATIONS

Long, C., Sebastian, R., Faria, A. V., & Hillis, A. E. (2018). Longitudinal Imaging of Reading and Naming Recovery after Stroke. Aphasiology, 32(7), 839–854. https://doi.org/10.1080/02687038.2017.1417538. PMID: 30127542; PMCID: 6097621.

Long, C. & Hillis, AE. (2017) Oxytocin as a Treatment Option in Right Hemisphere Stroke: Empathy and Social Learning. In C. Edwards (Ed.), Psychology of Empathy: New Research. New York: Nova Science Publishers, pp. 151- 172.

Sebastian, R., **Long, C.**, Purcell, J., Faria, A., Lindquist, M., Jarso, S., Race, D., Davis, C., Posner, J., Wright, A., & Hillis, AE. (2016). Imaging Network Level Language Recovery After Left PCA Stroke. Restorative Neurology and Neuroscience, 34(4):473-489. https://doi.org/10.3233/RNN-150621. PMID: 27176918; PMCID: 5003759.

Sebastian, R., Saxena, S., Tsapkini, K., Faria, A., **Long, C.**, Wright, A., Davis, C., Tippett, D., Mourdoukoutas, M., Celnik, P., & Hillis, AE. (2016). Cerebellar tDCS: A Novel Approach to Augment Language Treatment Post Stroke. Frontiers in Human Neuroscience, 10: 695.

AWARDS AND HONORS

First Place, Dana Foundation “Design a Brain Experiment” Competition	05/2013
Woodrow Wilson Undergraduate Research Fellowship, JHU	09/2013-05/2017
Departmental Honors, JHU Neuroscience Department	05/2017
Phi Beta Kappa Honor Society, JHU Chapter	05/2017
Leslie and Susan Gonda MSTP Fellowship	2018-2019, 2020-2021

FUNDING

UCLA Training Program in the Translational Neuroscience of Drug Abuse (TNDA), 5T32DA024635	07/2022-06/2023
UCLA-Caltech Medical Scientist Training Program, 5T32GM008042	07/2018-Current

CHAPTER 1: What is dopamine's function in the striatum?

1.1 Overview of current views

The striatal regulation of learning, movement, motivation, and decision-making is thought to critically depend on DA signaling [2]. Abnormal levels of striatal DA are implicated in several disorders including Parkinson's disease, addiction, and depression [3, 4]. Yet despite significant progress, our understanding of DA's complex modulatory functions in the striatum is incomplete and remains the subject of intense study and controversy. Nevertheless, the commonly accepted dogma is that striatal DA neurons serve two major functions, supporting associative learning and movement [5].

1.2 Dopamine's role in associative learning

Evidence for DA's role in learning spans several decades but much of the field's current conceptualizations descend from a seminal paper from Schultz and colleagues that described a phenomenon that would come to be known as reward prediction error (RPE) [6]. In this classic experiment, Schultz and colleagues recorded from midbrain DA neurons in monkeys while conducting a Pavlovian conditioning task. In this task, visual and auditory cues preceded the delivery of reward after a fixed delay. Early in learning, rewards caused DA neurons to fire, but later after learning, the anticipatory cues caused DA neurons to fire rather than the reward itself. In other words, predicted rewards did not cause DA neurons to fire while unpredicted rewards did still cause DA neuron firing. This change over the course of learning also appeared to coincide well with observations made from classical conditioning experiments where unconditioned stimuli are associated with conditioned stimuli, resulting in a transfer of appetitive value to the unconditioned stimuli. It is also worth noting that when reward was omitted, DA neurons briefly reduced their firing, resulting in a kind of pause close to when the reward would have normally been delivered. This provided evidence that these RPE signals also contain some degree of timing information, which the authors would further explore using a temporal difference algorithm.

All of this provided compelling evidence for a role for DA as a “teaching” signal that might be able to drive associative learning. However, this experiment did not establish a causational role for these signals in driving learning.

Of course, experiments done even before the original Schultz paper had already suggested a causational role for DA to some degree. For instance, it was known that drugs of abuse either caused the release of or lengthened the actions of DA, and in a sense drove learning in the form of addiction [7-9]. Moreover, electrical stimulation of brain regions where DA neurons reside also was apparently able to drive a sort of reward learning where animals would escalate their pressing of a lever that electrically stimulated those regions [10]. The reinforcement effects of this self-stimulation and other learning paradigms could also be disrupted by blocking DA neurotransmission, either pharmacologically or surgically through lesions [11]. However, these experiments were limited by the lack of specificity in the cells or pathways targeted, with experiments involving pharmacological interventions further limited by the timescale of the effects. In this sense, none of these experiments were truly able to approximate or recreate the DA neuron activity described by Schultz and colleagues [6].

Some of the strongest direct causational evidence would come with a set of experiments from Steinberg and colleagues that leveraged optogenetics to stimulate DA neurons [12]. Optogenetics does not face the same limitations as some of the earlier experiments because it can be specifically targeted to a genetically defined population of neurons (ie. DA neurons). Moreover, the effect of optogenetic manipulations can be time limited to the duration of laser exposure, providing a temporally precise intervention. The behavioral paradigm used was a blocking procedure that involved the introduction of a compound cue containing a previously learned reward predictive cue and a new sensory cue prior to the delivery of a reward. Animals generally will not learn to associate the new cue if the learned cue is present, so learning is “blocked”. Prior work had shown that DA neuron activity is similar between a learned cue alone and with the new compound cue, in theory diminishing an animal’s ability to learn the compound

cue [13]. It is important to note that the paradigm used here involved a cue that was continuous over the course of 30 seconds and that animals had to enter and stay in a nose poke for the duration of the cue to receive the maximum amount of reward. In a key experiment, the authors optogenetically stimulated DA neurons upon reward delivery during the presentation of a compound cue. When the novel component of the compound cue was presented alone, the animals that received DA neuron stimulation entered and stayed in the reward port more than controls, remarkably showing that learning had occurred. A subsequent set of experiments showed that optogenetic stimulation of DA neurons was able to slow the rate of behavioral extinction when the sucrose reinforcer was substituted for water or omitted completely. These experiments provided some of first definitive evidence for a causal role of DA, at least at the time of reward, in driving associative learning.

The cellular level mechanism for this DA driven associative learning was most directly demonstrated in a series of experiments from Yagishita and colleagues [14]. The authors hypothesized that DA would have effects on the dendritic spines of medium spiny projection neurons (MSNs), the principal neuron type in the striatum. They hypothesized this for two reasons. First, MSNs receive both glutamatergic inputs from cortical areas as well as dopaminergic inputs at their dendritic spines. Second, dendritic spines can undergo structural changes during long-term potentiation (thought be a neural substrate of learning) that facilitate more robust neurotransmission [15]. The authors conducted their experiments in acute coronal slice preparations, and so they needed a method that would be able to recapitulate DA release, glutamate release, and action potentials in a targeted MSN. This was accomplished using a combination of optogenetic stimulation of DA fibers, optical uncaging of glutamate in the surrounding extracellular fluid, and direct current injection into the spine-bearing MSN. The authors leveraged a method of repeated uncaging of glutamate and stimulation of the MSN in a manner than can induce long-term potentiation, a phenomena known as spike-timing-dependent plasticity (STDP) [16]. Consistent with prior research that demonstrated the necessity of DA for the

potentiation associated with intracranial self-stimulation of DA neurons, Yagishita and colleagues found that DA was necessary in order to observe robust dendritic spine enlargement (and presumably the plasticity such enlargement would support), with the STDP method alone achieving only about a quarter of the enlargement with DA [17]. Interestingly, the authors found that DA's enhancement of STDP was time limited to an approximately 2 second window, with DA release before the STDP stimulation or too long after being ineffective. Other experiments within the paper implicated protein kinase A, D1 DA receptors, protein synthesis, and calmodulin-dependent protein kinase II in this DA enhanced spine enlargement. This study was important because it provided a detailed cellular and protein-level mechanism that could explain decades of prior observations about DA and learning.

1.3 Dopamine's role in regulation of movement

At the same time, DA's role in movement has also been supported by several key studies spanning decades. The role of DA neurons in Parkinson's, a neurological disorder with well-documented motor impairments, had been identified years before the study from Schultz and colleagues in 1997, with pathologists noting the disappearance of nigrostriatal DA neurons as a hallmark of the disorder [11]. In addition, psychostimulants that were known to alter DA neurotransmission and pharmacologically induced DA neuron lesions had clear effects on motor activity and animal behavior, and these phenomena were known prior to the discovery of RPE [11]. Moreover, early work that examined electrical stimulation of DA-related neurons and networks proposed a separate "energizing" effect outside of its potential for reinforcement, such as in intracranial self-stimulation experiments where reinforcement is the focus [10]. In the early 2000s, there were also reports of DA release correlating with movements when measured with voltammetry [18]. Furthermore, theoretical frameworks emerged that dissociated reward seeking "wanting" behaviors from learning, arguing that fluctuations in DA might act on brain circuits to cause "wanting" [19]. Therefore, while our discussion thus far has focused on the historical evidence

behind DA's role in learning, there was an equally prevalent interest in DA's role in movement occurring at the same time. For this section, we will focus on contemporary evidence that has emerged since the widespread adoption of optogenetics.

Similar to its importance for establishing DA's causal role in learning, optogenetics has unlocked the ability to explore DA's causal role in regulating movement. Perhaps the most direct demonstration of this came from da Silva and colleagues in 2018 [20]. The work discussed above and investigations of Parkinson's had long established the importance of basal DA levels in regulating movement. Medications that raise overall levels of dopaminergic signaling, such as carbidopa-levodopa or DA receptor agonists, are current mainstays of treatment based on this fundamental concept [21]. While these agents raise DA signaling levels, they likely do not recapitulate the rapid, subsecond or phasic DA release activity that DA neurons normally perform. Models proposed prior to this study had hypothesized that tonic or resting levels of DA were principally responsible for DA's regulation of movement while rapid DA signaling events were not [22]. The authors of this study sought to challenge this view by showing that phasic DA events could in fact contribute to regulation of movement. If true, this would have implications for Parkinson's treatments as there would need to be a greater focus on reconstituting rapid DA transmission events. The authors began by identifying DA neurons within the substantia nigra pars compacta (SNc) using a method known as "photoidentification" or "opto-tagging". When using electrodes to record in vivo spiking activity from brain areas, it is often difficult to identify specific populations of neurons because electrodes cannot inherently discriminate between different types of neurons. While the identity of a detected neuron can be inferred from aspects of its spiking activity, this is not a perfect method and cannot distinguish between cell types that have very similar spiking properties (such as D1 and D2 MSNs). Photoidentification involves targeting a specific cell type using transgenic mouse lines and an appropriate virus bearing an excitatory opsin. By pairing recording electrodes with an optical fiber, DA neurons can be positively identified based on their response to light.

From this pool of DA neurons, da Silva and colleagues found that many of these positively identified neurons increased their firing prior to movement initiation [20]. When examining the movement initiations that occurred after DA neuron firing, the authors found no clear relationship between which DA neurons were active and the type of movement a mouse undertook, leading them to conclude that DA neurons contain a generalized movement initiation signal. They also found that the degree of DA neuron activity 300 milliseconds prior to movement initiation seemed to predict the vigor (in this case defined as body acceleration) of the subsequent movement. When DA neurons were optogenetically silenced for 15 second intervals, mice were more likely to remain immobile if they were already immobile. However, this inhibition did not seem to affect the vigor of ongoing movements and did not cause ongoing movements to halt. Conversely, optogenetic activation of DA neurons resulted in movement initiation but had minimal effect on ongoing movements. It is worth noting that there was a range of these effects, with some mice experiencing much greater enhancements in movement initiation than others. However, for the majority of their mice (5 out of 7), effects on action initiation were much more modest, not reaching mean accelerations across stimulated trials comparable to spontaneous movement. Further experiments examining action sequences found that inhibition of DA neurons only impacted the initiation but not continuation of sequences of actions. For these reasons, da Silva and colleagues concluded that consistent with their normal in vivo activity, these DA neurons have a causal role in movement initiation.

Critically, in explaining these results, the authors proposed an overall theoretical framework about how DA regulates movement that would be echoed in subsequent more comprehensive models of DA's effects on movement [5]. da Silva and colleagues hypothesized that DA-induced movement initiation occurred "presumably by modulating the excitability of striatal projection neurons". In making this statement, the authors referenced two papers that showed that the neuronal activity of the direct and indirect pathway was responsible for regulating movement, with the findings of those studies being informative for our understanding of

Parkinson's disease [23, 24]. Essentially, the claim was that DA alters the activity of striatal circuits that have been shown to control movements; therefore, DA can control movement. Although this was the top line message that would be referenced in later proposals about DA's role in the striatum, the point made by the authors was a bit more nuanced [5]. Rather than DA being a primary driver of movement or a primary driver of activity in the striatum, the authors proposed a more modest gate-like or permissive role for DA, whereby "planned" motor actions brought via glutamatergic inputs to the striatum would be more readily transmitted through striatal circuits in the presence of DA. This proposal made sense in the context of the presented data because the authors did not find that DA shaped ongoing movements but rather increased the likelihood that a movement would start. One could also speculate that the authors were cognizant of the small effects on movement initiation observed in many of their mice, further weakening the case for DA being a primary driver of movement. However, these ideas did raise a conundrum about the details of how such a gate could function and suggested that a simple model was not adequate. Of course, it is relatively straightforward to understand that if DA could change the excitability of striatal circuits, then that would upregulate those circuits and make action initiation more likely. However, it is far less straightforward to understand how that can selectively happen such that movements that have yet to occur are enhanced but ongoing movements are not similarly boosted. After all, a simple DA-induced increase in excitability should enhance the likelihood of action initiation and invigorate ongoing movements. Yet that is somehow not the case.

Contemporary and subsequent work was largely consistent with the idea that DA had a role in regulating movement but did recognize certain limits on this effect. For instance, Howe and Dombeck in 2016 showed that DA terminal activity within the dorsal striatum correlated well with head fixed mouse locomotion on a cylindrical treadmill [25]. They found that DA terminal activity increases tended to precede movement initiation, DA continued to rapidly fluctuate during movement, and that DA dropped during movement termination. However, Howe and Dombeck found that the fluctuations during movement were more strongly associated with prior movements

than with future movements, albeit still correlated with future movements. Howe and Dombeck also optogenetically stimulated DA neuron terminals in the dorsal striatum to determine whether DA had a causal effect on movement initiation and ongoing movements. Their results showed that DA could drive locomotion in a mouse that was at rest, but with important caveats. First, within single animals within single sessions, it appeared that the magnitude and timing of DA's effects varied, such that DA stimulation overall increased the likelihood of movement but that exactly how much and when movement would start was not predictable on a trial-by-trial basis, with stimulation sometimes not even having a behavioral effect. Moreover, there was a large amount of variation in the extent of this effect across different mice. For ongoing movements, they examined the power spectra of the movement accelerations of mice undergoing either 3 Hz or 6 Hz stimulation of their DA neurons. They found that there was a tendency to see a shift towards 6 Hz in the power spectrum of the animal's acceleration if a 6 Hz DA stimulation protocol was used compared to 3 Hz stimulation. This did provide some evidence that DA could exert an effect on ongoing movements, but the effect was variable and for most mice constituted a shift in the center of mass of the power spectrum of less than 1 Hz. Therefore, this outcome shows that DA might influence ongoing movement but that it was not quite a command signal, per se. The effects on initiation and ongoing movement taken together led Howe and Dombeck to carefully claim that the effects of DA were "strongly modulatory, rather than deterministic" and that DA was "unlikely to instruct movement alone", similar to claims discussed above [20]. Their ultimate point was that they felt that DA carried a meaningful, causal subsecond movement signal but that there were limits to its strength.

A third often cited piece of evidence of DA's effects on movement is a study from Hughes and colleagues in 2020 that used a novel approach that captured the forward and backward force exerted by a head fixed mouse [26]. It is worth noting at the outset that the animals in this study do not "move" in the way that a mouse would if it was on a circular treadmill or unrestrained. In that sense, this experimental preparation correlates the force the mouse applies with its head and

claims that this is a measure of vigor, motivation, and is a subtle aspect of movement that has previously been ignored. Whether or not impulse vectors derived from force sensors are comparable to other kinds of movement has not yet been definitively explored. The authors began by establishing that there were three separate populations of ventral tegmental area (VTA) DA neurons whose activity was correlated with forward (fast and slow) and backwards forces applied on the head. Next, the authors showed that DA neuron stimulation was able to cause the animal to increase the forward force it exerted, and that DA neuron inhibition caused the animals to increase the backward force it exerted. In addition, the authors showed that DA neuron manipulations could affect licking behavior when rewards were delivered at fixed intervals. Interestingly, they did not find the same optogenetic effects on movement in naive mice that were not trained on the task. The authors felt that this showed that the effect depended on a "motivational context" and that the effect was not "unconditional" (as it would be for a motor neuron, for instance). After all, if DA had a simple causational relationship to force generation, then the stimulation of DA neurons in a naïve animal should result in forward impulse vectors. All three of the studies discussed here seem to paint a somewhat consistent picture of DA as a paradoxically strong modulator of movement that somehow only functions in certain situations or contexts.

A number of recent studies have discussed the possibility that DA might affect aspects of motor timing or internal senses of time [27-29]. This is a somewhat interesting perspective because alterations in subjective perceptions of time would be consistent with the general concept that DA effects are inconsistent and indirectly impact motor output. Certainly, an internal representation of time would have an inherent level of inconsistency that would not alone drive motor movement if perturbed while not engaged in a task. One often referenced paper is a study conducted by Soares and colleagues in 2016 where an animal had to correctly differentiate between a short and a long delay to receive a reward [27]. In this task, Soares and colleagues varied the delay between two auditory tones and required mice to determine whether the delay was shorter or

longer than 1.5 seconds. The mice had to then correctly enter a nose poke corresponding to shorter or longer delay to receive a reward, effectively having to convert a judgement into a behavioral choice. This task also allowed the researchers to have trials with varying levels of difficulty depending on how close the delay was to 1.5 seconds. In an optogenetic experiment, Soares and colleagues inhibited or excited DA neurons during for the entire trial duration on 30% of trials. The authors found that DA neuron excitation resulted in bias towards choosing the shorter delay nose poke, with the greatest effect seen in more difficult trials where the delay was close to 1.5 seconds. No difference was seen for delays that were unambiguously shorter or longer than 1.5 seconds. Similarly, DA neuron inhibition resulted in a tendency to choose the longer delay nose poke, but the effect was much smaller. The authors interpreted these results to mean that DA neurons can control an animal's judgement of time, invoking the adage "time flies when you're having fun".

A later study from Hamilos and colleagues found somewhat conceptually similar results using a self-timed movement task [28]. In this study, the authors utilized a task where a mouse had to lick after a set interval following the presentation of a cue to receive a reward. If an animal licked too soon after the cue, the trial was not rewarded. Though both studies use tasks that involve timing in an abstract sense, this task is distinct from the one used by the earlier Soares et al. study, which was a discriminative task where movement timing itself was less important [27]. Despite these differences, Hamilos and colleagues found qualitatively similar results when conducting optogenetic manipulations of DA neurons. For their experiments, Hamilos and colleagues stimulated or inhibited DA neurons but critically halted optogenetic manipulations upon the animal's first lick on a given trial. Optogenetic stimulation of DA neurons resulted in earlier licking onset with a median advance of about 0.2 seconds. Optogenetic inhibition of DA neurons produced an opposite effect with licking onset occurring later with a median delay of around 0.2 seconds. Hamilos and colleagues noted that they chose optogenetic stimulation parameters that produced DA release that was similar in magnitude to naturally occurring release and that this

stimulation was unable to produce behavioral effects outside of the task context. Moreover, even when the laser power was greatly increased, optogenetic stimulation did not increase licking but rather caused “dyskinetic” body movements. Keeping in mind those limitations, the authors proposed a casual role for DA in the control of movement timing and that DA might modulate the probability of movements over time (but not drive movements alone). Although Soares et al. were more concerned with internal judgements of time and Hamilos et al. were more concerned with the timing of movements, both papers together established the idea that timing is an important aspect of how DA might impact movements [27, 28].

However, while interesting, DA effects on or via timing likely do not completely explain the role of DA in movements. This point was made most directly by Howard and colleagues in 2017 using a task very similar to the one used by Soares et al [27, 29]. Howard et al. trained mice to press one of two levers based on whether a time delay was long or short. In this task, the time delay that the mice had to judge was the amount of time the levers were temporarily retracted on each trial. The authors also conducted an optogenetic experiment, but their optogenetic manipulations involved limited 1 second manipulations immediately prior to lever presentation (end of retraction). Optogenetic stimulation resulted in a ~10% bias towards selecting the short duration lever for 4 second trials but not 2 second or 8 second trials. In another experiment, the authors stimulated or inhibited neurons for 1 second windows within the delay on 8 second trials. DA neuron excitation tended to increase the choice of the short delay lever, and DA neuron inhibition increased the choice for the long delay lever only at delay onset. Overall, these results are qualitatively similar to those of Soares et al. despite some differences in experimental design [27]. Where Howard et al. and Soares et al. greatly diverge is in their overarching interpretation of the results. The source of their disagreement is that the task structure used in both papers inherently requires two cognitive processes that are not easily separated, the judgement of time and behavioral choice. In theory, it could be argued that the effects they both observed are caused by either changes in timing judgement or behavioral choice. Indeed, they each present competing

computational models that support one explanation over the other. However, because our understanding of how DA relates to internal clocks and how these clocks might function is still incomplete, it is difficult to definitively litigate this issue one way or the other [30].

That being said, Howard and colleagues do present a number of interesting experiments where they strategically alter certain aspects of the tasks related to timing and choice while recording DA release with voltammetry. They were able to convert their task into a Pavlovian task, a forced choice task, a straightforward operant task, and a task where an auditory tone indicated the correct lever choice regardless of delay. At a broad level, the tasks all had very different behavioral requirements while still sharing similar time delays. The authors use this observation to argue that the DA release should be much more similar if DA was simply reflecting timing. They then further note that the last task involving a discriminative tone showed that DA fluctuation diverged much more at lever selection compared to tone onset when comparing the two stimuli or lever options. Since DA varied much more when comparing left or right lever presses than for the left and right tones, DA must have a larger role in action selection. While it is debatable the degree to which these outcomes prove that DA is an action selection signal, Howard and colleagues do succeed in convincingly showing that timing is not an all-encompassing explanation for DA dynamics and their role in movement. Like many of the other studies discussed, Howard and colleagues are careful to clarify that their data do not support a deterministic effect of DA on action selection, as this is not what they saw with their optogenetics experiments. They specifically use the term “bias” and emphasize the importance of behavioral context in driving these effects on action selection.

After examining several influential studies on DA’s role in movement, it becomes clear that the preponderance of evidence suggests that DA can have an effect on movement but that this effect is likely limited in its scope, context dependent, and is not comparable with that of motor neurons [20, 25-29]. In addition, these studies also seem to point to a greater role in action initiation rather than regulation of ongoing actions, with this interpretation being echoed throughout the

diverse spectrum of experiments discussed here. For these reasons, hypothesizing about the exact role of DA in movement and what circuit level mechanisms might underly this is challenging. To reiterate a point made earlier, a simplistic model likely would not be able to encompass the nuance that exists within the evidence. But it is worth stating outright that any viable theory would likely have to contend with the apparent scientific consensus around the idea that there are meaningful constraints on DA's impact on motor behaviors.

There have been some contemporary theoretical proposals that have attempted to provide overall frameworks for DA's effects on movement that account for these complexities [5]. Prior to the da Silva study, which was important because it showed DA's role in movement in the absence of any explicit reinforcer, research had shown that animals were more likely to initiate a reward-motivated behavioral task when DA neurons were optogenetically stimulated [31]. DA was proposed to be a kind of value signal that represented (and causally affected) the motivation of an animal to pursue rewards [31]. What separated this from prior theories about DA's effects on movement, like the other studies discussed thus far, was a focus on rapid or phasic DA signaling representing this information, rather than "slow" tonic signals. Subsequent work has claimed that these DA value signals are generated independently from RPE broadcast signals found in the somatic spiking activity of DA neurons, perhaps through the actions of local striatal interneurons [5, 32, 33]. It has also been speculated by some that interneurons might be able to selectively toggle DA's effects from learning to movement, explaining why DA only has movement effects under certain contexts [5]. Outside of that specific proposal, other important discoveries include work that established that DA neurons encode a wide range of psychomotor task variables pertinent to ongoing behaviors alongside canonical reward responses, adding to the case that DA might have a more expansive role in controlling movements than previously thought [25, 34]. Overall, this has coincided with an increased appreciation of the inherent complexity and intricacies of DA signaling throughout the striatum and the brain more generally, perhaps explaining why its effects on movement seem so complex [35]. Regardless of how exactly the inconsistency and context-

dependence is explained, most proposals of DA's role in movement generally hinge on an implicit assumption that DA impacts movements through exerting direct effects on striatal circuits [36]. This topic and its implications will be discussed in more detail in a subsequent section. At any rate, a large body of evidence utilizing optogenetics and measurements of DA release has established a role for DA in control of movements, although many questions remain.

1.4 A third view of dopamine: the learning primacy hypothesis

Though the case for DA's causal role in controlling movements has a cemented place in the literature with decades of scientific consensus supporting it, a dissenting view has emerged that argues that DA may only play a minor role in subsecond behavioral control, with its primary role being driving associative learning [37-40]. Crucially, this view reconciles itself with the body of literature discussed above by noting that DA has considerably enhanced effects when DA levels are artificially raised above physiological levels [28, 37]. In essence, this would mean that many of the discussed results where strong effects were observed may in fact only show the potential for DA to control movement, but that this generally does not occur to that extent under normal conditions. In subsequent discussion, I will be referring to this concept as the learning primacy hypothesis of DA. The term learning primacy hypothesis is one of my own creation and this set of ideas has not been called this previously.

The origin of the learning primacy hypothesis is generally traced back to a 2018 paper from Coddington and Dudman that investigated the activity of DA neurons across the early stages of learning of a Pavlovian task [37]. The authors played an auditory tone 1.5 seconds prior to reward delivery to a head fixed mouse placed in a basket with accelerometers. This setup allowed the measurement of both consummatory licking and general body movement. The authors strategically chose to focus on early learning, where hundreds of trials are generally necessary to train an animal to reliably respond to a Pavlovian task. Using opto-tagging, the authors performed in vivo single cell recordings from positively identified DA neurons within the VTA and SNc during

learning. In entirely naive animals, DA neurons showed a small excitation (in the form of a spike phase advance) followed by a more noticeable inhibition during the initiation of body movements. After the animals had been introduced to the task (and presumably understood they were in a reward context), the DA neuron excitation component for some neurons (particularly in SNc) became larger with the inhibition remaining similar in response to body movements. Across the first 300 trials of learning, DA neurons appeared to change their firing near, often prior to, the initiation of consummatory licking. During learning, DA neurons also developed an immediate response to reward delivery, prior there was little immediate response to reward delivery. This response also tended to precede lick onset. The authors also noted the coexistence of a learning related response to the anticipatory auditory cue that emerged over the course of training. Thus, the authors identified two separate components of DA activity, a learning component characterized by change across training and a movement component characterized by an excitation-inhibition sequence that was more stable across training. At first glance, this portion of Coddington and Dudman's results appear to be largely in agreement with the general idea that DA dynamics and movement are linked. However, Coddington and Dudman make several further observations throughout the paper that will directly challenge the core premise of that very idea.

Coddington and Dudman then proceeded to more directly examine the excitatory DA neuron dynamics that they found were associated with movement. It is worth pointing out that although they did find consistent DA neuron inhibition after action initiation, even in untrained animals, this itself would not align with the proposals made about DA's role in movement. This is because many key optogenetics experiments and their accompanying hypotheses suggested that DA release, not inhibition, is the key to action initiation [20, 25]. They noted that the overall population activity of DA neurons, even in trained animals, was on average neutral or even net negative around the time of action initiation. Of course, Howe and Dombeck (and later work) observed a degree of diversity in the information encoded in DA neuron activity and suggested that only a subset might be responsible for its effects on movement [25, 34, 35]. Therefore, Coddington and

Dudman chose to focus on the excitatory component of DA neuron activity, even if not all neurons robustly carried this signal. Building off the observation that the introduction of rewards boosted the excitatory activity of DA neurons, they segregated movements into three categories with ascending “reward expectation”: body movements with no licks, body movements with licks (but not within a trial), and body movements made in response to the cue (that likely included anticipatory licking). Of the neurons that were excited by movement, the authors found that movements that included licks and cued movements had dramatically larger excitatory components, particularly cued movements. They found no general increase in DA neuron population firing across training (due to a large number of DA neurons that were inhibited during movement), but they did find a significant correlation between the DA neuron response to cue and the movement related response of individual neurons. This led them to conclude that the DA neuron excitation signal was a reward expectation signal related to learned cue associations rather than a movement signal per se.

Nevertheless, it could still be argued that this excitatory signal causes movements, leading Coddington and Dudman to perform a deeply important optogenetic experiment that has left a lasting mark on the literature [28, 35, 39-41]. The authors used a combined optogenetic and fiber photometry approach, where they were able to monitor the activity of DA neuron terminals in the striatum using a calcium sensor while simultaneously optogenetically stimulating DA neurons in the SNc or VTA. This granted them the ability to compare the evoked activity from optogenetic stimulation to the activity evoked by natural rewards. By varying the laser power (1 mW or 7 mW) and the time the laser was active (150 ms or 500 ms), they were able to evoke a range of activity levels while retaining valuable context when compared to rewards. Critically, only SNc DA neurons were able to induce movement, and this only occurred when the magnitude of the optogenetically evoked signal was about 5 times larger than the natural reward signal. The authors then showed that reward-matched DA neuron stimulation, in either SNc or VTA, was able to induce conditioned place preference for an animal moving in an open field, demonstrating that

DA driven learning was intact at these magnitudes. This experiment was highly consequential for two reasons. First, it is credited as the earliest DA-related study to establish the importance of calibrating optogenetic manipulations against a physiological reference point [35]. Second, the study raised the prospect that DA's effects on movement may have been a product of optogenetic overstimulation that may not truly reflect the typical function of DA. By extension, this would imply that DA's role in learning is its "true" function while movement is not, or at least only a minor function. Through this study, Coddington and Dudman laid the foundation for the learning primacy hypothesis.

In the wake of the original Coddington and Dudman study, a paper from Lee and colleagues would further bolster the case for the learning primacy hypothesis [38]. Lee and colleagues exploited a feature of Pavlovian conditioning tasks that allowed for learning and motor execution to be experimentally separated. They trained head fixed mice on a task where an odor cue was presented at a fixed interval prior to the delivery of a food reward, with mice often beginning licking behaviors prior to reward delivery in anticipation. Because mice began licking prior to the delivery of the food reward (reinforcer), the task can be broadly separated into two epochs. The first epoch spans the cue presentation until just before the reward is delivered and represents a period where motor execution starts but reinforcement does not occur; the second epoch spans the period following the delivery of reward where reinforcement occurs. By optogenetically inhibiting DA neurons in each of these epochs (4 seconds for pre-reward, 2 seconds for post-reward), they were able to compare the contribution of DA to movement and reinforcement with changes in licking behavior being the outcome variable of interest. Throughout the study, the authors used a 3-block experimental design where the animals completed a block of 40 trials of the task without any optogenetic manipulations, followed by another block where optogenetic manipulations were performed, followed by a third block that was identical to the first.

In a key set of experiments, they inhibited DA neurons in either the pre-reward or post-reward periods to test the effects on motor generation and reinforcement, respectively. They

performed these experiments while targeting VTA or SNc neurons. Interestingly, they found that inhibiting DA neurons during the pre-reward period had little to no effect on the probability that a mouse would engage in anticipatory licking during the inhibition block. However, the anticipatory licking probability was greatly decreased if the DA neurons were inhibited in the post-reward period when reinforcement occurs, with this decrease occurring over the course of about 10 trials. They were able to see a small decrease in licking probability with pre-reward SNc inhibition but the effect was much smaller than for the post-reward inhibition. Their experimental design was also able to confirm that there was no lasting motor deficit with licking largely returning to normal in the third block for all mice, which is perhaps why the authors chose the 3-block design. They then also performed an important comparison experiment where they performed similar inhibition on M2 neurons. In that case, pre-reward inhibition decreased licking probability but post-reward inhibition had no effect on anticipatory licking. Using the 3-block experimental design, the authors showed that mice decreased their anticipatory licking in response to a sudden drop in the reward volume in a similar fashion to post-reward optogenetic inhibition. This suggests that post-reward inhibition indeed did decrease the reinforcement value of the reward and that this explains the behavioral effects. Through a series of confirmatory experiments, the authors ruled out the possibility that their observed effects were caused by ineffective suppression during the pre-reward period due to much higher DA activity during that period. If anything, the authors found that post-reward DA neuron activity was larger, and the degree of suppression appeared similar for both epochs. They also showed that extended inhibition spanning both the pre-reward and post-reward period had identical behavioral effects to post-reward inhibition, simultaneously allaying concerns about the differences in duration of inhibition between the two epochs and concerns about DA neuron rebound excitation upon laser offset.

Lee and colleagues then continued with three further notable experiments that expand on this initial finding to demonstrate the importance of the reinforcement role of DA neurons. The authors decided to investigate whether the reinforcement effects they observed were limited to

specific associations or if the effects were instead a generalized motivational deficit. To do this, they trained mice on two distinct olfactory cues, providing identical rewards after both. They then randomly interleaved trials with both cues but only suppressed DA neurons during post-reward period for one of the cues but not the other. They did see a decrease in the anticipatory licking probability for both cues (even though only one had post-reward inhibition), but the cue that had post-reward inhibition decreased by much more than the non-inhibited cue. This experiment did not indicate a strong generalized motivational deficit and instead was more consistent with a reinforcement effect as it was much stronger for the inhibition targeted cue. To further make the case for reinforcement, the authors then conducted an experiment where reward was omitted for the second block in the 3-block design and DA neurons were optogenetically excited at the normal time of reward delivery. This was designed to see if optogenetically induced DA release alone could maintain or at least slow the extinction of the conditioned anticipatory licking response, conceptually similar to the foundational work of Steinberg and colleagues discussed earlier [12]. Control animals showed a rapid, substantial decline in licking after the removal of reward reinforcement. Remarkably, animals that received 2 seconds of continuous DA neuron excitation without rewards did not show a decline in their anticipatory licking responses. Even when the authors titrated this DA neuron stimulation down to 0.5 seconds of 20 Hz stimulation, they saw nearly identical results. In a final additional experiment, the authors determined through systematically increasing the delay between reward delivery and onset of optogenetic inhibition of DA neurons that the effects on anticipatory licking were strongest within 0.5 seconds of reward delivery with the trend line of effects reaching 0 at about 1.6 seconds post-reward. This result appeared to be consistent with the “time window” of DA-related plasticity from in vitro work conducted by Yagishita and colleagues [14].

The results of this study from Lee and colleagues were noteworthy and offered key support for the learning primacy hypothesis. Prior work had not extensively directly compared motor effects to learning effects, despite both effects being present in a large amount of literature

(discussed above in detail). Here, the authors were able to examine the relative effects on anticipatory licking within the same task, allowing for a novel apples-to-apples comparison. The bias in the effects of DA neurons towards reinforcement provided compelling evidence that the learning effects had primacy over the potential motor execution effects (with this being particularly pronounced for VTA DA neurons). By showing that M2 neurons had the opposite bias, the authors highlighted the inconsistency of even SNc DA neuron response profiles (which still had bias towards reinforcement) with a population of neurons that controls movement. This point made about the SNc had the further effect of challenging views expressed by Howe and Dombeck as well as others that the SNc was fundamentally functionally distinct from the VTA [25, 42, 43]. Instead, these results suggested that the distinction in DA's downstream actions may be less stark than might have been previously thought, even if the DA dynamics from SNc and VTA are different. The optogenetic excitation experiment conducted by Lee and colleagues also clearly demonstrated that DA release alone was adequate to maintain a learned association and its behavioral correlate. Considered as a whole, these results suggest that motor drive is derived from cortical sources with DA primarily playing a reinforcement role.

The core idea expressed in the paper from Lee and colleagues would be independently affirmed and expanded on by another contemporaneous study from Pan and colleagues [38, 39]. Pan and colleagues were able to replicate the DA neuron excitation experiments in the Lee et al. paper in an experiment involving an auditory cue and a rewarded nose poke in a freely moving mouse. They found that physiologically calibrated DA stimulation was able to substitute for an unconditioned stimulus and maintain cue-driven approach behavior and licking behavior. They then built on those experiments by substituting DA neuron excitation for the auditory cue to see if optogenetic stimulation alone could induce approach behavior. The concept here is that if DA plays a causal role in cued conditioned responses, then DA release ought to be adequate to induce conditioned response behavior. They found that optogenetic DA release could not induce approach behavior when it substituted a cue even after many training sessions. They tried adding

a degree of natural jitter to the exact timing of the laser onset and still could not induce conditioned responses. They confirmed using an overt visual cue paired with optogenetic excitation that the animals were still able to learn, and they confirmed using DA neuron recordings that reward responses after DA neuron stimulation did not resemble cued reward responses. Similar in concept to the point made by Lee and colleagues, Pan and colleagues showed that DA neurons appeared to have a much stronger role in reinforcement and did not seem to have a causal role in driving cue responses. They further confirmed in a 2-port choice task that DA neurons stimulation during cue presentation could not confer discriminatory information. In a final experiment, Pan and colleagues showed that optogenetic excitation of medial frontal cortex neurons and excitatory axon terminals within the VTA were both a viable substitute for sensory cues, highlighting DA's reinforcement role in comparison to these inputs to the striatum. Even though DA neurons might co-vary with or encode relevant task information, Pan and colleagues definitively demonstrated that downstream regions do not apparently use these DA dynamics to inform ongoing actions nor do they transmit a usable copy of sensory information.

Both Pan et al. and Lee et al. examined the effects of DA within the context of Pavlovian conditioning tasks, but influential work, namely from da Silva et al., showed the effects of DA on movement in the absence of explicit reinforcers [20, 38, 39]. This left open the question of whether the learning primacy hypothesis could be generalized to that context. A recent study from Markowitz and colleagues sought to do just that by performing photometric measurements of the genetically encoded DA sensor dLight alongside optogenetic manipulations of DA neurons for mice freely moving within an open field [40, 44]. The authors relied on machine learning driven motion sequencing, which was a method that allowed them to decompose mouse movements into subsecond components called "syllables". These syllables included actions such as "scrunch", "rear", "body lick", and "orient left". The authors began by affirming several key observations that had been made throughout the literature on the correlation of DA dynamics with movement parameters. They confirmed that DA release within the dorsolateral striatum was positively

correlated with translational velocity and that DA showed burst-pause fluctuations with movement onset after a movement pause [20, 33]. They further showed that DA fluctuations were on average distinct for individual syllables, suggesting that DA might predict what syllable is used. However, they made several further observations that were inconsistent with the idea that DA might directly control movements. For instance, very kinematically similar syllables often had vastly different associated DA fluctuations, and very different syllables could have similar DA fluctuations. Adding to this complexity, individual instances of syllables often could have a fair amount of variability in the amplitude and shape of the associated waveform. For these reasons, machine learning classifiers were unable to predict what syllable was used based on the DA waveform. They further found that many aspects of the execution of behavioral syllables were not correlated with DA activity. For these reasons, the authors concluded that DA signals do not carry information that can specify movements or their kinematics.

Markowitz and colleagues then presented several experiments to build their case for a DA reinforcement driven understanding of animal behavior within the open field. The basic premise of their argument was that DA fluctuations could be reinforcing individual behavioral syllables. They made the important observation that the magnitude of DA release within the dorsolateral striatum to surprise rewards was similar in magnitude to the DA signals associated with spontaneous behavior. This implied that reinforcement was possible in this context even if there were no explicit rewards and raises the interesting proposition that the brain produces intrinsic reinforcement. The authors noticed that syllables associated with the largest DA release tended to be used more often during a given session and that large DA release amplitudes at the time a certain syllable was used predicted increased use of that syllable for several minutes after. The converse was also true with low DA amplitude predicting less syllable use. This effect also appeared to be specific to individual syllables and did not generalize to adjacent syllables. Interestingly, they also observed that large DA transients made the following syllables less predictable and smaller DA transients made following syllables more predictable. In other words,

DA was able to impact behavioral sequence entropy in addition to its effects on individual syllable usage. Like many other studies, Markowitz and colleagues leveraged optogenetics to further bolster their claims and establish causality. Specifically, the authors used a regime of physiologically calibrated optogenetic DA terminal stimulation in dorsolateral striatum paired with targeted behavioral syllables in a closed-loop system. This resulted in a "stable" (non-exponential) upregulation in the use of the targeted syllable over time that persisted even after stimulation was halted, consistent with a reinforcement model. They were also able to recapitulate the observed effects of DA release on behavioral sequence variability, but this effect did not persist after stimulation was stopped. The authors then restricted optogenetic stimulation to particularly fast or particularly slow executions of syllables, resulting in upregulation or downregulation of the speed of that syllable's future executions, respectively. This implied that previous reports of DA's effects on aspects of vigor or motor execution could be explained by a reinforcement effect. Overall, Markowitz and colleagues concluded that DA acts a "moment-to-moment teaching signal" that shapes behavior even in the absence of explicit extrinsic reinforcers.

While there is strong evidence supporting the DA learning primacy hypothesis, there are some potential concerns when attempting to reconcile this idea with the existing literature. The first and most obvious concern has to do with the definition and analysis of animal behavior in these studies, with implications for how effects on motor actions are then evaluated. The papers that have been discussed in this section essentially involve performing calibrated optogenetic perturbations on DA neurons and then determining effects on behavior as a result. At the most basic level, the way that behavior is defined or measured could obviously impact the results. For instance, one could claim that the behavioral metrics used by Coddington and Dudman, Lee et al., and Pan et al. were inadequately detailed to capture the effects of DA on movement [37-39]. Certainly, Hughes et al. claimed this was a major advantage of their use of force sensors in the context of head fixed mice over the accelerometer paired basket used by Coddington and Dudman [26]. In general, the idea would be that even if effects weren't seen on gross metrics of

movement or licking, there could still be subtle differences that simply weren't measured or lost due to using a head fixed paradigm. For example, a mouse might be less forcefully extending its tongue in a way that isn't reflected in lick probability and lick count, or a mouse might perform a number of subtle forward and backward postural or positioning shifts that can only be detected through more sophisticated methods. While many studies that do claim dopaminergic effects on motor behavior did not use particularly subtle measurement methods (in particular [20, 25, 28, 29]), it is true that the studies discussed in this section did not perform truly exhaustive behavioral measurements and analyses. A similar logic could be applied to the syllables examined by Markowitz et al. in that the machine learning algorithms could be claimed to erase a large degree of behavioral complexity when categorizing actions into syllables or that their systems simply couldn't detect these subtleties [40]. Regardless of how plausible one personally finds this line of reasoning, it nonetheless means that consideration of behavioral evidence (even if optogenetics was used) may leave some lingering doubt about the extent of the effects on behavior.

A second related concern is that the papers that support the DA learning primacy hypothesis still show at least some degree of rapid effects on movement. The two best examples of this are the fact that Lee et al. showed a small effect of pre-reward inhibition of SNc DA neurons on movement and Markowitz et al. showed transient increases in behavioral sequence variability following DA neuron stimulation [38, 40]. Ultimately, even though all these papers conduct data analysis involving direct comparisons and statistical modeling to illustrate that the effects of DA on reinforcement are stronger (or that reinforcement explains the data better), these are ultimately judgements about behavior that have room for subjective interpretation. When combined with the first criticism of the learning primacy hypothesis, one could say that the true extent of the behavioral effects is simply not explored and that the observed behavioral effects are deeply underappreciated. If we then take the perspective, as is reflected in much of the literature in the prior section, that the effects of DA on movement are complex, probabilistic, at times mixed, and not yet fully understood, it may seem unfair to perform head-to-head comparisons between the

learning and movement functions of DA. In other words, declaring that one function has “primacy” over the other on the basis of this evidence may be premature.

A third issue with the learning primacy hypothesis has to do with the relatively homogenous actions of optogenetics on a heterogeneous population of DA neurons. There has been a growing appreciation over decades of research on DA that there is a large diversity in the targets, release dynamics, encoded information, and electrophysical properties of DA neurons within the brain [25, 34, 45-47]. While optogenetics may be able to specifically target DA neurons within a certain area, the individual neurons are often not parsed or subdivided or functionally characterized beyond this in much of work we have discussed in this section. Whether one ascribes to the view that DA neurons can be coherently grouped based on target regions or the view that DA neurons and their targets ought be considered at the individual neuron level, researchers appear to generally agree that optogenetic manipulations likely affect a broad group of neurons with diverse properties [35, 46]. It could then be possible that the reason why strong behavioral effects were not observed in the studies supporting the learning primacy hypothesis is because optogenetic stimulation did not properly reproduce the detailed dynamics needed for DA to exert its normal behavioral control. If one ascribes to the view that DA encodes and faithfully transmits complex information to downstream targets, optogenetics would likely not evoke activity that carries any kind of coherent information, leading to a lack of convincing effects even if carefully calibrated. In other words, optogenetic DA neuron stimulation, even if calibrated to the average population magnitude associated with natural rewards, may not in fact faithfully recapitulate the complex single neuron level dynamics of normal signaling. If DA’s ability to drive learning was less subject to degradation by the neuronal “gibberish” evoked by optogenetics, then it could lead to the mistaken impression of learning’s primacy over movement.

Thus, there are three main objections to the evidence supporting the learning primacy hypothesis: that subtle behavioral effects were missed, that observed behavioral effects were unfairly assessed, and that optogenetics (even if calibrated) is too blunt a tool. While these

concerns may seem disparate, they all surprisingly hinge on a single question: to what degree and in what way does DA shape the activity of striatal neurons? It is well understood that striatal circuits have an important role in controlling behavior and that changes in these circuits and their outputs can have clear, rapid behavioral consequences [23, 24, 48, 49]. Because DA release occurs extensively throughout the striatum, many of the studies that have supported a role for DA in subsecond control of movement have logically proposed that rapidly altering these striatal circuits must be the mechanism by which DA does this (previously discussed examples include [20, 25, 26]). If DA rapidly and strongly shapes neuronal activity in the striatum, then this would lend some credence to the concerns discussed above with the learning primacy hypothesis. It would then be conceivable that physiologically calibrated levels of DA still play a major role in rapidly shaping neural activity and that the weakness of observed behavioral effects are a reflection of the limitations of optogenetics rather than a true reflection of DA's function. Therefore, proponents of this position would opine that Markowitz et al. and Lee et al. would have seen rapid perturbations in striatal dynamics upon optogenetic inhibition even in the absence of behavioral effects had they attempted to measure them [38, 40]. However, if DA does not strongly and rapidly shape activity in the striatum, then this would largely vindicate the core claims of the learning primacy hypothesis. This is because it would raise confidence in the characterization of DA's rapid behavioral effects as truly weak (relative to reinforcement effects), and it would undermine arguments dependent on intricate and faithful propagation of diverse informational inputs by DA neurons to the striatum. After all, it becomes much more difficult to dismiss the learning primacy hypothesis in favor of increasingly granular examinations of behavior and DA neuron activity if the basis of the proposed causal link between the two is found to be inadequate to support rapid, precise control.

1.5 Drugs of abuse and supraphysiological DA signaling

Before continuing, I feel it is worthwhile to briefly discuss the relevancy of supraphysiological DA signaling in the learning primacy hypothesis. As stated earlier, a major component of the learning

primacy hypothesis is that DA is predicted to have rapid effects on striatal activity if DA greatly exceeds a physiological range associated with natural rewards. One pathological scenario of interest where this is expected to occur is in the context of addiction and the effects of drugs of abuse. Many drugs of abuse greatly increase DA signaling, including baseline levels (tonic) and rapid spikes in DA release (phasic), and this alteration in DA signaling is thought to contribute to the development of addiction [50, 51]. This has led to the overall view that alterations in DA signaling are a "final common pathway" for addiction, although the effects of drugs of abuse and the mechanisms by which they may contribute to addiction are not exclusively dopaminergic [51-54]. Over time, these effects on DA neurotransmission can lead to chronic, widespread changes in the brain and cause addiction [54].

In recent years, opioid use disorder and opioid-associated overdose deaths in particular have emerged as severe and persistent public health challenges [55]. Studies utilizing microdialysis, fast-scan cyclic voltammetry, and fluorescent DA sensors have found that morphine and other mu opioid receptor agonists significantly increase tonic and phasic DA release [56-61]. Mu opioid receptors typically have an inhibitory effect, so one of the most long-standing models argues that opioids disinhibit VTA DA neuron firing principally through inhibiting midbrain GABA neurons [62, 63]. Studies have found that opioids tend to increase the amplitude and frequency of phasic, subsecond DA release events independent of increased basal DA release [58, 61]. These increased transients were observed in animals at rest and were not evoked by any particular sensory stimulus. However, this raises the possibility that phasic DA responses to sensory stimuli of different valence might also be altered by systemic morphine administration, perhaps amplified to supraphysiological levels [64].

1.6 Dopamine's effects on striatal circuits: electrophysiological evidence

The lack of clear consensus about DA neurons' role in subsecond behavioral control, suggests the need to reconsider the electrophysiological evidence on which this argument hinges – that

striatal spiking activity is strongly and rapidly influenced by physiological DA signals. Superficially, this dependence appears to have been already established, as there is a large body of literature using in vivo and ex vivo measurements to show that DA indeed influences striatal activity. Based on that work, many of the papers supporting a role for DA in subsecond behavioral control have treated DA's ability to rapidly influence activity in the striatum as a scientific fact. However, as we will see, while in broad terms the claim that "DA can impact striatal firing rates" is true, DA's capacity to support subsecond, complex regulation of behavior is not a forgone conclusion. In this section, we will explore previous attempts to study the effects of DA on striatal neurons and discuss why limitations of those studies have prevented a truly definitive settlement of this issue.

Historically, early efforts to understand DA's modulatory effects on striatal activity leveraged iontophoresis and electrical stimulation of dopaminergic neurons and formed the basis for a number of ideas that would persist in the literature. Iontophoretic experiments generally involved applying current on a small glass micropipette filled with a solution that contained DA, which caused the expulsion of DA from the tip of the micropipette [65-67]. These experiments generally occurred in anesthetized animals and the spiking responses of striatal neurons were measured to assess effects of DA. However, iontophoretically applied DA generally was not calibrated to physiological levels and it is unclear the degree to which it recapitulated physiological DA release events [65-67]. Another early approach came in the form of electrical stimulation of the midbrain or the medial forebrain bundle [68-70]. These experiments involved driving activity in DA pathways using direct electrical stimulation and examining the activity of striatal neurons in response. However, it became clear over time, particularly for medial forebrain bundle stimulation, that these techniques were not clean dopaminergic manipulations and recruited multiple neurotransmitter systems and circuits [70]. In fact, one key study found that the striatal effects of stimulating the medial forebrain bundle may be largely driven by GABAergic signaling rather than dopaminergic signaling [70]. While these approaches were influential in establishing the concept

that DA and DA circuits could strongly modulate activity in the striatum, their limitations meant that further work was needed to explore these questions.

Following these earlier approaches, other methods included pharmacological manipulation of DA neurons or receptors [71, 72], as well as chronic DA lesions [73-75]. These approaches are primarily intended to examine the protracted modulatory contributions of DA (e.g., synaptic plasticity) that occur over multiple seconds to weeks. But they are less suited for identifying rapid electrophysiological effects on subsecond timescales. One influential 2014 study from Hoffmann and Nicola that utilized pharmacological blockade of DA receptors is one of the most prominent that is often referenced [5, 72]. These authors performed in vivo electrophysiological recordings with an electrode array that was paired with a drug injection cannula, which allowed the authors to record from striatum while locally applying DA receptor antagonists. The headline result of this study was that both D1 and D2 receptor antagonists broadly reduced the spiking response of nucleus accumbens neurons to a conditioned stimulus. Even when the antagonists were applied ipsilaterally and no behavioral effects were observed, the effects of the DA receptor antagonists on local striatal activity were still seen. The results did generalize somewhat to neutral stimuli, but D1 antagonists notably failed to alter striatal responses to neutral stimuli. As mentioned above, Hoffmann and Nicola recognized that it was debatable the degree to which these effects can be attributed to alterations in shorter timescale or longer scale dopaminergic effects such as plasticity. Nevertheless, the paper is remembered as a direct in vivo demonstration of DA's ability to depress neuronal activity in the striatum. A well-known DA lesion study that performed in vivo electrophysiology is the 2017 paper from Ketzef and colleagues. These authors unilaterally injected 6-hydroxydopamine into the medial forebrain bundle, which resulted in the degeneration of dopaminergic neurons within the ipsilateral striatum. They used optogenetic identification to distinguish between D1 and D2 MSNs. Ketzef et al. were focused on the impacts of DA depletion on sensory processing in the striatum, and they observed several interesting changes in striatal neurons including depolarized resting membrane potentials, lower

D2 spontaneous activity, and smaller striatal sensory responses. Because of the long time scale of this DA lesion and possible compensatory effects, this study only indirectly supports the contention that DA could be rapidly shaping striatal activity. Both lines of evidence from pharmacological or lesion studies justify the basic premise that DA impacts the activity of the striatum, but neither lines of evidence provided definitive clarity on the issues surrounding the learning primacy hypothesis.

Of course, optogenetic manipulations address both the need for fast temporal control and cellular specificity, and they have been as impactful here as they were for the other lines of neuroscience investigation we have discussed [4, 36, 76, 77]. In one of the clearest examples, Wang et al. performed optogenetic stimulation of the VTA while performing in vivo electrophysiology in the ventral striatum [76]. They found that nearly half of their recorded neurons were rapidly excited to this stimulation within 200 milliseconds, with another 13 percent inhibited in this same timeframe. Even when Wang and colleagues repeated these experiments in a dopamine neuron selective glutamate transporter knockout mouse, they attenuated but did not eliminate this rapid effect, showing that it was not solely due to glutamate (a “fast” neurotransmitter). Similarly, Ferenczi et al. used functional magnetic resonance imaging paired with optogenetic stimulation of DA neurons and found robust neuronal responses in the striatum, albeit this study did not leverage electrophysiology [77]. Another notable study from Tye et al. involved a similar experimental paradigm to Wang et al. with a focus on depression-related escape behaviors [4]. Using in vivo electrophysiology, they saw neurons in the nucleus accumbens that were responsive to VTA stimulation, and they also found that VTA stimulation could alter striatal neuron responses to escape-related behavior. To be clear, the nature of the effects found in these studies, particularly from Wang et al. and Tye et al., would argue against the learning primacy hypothesis as DA’s effects on striatal neurons were robust and rapid. However, these three previous studies did not adequately calibrate their optogenetic stimulation nor consider the potential distinction between physiological and supra-physiological DA levels, which is a core issue at the heart of the

learning primacy hypothesis [37]. Still, there has been electrophysiological work done after the revelations of Coddington and Dudman that has taken these issues into account [36, 37].

Perhaps some of the strongest electrophysiological evidence that considered the issue of the magnitude of optogenetically evoked DA release is the 2020 paper from Lahiri and Bevan that examined the effects of DA release on D1 MSNs in ex vivo brain slices [36]. At the outset, there are a couple of critical pieces of context needed to fully understand the experimental decisions made by Lahiri and Bevan as well as appreciate the impact of their work. Although a comprehensive examination is outside the scope of the current literature overview, suffice it to say there have been decades of extensive work done at the single cell level as well as at the protein level to characterize the signaling pathways and effects evoked by DA binding to its receptors [78, 79]. DA receptors are GPCRs and can impact cAMP-regulated protein kinase A (PKA) pathways or DAG-IP3-regulated protein kinase C pathways. Of course, we have already discussed findings from Yagishita and colleagues that show the importance of these actions in facilitating long term potentiation and changes in dendritic spine morphology [14]. However, these pathways may also act on targets that could have distinct and decidedly more rapid effects. Particularly through PKA-dependent actions, DA receptors can modulate voltage-gated K+, Na+, and Ca2+ channels in the neurons that bear these receptors, which could change the ease with which a neuron achieves and action potential [78]. However, it is worth pointing out that DA receptors do not act on G-protein gated inwardly rectifying potassium channels or hyperpolarization-activated cyclic nucleotide gated channels, both of which would quickly and directly change membrane potential [35]. Because of DA binding to its receptor does not directly result in ion channels being opened or closed, DA will not change the membrane potential upon binding its receptor because the overall conductance of the membrane to ions does not change [79]. This is why Lahiri and Bevan focus on excitability and must inject current into their D1 receptor bearing target neurons; they never expected DA alone to have an impact on membrane potential [36]. Lahiri and Bevan also understood that the effects that they were looking for would have been highly dependent on second

messenger cascades, which could be disrupted by whole cell patching techniques where the fluid inside an electrode is made continuous with the intracellular space. For this reason, they utilized an approach involving the antibiotic gramicidin that forms a pore that is permeable to cations and small uncharged molecules but not other molecules [80]. This allowed Lahiri and Bevan to record the membrane potential of a target neuron without adversely affecting second messenger or protein kinase concentrations because these molecules wouldn't be able to diffuse through gramicidin pores into the electrode space. They suggested that this is why finding rapid effects of DA on membrane excitability, particularly through PKA pathways, have historically proven difficult and controversial [78, 81].

With this background in mind, Lahiri and Bevan sought to settle this issue by using the gramicidin perforated patch-clamp method focusing on dorsolateral striatum D1 MSNs, neurons that are thought to be the most relevant to movement onset [25, 36, 40]. The authors used a mouse line that enabled them to selectively express an excitatory opsin in DA neurons and identify D1 MSNs based on the expression of tdTomato. In ex vivo slices, they were able to record from positively identified D1-MSNs, optogenetically evoke DA release, and record the magnitude of DA release with voltammetry. Using 4 Hz and 20 Hz laser frequencies, the authors were able to evoke distinct levels of DA that they claim were physiologically relevant, referencing Coddington and Dudman [37]. The evoked DA concentrations were on average ~300 nM for the 4 Hz stimulation and ~500 nM for the 20 Hz stimulation, with some 20 Hz pulses producing micromolar DA release. Whether or not these concentrations ought to be considered physiological is debatable, but they are a great deal higher than the in vivo 10-100 nM range that has been typically associated with food rewards and related motivated behavior [29, 31]. At any rate, the authors verified that their optogenetically evoked releases appeared reasonably similar to previous work in magnitude and decay. The authors conducted their experiments in intentionally silenced slices with a cocktail of receptor antagonists for GABA, glutamate, and acetylcholine receptors blocking the actions of any neurotransmitters that might be co-released with DA. Because spontaneous circuit activity

would have been absent, the authors had to induce firing in targeted D1 MSNs through the injection of current. The authors used a “long up-state” electrical stimulation protocol that caused recorded neurons to fire a sustained train of action potentials over several seconds. After a few seconds of initial depolarization had elapsed, the authors continued this membrane depolarization and then optogenetically induced DA release onto the target neuron and examined changes in the neuron’s firing rate. Comparing the firing rate before and after DA release, the authors found a robust D1-PKA dependent increase in firing rate of D1 MSNs. The overall effect size, about 5-10 Hz, was the same for a single 2 millisecond laser pulse as well as for their 4 Hz and 20 Hz laser stimulations. The onset time of this effect was generally subsecond with a median of 458 milliseconds, and it could last for over 10 minutes following a single laser stimulation. They also found that DA could decrease the latency of neuron firing after current injection and that this could facilitate the transition of striatal neurons to upstates. Thus, Lahiri and Bevan were able to uncover a rapid and robust mechanism by which DA might be able to increase excitability of D1 MSNs and ostensibly invigorate movements.

However, there were a number of limitations of these ex vivo slice experiments and details of the results that meant that it could not necessarily be readily generalized to an in vivo context. It was contemporaneously understood that Lahiri and Bevan’s work could not definitively sway the debate about DA’s effects in vivo [36, 82]. In an accompanying commentary from McGregor and Nelson, there was recognition that the effect uncovered by Lahiri and Bevan may play an overall regulatory role in balancing striatal circuit excitability without necessarily playing a direct role in movement onset [82]. Of course, MSNs in vivo constantly receive dynamic excitatory and inhibitory inputs from many brain areas, and the overall impact of this dopaminergic excitability effect is not necessarily easily predictable from Lahiri and Bevan’s experiments. McGregor and Nelson further made the point that the 10+ minute duration of the effect and the fact that the effect’s ceiling could be reached even with minimal optogenetic excitation (i.e. a single 2 ms pulse) meant that this might be more relevant for the “tonic” functions of DA. Indeed, MSNs in vivo undoubtedly receive

more than one pulse of DA in a 10-minute span, raising the possibility that this effect is constitutively online [35]. If the average DA activity level fluctuated over minutes long timescales, then this could result in this excitability effect being titrated up or down over that span. Of course, there might be compensatory striatal circuit mechanisms that could counteract and "cancel out" such an effect. A later literature review from Sippy and Tritsch largely affirmed these points about Lahiri and Bevan's work but also introduced other caveats regarding ex vivo preparations that are worth considering [35]. Sippy and Tritsch point out that the resting levels of DA within striatal brain slices is generally very low as they only contain the severed axons of DA neurons, and they would be expected to be near zero in Lahiri and Bevan's preparation where the slice was pharmacologically silenced. While Lahiri and Bevan focus on the magnitude of their evoked DA release as being physiologically relevant (which is itself potentially objectionable), the presumable absence of background DA release in their slice is likely not physiologically representative and is arguably closer to a Parkinsonian DA depletion state. Viewed this way, Lahiri and Bevan's results might be better interpreted as a mechanism for the therapeutic benefits of levodopa, a Parkinson's medication that increases DA availability. Or their results might explain the long timescale effects caused by blocking DA receptors seen in Hoffman and Nicola's work [72]. Sippy and Tritsch also mention that there are other pathways by which DA might impact striatal firing other than DA binding to receptors on MSNs. While circuit activity outside of MSNs and compensatory mechanisms could cancel out DA's effects, they could just as easily be the mechanism by which DA impacts striatal circuits. For instance, there is interplay between dopaminergic and cholinergic activity, and cholinergic interneurons could in turn shape striatal activity [83, 84]. Slice preparations, such as those used by Lahiri and Bevan, understandably block neurotransmission that is not the target of their investigations, but this means that the data cannot rule in or out the contributions of those neurotransmitter systems.

In summary, due to a variety of limiting factors, previous studies have fallen short of directly addressing to what extent reward-matched or other physiologically calibrated DA signals

control in vivo striatal spiking activity on subsecond timescales. While there is evidence that suggests it is plausible that DA might rapidly influence striatal activity, there is a clear need for this question to be definitively explored in vivo. To reiterate, clarity on this issue is critical to determine the viability of the DA learning primacy hypothesis, which thus far has been primarily based on behavioral results without strong electrophysiological backing.

1.7 The current study: to what degree does DA rapidly alter in vivo striatal activity?

To address this significant gap, we examined the ability of DA neurons to elicit strong and rapid changes in striatal spiking activity. We focused on the role of DA in the ventral striatum during the delivery of unexpected rewarding stimuli, precisely because these stimuli elicit large and rapid DA responses in the ventral striatum [85] – and thus, in principle, appear ideally suited to drive strong electrophysiological effects. Since unexpected rewards induce concurrent changes in striatal firing patterns [86], we tested whether DA signals are necessary or sufficient for driving reward-evoked neural activity in the ventral striatum. Experiments were carried out by optogenetically adding or subtracting DA activity while simultaneously monitoring ventral striatal DA and electrophysiological activity. This combined approach enabled a systematic study of striatal dynamics under calibrated reward-matched and supra-reward DA signaling conditions. We also performed a set of pilot experiments to examine the effects of morphine administration on DA responses to rewards and other stimuli.

Our main finding was that reward-matched DA levels had only a small, inconsistent effect on spontaneous and reward-evoked striatal firing rates, whereas artificially elevating DA to supra-reward magnitudes produced robust electrophysiological effects. The small effect size stood in contrast to the markedly stronger electrophysiological changes caused by manipulating non-DA inputs to the striatum. Together, these results have three key implications. First, these observations indicate that DA signals at physiological levels (i.e., levels evoked by food rewards) play a relatively minor role in rapidly shaping neural activity in the striatum. This is surprising given

that DA signals appear to correlate well with moment-to-moment behavioral variables such as movement kinematics and vigor, and this correlation has often been assumed to suggest strong causal relevance of DA in performing those behaviors. Thus, on rapid timescales, striatal spiking is likely to be primarily influenced by non-DA inputs. Indeed, previous work has demonstrated the effectiveness of glutamatergic terminals from cortex and thalamus in driving striatal reward activity [87]. Second, these data provide a plausible mechanism to explain previous negative behavioral results, in which reward-matched DA stimulation failed to meaningfully alter performance [28, 37, 40]. Finally, the significant enhancement of electrophysiological effects seen at supra-reward DA levels illustrates the potential for overestimating the importance of DA if the supra-physiological regime is not identified as such. To avoid this, there is a need to adopt standardized calibration practices for manipulating DA neurons. Collectively, this work places important constraints on the magnitude of DA neuron's contribution to striatal dynamics on subsecond timescales.

CHAPTER 2: Experimental results

2.1 Simultaneous monitoring of DA and electrophysiological activity in the ventral striatum

In order to measure neural population activity in the ventral striatum under calibrated levels of change in DA signaling, we constructed an opto-probe device consisting of a multielectrode recording array on a silicon microprobe, coupled to an optical fiber for photometry (**Fig. 1a**). The optical fiber was used for fluorescence-based DA monitoring in the vicinity of the electrodes with the genetically encoded sensor dLight1.2 [44]. The opto-probe was implanted in the nucleus accumbens area of the ventral striatum in head-restrained food-restricted mice. In each animal, the electrodes captured the spiking response of tens of single-units (mean ± SD: 95 ± 44 neurons) alongside photometric data on local DA signaling changes. Another optical fiber was implanted in the VTA to allow for optogenetic manipulation of DA neurons using virally mediated Cre-dependent opsin expression in DAT-Cre mice (**Fig. 1b**). Unconditional rewarding stimuli (sweetened milk) were delivered at random intervals to elicit robust increases in licking and DA release (**Fig. 1c**). Concurrently, the majority (58 %) of recorded neurons in the ventral striatum were modulated by unexpected rewards (**Fig. 1d,e**) [86]. Neurons showed a predominantly excitatory subsecond-scale initial reward response, which was sometimes followed by a more prolonged period of inhibition. A strong reward response was observed in all three major putative classes of electrophysiologically identified cell types, corresponding to medium spiny projection neurons (MSNs), fast spiking interneurons (FSIs), and tonically active projection neurons (TANs) (**Extended Data Fig. 1**). However, to avoid possible bias arising from cell-type specific modulatory effects or classification errors, our primary analysis used spiking data from every recorded neuron. These combined measurements revealed temporally correlated changes in ventral striatal DA and spiking activity during reward delivery. Based on the view that DA neurons drive changes in spiking on subsecond timescales, it has been posited that these two forms of

reward signals are causally related [70, 88]. We therefore bidirectionally manipulated DA activity in order to directly test this prediction.

2.2 Small effect of inhibiting DA reward transients on striatal spiking activity

First, we examined whether DA neuron signaling is necessary for reward-evoked striatal spiking activity, through a transient loss of function approach involving opto-inhibition (**Fig. 2a**). We expressed eNpHR3 in VTA DA neurons, and timed laser stimuli to occur together on 50 % of reward trials (5 mW, 0.5 s continuous laser). Trials consisting of only reward (R) and laser-paired reward (R+L) were randomly interleaved. As expected, reward-evoked DA release in eNpHR3 expressing animals, but not opsin-free controls, was significantly inhibited and even transiently fell below baseline levels (**Fig. 2b,c**). Miniscope imaging of dLight signal in the ventral striatum appears to confirm that the transient reduction in DA levels was widespread (**Extended Data Fig. 2**). The duration of opto-stimulation was intentionally matched to previous studies supporting the involvement of DA neurons in rapid behavioral control [20, 31]. Consummatory licking was not significantly altered by DA neuron inhibition, which was applied unilaterally (**Fig. 2d**). This allowed us to examine changes in neural dynamics in the absence of potentially obfuscating behavioral changes [72]. According to a prominent view of DA function, transiently reducing DA levels should cause strong and rapid changes in reward-evoked spiking responses in the ipsilateral ventral striatum. To test this we directly compared differences in individual neuron firing profiles between reward and laser-paired reward trials (**Fig. 2e,f**). The average change in firing rate was small for all three putative cell types, with only FSIs showing a significant difference relative to controls (**Extended Data Fig. 3a**). We checked if further subdividing the different cell types by their spatial location in the ventral striatum would reveal hotspots with high firing rate changes. While MSN activity did show significant spatial variations, these changes had no clear pattern, and were consistently small in magnitude (**Extended Data Fig. 3b,c**). There was also no significant change in the duration of the TAN pause following reward delivery (**Extended Data Fig. 3d**).

Due to the small magnitude of firing rate differences, optogenetically induced changes in neural activity were further quantified by applying a statistical test to identify time bins with significant differences in firing (see Methods, **Fig. 2g**). Overall, DA neuron inhibition altered the reward response of 8 % of neurons, compared to 4 % in control animals – a statistically significant but modest difference (**Fig. 2h**). We also calculated a selectivity index based on the area under the receiver operating characteristic (ROC) curve (range of index: ±1). Positive/negative index values denote higher/lower mean spike rates during DA inhibition, respectively. On average the magnitude of the selectivity index differed significantly from controls, but again by a small amount (**Fig. 2i**).

The above results reveal a small but statistically significant effect of DA neuron inhibition on individual striatal neuron firing properties. But it is still conceivable that these subtle electrophysiological changes may be important for regulating striatal dynamics and, in turn, behavior, on subsecond timescales. Indeed, if DA inhibition weakly but consistently alters the activity pattern of even a small group of striatal neurons, in principle, downstream brain areas could reliably read out this information to guide behavior. To test this possibility, we attempted to distinguish population-level dynamics observed between reward and laser-paired reward trials (i.e., R versus R+L), using a machine learning-based decoder tasked with performing binary classification. Decoding accuracy, reflecting the percentage of correctly classified trials, was briefly elevated following reward delivery, but differences between eNpHR3 and opsin-free control mice were still modest (**Fig. 3a**). It is also notable that the decoder never performed significantly above chance levels, as defined by the 95th percentile of decoders trained on trial-shuffled data. Decoder performance using data from eNpHR3-expressing mice improved by training the classifier with higher numbers of neurons, and eventually surpassed control data (**Fig. 3b**). However, even under the most favorable decoder training conditions using spiking data from 400 neurons, only a 13 % improvement in performance was observed (75 % accuracy from eNpHR3 group data, compared to 62 % from control group data). Thus, even with the aid of population-

level recordings and highly sensitive machine learning tools, we could not consistently distinguish reward from laser-paired reward trials.

On a subset of trials and animals in the same experiment, we also examined DA contributions to spontaneous striatal spiking activity, by applying identical laser stimulation but without accompanying rewards (**Extended Data Fig. 4a**). This caused DA levels to transiently fall below baseline levels (**Extended Data Fig. 4b,c**). We looked for changes in neural activity relative to a baseline period preceding laser stimulation. Neither the firing rate differences, the fraction of selective cells, nor the mean selectivity index were significantly different from controls (**Extended Data Fig. 4d-g**). A statistically significant difference arose when testing a decoder trained to distinguish laser-evoked population activity from baseline activity, but again the size of this effect was small (**Extended Data Fig. 4h,i**). Taken together, the results show that DA neuron inhibition has a weak effect on both reward-evoked and spontaneous spiking activity in the ventral striatum.

To test whether these results may generalize to more complex behaviors, and to other striatal regions, we conducted an additional experiment where DA was inhibited during the reward anticipation phase of a Pavlovian conditioning task (**Extended Data Fig. 5a**) [38]. We recorded from the ventral striatum while inhibiting DA neurons in the VTA, or from the dorsal striatum while inhibiting DA neurons in the SNc. DA neuron inhibition was performed for half the trials and caused a temporary decrease in DA levels below baseline (**Extended Data Fig. 5b,c**). Anticipatory licking was not changed by unilateral DA neuron inhibition (**Extended Data Fig. 5d**). When comparing spiking data from inhibited trials (C+L) to uninhibited trials (C), the fraction of neurons that were selective for each trial type was not significantly different in ventral or dorsal striatum compared to controls (**Extended Data Fig. 5e,f**). A small number of neurons that were positively modulated in the eNpHR3 group drove a brief increase in decoding accuracy when trained on combined ventral and dorsal striatum data; however, a similar brief increase in decoding accuracy was seen in controls and the overall difference from controls was similar in magnitude to our other opto-inhibition experiments (**Extended Data Fig. 5g,h**). Decoding

accuracy when compared to controls was higher in the ventral striatum than in the dorsal striatum (**Extended Data Fig. 5i**). These results suggest that DA's small effects on striatal neuron spiking also apply in the dorsal striatum and during reward anticipation behaviors.

To better put our findings into context, and ensure that the lack of strong electrophysiological effects cannot be attributed to problems with our recording or analysis approach, we carried out a positive control experiment involving activation of VTA GABA neurons, which locally regulate DA neuron firing, but also directly project to other areas including the ventral striatum [63, 89, 90] (**Fig. 4a**). Since these neurons release an inhibitory neurotransmitter with rapid postsynaptic effects, we expected their contribution to striatal spiking activity would substantially exceed those of DA neurons. We expressed the excitatory opsin Chrimson in VTA GABA neurons, and delivered pulsed laser stimulation on a subset of reward trials. Activating these neurons produced a marked reduction in both reward-evoked striatal DA release and spiking activity, consistent with their inhibitory action (**Fig. 4b,c**). Similarly, at the population level, neural decoding performance reached 95 % accuracy with as few as 25 neurons used for training the classifier (**Fig. 4d**). VTA GABA neuron activation without reward delivery also had a large effect on spontaneous striatal firing rates (**Extended Data Fig. 6a-d**). These data clearly show that our analysis methods are capable of detecting strong changes in neural activity if they exist. Furthermore, the stark contrast between the effect of VTA GABA neuron activation and DA neuron inhibition on striatal spiking activity, supports the view that VTA GABA neurons do not only serve to inhibit local DA neurons, but to directly regulate striatal microcircuits [89].

2.3 Small effect of artificial reward-matched DA transients on spontaneous striatal activity

While our results suggest that DA signals during natural reward delivery only weakly contribute to striatal spiking activity, we cannot rule out the presence of enhanced effects under artificially elevated DA signaling regimes. To address this, we next used opto-activation to examine whether certain levels of DA are sufficient to produce strong and rapid changes in striatal spiking. The

excitatory opsin Chrimson was expressed in VTA DA neurons, and pulsed laser stimulation was applied as before for a total duration of 0.5 s, both in isolation and paired with a subset of reward trials (**Fig. 5a**). To systematically study the effect of different DA levels on spontaneous striatal dynamics, we presented isolated laser stimuli and varied the stimulation frequency from 4 to 40 Hz in order to evoke increasing magnitudes of DA release (**Fig. 5b**). We then compared the peak level of optogenetically evoked to behaviorally evoked DA during reward delivery. The ratio between these values indicated the factor increase in DA above reward-matched levels [37]. Across all sessions tested, this factor varied between 0.2 and 23, representing two orders of magnitude change in DA levels relative to reward. We next looked for laser-evoked changes in the spontaneous activity of simultaneously recorded neurons, relative to a pre-laser baseline period. Under certain conditions we found alterations in spontaneous firing upon DA neuron activation (**Fig. 5c**), but it was evident that the size of this effect depended on the level of DA release. Specifically, spontaneous spiking activity was only weakly sensitive to reward-matched DA levels, but appeared more strongly inhibited with higher, supra-reward DA levels (**Fig. 5d**). A similar trend emerged when testing a decoder trained to distinguish laser stimulation-evoked population activity from pre-laser baseline activity. Across all the data collected, we found an approximately linear change in the percentage of laser-modulated cells, selectivity index, and decoder performance as a function of the factor increase in DA (**Fig. 5e-g**). The firing rate changes of different cell subtypes showed qualitatively similar trends, with MSN activity displaying a linear dependence on DA levels (**Extended Data Fig. 7a-c**). Together, these results suggest that artificial DA signals comparable to the magnitude of rewards are not on their own sufficient to cause large, rapid changes in striatal spiking. But at the same time, supra-reward DA transients are clearly capable of driving strong changes in neural activity.

Our estimate of the factor increase in DA may depend on several variables such as recording location. A large overestimate of this factor would be problematic, as it would weaken the conclusion that only supra-reward DA signals are sufficient to produce sizable

electrophysiological effects. Spatial variations in dLight photometry signals are one potential source of uncertainty. Our opto-probe was configured such that the photometry fiber was placed above the most dorsal electrode recording sites, and thus was separated by over 1 mm from the most ventral recording sites (**Fig. 1a**). Since it is known that DA reward signals vary substantially across the striatum [85], the factor increase in DA cannot be assumed to be constant over the entire span of the electrode array along the dorsal to ventral axis. To quantify how this factor varies with depth, we performed a dLight photometry experiment with the optical fiber gradually moved more ventrally (6 different locations spanning 1 mm; **Extended Data Fig. 8a**). At each location we measured the magnitude of DA signals during reward delivery and optogenetic stimulation. While both of these signals varied considerably across different locations (**Extended Data Fig. 8b,c**), overall the factor increase in DA displayed a moderate increase with depth (**Extended Data Fig. 8d**). In experiments involving combined electrophysiology and photometry, access to DA signals from only near the most dorsal optical fiber position was available. Thus, these data suggest we may have underestimated the value of the factor increase in DA. If this is the case, it would actually strengthen our conclusion. Another source of variability may relate to changes in DA reward signaling across successive trials. In particular, it is possible that the initial set of reward trials are more unexpected than latter trials, and therefore induce a stronger RPE response. To check if this impacts our estimate, we calculated a recalibrated factor increase in DA using the average of only the first 10 % of reward trials. On average the recalibrated factor was 17 % smaller than the original estimate, but this did not appear to fundamentally alter our conclusions (**Extended Data Fig. 8e,f**).

2.4 Supra-reward DA signals modulate reward-evoked striatal spiking activity

Since the modulatory effects of DA are thought to be state-dependent [2], we also examined how amplifying DA reward signals alters ventral striatal spiking activity. Experiments involving DA neuron activation included a block of randomly interleaved reward and laser-paired reward trials.

Pairing optogenetic stimulation with reward amplified the dLight fluorescence signal (**Fig. 6a,b**), but the consummatory licking response was not significantly altered (**Fig. 6c**). Consistent with our analysis of spontaneous activity, changes in reward-evoked neural activity were only clearly distinguished with high DA amplification (**Fig. 6d,e**). We next characterized how reward-evoked activity is altered under conditions when R and R+L trials could be reliably decoded (greater than 80 % decoder accuracy). In this supra-reward regime, neurons displayed a mixture of excitatory and inhibitory changes in firing (**Fig. 6f,g**). All three putative cell types exhibited changes in activity (**Extended Data Fig. 9**). MSNs showed mixed excitatory and inhibitory effects and the weakest overall changes, while the main influence on FSIs appeared to be inhibitory. A subset of TANs showed a significantly longer reward pause duration, in qualitative agreement with previous work attributing the pause to DA modulation [91]. Effects on neuron firing rates, particularly for MSNs, appeared to be strongest within a region roughly corresponding to the nucleus accumbens core (**Extended Data Fig. 10**).

In order to examine DA's modulatory effects during non-rewarding stimuli, in a subset of animals we included a block of trials consisting of randomly interleaved neutral auditory tone and laser-paired tone stimuli. As before, neural activity could only reliably distinguish these conditions when DA signals were amplified above tone-matched levels (**Extended Data Fig. 11**). Together, the DA opto-activation experiments suggest that only supra-reward DA signals are capable of producing consistent changes in either spontaneous, reward-evoked, or auditory stimulus-evoked spiking in the ventral striatum.

2.5 Morphine alters stimulus-evoked DA responses

Drugs of abuse, such as opioids, are thought to increase DA release, which in turn contributes to the development of addiction [54]. Certainly, this likely involves an overall increase in basal DA levels and rapid supra-reward DA release events, but it has been unclear precisely what effects opioids may have particularly on rapid DA dynamics. Therefore, we conducted several pilot

experiments involving subcutaneous systemic administration of morphine while recording stimulus-evoked DA release events using fiber photometry with dLight (**Fig. 7a**). Low dose morphine (4 mg/kg) in general had very little effect on reward-evoked DA signals, but it attenuated or even reversed decreases in DA evoked by neutral auditory tones or air puffs (**Fig. 7b,c**). Low dose morphine had no statistically significant effect on reward-evoked consummatory licking although in some instances appeared to slightly attenuate it (**Extended Data Fig. 12a,c**). When examining the eye blink response caused by air puffs, low dose morphine appeared to slightly increase the speed and consistency with which mice opened their eyes after an air puff (**Extended Data Fig. 12c**). High dose morphine (8-16 mg/kg) tended to cause much more marked increases in stimulus-evoked DA release, also causing near extinction of consummatory licking (**Extended Data Fig. 12b,13**). We also performed experiments where we administered daily injections of morphine over the course of 7 days and examined whether the effects changed over time (**Extended Data Fig. 14**). In general, the effects of morphine on the DA response to air puff intensified over the course of the 7 days with a statistically significant interaction between timepoint and drug effect (**Extended Data Fig. 14b**). In a final set of experiments, we administered low dose morphine (2 mg/kg) while an animal completed a Pavlovian conditioning task (**Extended Data Fig. 15**). This dose of morphine had minimal effects on licking behavior and slightly decreased the cue-evoked DA response while increasing the reward-evoked response on rewarded trials (**Extended Data Fig. 15c**). Interestingly, on unrewarded trials, there was a small delay in the licking behavior with a coinciding delay and widening in the DA dip associated with this type of trial (**Extended Data Fig. 15c**).

CHAPTER 3: Discussion and conclusions

3.1 DA's role in rapidly shaping striatal activity – interpretation and implications

This study sought to clarify a longstanding question about the magnitude of the dependence of striatal spiking activity on DA neuron input on subsecond behavioral timescales. This issue is at the heart of popular models that propose DA plays a major role not only in shaping future behavior (i.e., via learning-dependent plasticity mechanisms), but in imminent or ongoing actions [5, 18, 20, 25, 26, 28, 31, 33, 34, 92, 93]. Our findings challenge the idea of a strong influence of DA on moment-to-moment striatal activity, unless potentially non-physiological conditions are met. The optogenetic inhibition experiments show that striatal neural dynamics are largely uncoupled from rapid increases in DA signaling that accompany unexpected reward events and more generally, during reward anticipation [88]. These findings were confirmed using multiple complementary experiments and analyses. By contrast, the optogenetic activation experiments reveal that under sufficiently elevated, supra-reward DA signaling conditions, striatal dynamics are indeed strongly influenced by DA neuron input. To estimate the transition point from small to large electrophysiological effects, we refer to the population decoding analysis, which provided a sensitive method for detecting changes in neural activity. Interpolating these results, the transition to the high decoding accuracy (80 %) regime occurs when DA signals are amplified around three to four times above reward-matched levels (**Fig. 5g, 6e**).

There are three major implications of this study. First, it appears that on short timescales striatal dynamics are primarily driven by non-DA neuron inputs, such as VTA GABA neurons and glutamatergic projections from cortex and thalamus. While the suppression of DA signaling did produce statistically significant changes in striatal dynamics, this effect's magnitude is greatly surpassed when perturbing other inputs to the striatum. For example, we have previously shown that suppressing cortical and thalamic input can effectively dampen striatal reward activity by over 50 % [87]. Of course, a striatal network whose dynamics on fast timescales are largely uncoupled

from DA may appear counterintuitive, given that DA neurons encode the performance of imminent or ongoing actions [20, 34], and the assumption has been that this information drives behavior by modulating striatal activity. However, the influence of this information on downstream neural dynamics has not been rigorously demonstrated until now, and our work suggests that strong DA signals do not translate on a subsecond basis to strong changes in striatal firing. One alternative explanation is that the primary role of such DA signals may be to reinforce or weaken future actions by modulating synaptic plasticity [38, 40].

Second, these findings provide a possible electrophysiological mechanism for the results of some behavioral studies, which showed that performance is markedly altered only by stimulating DA above physiologically calibrated levels [28, 37, 40]. Our results indicate that large, supra-reward DA release evokes a distinctly strong response in striatal neurons compared with the response produced under reward-matched conditions. These distinct electrophysiological effects could underlie the different magnitude of behavioral effects reported in the literature. In terms of the neural dynamics which were observed, multiple synaptic and microcircuit-level mechanisms could potentially explain how DA release modulates striatal spiking activity on subsecond timescales [2]. Disentangling these contributions is the subject of other work and was not the purpose of this study, but we can speculate on potential mechanisms. Previous work has shown that stimulating DA terminals can rapidly enhance the firing of D1 receptor (D1R) expressing MSNs via a D1R-dependent mechanism, although this effect was easily saturated with minimal DA release as discussed earlier [36]. DA also modulates lateral inhibition between MSNs, which may alter spiking activity [94]. In addition to DA itself, there may be a role for DA neuron co-release of glutamate and GABA [95, 96], as both of these neurotransmitters are well suited to drive rapid changes in striatal firing. There is evidence that blocking glutamate co-release attenuates DA stimulation-evoked spiking in the ventral striatum [76]. However, since these mechanistic studies did not appear to discriminate between reward-matched and supra-reward (or physiological and potentially supra-physiological) DA signals and the degree to which this

affects co-release, it remains unclear to what extent these mechanisms contributed to our observed results.

Third, our study points to an urgent need to carefully distinguish between the modulatory function that DA neurons normally have in shaping striatal dynamics, and the role they can serve under potentially supra-physiological conditions. One possible approach to circumvent this issue is to measure and calibrate DA stimulation levels on a per animal basis with respect to a known reference signal [28, 31, 37, 40, 41, 97]. However, we suggest that to avoid potential calibration errors, it is essential to include experiments that assess in vivo DA function through inhibitory manipulations. It is further important to note that this outcome was achieved using a relatively low laser power (5 mW), showing that this is a broadly relevant issue for researchers who utilize optogenetic excitation of DA neurons in their work. The adoption of standardized practices surrounding optogenetic manipulation of DA is critical to ensure the future reproducibility and consistency of scientific work in this field.

3.2 Validity of chosen analysis methods

Before continuing the discussion of our results, it is worth examining in depth the rationale and validity of the main chosen analysis methods, in particular selectivity index and SVM. Although we did examine firing rates and they did not reveal anything that would contradict our main results (see **Extended Data Fig. 3, 4f, 7, 9e, 10**), selectivity index and support vector machine (SVM) decoding accuracy may nevertheless be seen as somewhat opaque metrics that could obfuscate DA's striatal effects. Therefore, it seems necessary to describe the general mathematical rationale behind these metrics to ensure that all readers, including those who may not be familiar with these methods, are able to interpret the results. Further specific technical details can be found in the methods section.

As an aside, I want to discuss a problem I feel exists with the way that computational methods are sometimes written about in scientific literature. From what I have seen, work in the

DA field that utilizes machine learning or other modeling methods has tended to describe its models only very generally in the main text with a more technical or formula-based mathematical description in the methods section (see examples [29, 40]). This is not unreasonable; it would be unrealistic to burden authors with providing granular descriptions especially when modeling is not the focus of the paper and when experts would not require such explanations. However, it ought to be recognized that currently there is poor general scientific literacy in advanced computational and mathematical methods when compared to more "traditional" neuroscience topics such as second messenger cascades or the role of ions in establishing membrane potential [98-100]. A simple consequence of this fact is that mathematical or highly technical explanations are frankly opaque to a large proportion of neuroscientists who would benefit from but are unable to meaningfully engage with the work. This high barrier to entry prevents (and discourages) many neuroscientists from rigorously evaluating computational results in the same way that they can evaluate core experimental logic, and I believe that this impedes progress. One example of this is the disagreement between Howard et al. and Soares et al. over whether DA affects timing or action selection [27]. In my opinion, the dialogue between the two would have benefitted from plain language descriptions of their modelling approaches and candid arguments about the limitations of each. Without this, many scientists are left unable to adjudicate between the two positions, relegating contrasting computational approaches to the position of differences in opinion dressed in mathematical clothing. I am certainly not the first to have felt this way about the need to provide plain language commentary alongside mathematical argument, even if not formally included within a paper itself. Scientists in other branches of biology and neuroscience have made efforts to overcome these problems through disseminating useful and broadly understandable explanations of computational methods to their colleagues [101, 102]. In this spirit, Wolfram Schultz, famously associated with the original paper that demonstrated reward prediction error, followed his landmark 1997 paper with a review where he made the choice to discuss reward learning ideas using simplified equations and plain language [6, 103]. One can only speculate that he

appreciated that the impact of the idea of reward prediction error would depend on how widely it could be truly understood. Understanding this, computational neuroscientists have mirrored this approach to make their complex modelling approaches as accessible as possible to readers outside of that subfield (see example [104]). Considering all of this, I will now attempt to describe our analysis methods in a manner that I hope will make our results more readily understandable, albeit impractical in the setting of scientific publication.

In broad terms, the selectivity index measurement can be thought of as the degree of separability of an individual neuron's spike count in a given time bin for two trial types of interest. For instance, let us imagine a hypothetical scenario where we are comparing a neuron's integer spike counts between R and R+L trials for a specific time bin. Let us further imagine that the spike count in the time bin we are examining is always 5 spikes for R trials and always 6 spikes for R+L trials. Because the R+L trials have a higher spike count than the R trials on 100% of trials, the selectivity index is the maximum of 1. It would make no difference if one of the R+L trials had a spike count of 20 or some other permutation; as long as all spike counts are higher than those observed for R trials, then the selectivity index would still be equal to 1. If the situation were inverted such that R trials had a spike count of 6 and R+L trials had a spike count of 5, then the selectivity index would be equal to -1 (the minimum). If both trial types had perfectly overlapping distributions of spike counts, then the selectivity index would be 0. It should be further appreciated that if the spike counts were all multiplied by a factor of 10, this would not impact the selectivity index in any way. Of course, real spiking data involves a distribution of spike counts for each trial type with some degree of overlap in the distributions. These would be intermediate cases where the value would be between -1 and 1 depending on the extent of overlap.

In this way, selectivity index most directly reflects the degree of overlap between two spike count distributions with an indication of whether one distribution is higher or lower than the other. Strictly speaking, selectivity index does not reflect the magnitude of modulation per se. For this reason, selectivity index is a highly sensitive ground truth metric for determining whether two

distributions of spike counts are different and is most appropriate for scenarios where effects are small and variable, which was the case here. Furthermore, because selectivity index is purely based on spike count distribution overlap, it is a broadly applicable metric that can be reasonably applied across different neuron types with differing baseline firing rates or characteristics because each neuron is ultimately contextualized against itself. Put another way, while a +3 Hz modulation may be substantial or inconsequential depending on the cell type that is being examined, selectivity index and the distribution overlap it reflects is a far more universally interpretable concept.

That said, the selectivity index used here does have three key shortcomings that should be kept in mind. First, this index is best suited for a scenario where we expect that a neuron's spike counts are shifted up or down, and it is not suited to a scenario where the effects on a single neuron's spike rates are mixed. For instance, imagine a scenario where spike counts are always 3 for one trial type but either 2 or 4 for the other trial type. This would not be an ideal scenario for the use of this selectivity index. To be clear, because selectivity index is calculated separately for each neuron, it would not be an issue if one neuron was upregulated while another neuron was downregulated. A second potential weakness revolves around the potential for very strong effects to saturate a selectivity index. The usefulness of this selectivity index declines if the effects are so strong that there is very little to no overlap in the distributions. Therefore, it is not as well suited to characterize very strong effects should they exist as it would not offer meaningful information in that scenario beyond confirming that modulation had indeed occurred. A third shortcoming of selectivity index is that it only considers the integer spike counts but not the distribution of those spikes within the time bin. Hypothetically, a trial where 3 spikes occurred quickly at the beginning of the time bin and then 1 spike occurred near the end would be considered equivalent to a trial where 1 spike occurs quicky followed by 3 late spikes. We might consider these two trials meaningfully different, but this information is lost when considering only integer spike counts. This is most acutely problematic if the time bins used are particularly long.

While these issues indicate that selectivity index is not always an ideal approach, they are not major sources of concern for the determination of the extent of rapid dopaminergic effects on striatal spiking. First, to our knowledge there are no prior claims or known mechanisms by which individual neurons would see mixed dopaminergic effects. Even if effects are expected to be different amongst different cell types such as D1 MSNs and D2 MSNs, we did not come across any claims or viable mechanisms whereby individual neurons would be upregulated and downregulated by DA [35]. Second, as can be seen in our results, the effects of DA neuron inhibition and calibrated excitation are clearly subtle and do not appear to saturate the selectivity index (**Fig. 2, 5**). Furthermore, when considered in the context of our VTA GABA positive control experiments, it becomes clear that there was adequate sensitivity and dynamic range to capture strong modulatory effects if they occurred (**Fig. 4**). Third, the issue of time bin size was considered and optimized during the early stages of analysis and the final bin size of 50 milliseconds is short enough that variability of spike timing within the bin should not be a major concern. It also bears mentioning that if optogenetic manipulations had caused a sustained ramp up or down in striatal firing, then it would have become clear over the course of several consecutive time bins with selectivity index progressively increasing. Time bin size would only have influenced detection of the earliest signs of spiking activity changes in the first couple bins, but not later detection. Based on the above reasoning, we found selectivity index to be an ideal tool for this scientific application without any major reservations regarding its suitability.

Turning our attention to the application of SVM to this data set, our specific implementation of SVM-based decoders has similar overall aims to selectivity index but with some important differentiating features. In plain language, our SVM decoder is an algorithm trained on neural spiking data from two compared trial types (ex. R vs. R+L) that attempts to guess the trial type based on the neural spiking data for a group of trials that it was not trained on. It should be noted that each neuron's data is labelled with the neuron's identity and the trial type during the training phase, and the same neurons are considered in the guessing phase with the neuron identity

labels but not the trial type labels. Therefore, the algorithm “understands” and individually differentiates between neurons. Decoding accuracy refers to the accuracy of these guesses, which has a base 50% probability of being correct due to random chance. Rather than a spectrum of overlap expressed by the selectivity index, decoding is a binary classification process. Another important distinction from selectivity index is that the spiking data used for the decoder is a smoothed spike rate rather than a pure integer spike count. This means that the data considered for decoding at a certain time bin in theory contains information about the spiking of that neuron in the adjacent bins, but this does of course mean than differences between adjacent bins are decreased (or “smoothed out”) as well. Like selectivity index, the goal here is to ask “how different” spiking activity is, but with the added step of attempting to meaningfully use the differences to separate data into its two trial types. We will now discuss some further general principles of SVM and then return to details of our implementation.

SVM is neither experimental, untested, obscure, nor new to biological or neuroscience applications [101, 102]. It is also not a model of the brain in any sense. It is simply a powerful tool that attempts to split two collections of numbers into piles (in our case trial types), and it does not require context to do this. It had previously been used decades ago to nearly-perfectly differentiate between similar cancer types when given expression levels of thousands of genes, a task and a dataset we can all agree is comparably complex to our use case here [102]. So how did SVM accomplish that? Because we used a linear kernel SVM, the first assumption that the SVM makes is that the data is linearly separable. Similar to the selectivity index, this essentially would mean that DA effects are assumed to be unidirectional on a per neuron basis, which we discussed was a reasonable starting assumption based on prior literature. We did test a radial kernel that does not rely on this assumption, which we will discuss later. At any rate, the SVM attempts to draw a boundary (a hyperplane) between spiking data of the two trial types to maximize its ability to discriminate between the two trial types. It is important to note that the SVM need not consider every neuron in drawing this hyperplane. In other words, if a small number of neurons made the

whole dataset easily linearly separable, then the SVM would “ignore” the other neurons. Neurons that are redundant or simply aren’t different between the two trial types are also “ignored”. SVM will naturally optimize for the solution that requires the fewest neurons and yields the highest accuracy on the training data, though its willingness to accept more complex solutions for incremental gains in accuracy can be adjusted by users [101]. After the hyperplane is set, the SVM is then tested with the trial data to determine its accuracy at guessing the trial type.

As with many machine learning methods, there is a real risk of a phenomenon called overfitting. This occurs when an algorithm essentially memorizes the source data set and therefore has poor generalization performance when tested on data it was not trained on. Of course, we do practically demonstrate using our positive control experiments that this is not a concern, and we credit this to a couple of key decisions. First, linear kernels generally have the lowest chance of overfitting compared to other kernels because more complex kernels facilitate more complex boundaries that are more prone to that problem [102]. Second, we perform multiple random shuffles of the trials designated for training and testing and sample from many random drawings of neurons. This ensures that the decoding results obtained are a true representation of the overall decoder performance on the data and minimizes the influence of outlier runs.

To verify our assumption that the data is linearly separable, we implemented a SVM that used a radial basis function kernel with a systematic optimization of how flexibly the SVM would draw the boundary and how complex a solution it was willing to accept. We tested this on the data used in Figure 3 with drawings of 400 neurons. This yielded a nearly identical result (0.84% more decoding accuracy for radial kernel), which was not statistically significant when compared to the linear kernel’s performance (unpaired t-test, $t = 1.315$, $P = 0.42$). This is a definitive result that supports the use of the linear kernel and would indicate that there is not meaningful information in the data that we simply didn’t extract.

So how should our SVM results be interpreted? First, it is notable that the decoding performance appears to asymptote with drawings of 300 neurons with very minimal gains from

exceeding that (**Fig. 3**). This strongly indicates that the SVM is not receiving any further useful information, and this is not necessarily surprising even though the mouse striatum contains millions of neurons. The reality is that there are only so many meaningful permutations of spike counts that are possible, despite spiking activity being smoothed which adds some granularity to the data. While the brain may leverage its many neurons to perform a great number of operations, this number clearly does not add to the predictive power of the SVM. Second, the failure of the SVM to achieve robust decoding accuracy as a result of DA neuron inhibition is itself a surprising outcome (**Fig. 3**). SVM is an exquisitely sensitive tool capable of handling bioinformatic data sets far larger than even the ones used here. To reiterate a point made earlier, this analysis is not a proposal of a model of DA function nor is it an argument that the data is better explained by one theory over another. Rather, these SVM findings should themselves be considered a kind of experimental result. Lastly, SVM performance cannot be considered in a vacuum; it must always be contextualized. The greatest risk for misinterpretation with our SVM results is that someone may observe that 75% decoding accuracy was achieved in the eNpHR3 expressing animals and conclude that the decoding was strong because 75% would be favorable odds at a casino (**Fig. 3**). In that case, it would be important to compare that result with the opsin-free controls and the VGAT positive control experiments. When properly contextualized, it is clear that the increase in decoding compared to opsin-free controls is small and weak compared to the decoding seen in positive control experiments. Again, it bears mentioning that the expected performance of an SVM for meaningfully separable data is near 100% [101, 102].

3.3 Generalizability of findings

An outstanding question is whether the electrophysiological findings generalize across different behaviors and striatal subregions. We primarily examined rapid DA modulation of two highly distinct modes of striatal activity – spontaneous and reward-evoked firing. In addition to examining unexpected rewards, we carried out experiments in mice trained on a Pavlovian conditioning task,

allowing us to assess the contribution of DA neurons during periods of reward anticipation (**Extended Data Fig. 5**). Since similar trends were observed under all these conditions as well as during presentation of neutral auditory stimuli, our findings do indeed appear to generalize to a large extent (**Extended Data Fig. 11**). Of course, we cannot rule out the presence of stronger modulatory effects during other types of tasks with more complex behavioral requirements – as has been suggested previously [31]. However, the existence of mechanisms by which a different reward-guided task could cause a dramatically higher dopaminergic effect size remains hypothetical. There is also considerable interest in the role of DA in the initiation and performance of movements in the absence of explicit rewards [20, 34, 93]. But, again, there is no direct evidence that these DA motor signals are fundamentally different in their actions on striatal function than reward signals. In fact, evidence suggests that these intrinsically generated motor-related DA signals in the absence of reward are functionally similar to natural rewards and play a reinforcement role [40].

Our experiments were mostly restricted to measuring ventral striatal dynamics during somatic VTA DA manipulation. While our Pavlovian conditioning experiment suggests generalization to the dorsal striatum, this was not stringently tested across all our experiments. Turning to the literature, one study suggested that DA has a stronger effect on synaptic transmission in the ventral compared to the dorsal striatum consistent with our decoding accuracy comparison [105]. There is also evidence that DA glutamate co-release mechanisms may be biased toward the nucleus accumbens [95]. Additionally, DA drives heterogeneous synaptic responses on cholinergic interneurons in different striatal subregions [91]. It is therefore conceivable that DA neurons may differentially modulate dorsal and ventral striatal dynamics on short timescales. But there is no known mechanism that would suggest these effects are enhanced in the dorsal striatum. A related issue is the specificity of our DA neuron manipulations to the striatum. In order to maximize the likelihood of altering neural dynamics, we intentionally targeted all types of VTA DA neurons instead of only the population projecting to the ventral striatum. Many of the cited

behavioral studies also did not exclusively target striatal projecting DA neurons. This raises the potential pitfall of off-target network effects of opto-stimulation [77]. For example, under certain stimulation conditions VTA DA neurons can rapidly alter the dynamics of prefrontal cortical neurons [106, 107], which in turn project to the striatum. However, even if such indirect effects are present, then the direct contribution of DA neurons to striatal dynamics would likely be even smaller than our current approach suggests, thus further strengthening the main conclusions of this work.

3.4 Validity and limitations of the DA calibration method

To date, most experiments incorporating optogenetically stimulated DA calibration have done so using food rewards as the reference point and photometry as the means of measuring DA release [28, 37, 40, 41]. To maintain consistency with prior work in the field, our calibration method relied on photometry to compare optogenetically evoked DA levels to those evoked by unexpected rewards. On a practical level, the data clearly show a markedly higher electrophysiological effect size with elevated DA signaling. This is a key result that ought to be considered in the design of future experiments involving DA manipulations regardless of the terminology used to refer to them. Several terms have been previously used in the literature to describe the high amplitude DA regime including supra-reward, supra-physiological, and uncalibrated [28, 37, 41]. Many studies, including those explicitly proposing a role for DA in rapidly performing actions, use food rewards as reinforcers, which suggests that calibration of DA manipulations to food rewards is physiologically relevant for a wide range of behavioral tasks. Despite our adherence to previously accepted calibration practices, one potential limitation is that the significant spatial variability of DA signals is not captured by photometry [35]. Reassuringly, the calibration factor was relatively consistent across a substantial extent (i.e., 1 mm depth) of ventral striatum (**Extended Data Fig. 8**). This appears to suggest that, despite the variability of DA signals, the photometrically determined calibration method is reasonably sound. We also sought additional characterization

of DA signaling through the use of miniscopes, a technique with greater spatial resolution (**Extended Data Fig. 2**). The miniscope experiment revealed that while the amplitude of DA reward responses varied across the field of view at the single pixel level, optogenetic inhibition of DA reward signals was broadly effective. This further suggests that the variability of DA signaling at cellular resolution was not a major confounding factor in this study.

Another potential limitation of this work is that, even though unexpected rewards evoke some of the most robust DA responses [34], DA signals may exceed the proposed supra-physiological threshold under certain behavioral conditions, such as for highly threatening or painful stimuli [108]. Furthermore, experiments involving altered DA signaling such as those concerned with drugs of abuse or addiction may also regularly exceed the natural reward-evoked threshold. Indeed, in our pilot experiments, morphine at high doses caused DA release events that far exceeded typical reward-evoked amplitudes even for aversive stimuli, which normally cause decreases in DA. That said, given that food is a primary survival need and is widely used to incentivize task performance, it would seem dubious to suggest that the DA release evoked by food rewards in a hungry mouse is somehow not a reasonably representative physiological comparison point. Nevertheless, despite caveats regarding the spatial resolution of photometry and the use of natural rewards as a reference point, the standardization of our DA calibration method with several other studies helps place these results into context within the broader literature. Further research and scientific discussion will likely be needed to arrive at a consensus on best calibration practices and the exact terminology demarcating the transition to the supra-physiological regime.

3.5 Reconciling findings with previous evidence

A major question is whether this work can be reconciled with a sizable body of evidence that appears to contradict our findings either directly via electrophysiological measurements [36, 76], or indirectly via behavioral measurements [20, 25, 31]. On one hand, a subset of those experiments may

have inadvertently over-stimulated DA neuron cell bodies or terminals, causing supra-physiological neurotransmitter release in the striatum. There is further support for this possibility from a study showing that animal motion effects are only evident when SNc DA levels are raised five times above reward-matched levels [37]. This five-fold behavioral threshold is remarkably similar to our electrophysiologically determined threshold estimate of three to four times reward-matched levels, suggesting the same underlying modulatory mechanisms. While it is tempting to speculate on the prevalence of engaging potentially supra-physiological mechanisms within the literature, it is noted that since optogenetically evoked DA signals are often not measured let alone calibrated, this possibility can neither be firmly proved nor disproved. On the other hand, a number of studies used either DA neuron inhibition or a pre-calibrated level of activation, thereby mitigating the risk to engage supra-physiological DA release mechanisms. In one such experiment, optogenetically manipulating VTA DA neurons did alter the probability of initiating reward-motivated approach behavior [31]. Two separate studies showed that SNc DA manipulations biased the probability of selecting a particular action over another [27, 29]. Elsewhere, optogenetic SNc DA manipulations modulated the probability of initiating an action at a given moment in time [28]. Moreover, an earlier study from our group showed that inhibiting SNc DA neurons reduced the probability of initiating reward-anticipatory movements [38]. These statistically significant behavioral changes may appear to contradict the findings in this work, but a crucial observation is that the average change in movement or action probability seen in those studies is generally small often inconsistent, especially in the case of DA inhibition. By comparison, our neural decoding analysis suggests a downstream area can theoretically distinguish reward-evoked striatal population dynamics with or without DA signals in only 13 % more trials relative to controls (**Fig. 3b**). Thus, when considering inhibitory manipulations, the size of previously reported behavioral effects, and the upper bound on the size of electrophysiological effects shown here, are comparable. The most parsimonious, unified interpretation of these observations is that under physiological

conditions DA neurons modulate striatal activity, and consequently behavior, in a weak probabilistic manner, i.e., on a small minority of trials.

Still, there is an alternative interpretation of the data that should also be considered. It is true that we did find a statistically significant effect of physiological DA dynamics on neuronal activity in the ventral striatum, albeit small and inconsistent. Therefore, it could be argued that the data support the possibility that physiological DA can shape activity in the striatum and in turn could presumably regulate behavior. That said, the observed effect sizes and decoding results must also be considered as well. Around 58 % of cells were reward modulated in our experiments with around 4 % of those cells having their reward response altered by DA inhibition (compared to controls). Within this small pool of neurons, the alteration in the reward response was quite subtle and failed to achieve robust decoding accuracy. The small magnitude of this effect was particularly evident when compared to the essentially perfect decoding accuracy achieved by manipulating another input to the striatum, VTA GABA neurons (**Fig. 4d**). Cortical and thalamic input is also known to greatly influence striatal spiking [87, 109].

To further expand on the characterization of the rapid effects of DA neuron inhibition as weak, we performed a reanalysis of data from our lab's prior work (**Extended Data Fig. 16,17**) [83, 87]. In these experiments, an opto-probe was used to perform striatal electrophysiology with local optogenetic inhibition of either VGAT-expressing striatal cells or bilateral projections from area M2 to the striatum. This inhibition took place in the context of a Pavlovian conditioning task with inhibition starting 2 seconds prior to cue onset (**Extended Data Fig. 16a,e; 17a,e**). As previously reported, directly inhibiting VGAT-expressing striatal cells led to rapid, robust, and persistent decreases in MSN spiking activity (see Fig 4 in ref [83]). Nevertheless, we wanted to take this opportunity to make a more direct comparison between the data in that paper and the current manuscript, by applying the same measures of neural activity used here. We reanalyzed the data during two crucial epochs of the task, with consistently strong effects in both epochs. The laser for opto-inhibition was started prior to cue presentation, allowing us to examine the effects on

spontaneous striatal activity as well as during task engagement (including anticipatory lick performance epochs). We first analyzed the effect of optogenetically inhibiting striatal GABA neurons on spontaneous striatal activity, i.e., before the reward-predicting cue and reward were delivered (**Extended Data Fig. 16a-d**). Next, we analyzed the effect of inhibiting striatal GABA neurons during the performance and reward periods of a Pavlovian conditioning task (**Extended Data Fig. 17a-d**). It is apparent in both cases that the optogenetic manipulations produce strong changes in striatal spiking activity; for example, Extended Data Fig. 13d reveals ~100% decoding ability with just 50 cells. It is perhaps unsurprising that inhibiting the same cells we are recording (i.e., striatal GABA neurons) produces huge electrophysiological effects, so we wanted to expand our "positive control" analysis to another scenario – inhibition of cortical input. Again, we reanalyzed a previously collected and published dataset from our lab which targeted a bilateral projection from area M2 to the striatum [87]. Again, our reanalysis of the data reveals that optogenetically inhibiting this cortical projection to the striatum produced marked alterations in both spontaneous (**Extended Data Fig. 16e-h**) and performance/reward epoch striatal spiking (**Extended Data Fig. 17e-h**).

Critically, apart from differences in the cell type or pathway being inhibited, the experimental approaches were similar across our previous papers from which these data originated [83, 87], and the current manuscript. This enabled a fairly reliable "apples-to-apples" comparison of the effect of inhibiting different inputs on striatal spiking (**Extended Data Fig. 16i,j**). Intriguingly, these plots demonstrate that inhibiting either striatal GABA, VTA GABA, or corticostriatal projections from area M2 substantially outweigh the effect of inhibiting DA neurons. In fact, when viewed in this context, the DA inhibition data are not just more subtle – they are virtually indistinguishable from negative controls. Even when examining analogous comparisons from the performance epoch of the Pavlovian task, the effects of DA inhibition are similarly overshadowed by the effects of inhibiting striatal GABA or M2 corticostriatal projections (**Extended Data Fig. 17i,j**). These results all further bolster a key implication of this study, that

physiological DA signals play a relatively minor role in rapidly shaping neural activity in the striatum.

Thus, the comparably small effect size of DA inhibition is incompatible with the idea that DA is a major modulator of neural dynamics in the striatum on subsecond timescales. The most immediate implication is that proponents of physiological DA's role in rapid behavioral performance will either need to determine the mechanisms by which downstream regions might read out subtle striatal effects into strong, meaningful outputs – or, to contextualize the importance of DA in relation to other inputs.

CHAPTER 4: Methods

4.1 Animals. DAT-Cre mice (Strain # 006660, Jackson Laboratories) and VGAT-Cre mice (Strain # 028862, Jackson Laboratories), 8-18 weeks old at the time of the first surgery, were used for experiments involving optogenetic manipulation of dopaminergic neurons or GABAergic neurons, respectively. Both male and female mice were used, and the mice were maintained as heterozygotes on a C57BL/6J background. Animals were kept on a 12 hour light/dark cycle, and were group housed until the first surgery. All procedures were approved by the University of California, Los Angeles Chancellor's Animal Research Committee.

4.2 Surgical procedures. Surgery was performed under isoflurane anesthesia and aseptic conditions using a stereotaxic apparatus (Model 1900, Kopf Instruments). Surgery began with anesthetizing the mice in an induction chamber. This was then followed by removal of hair over the scalp using a combination of commercial electric razor and small scissors. Anesthesia was maintained during this procedure using a nose cone kept separate from the stereotaxic apparatus in order to avoid contamination. At this stage the animal was then mounted to the stereotaxic nose cone and then further secured using stainless steel ear bars. The ear bars used were small in order allow for maximum clearance for later placement of head bars that will be secured to the skull. There was a risk if the ear bars were over-tightened to cause skull fracture, intracranial bleeding, and/or high intracranial pressure, which are all detrimental to survival and health of the animal. Care was taken to ensure that the animal's tongue did not obstruct its airway and was not accidentally placed between the nose cone bite bar and animal's bottom teeth. This was accomplished by carefully pulling the tongue to the side using a pair of clean forceps. Airway obstruction was further avoided by taking care to prevent over-tightening of the nose cone by sliding the nose cone too far forward, potentially putting pressure on the airway. After fixing the ear bars, skull fixation was tested by gently pressing on the animal's skull or very lightly tugging

on the front nose cone assembly. For smaller mice, it was sometimes necessary to gently hold up the skull by tugging on the ears. At this stage, the skull surface was then cleaned using alternating swabs of 75% ethanol and betadine solution with sterilized cotton tipped applicators. A few drops of bupivacaine solution were then applied to the skull surface using a sterile needle to act as a local anesthetic. Following this, a 5 mg/kg dose of carprofen diluted in saline was administered subcutaneously to the mouse for further pain relief and to provide additional fluids to the mouse.

The initial incision was then made using a pair of sterile scissors while also using sterile tweezers to pinch up a portion of the scalp. This procedure should not be done with a traditional caudal to rostral scalp incision, as this leaves too much skin behind and will interfere with later head bar attachment. After the incision was made, care was taken to remove all debris using a cotton applicator moistened with sterile saline. To further clear the surgical fields and expose the skull, a sterile cotton tipped applicator was used to carefully push out and widen an oval shaped window exposing the skull. The exposed area extended from the bifurcation of the frontal fork to behind the caudal lambdoid sutures. The skull near the skin perimeter was then superficially scored in a crosshatch pattern using a sterile scalpel, avoiding suture lines or other areas prone to bleeding. The crosshatch scoring of the skull enhanced the adhesion of dental cement, head bars, and other implants.

Next, the skull was aligned along the three principal axes using roll, yaw, and pitch adjustments. First, bregma and lambda were located using a surgical microscope. Then, yaw adjustments were made such that bregma and lambda were aligned along the AP axis of stereotaxic apparatus. Following this, pitch was adjusted such that bregma and lambda were at the same height along the DV axis of the stereotaxic apparatus. Finally, roll was adjusted by ensuring that two points equidistant along the ML axis of the stereotaxic apparatus are at equal heights along the DV axis of the stereotaxic apparatus. To remove the periosteum, remaining subcutaneous tissue, and further sterilize the area, a diluted hydrogen peroxide solution was

applied using a sterile cotton tipped applicator to the outer rim near the interface of the skull and skin. Excess hydrogen peroxide was then removed using a dry cotton tipped applicator, followed by a rinse of the area with sterile saline and the subsequent removal of that saline using another applicator. These steps were completed in quick succession as leaving hydrogen peroxide on the skull surface for too long or failing to remove it may interfere with later steps. A small superficial hole was then drilled at bregma using a stereotaxic drill and the hole was then further marked with a fine tipped permanent marker. The marking ensured that bregma would remain clearly visible during future procedures.

The remaining steps in the procedure consisted of three principal components: injection of adeno-associated virus (AAV) into the ventral tegmental area (VTA) and ventral striatum, placement of a fiber-optic implant in the VTA, and attachment of stainless-steel head fixation bars to the skull. In order to allow for injection of AAV into the VTA and ventral striatum, two ipsilateral holes were drilled into the skull (VTA coordinates relative to bregma: 3.30 mm posterior, 0.50 mm lateral; ventral striatum coordinates relative to bregma: 1.30 mm anterior, 1.25 mm lateral). The holes were subsequently slightly widened using a handheld dental drill handpiece, and the exposed brain area irrigated with sterile artificial cerebrospinal fluid. Sterile fine forceps were used to remove any remaining meningeal layers above the sites or local debris. In situations where bleeding occurred at drilling sites, hemostatic collagen foam was placed and wetted with artificial cerebrospinal fluid.

For injections, a pulled glass pipette was loaded with mineral oil and placed into a Nanoject 3 programmable nanoliter injector (Drummond). Mineral oil was expelled from the pipette and replaced by the uptake of undiluted AAV stock, using a permanent marker to mark the AAV fluid level on the pipette. AAV for optogenetics (UNC Vector Core) was unilaterally injected into the VTA, consisting of 500 nl of Cre-dependent AAV expressing either AAV5/EF1a-DIO-eNpHR3-mCherry, AAV5/Syn-Flex-ChrimsonR-tdTomato, or AAV5/EF1a-DIO-mCherry (coordinates relative to bregma: 3.30 mm posterior, 0.50 mm lateral, 4.30 mm ventral). For experiments

involving the substania nigra pars compacta different coordinates were used (coordinates relative to bregma: 3.16 mm posterior, 1.50 mm lateral, 4.30 mm ventral). AAV expressing the fluorescent DA sensor dLight1.2 under the synapsin promoter (pAAV-hSyn-dLight1.2) was then also unilaterally injected into the striatum (ventral striatum coordinates relative to bregma: 1.30 mm anterior, 1.25 mm lateral, 4.25 mm ventral; dorsal striatum coordinates relative to bregma: 0.9 mm anterior, 1.5 mm lateral, 2.8 mm ventral). pAAV-hSyn-dLight1.2 was a gift from Lin Tian (Addgene plasmid # 111068) [44]. Generally, pipettes were lowered to about 0.05 mm lower than the intended coordinates and then raised back to the intended coordinates prior to beginning injection. Before slowly removing the pulled glass pipette from the brain, 10 minutes were allowed to elapse so that the virus could diffuse from the injection site and to prevent backflow.

After viral injections, stainless steel head bars were secured to the skull surface. In order to accomplish this, autoclaved stainless steel head bars were submerged in 75% ethanol for at least 15 minutes. They were then secured to a customized head bar stereotaxic adapter using 4 small hex screws. The entire assembly was then further sprayed with 75% ethanol before being allowed to dry completely before proceeding. Using the stereotaxic apparatus, the head bars were slowly positioned such that they were centered relative to the bregma-lamba axis. Care was taken to ensure that the head bars were able to make contact with the skull surface. This was important to ensure because the head bars would not be stable if they rested entirely on skin, which is also why classic longitudinal incisions were inappropriate for this procedure. After the position that satisfied these requirements was found, the head bar assembly was then removed prior to the adhesion process. At this point, metabond quick adhesive cement (Parkell) was prepared in a ceramic mixing dish using a 1:3:1 ratio corresponding to scoops of radiopaque L-powder, drops of “B” Quick base, and drops of “C” catalyst. The ceramic mixing dish was occasionally kept on ice in order to extend the curing time of the metabond. A sterile wooden applicator was used to form a thin ring around the border between the skin and exposed skull, with a small amount of metabond also applied to head bars where they contact the skull. The head bar assembly was

then resecured into its original position with the head bars making contact with the skull. Metabond was then slowly applied around the outer perimeter until the head bars were secured to the skull such that all exposed subcutaneous tissue was also covered in a layer of metabond. The final result was a shallow ring of metabond surrounding exposed skull with medial portions of the head bar embedded in this outer ring. Metabond was allowed to fully dry and harden before the head bar screws were removed from the stereotaxic adapter and the adapter fully decoupled from the head bars. Because head bar application can result in a change in skull alignment, skull alignment and security were verified and adjusted prior to proceeding to the next step.

Following viral injection, a ferrule-coupled optical fiber implant (0.2 mm diameter, 0.22 NA, Thor Labs) was placed with the tip approximately 0.2 mm above the viral injection site. For experiments involving optogenetic manipulation of DA neurons, this cannula was targeted to the caudal site, whereas experiments involving gradient-index (GRIN) lens implantation also targeted the rostral site. The implant was secured to a custom cannula holder and then mounted to the stereotaxic apparatus. Prior to implantation, the implant was soaked in ethanol for as long as was practical during the procedure to ensure as much sterility as could be managed. This ferrule was constructed by cutting length of optical fiber such that ~0.8 mm of exposed fiber remained after the cannula was placed at one end of the fiber. The fiber was manually polished on both ends using diamond lapping sheets and secured in the cannula using a small amount of epoxy. For the miniscope imaging experiment, a 0.5 mm diameter GRIN lens (Inscopix, 8.4 mm length) was additionally implanted into the striatum (coordinates relative to bregma: 1.30 mm anterior, 1.25 mm lateral, 4.5 mm ventral). The implants were all lowered slowly to prevent unnecessary intracranial pressure and brain tissue damage. After the implants were lowered to the desired depth the remaining exposed brain was sealed using a small amount of low viscosity silicone kwik-cast sealant (World Precision Instruments). Care was taken to limit the amount of sealant used as this could have affected the stability of the implant. At this stage the implant was secured using either a dental acrylic base plate material (Fastray, Keystone Industries) or a UV-cured

dental cement (RelyX Unicem 2, 3M). If a UV-cured dental cement was used, care was taken to protect the animal's eyes from exposure to UV light. After the compound surround the implant hardened, the custom cannula holder was carefully removed with extra care taken to avoid excess pressure on the implant. Animals that would later undergo electrophysiology had an area of skull over the cerebellum still exposed for later placement of a ground wire and an area of skull around the rostral injection site also exposed. These exposed brain areas were covered with removable kwik-cast sealant and covered further with a thin layer of metabond to allow for later access.

After surgery, all animals were individually housed and were given daily carprofen injections (5mg/kg, subcutaneous) for the first three post-operative days. Over the first post-operative week, animals were administered ibuprofen and amoxicillin dissolved in their drinking water. The mice were given a recovery period of at least two weeks prior to behavioral training.

4.3 Behavioral task. Mice were food restricted but given ad libitum access to water to maintain around 90 % of their baseline weight. After the first 24 hours of food restriction, animal habituation and introduction to rewards began. This was accomplished by filling a 1 ml syringe with a small blunt tipped stainless steel feeding tube and extruding a single drop of reward that would hang on the end of the tube (10% sweetened condensed milk). This syringe was positioned to rest in the water port of the animal's cage and the animal was allowed to interact with the feeding tube. Taking care not to startle the animal, drops were gradually extruded one at a time to slowly encourage the mouse to consume the reward. This was gradually repeated over the course of 2-3 days until the animal was willing to extend its snout into the water port and extend its tongue to reach the feeding tube. To habituate the animal to handling, mice were introduced to a clear plexiglass handling tube (Braintree Scientific) and allowed to freely interact with it. Using non-aversive animal handling techniques, the mouse was gently encouraged to enter the tube and the tube was lifted to pick up the mouse. The mouse could then be gently removed from the tube onto

a gloved hand by tilting the tube backwards relative to the snout and lifting the tube. Over the course of animal training, this was done daily to reduce animal stress resulting from handling.

Following initial habituation, animals were habituated to head fixation and trained to reliably consume uncued rewards (6 µl, 10 % sweetened condensed milk) delivered via an audible solenoid valve. Each day, mice were head fixed by securing their head bars to a stainless steel head bar holder which was then held in place within the behavioral apparatus. Rewards were delivered via a tube within an infrared lick detection port located approximately 5 mm from the mouth. Mice had to extend their tongue out of the mouth multiple times to consume the reward and these tongue extensions were detected as licks. Mean lick rates were calculated based on the average number of consummatory licks from 0-1 s following reward delivery. During behavioral training, mice were given rewards (20-30 s inter-trial interval, ITI, 100 trials per day). Head fixed licking training took place over the course of 5 total days. A subset of animals was also habituated to an audible tone on the last two days of training (12 kHz tone, 0.5 s duration, 20-30 s ITI, 100 trials per day). Another subset of animals was further trained for 5 days on a Pavlovian conditioning task using the same overall apparatus with an olfactory cue presented through an olfactometer (10 % dilution isoamyl acetate in mineral oil, 1.5 L/min total air flow). Each trial consisted of a 1 second odor cue followed by a 1 s delay until reward was finally delivered (20-30 s ITI, 100 trials per day). For this task, anticipatory licking rate was calculated during the delay period before reward delivery. After the training period, mice then proceeded to the experiment involving electrophysiology, photometry, and optogenetics.

For the morphine pilot experiments, similar training and experimental parameters were used for rewards (6 µl, 10 % sweetened condensed milk), auditory cues (12 kHz tone, 0.5 s duration), and air puffs (0.5 second duration). The experiment was broken into two halves with 3 blocks (reward, auditory cue, airpuff, 30 trials each, 20-30 s ITI) in each half. Between each set of 3 blocks, the animal was administered a subcutaneous injection of saline or morphine. For Pavlovian experiments, each trial consisted of a 1 second odor cue followed by a 1 s delay until

reward (20-30 s ITI, 90% rewarded, 100 trials before and 100 trials after injection). Between each block of 100, the animal received an injection of saline or morphine. All saline injections occurred in morphine naïve animals one day prior to receiving morphine.

4.4 Optogenetics. Optogenetic manipulations were organized into two principal sections, with the first block consisting of reward trials and laser-paired reward trials (R and R+L), and the second block consisting of unpaired laser stimulation (589 nm, 5 mW power, MGL-F-589-100 mW, CNI Laser). For eNpHR3-mediated opto-inhibition, laser stimulation was applied continuously for 0.5 s. The first block consisted of 40 R and 40 R+L trials randomly interleaved (20-30 s ITI). The second block consisted of 80-100 trials where laser was delivered alone (20-30 s ITI); in order to maintain a consistent number of trials across different analyses, only the last 40 trials were analyzed. For experiments involving Chrimson-mediated opto-activation, laser stimulation was pulsed at either 4, 20, or 40 Hz over a 0.5 s period (10 ms pulse width). The first block was similar to the inhibition experiments consisting of 38-40 each of R and R+L trials (laser was pulsed at 40 Hz). The second block consisted of 40 trials each of 4, 20, and 40 Hz unpaired laser stimulation presented in random order. One animal did not undergo 4 Hz laser stimulation. A subset of the animals used for the activation experiments also completed an additional block where an auditory tone was presented with half the trials randomly paired with optogenetic stimulation (40 T and 40 T+L trials). For the subset of animals who completed the Pavlovian task, half the trials were randomly paired with 2 s of opto-inhibition from cue onset to reward delivery (40 C and 40 C+L trials).

4.5 Electrophysiology. To measure reward and optogenetically evoked rapid changes in DA signaling alongside neural spiking responses in the ventral striatum, an opto-probe containing a 256 electrode silicon microprobe (type 256AS, developed by our lab [110]) attached to a low autofluorescence photometry fiber positioned about 0.15 mm above the nearest electrode

recording sites (0.43 mm diameter, 0.48 NA, Doric Lenses). The silicon microprobe contained four shanks separated by 0.4 mm. Each shank contained 64 electrodes spanning a total length of 1.05 mm. This device allowed recordings from a wide area of the nucleus accumbens in the ventral striatum. Prior to electrophysiological recordings, a second surgery was performed under isoflurane anesthesia in which a rectangular craniotomy over the striatum was opened and then covered with a silicone sealant (Kwik-Cast, World Precision Instruments). During this surgery, the distance between bregma and a reference point on the head bar was recorded to aid in later targeting of the opto-probe. The animal was allowed to recover for at least 3 hours before inserting the opto-probe into the nucleus accumbens region of the ventral striatum (ventral tip of probe shanks: 1.30 mm anterior, 0.65-1.85 mm lateral, 5.1 mm ventral). For recordings in the dorsal striatum, the coordinates were shifted (ventral tip of probe shanks: 0.9 mm anterior, 0.9-2.1 mm lateral, 3.8 mm ventral). The shanks of the opto-probe were coated with a fluorescent dye prior to insertion (DiI or DiD, Thermo Fisher Scientific). The exposed brain was irrigated with artificial cerebrospinal fluid (Harvard Apparatus) before being further covered in a layer of mineral oil to prevent evaporative loss of fluid. Electrophysiological recordings were conducted at a sampling rate of 25 kHz per electrode using a commercial multichannel data acquisition (DAQ) system (C3316 and C3004, Intan Technologies). Spike sorting was carried out with Kilosort [111], using bandpass filtered data (3 pole Butterworth filter with a passband of 0.6-7 kHz). After identifying single-units, raw data were refiltered with a wider passband (0.3-7 kHz) and then classified as putative MSNs, FSIs, and TANs. The classification procedure was based on spike trough-to-peak duration and baseline firing rate [112]. Putative FSIs were defined by a narrow spike waveform (0.45 ms maximum duration). MSNs and TANs were both defined by wider waveforms (0.475-2 ms duration). TANs were separated from MSNs by the regularity of their baseline firing (maximum TAN coefficient of variation: 1). The minimum baseline firing was defined as 0.01 Hz for MSNs and FSIs and 2.5 Hz for TANs, and the maximum was defined as 20 Hz for MSNs and TANs. TAN pause responses were calculated for TANs that had at least 4 adjacent timebins below -0.5

Z scores compared to baseline from 0-1 s after trial onset. Unless specified, all analysis of neural activity used all cell types including unclassified units, in order to avoid possible bias arising from cell-type specific modulatory effects as well as classification errors. For assessing changes in neural activity (see the Selectivity Index and Population decoding sections below) firing rates were calculated by binning spikes in 50 ms increments using a three bin sliding window average. For spatial analysis of firing rate changes, ML position was determined based on the shank and DV position was based on binned positions (4.02-4.245 mm, 4.27-4.52 mm, 4.545-4.795 mm, 4.82-5.07 mm). For calculating average firing rates changes by neuron type for experiments comparing two different trial types (R vs. R+L, etc.), the mean firing rate for each neuron from 0-0.5s for the non-laser paired trials was subtracted from the mean firing rate for each neuron from 0-0.5s for the laser paired trials. For calculating average firing rates by neuron type for experiments comparing that compared a trial to baseline activity, the mean firing rate for each cell for 1 second of baseline was subtracted from the mean firing rate for each neuron from 0-0.5s for the trial type of interest.

4.6 Fiber photometry. Photometry of dLight fluorescence was conducted synchronously with electrophysiological recordings using the opto-probe described above. The photometry fiber was coupled via a fiber patch cord to a four-port connectorized fluorescence mini cube (FMC4_AE(405)_E(460-490)_F(500-550)_S, Doric Lenses), with one excitation port (460-490 nm) used for dLight1.2 fluorescence, and a 500-550 nm detection port. Fluorescence excitation light was delivered via a 465 nm light emitting diode oscillated at 211 Hz. The emitted signal was then sent to a low noise femtowatt photoreceiver (Model 2151, Newport) connected to a lock-in amplifier (SR810, Stanford Research Systems). The demodulated signal was sampled at 25 kHz by a DAQ (Intan Technologies), and then downsampled to 1 kHz for storage. Offline analysis involved downsampling the signal again to 20 Hz. The fractional change in fluorescence ($\Delta F/F$) was calculated with respect to the average baseline signal 1 s prior to the onset of a trial. The

average fractional change in fluorescence was determined by calculating the average value of ΔF/F for a given trial condition from 0-1 s post-stimulus onset. For the Pavlovian task, the average value of ΔF/F was calculated from 0-2 s post-cue onset. The factor increase in DA was calculated as the ratio of the maximum laser-evoked ΔF/F to the maximum reward-evoked ΔF/F signal. The DA reward amplification factor was calculated as the ratio of the maximum ΔF/F on R+L trials to the maximum ΔF/F on R trials. These factors were both calculated on a per-animal basis and this was used as a calibration method.

4.7 Photometry depth test. To determine whether the estimated factor increase in DA was consistent across different depths spanned by the electrode array, a group of DAT-Cre mice were injected with AAV5/Syn-Flex-ChrimsonR-tdTomato in the VTA and pAAV-hSyn-dLight1.2 in the ventral striatum in a procedure identical to that described in the Surgical procedures section. The mouse also underwent identical behavioral training and craniotomy procedures. However, for recording, instead of using an opto-probe, a photometry fiber was lowered to a site corresponding to the top of the area targeted by the electrodes (1.30 mm anterior, 1.25 mm lateral, 4.1 mm ventral). It was then lowered in 0.2 mm increments to target five additional depths (maximum depth: 5.1 mm). At each depth, the mouse was given a randomly interleaved combination of 10 reward trials and 10 trials of 40 Hz pulsed optogenetic stimulation. Between each depth measurement, a 10 minute waiting period ensued for tissue surrounding the optical fiber to stabilize. Before each recording commenced, mice were given 2-3 rewards to ensure that licking behavior was maintained.

4.8 Miniscope recordings. For the animals with an implanted GRIN lens in the ventral striatum, a miniscope (OpenEphys, v4.4) was lowered above the GRIN lens using a micromanipulator. The task completed by these mice was otherwise identical to the reward-paired opto-inhibition experiments. The miniscope was powered and imaging data was acquired by using a Miniscope

DAQ (OpenEphys, v3.3). Imaging data acquisition was synchronized with behavioral data via a 5V trigger connection to the Miniscope DAQ, with the excitation light on from -2 to 3 s relative to reward delivery. Following imaging data acquisition, the data was processed using the Minian [113] pipeline to perform glow removal and denoising (median filter, ksize = 3) without background removal or motion correction. The data was then exported to Matlab for pixel binning (binning was performed by averaging a 5 x 5 array of adjacent camera pixels, corresponding to an effective field size of approximately 8 μm x 8 μm per pixel) and removal of pixels outside of the GRIN lens field of view. Pixel removal was accomplished by setting a threshold for average fluorescence of each pixel and excluding those below a per recording determined percentile cutoff. The per pixel ΔF/F was calculated with respect to the average baseline signal 1 s prior to the onset of a trial. The per pixel peak signal was determined from the average ΔF/F values from 0 to 0.5 s after the trial onset for each trial type (R or R+L). The overall mean signal was produced from the average of each per pixel ΔF/F for a given trial type (R or R+L). For normalized comparisons, the average peak signal for each pixel across both trial types was compared to the overall mean peak signal for R trials.

4.9 Immunohistochemistry. Mice were anesthetized with isoflurane and transcardially perfused with phosbate buffered saline followed by 10 % neutral buffered formalin. Brains were stored in the neutral buffered formalin solution for at least 1 day at 4 ºC before being sectioned at a thickness of 100 μm on a vibratome. Sections were blocked using donkey serum before antibody incubations. VTA slices were incubated with rabbit anti-dsRed polyclonal antibody (632496, Takara) as the primary antibody (1:1000 dilution) overnight at 4 ºC. Striatal slices were incubated with chicken anti-GFP (ab13970, Abcam) as the primary antibody (1:1000 dilution) overnight at 4 ºC. VTA slices were then incubated with Alexa Fluor 647-conjugated donkey antibody to rabbit IgG (711-605-152, Jackson ImmunoResearch) as the secondary antibody (1:200 dilution) for 4 hours. Striatal slices were incubated with Alexa Fluor 488-conjugated donkey antibody to chicken

IgG (703-545-155, Jackson ImmunoResearch) as the secondary antibody (1:200 dilution) for 4 hours. Sections were then mounted using tissue-mounting medium and imaged under a confocal or epifluorescence microscope.

4.10 Selectivity index. Single-neuron selectivity indices were used in two different ways. The first selectivity index was used to determine a neuron's discrimination of R+L versus R trials, C+L versus C trials, or T+L versus T trials. The second selectivity index was used to determine a neuron's discrimination of unpaired reward or laser trials versus a preceding baseline period. In both cases the selectivity index was calculated by subtracting 0.5 from the area under the ROC curve and then multiplying the result by 2. The selectivity index was calculated separately for each time bin from -1 to 3 s post-stimulus. Prior to calculating selectivity index, the spike counts in each time bin were corrected for smoothing and returned to integer spike counts per time bin. For each neuron, an integer list of thresholds is generated based on the maximum response ranging from zero to the maximum number of spikes seen in that neuron for a given time bin. At each threshold, the proportion of trials with spike counts higher than the tested threshold was determined for the two trials types or conditions being compared. Across all tested thresholds, a receiver operator curve was then constructed with the probability of the second condition being higher than the threshold plotted against the probability of the first condition being higher than the threshold. An area under this curve (auROC) greater than 0.5 would imply that a particular neuron tended to spike more in the first condition compared to the second, and an auROC below 0.5 would imply that a neuron spiked less in the first condition compared to the second. By subtracting 0.5 from auROC and multiplying the result by 2, a selectivity index that reflects upregulation and downregulation that ranges from -1 to 1 was created, with negative values indicating downregulation and vice versa. The significance level of neural responses was assessed by comparing the observed difference in mean to the resampled difference in mean using randomly shuffled trial assignments (400 permutations). A neuron was defined as being significantly

modulated by a stimulus, or selective between two stimulus conditions, if three or more consecutive time bins from 0-1 s showed a statistically significant change in firing (observed difference in mean was greater than 95 % of resampled differences in mean). This was assessed from 0-2 s for the Pavlovian task.

4.11 Population decoding. As with the selectivity index described above, population decoding was used in two different ways. The first type of decoder was trained to distinguish population dynamics on R+L versus R trials, C+L versus C trials, or T+L versus T trials. The second type of decoder was trained to distinguish population dynamics on unpaired laser trials versus a preceding baseline period. For analysis of DA neuron opto-inhibition and GABA neuron opto-activation effects, the decoder was trained on data pooled across all animals, with the number of neurons varied from 5 to 400 or 300, respectively. For analysis of DA neuron opto-activation effects, the decoder was trained on 50 simultaneously recorded neurons. The decoder was based on a linear support vector machine (SVM) learning algorithm [114], and utilized 80 % of trials for training and the remaining 20 % for testing the accuracy. For the DA neuron opto-inhibition experiment, we also used a radial basis function nonlinear SVM that included a grid search during training to optimize the C and gamma parameters, with an identical train-test split; this yielded statistically similar results as the linear SVM (not shown). The mean decoder accuracy was calculated from the average performance of 50 random neuron selections, and 100 random train/test trial selections. The decoder was separately trained using randomly shuffled trial assignments. The 95 % confidence interval was calculated from testing decoder performance on these shuffled datasets [115].

4.12 Statistics. Statistical analysis utilized standard functions in Matlab (MathWorks) and Prism (GraphPad Software). Data collection and analysis were not performed with blinding to the experiments. No statistical methods were used to predetermine sample sizes, but our sample

sizes are similar to those used in previous publications. The sample size, type of statistical test, and P values are indicated in figure legends. Data distribution was assumed to be normal, but not specifically tested. T-tests were two sided. One-way ordinary or RM ANOVA was followed by Tukey's post hoc test for multiple comparisons. Two-way RM ANOVA was followed by Sidak's post-hoc test for multiple comparisons. For the morphine pilot experiments examining day x drug effects, a mixed-effects model was used.

Figures:

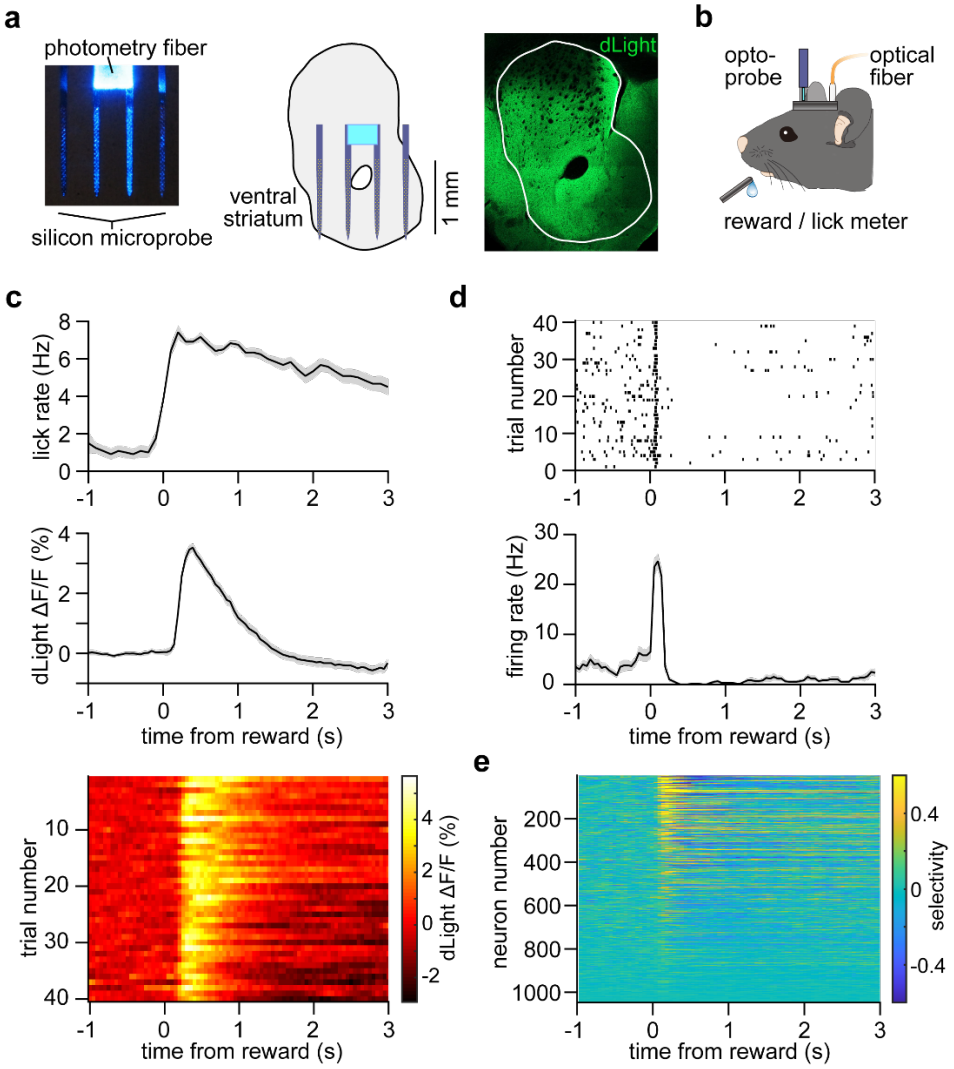


Fig. 1: Simultaneous monitoring of DA and electrophysiological activity in the ventral striatum.

a. Left: Opto-probe device comprised of a 256 electrode array distributed on four shanks, together with an optical fiber for photometry. Middle: Illustration of the opto-microprobe in the targeted area of the ventral striatum. Right: Confocal image of dLight1.2 expression.
b. Mouse recording apparatus during reward delivery and optogenetic stimulation.
c. Reward-evoked licking and dLight fractional fluorescence change data from one animal.
d. Spike raster and mean firing rate of one reward-responsive MSN.
e. Selectivity index of 1047 ventral striatal neurons during reward delivery. Selectivity of neural activity is calculated relative to baseline. All data in this figure represent mean ± SEM.

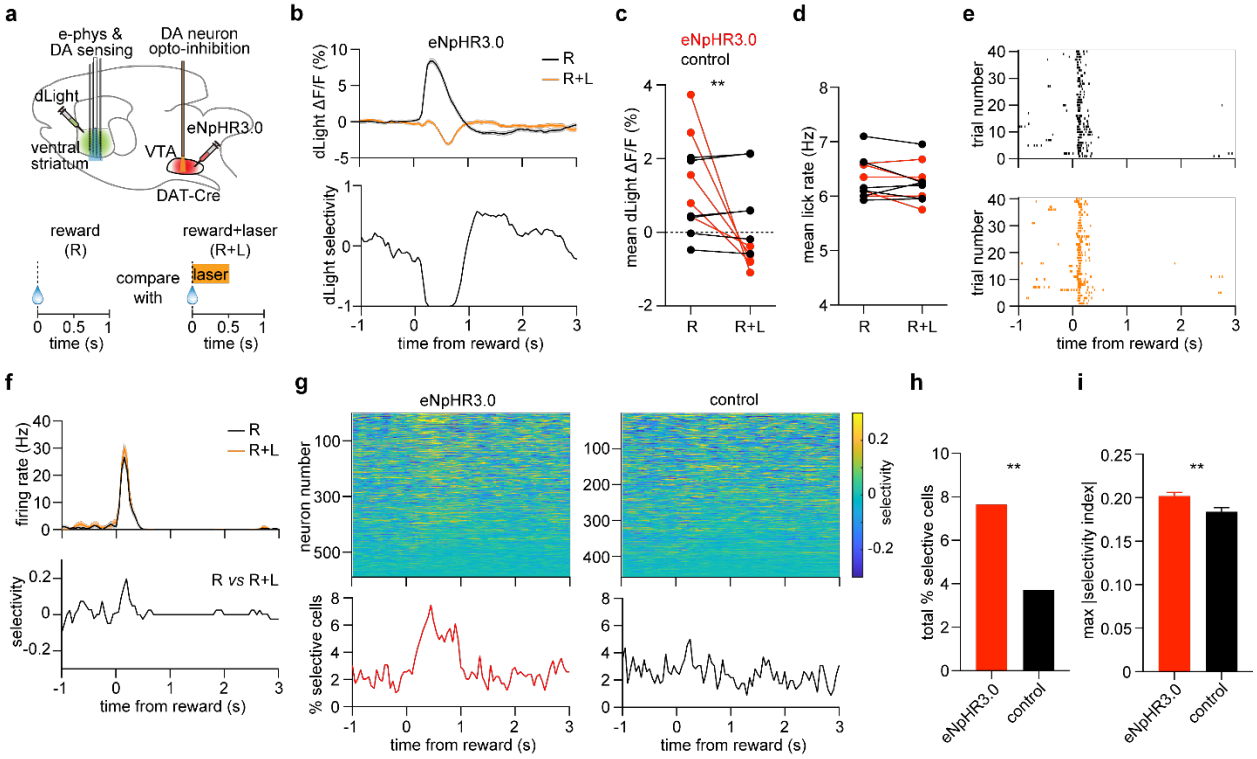


Fig. 2: Small effect of inhibiting VTA DA neurons on reward-evoked striatal spiking activity.

a. Top: Experimental approach. Bottom: task schematic in which reward trials (R) are compared to laser-paired reward trials (R+L).

b. dLight fractional fluorescence change signal from one eNpHR3-expressing mouse (top) and selectivity index plot (bottom). Selectivity is calculated between R and R+L trials.

c. Mean reward-evoked dLight fluorescence on R and R+L trials (n = 5 eNpHR3 and 6 control mice, two-way RM ANOVA, group effect: $F_{1,9} = 0.13$, $P = 0.7$, trial effect: $F_{1,9} = 15$, $P = 0.004$).

d. Mean reward-evoked lick rate on R and R+L trials (n = 5 eNpHR3 and 6 control mice, two-way RM ANOVA, group effect: $F_{1,9} = 0.008$, $P = 0.9$, trial effect: $F_{1,9} = 1.9$, $P = 0.2$).

e. Spike raster of one reward-responsive neuron on R (top) and R+L (bottom) trials. Data are from an eNpHR3-expressing animal.

f. Mean firing rate (top) and selectivity index (bottom) of the same neuron as panel e. Selectivity is calculated between R and R+L trials.

g. Top: Selectivity index of 589 neurons pooled across 5 eNpHR3-expressing mice, and 458 neurons pooled across 6 control mice. Bottom: Percentage of neurons that were significantly selective for R versus R+L trials, as a function of time.

h. Total percentage of neurons that were selective for R versus R+L trials (n = 45 out of 589 cells (7.6 %) in the eNpHR3 group, n = 17 out of 458 cells (3.7 %) in the control group, chi square test for proportions, $\chi^2 = 7.1$, $P = 0.008$).

i. Maximum absolute value of the selectivity index per neuron (n = 589 cells in the eNpHR3 group, n = 458 cells in the control group, unpaired t-test, $t = 2.9$, $^{**}P = 0.004$). All data in this figure represent mean ± SEM.

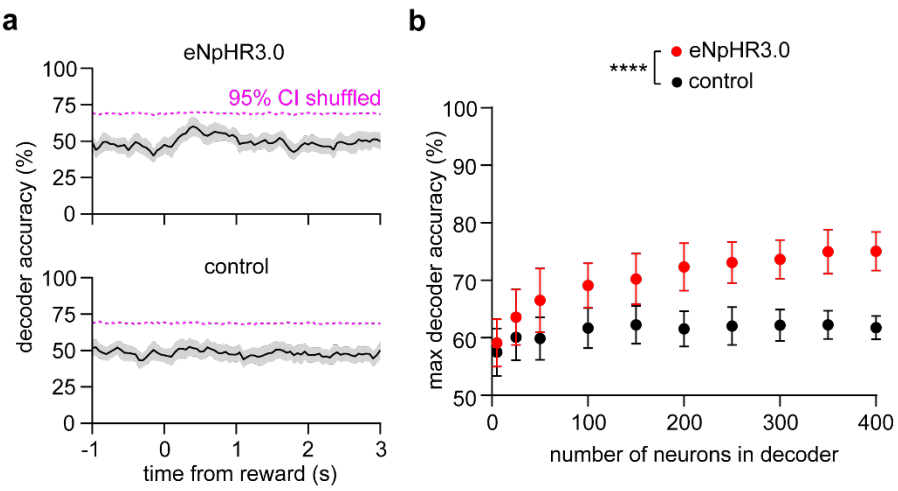


Fig. 3: Population-level decoding discriminates striatal neuron reward responses with and without DA.

a. Mean accuracy of an SVM decoder trained using 50 neurons to discriminate R from R+L trials. Magenta dashed line indicates the 95 % confidence interval of decoder performance trained on trial-shuffled data. Top: neurons selected from the eNpHR3 group. Bottom: neurons selected from the control group. Shaded area represents represent the SD across 50 random drawings of neurons.

b. Maximum decoder accuracy as a function of neuron number (two-way ANOVA, group effect: $F_{1,980}$ = 1322, P < 0.0001, neuron number effect: $F_{9,980}$ = 80, P < 0.0001). Post-hoc Sidak's test: ****P < 0.0001 for all data points with n > 5 neurons. Data represent the mean and SD across 50 random drawings of neurons.

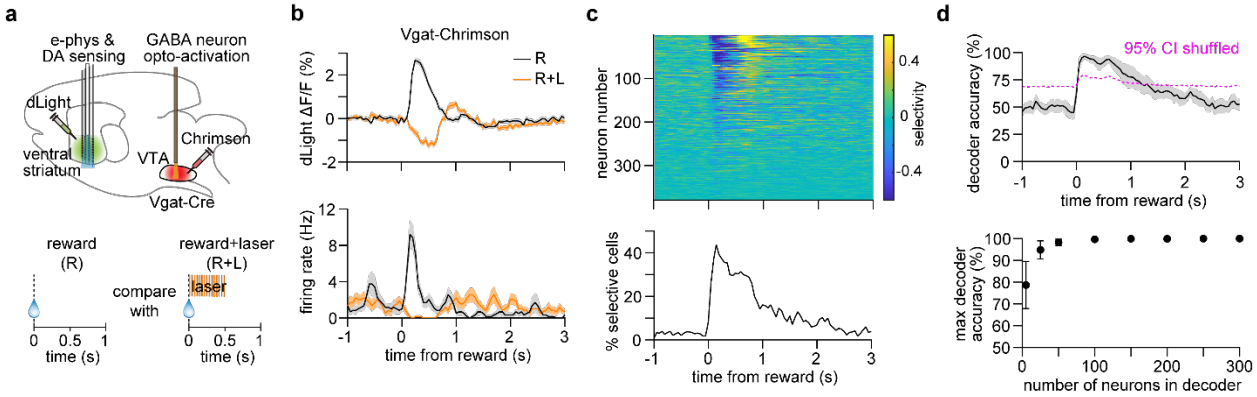


Fig. 4: Strong effect of activating VTA GABAergic neurons on reward-evoked striatal spiking activity.

a. Top: Experimental approach. Bottom: task schematic in which reward trials (R) are compared to laser-paired reward trials (R+L). Laser is pulsed at 40 Hz for 0.5 s total duration.

b. Top: dLight fractional fluorescence change signal on R and R+L trials from one Vgat-Chrimson mouse. Bottom: Mean firing rate of one neuron on R and R+L trials. Data represent mean ± SEM.

c. Top: Selectivity index of 379 neurons pooled across 3 Vgat-Chrimson mice. Bottom: Percentage of neurons that were significantly selective for R versus R+L trials, as a function of time.

d. Top: Mean accuracy of an SVM decoder trained using 50 neurons to discriminate R from R+L trials. Shaded area represents the SD across 50 random drawings of neurons. Magenta dashed line indicates the 95 % confidence interval of decoder performance trained on trial-shuffled data. Bottom: Maximum decoder accuracy as a function of neuron number (one-way ANOVA, neuron number effect: $F_{7,392} = 158$, $P < 0.0001$). Post-hoc Sidak's test: $P < 0.0001$. Data represent the mean and SD across 50 random drawings of neurons.

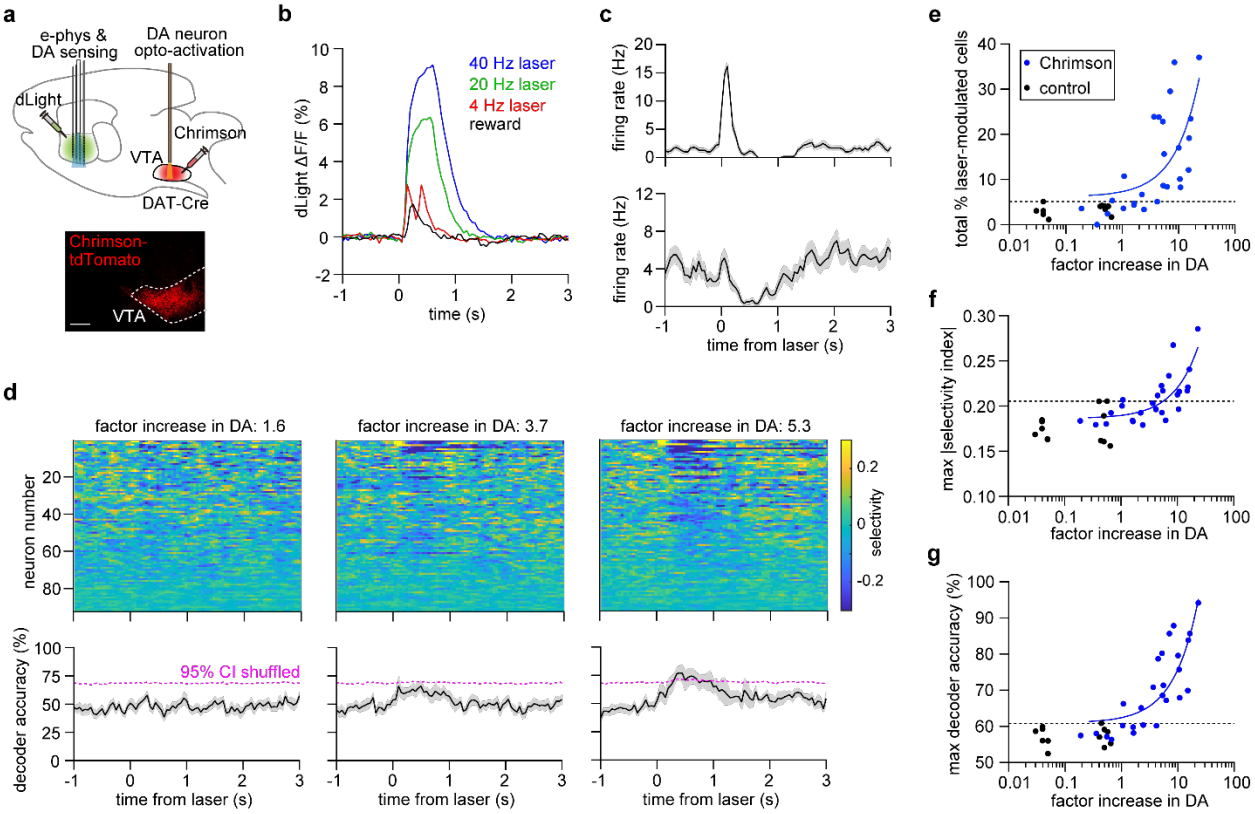


Fig. 5: Small effect of artificial reward-matched DA transients on spontaneous striatal activity.

a. Top: Experimental approach. Bottom: Confocal image of Chrimson-tdTomato expression in the VTA.

b. dLight fractional fluorescence change signal from one Chrimson-expressing mouse. The data represent reward-evoked (black) and laser-evoked signals at 4 Hz (red), 20 Hz (green), and 40 Hz (blue) stimulation for 0.5 s.

c. Mean firing rate of two laser-modulated MSNs.

d. Top: Selectivity index of 92 simultaneously recorded neurons from one Chrimson-expressing mouse under 4 Hz (left), 20 Hz (middle), and 40 Hz (right) laser stimulation. Bottom: Mean accuracy of an SVM decoder trained using 50 neurons to discriminate laser-evoked from baseline activity. Shaded area represents the SD across 50 random drawings of neurons. Magenta dashed line indicates the 95 % confidence interval of decoder performance trained on trial-shuffled data.

e. Percentage of neurons per recording session that were significantly modulated by the laser, as a function of factor increase in DA. n = 26 sessions from 9 Chrimson-expressing mice, and n = 12 sessions from 4 control mice. Chrimson group data are positively correlated with the factor increase in DA (Pearson r = 0.6, P = 0.0005). Dashed line represents the 95 % confidence interval of the control data. Blue line represents the linear fit to the Chrimson data.

f. Mean of the maximum absolute value of the selectivity index per recording session, as a function of factor increase in DA. Selectivity of neural activity is calculated relative to baseline. Chrimson group data are positively correlated with the factor increase in DA (Pearson r = 0.8, P < 0.0001). Dashed line represents the 95 % confidence interval of the control data. Blue line represents the linear fit to the Chrimson data.

g. Maximum decoder accuracy per recording session, as a function of factor increase in DA. The decoder was trained using 50 neurons to discriminate laser-evoked from baseline activity.

Chrimson group data are positively correlated with the factor increase in DA (Pearson $r = 0.8$, $P < 0.0001$). Dashed line represents the 95 % confidence interval of the control data. Blue line represents the linear fit to the Chrimson data.

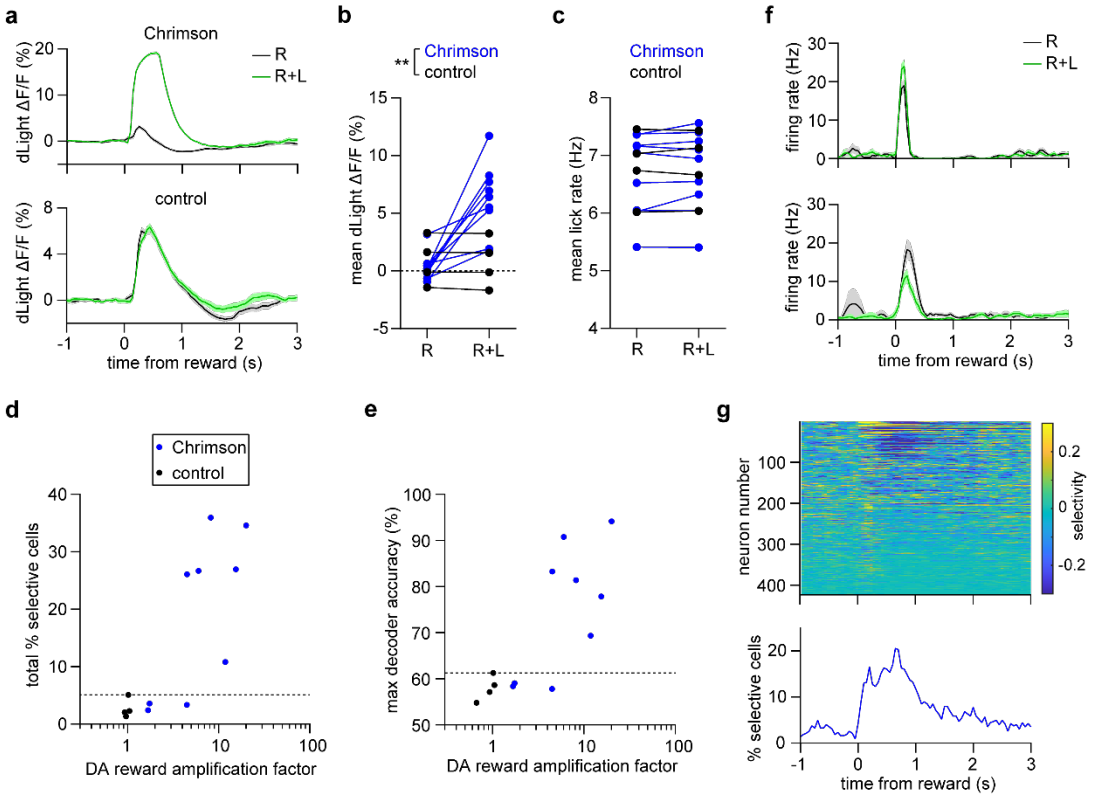


Fig. 6: Supra-reward DA signals modulate reward-evoked striatal spiking activity.

a. dLight fractional fluorescence change signal from one Chrimson-expressing mouse (top) and one opsin-free control mouse (bottom) on reward (R) and laser-paired reward (R+L) trials. Data represent mean ± SEM.

b. Mean reward-evoked dLight fluorescence on R and R+L trials (n = 9 Chrimson and 4 control mice, two-way RM ANOVA, group effect: $F_{1,11} = 5.3$, $P = 0.04$, trial effect: $F_{1,11} = 11$, $P = 0.007$). Post-hoc Sidak's test: $^{**}P = 0.001$ for R+L trials.

c. Mean reward-evoked lick rate on R and R+L trials (n = 9 Chrimson and 4 control mice, two-way RM ANOVA, group effect: $F_{1,11} = 0.1$, $P = 0.8$, trial effect: $F_{1,11} = 0.6$, $P = 0.5$).

d. Percentage of neurons per animal that were selective for R versus R+L trials, as a function of DA reward amplification factor. n = 9 Chrimson-expressing and 4 control mice. Chrimson group data show a trend for being positively correlated with the DA amplification factor (Pearson $r = 0.6$, $P = 0.075$). Dashed line represents the 95 % confidence interval of the control data.

e. Maximum decoder accuracy per animal, as a function of DA reward amplification factor. The decoder was trained using 50 neurons to discriminate R from R+L trials. (Pearson $r = 0.6$, $P = 0.08$). Dashed line represents the 95 % confidence interval of the control data.

f. Mean firing rate of two MSNs showing an excitatory (top) and inhibitory (bottom) response to DA neuron activation. Data are from the supra-reward condition (animals exhibiting over 80 % decoding accuracy in panel e). Data represent mean ± SEM.

g. Top: Selectivity index of 423 neurons pooled across 4 Chrimson-expressing mice. Bottom: Percentage of neurons that were significantly selective for R versus R+L trials, as a function of time. Data are from the supra-reward condition (animals exhibiting over 80 % decoding accuracy in panel e).

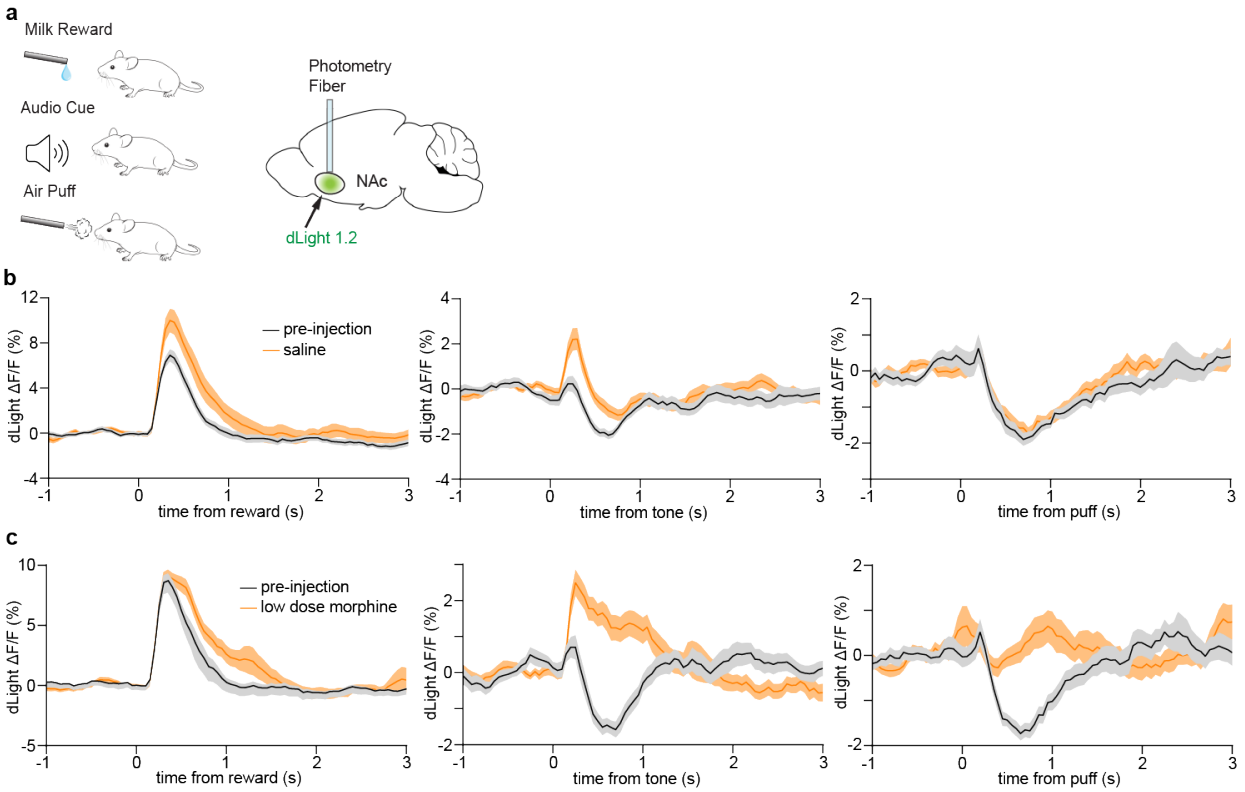
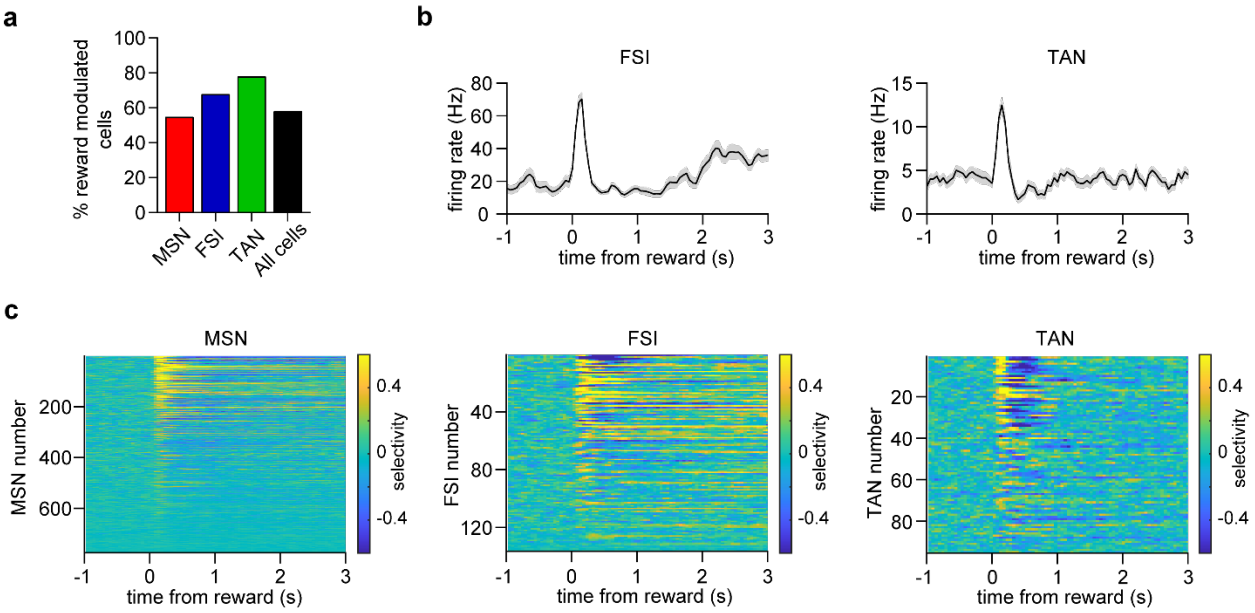


Fig. 7: Effects of subcutaneous low dose morphine administration on DA responses to rewarding, neutral, and aversive stimuli.

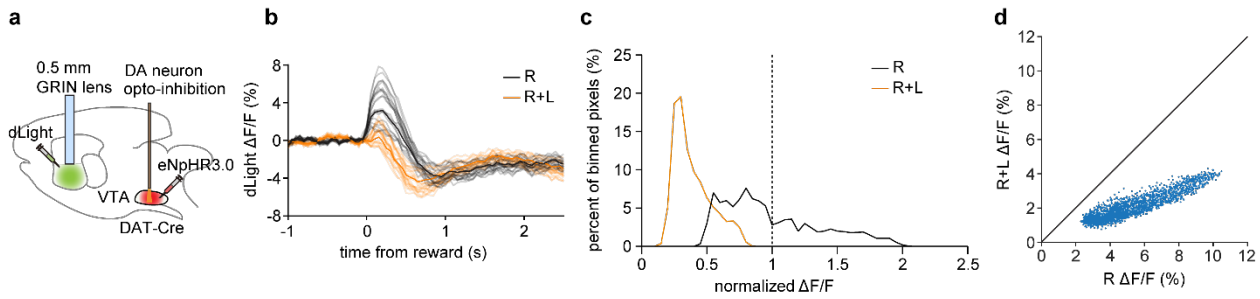
a. Left: Task schematic showing various stimuli given to mice. Right: Experimental approach.
b. dLight fractional fluorescence change signal in response to rewarding (left), neutral (center), and aversive (right) stimuli from one mouse before and after saline injection.
c. dLight fractional fluorescence change signal in response to rewarding (left), neutral (center), and aversive (right) stimuli from one mouse before and after low dose morphine injection.

Extended Data Figures:



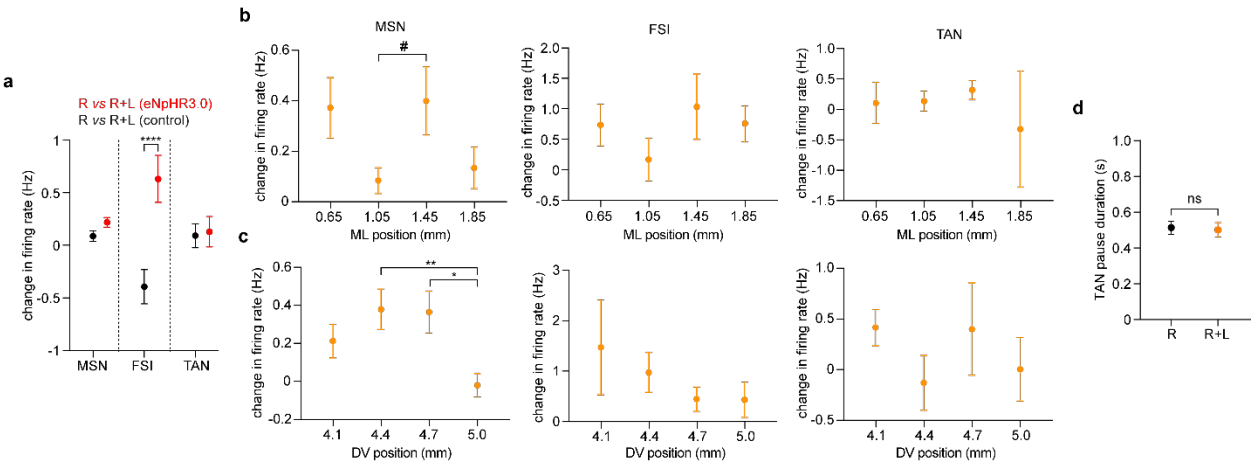
Extended Data Fig. 1: Reward response of electrophysiologically classified cell types in the ventral striatum.

a. Percentage of each cell type that showed a significant reward response.
b. Mean firing rate of one putative FSI (left) and TAN (right) unit during reward delivery.
c. Selectivity index of a population of 774 MSNs, 136 FSIs, and 95 TANs during reward delivery. Selectivity of neural activity is calculated relative to baseline. All data in this figure represent mean ± SEM.



Extended Data Fig. 2: GRIN lens imaging of DA dynamics in the ventral striatum.

a. Experimental approach. Miniscope placed above GRIN lens for imaging (not pictured).
b. dLight fractional fluorescence change for R trials (black) and R+L trials (orange). Translucent traces represent a random sample of 20 individual pixel responses for each trial type; opaque traces represent the mean signal across all pixels.
c. Frequency plot of pixel mean peak fluorescence change for R (black) and R+L (orange) trials relative to mean reward response (black dashed line) (n = 2339 pixels, paired t-test, t = 117, $P < 0.0001$).
d. Comparison plot of pixel mean peak fluorescence in R and R+L trials (n = 2339 pixels). Line of identity in black.



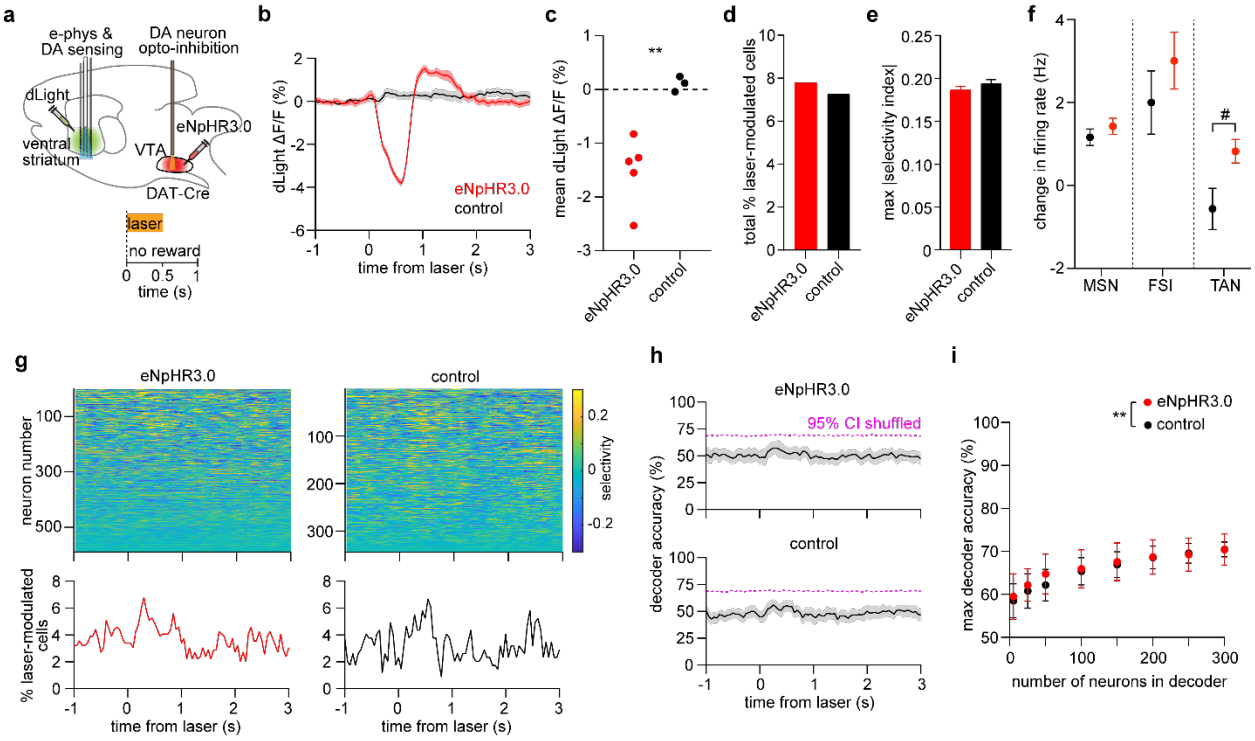
Extended Data Fig. 3: Effect of inhibiting VTA DA neurons on reward-evoked striatal firing rates.

a. Mean firing rate changes for neurons of each cell type (eNpHR3 animals: n = 435 MSN, 73 FSI, 58 TAN; control animals: n = 339 MSN, 63 FSI, 37 TAN; one-way ANOVA, cell type x group effect: $F_{5,999}$ = 6.937, P < 0.0001). Post-hoc Sidak's test comparing eNpHR3 animals and controls: MSN (P = 0.25), FSI (P < 0.0001), TAN (P = 0.99). All data in this figure represent mean ± SEM.

b. Mean firing rate changes for neurons of each cell type group by ML position. One-way ANOVA: MSN ML position effect ($F_{3,431}$ = 3.193, P = 0.02), FSI ML position effect ($F_{3,69}$ = 0.869, P = 0.46), TAN ML position effect ($F_{3,54}$ = 0.474, P = 0.70). Post-hoc Tukey's test for MSN: P = 0.0529 for 1.05 mm vs. 1.45 mm.

c. Mean firing rate changes for neurons of each cell type group by DV position. One-way ANOVA: MSN DV position effect ($F_{3,431}$ = 4.557, P = 0.004), FSI DV position effect ($F_{3,69}$ = 0.904, P = 0.44), TAN DV position effect ($F_{3,54}$ = 1.002, P = 0.40). Post-hoc Tukey's test for MSN: P = 0.005 for 4.4 mm vs. 5.0 mm, P = 0.02 for 4.7 mm vs. 5.0 mm.

d. Comparison of TAN pause duration in eNpHR3 expressing animals across R and R+L trial types (n = 29 TANs, paired t-test, t = 0.3825, P = 0.7).



Extended Data Fig. 4: Small effect of inhibiting VTA DA neurons on spontaneous striatal spiking activity.

a. Top: Experimental approach. Bottom: task schematic in which 0.5 s laser stimulation is delivered in the absence of reward or other stimuli.

b. dLight fractional fluorescence change signal from one eNpHR3-expressing mouse (red) and one opsin-free control mouse (black).

c. Mean laser-evoked dLight fluorescence (n = 5 eNpHR3 and 3 control mice, unpaired t-test, t = 4.2, $^{**}P$ = 0.005).

d. Total percentage of neurons that were significantly modulated by the laser relative to baseline (n = 46 out of 589 cells in the eNpHR3 group, n = 25 out of 344 cells in the control group, chi square test for proportions, χ^2 = 0.09, P = 0.8).

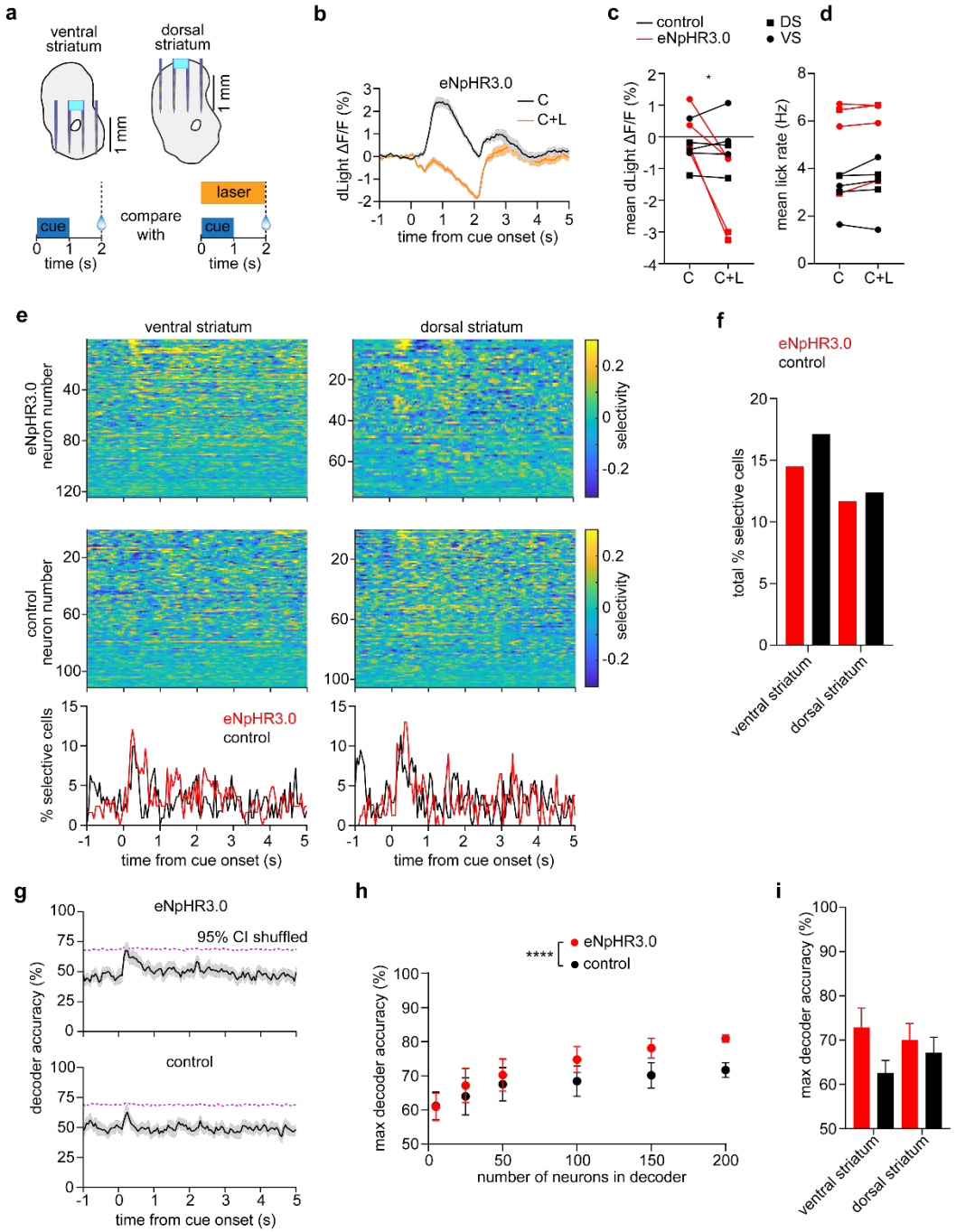
e. Maximum selectivity index per neuron (n = 589 cells in the eNpHR3 group, n = 344 cells in the control group, unpaired t-test, t = 1.1, P = 0.3). Selectivity of neural activity is calculated relative to baseline. Data represent mean ± SEM.

f. Mean firing rate changes for neurons of each cell type (eNpHR3 animals: n = 435 MSN, 73 FSI, 58 TAN; control animals: n = 261 MSN, 44 FSI, 25 TAN; one-way ANOVA, cell type x group effect: $F_{5,890}$ = 4.303, P = 0.0007). Post-hoc Sidak's test comparing eNpHR3 animals and controls: MSN (P = 0.78), FSI (P = 0.45), TAN (P = 0.37). All data in this figure represent mean ± SEM.

g. Top: Selectivity index of 589 neurons pooled across 5 eNpHR3-expressing mice, and 344 neurons pooled across 3 control mice. Bottom: Percentage of neurons that were significantly modulated by the laser relative to baseline, as a function of time.

h. Mean accuracy of an SVM decoder trained using 50 neurons to discriminate laser-evoked from baseline activity. Magenta dashed line indicates the 95 % confidence interval of decoder performance trained on trial-shuffled data. Top: neurons selected from the eNpHR3 group. Bottom: neurons selected from the control group. Shaded area represents the SD across 50 random drawings of neurons.

i. Maximum decoder accuracy as a function of neuron number (two-way ANOVA, group effect: $F_{1,784}$ = 8, P = 0.004, neuron number effect: $F_{7,784}$ = 116, P < 0.0001). Post-hoc Sidak's test: **P = 0.005 for n = 50 neurons; all other comparisons between the eNpHR3 and control group were not significant. Data represent the mean and SD across 50 random drawings of neurons.



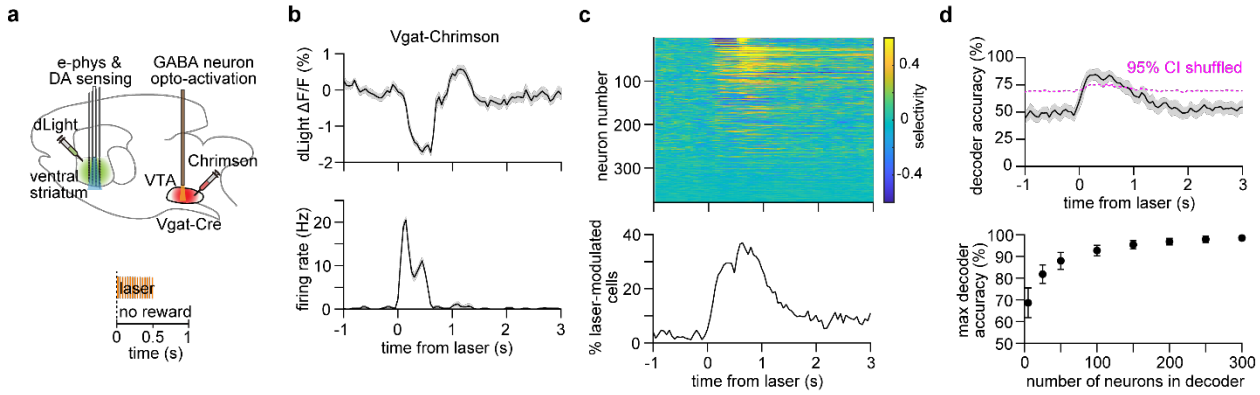
Extended Data Fig. 5: Small effect of inhibiting VTA DA neurons during reward anticipation on striatal spiking activity.

a. Top: Illustration of the opto-microprobe in the targeted area of the ventral striatum and dorsal striatum. The corresponding DA neuron populations are optogenetically targeted (not pictured). Bottom: task schematic in which cued reward delivery trials (C) are compared to laser-paired trials where DA neurons are inhibited for 2 s (C+L).

b. dLight fractional fluorescence change signal from one eNpHR3-expressing mouse.

c. Mean cue-evoked dLight fluorescence on C and C+L trials (n = 4 eNpHR3 and 5 control mice, two-way RM ANOVA, group effect: $F_{1,14}$ = 1.6, P = 0.2, trial effect: $F_{1,14}$ = 5, P = 0.04).

d. Mean lick rate during reward anticipation on C and C+L trials (n = 5 eNpHR3 and 6 control mice, two-way RM ANOVA, group effect: $F_{1,14}$ = 15, P = 0.0016, trial effect: $F_{1,14}$ = 0.1, P = 0.8).
e. Top: Selectivity index of neurons pooled across eNpHR3 expressing mice (n = 124 ventral striatal neurons from 2 mice, n = 77 dorsal striatal neurons from 2 mice). Middle: Selectivity index of neurons pooled across control mice (n = 111 ventral striatal neurons from 3 mice, n = 105 dorsal striatal neurons from 2 mice). Bottom: Percentage of neurons that were significantly modulated by the laser relative to baseline, as a function of time. Red is eNpHR3 expressing mice and black is control mice.
f. Total percentage of neurons that were selective for C vs. C+L trials pooled by recording location (n = 18 out of 124 ventral striatal cells in the eNpHR3 group and n = 19 out of 111 ventral striatal cells in the control group, chi square test for proportions, χ^2 = 0.3, P = 0.6; n = 9 out of 77 dorsal striatal cells in the eNpHR3 group and n = 13 out of 105 dorsal striatal cells in the control group, chi square test for proportions, χ^2 = 0.02, P = 0.9).
g. Mean accuracy of an SVM decoder trained using 50 neurons to discriminate C vs. C+L trials. Magenta dashed line indicates the 95 % confidence interval of decoder performance trained on trial-shuffled data. Top: neurons selected from the eNpHR3 group. Bottom: neurons selected from the control group. Shaded area represents the SD across 50 random drawings of neurons. The analysis in this panel pooled cells across the ventral and dorsal striatal groups.
h. Maximum decoder accuracy as a function of neuron number (two-way ANOVA, group effect: $F_{1,588}$ = 214, P < 0.0001, neuron number effect: $F_{5,588}$ = 194, P < 0.0001). Post-hoc Sidak's test: P = 0.0006 for n = 25 neurons, P = 0.006 for n = 50 neurons, P < 0.0001 for n $\geq$ 100 neurons. Data represent the mean and SD across 50 random drawings of neurons. The analysis in this panel pooled cells across the ventral and dorsal striatal groups.
i. Maximum decoder accuracy using 50 neurons to train the decoder separated by recording location (one-way ANOVA, group x recording location effect: $F_{3,196}$ = 72, P < 0.0001). Post-hoc Sidak's test: P = 0.0008 for ventral striatum decoding compared to dorsal striatum for the eNpHR3 group, P = 0.0007 for dorsal striatum decoding comparing eNpHR3 group to controls, all other comparisons P < 0.0001. Data represent the mean and SD across 50 random drawings of neurons.



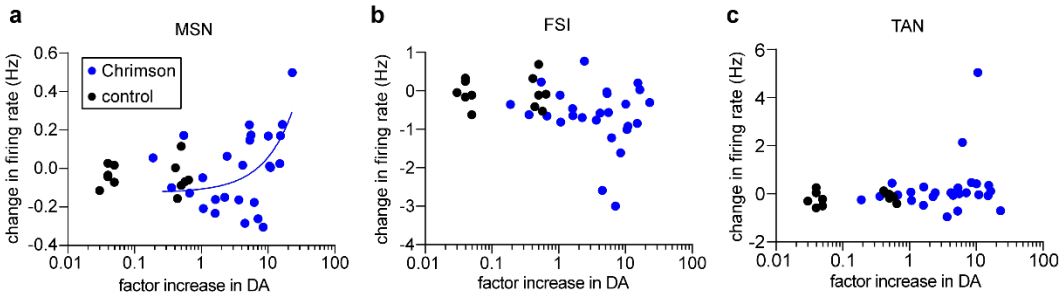
Extended Data Fig. 6: Strong effect of activating VTA GABAergic neurons on spontaneous striatal spiking activity.

a. Top: Experimental approach. Bottom: task schematic in which 0.5 s pulsed laser stimulation is delivered in the absence of reward or other stimuli.

b. Top: dLight fractional fluorescence change signal during laser stimulation from one Vgat-Chrimson mouse. Bottom: Mean firing rate of one neuron. Data represent mean ± SEM.

c. Top: Selectivity index of 379 neurons pooled across 3 Vgat-Chrimson mice. Bottom: Percentage of neurons that were significantly modulated by the laser relative to baseline, as a function of time.

d. Top: Mean accuracy of an SVM decoder trained using 50 neurons to discriminate laser-evoked from baseline activity. Shaded area represents the SD across 50 random drawings of neurons. Red line indicates the 95 % confidence interval of decoder performance trained on trial-shuffled data. Bottom: Maximum decoder accuracy as a function of neuron number (one-way ANOVA, neuron number effect: $F_{7,392}$ = 441, P < 0.0001). Post-hoc Sidak's test: P < 0.0001. Data represent the mean and SD across 50 random drawings of neurons.

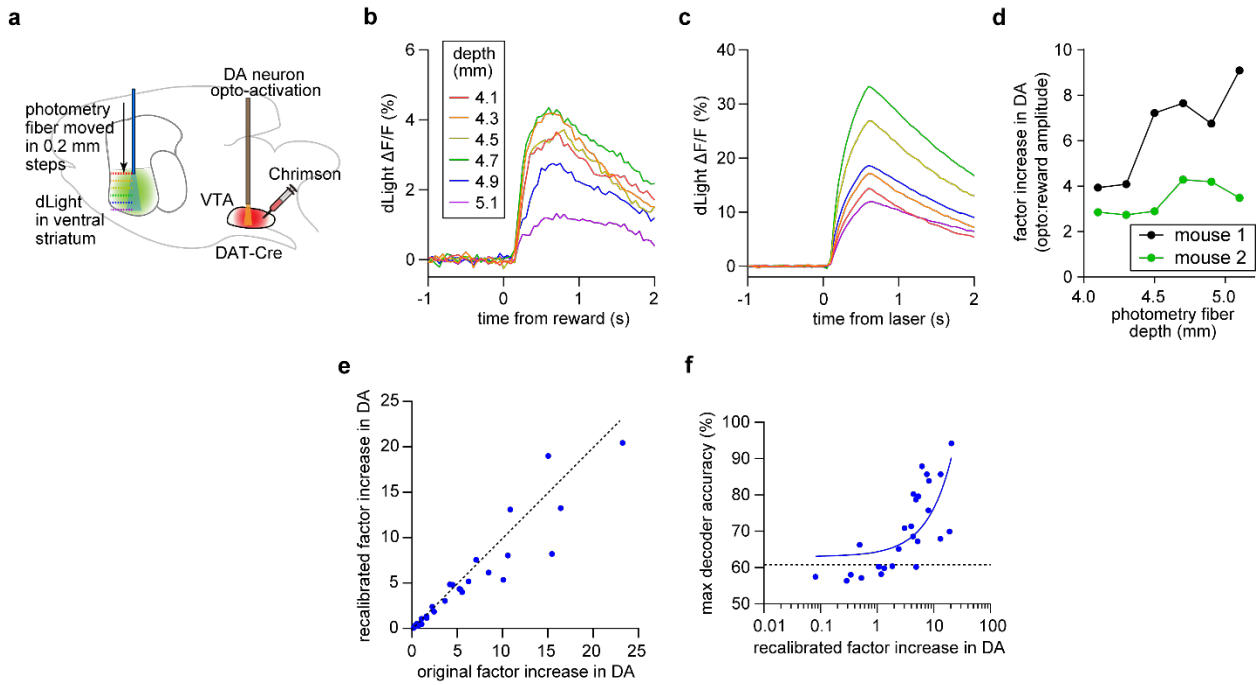


Extended Data Fig. 7: Factor increase in DA effects on neuron firing rates.

a. Changes in MSN firing rates compared to baseline as a function of factor increase in DA. n = 26 sessions from 9 Chrimson-expressing mice, and n = 12 sessions from 4 control mice. Chrimson group MSN data are positively correlated with the factor increase in DA (Pearson r = 0.6, P = 0.0033). Dashed line represents the 95 % confidence interval of the control data in all plots in figure. Blue line represents the linear fit to the Chrimson data.

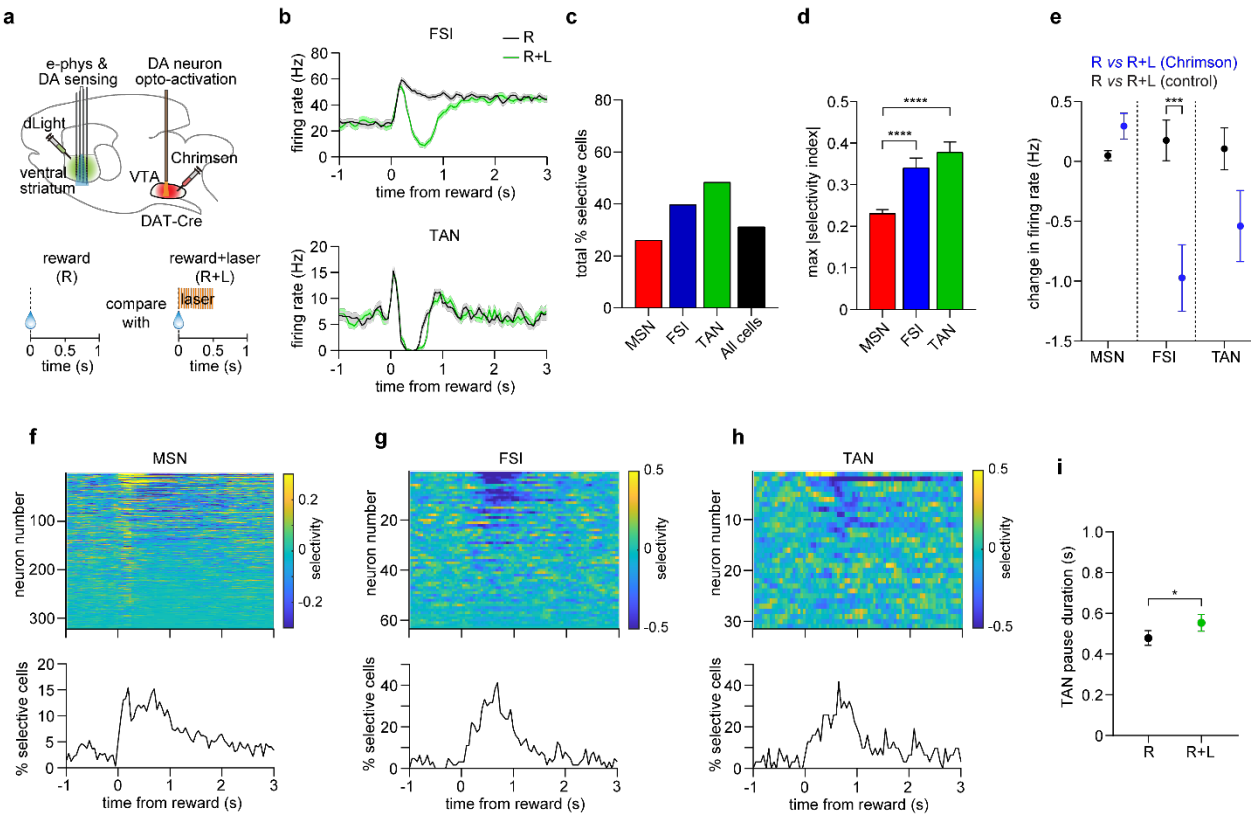
b. Changes in FSI firing rates compared to baseline as a function of factor increase in DA (Pearson r = 0.005, P = 0.98).

c. Changes in TAN firing rates compared to baseline as a function of factor increase in DA (Pearson r = 0.12, P = 0.55).



Extended Data Fig. 8: Assessing sources of variability in estimating the factor increase in DA.

a. Experimental approach of the dLight photometry depth test.
b. Reward-evoked dLight fractional fluorescence change signal from one Chrimson-expressing mouse. The plots are color-coded by photometry fiber depth from bregma.
c. Laser-evoked dLight fractional fluorescence change signal from one Chrimson-expressing mouse. The plots are color-coded by photometry fiber depth from bregma.
d. Factor increase in DA (the ration of laser-evoked to reward-evoked dLight signal amplitudes) calculated from data at different photometry fiber depths. Data are from n = 2 mice.
e. Recalibrated versus original factor increase in DA. The recalibrated factor used only the first 10 % of reward trials (n = 26 sessions from 9 Chrimson-expressing mice, paired t-test, $t = 2.1$, $P = 0.04$).
f. Maximum decoder accuracy per recording session, as a function of recalibrated factor increase in DA. The decoder was trained using 50 neurons to discriminate laser-evoked from baseline activity. Chrimson group data are positively correlated with the factor increase in DA (Pearson $r = 0.9$, $P < 0.0001$). Dashed line represents the 95 % confidence interval of the control data. Blue line represents the linear fit to the Chrimson data.



Extended Data Fig. 9: Effect of supra-reward DA signals on reward response of electrophysiologically classified cell types in the ventral striatum.

a. Top: Experimental approach. Bottom: task schematic in which reward trials (R) are compared to laser-paired reward trials (R+L). All data in this figure are from the supra-reward condition (4 mice exhibiting over 80 % decoding accuracy in Fig. 6e).

b. Mean firing rate of a putative FSI (top) and TAN (bottom) MSNs on R and R+L trials. Note the broadening of the TAN post-reward pause response. Data represent mean ± SEM.

c. Total percentage of each cell type that was selective for R versus R+L trials.

d. Mean of the maximum absolute value of the selectivity index per neuron, grouped by cell type (n = 322 MSN, 63 FSI, 31 TAN, one-way ANOVA, cell type effect: $F_{2,413}$ = 18, P < 0.0001). Post-hoc Sidak's test: ****P < 0.0001. Data represent mean ± SEM.

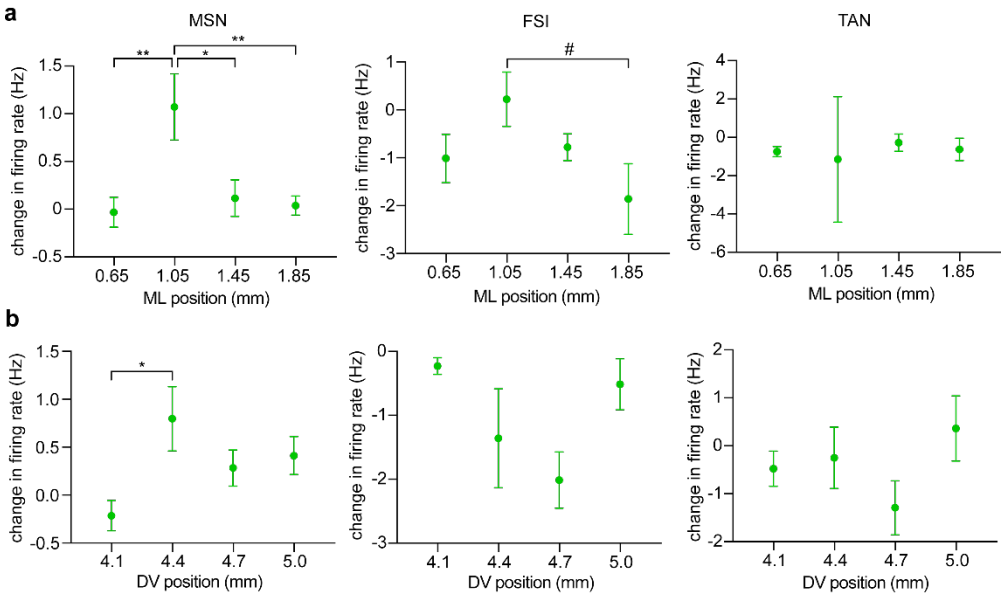
e. Mean firing rate changes for neurons of each cell type (Chrimson animals: n = 322 MSN, 63 FSI, 31 TAN; control animals: n = 308 MSN, 42 FSI, 17 TAN; one-way ANOVA, cell type x group effect: $F_{5,777}$ = 8.416, P < 0.0001). Post-hoc Sidak's test comparing eNpHR3 animals and controls: MSN (P = 0.12), FSI (P = 0.0005), TAN (P = 0.40). All data in this figure represent mean ± SEM.

f. Top: Selectivity index of 322 MSNs pooled across 4 Chrimson-expressing mice. Bottom: Percentage of neurons that were significantly selective for R versus R+L trials, as a function of time.

g. Top: Same as above for 63 FSIs. Bottom: Same as above.

h. Top: Same as above for 31 TANs. Bottom: Same as above.

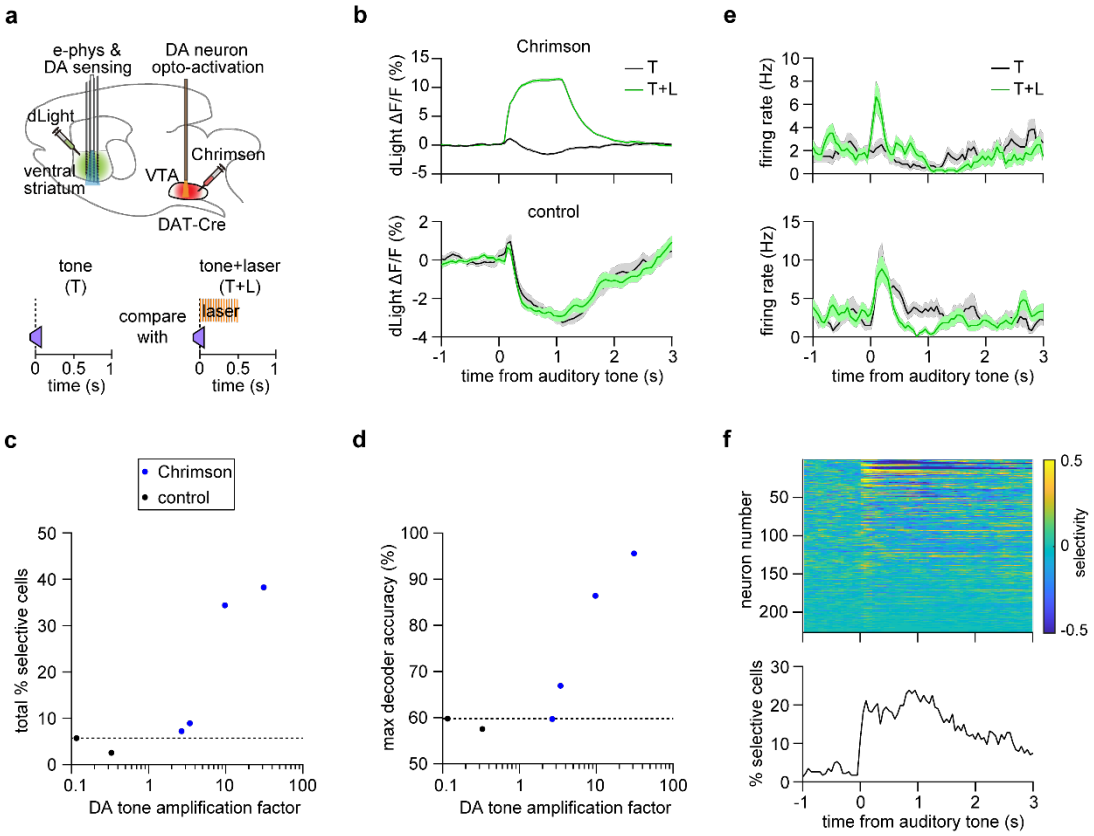
i. TAN post-reward pause responses (n = 24 TANs, paired t-test, t = 2.480, P = 0.0209).



Extended Data Fig. 10: Effect of supra-reward DA signals on neuron firing rates in the ventral striatum.

a. Mean firing rate changes for neurons of each cell type group by ML position. One-way ANOVA: MSN ML position effect ($F_{3,318}$ = 6, P = 0.0005), FSI ML position effect ($F_{3,59}$ = 2, P = 0.12), TAN ML position effect ($F_{3,27}$ = 0.22, P = 0.88). Post-hoc Tukey's test for MSN: P = 0.004 for 1.05 mm vs. 0.65 mm, P = 0.001 for 1.05 mm vs. 1.45 mm, P = 0.004 for 1.05 mm vs. 1.85 mm, all other comparisons were not significant.

b. Mean firing rate changes for neurons of each cell type group by DV position. One-way ANOVA: MSN DV position effect ($F_{3,318}$ = 3.027, P = 0.03), FSI DV position effect ($F_{3,59}$ = 1.848, P = 0.15), TAN DV position effect ($F_{3,27}$ = 1.599, P = 0.21). Post-hoc Tukey's test for MSN: P = 0.02 for 4.1 mm vs. 4.4 mm, all other comparisons were not significant.



Extended Data Fig. 11: Effect of activating DA neurons during auditory stimulus processing in the ventral striatum.

a. Top: Experimental approach. Bottom: task schematic in which neutral auditory tone trials (T) are compared to laser-paired tone trials (T+L).

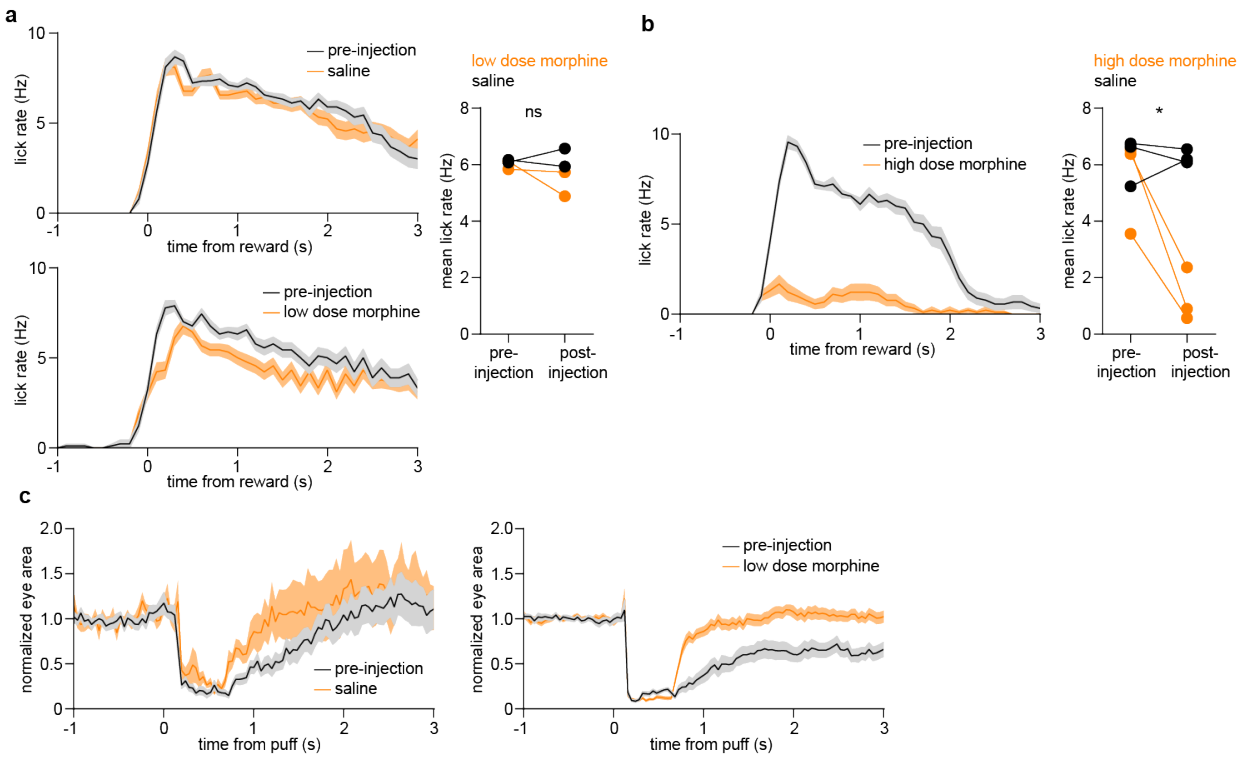
b. dLight fractional fluorescence change signal from one Chrimson-expressing mouse (top) and one opsin-free control mouse (bottom) on tone (T) and laser-paired tone (T+L) trials. Data represent mean ± SEM.

c. Percentage of neurons per animal that were selective for T versus T+L trials, as a function of DA tone amplification factor (an amplification factor of one corresponds to tone-matched DA levels). n = 4 Chrimson-expressing and 2 control mice (Pearson r = 0.8, P = 0.18). Dashed line represents the 95 % confidence interval of the control data.

d. Maximum decoder accuracy per animal, as a function of DA tone amplification factor. The decoder was trained using 50 neurons to discriminate R from R+L trials. n = 4 Chrimson-expressing and 2 control mice (Pearson r = 0.9, P = 0.12). Dashed line represents the 95 % confidence interval of the control data.

e. Mean firing rate of two putative MSNs on T and T+L trials. Data are from the supra-reward condition (mice exhibiting over 80 % decoding accuracy in panel d). Data represent mean ± SEM.

f. Top: Selectivity index of 226 neurons pooled across 2 Chrimson-expressing mice. Bottom: Percentage of neurons that were significantly selective for T versus T+L trials, as a function of time. Data are from the supra-reward condition (mice exhibiting over 80 % decoding accuracy in panel d).

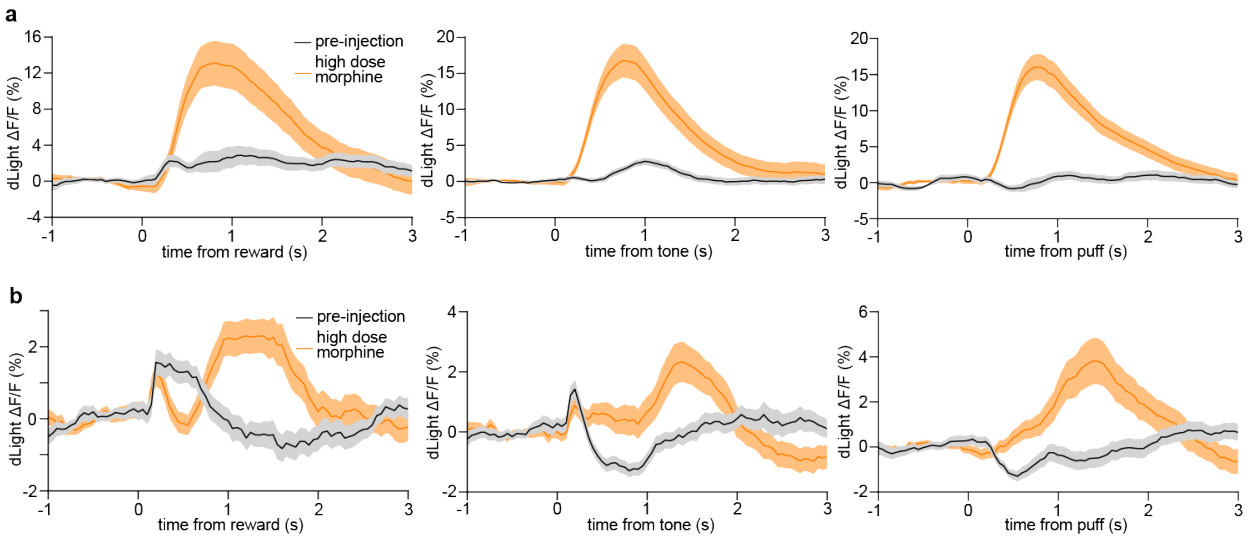


Extended Data Fig. 12: Effects of subcutaneous morphine administration on consummatory licking and eye blinking response to air puff.

a. Left: Reward-evoked licking before and after saline injection (top) or low dose morphine injection (bottom) from individual animals. Right: Mean reward-evoked lick rate before and after injection of saline or low dose morphine (n = 2 mice with 1 saline and 1 morphine session per mouse, two-way RM ANOVA, drug effect: $F_{1,2}$= 7.6, P = 0.11, pre-post injection effect: $F_{1,3}$ = 0.6, P = 0.5).

b. Left: Reward-evoked licking before and after high dose morphine injection from one animal. Right: Mean reward-evoked lick rate before and after injection high dose morphine (n = 3 mice with 1 saline and 1 morphine session per mouse, two-way RM ANOVA, drug effect: $F_{1,4}$= 15, P = 0.02, pre-post injection effect: $F_{1,5}$ = 4, P = 0.1).

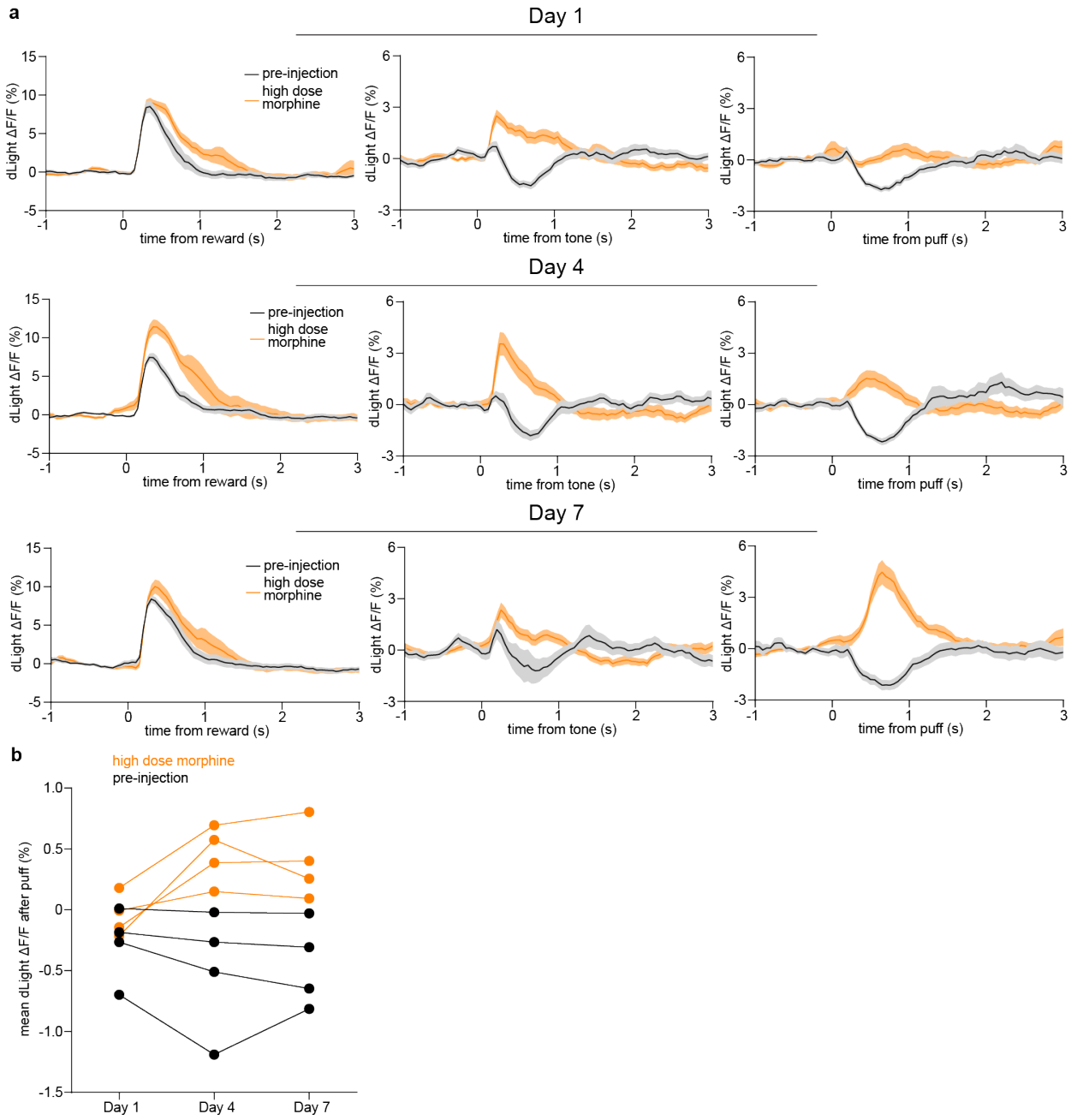
c. Eye blink response to air puff before and after saline (left) or low dose morphine injection (right).



Extended Data Fig. 13: Variability in effects of subcutaneous high dose morphine administration on DA responses to rewarding, neutral, and aversive stimuli.

a. dLight fractional fluorescence change signal in response to rewarding (left), neutral (center), and aversive (right) stimuli from a male mouse before and after high dose morphine injection.

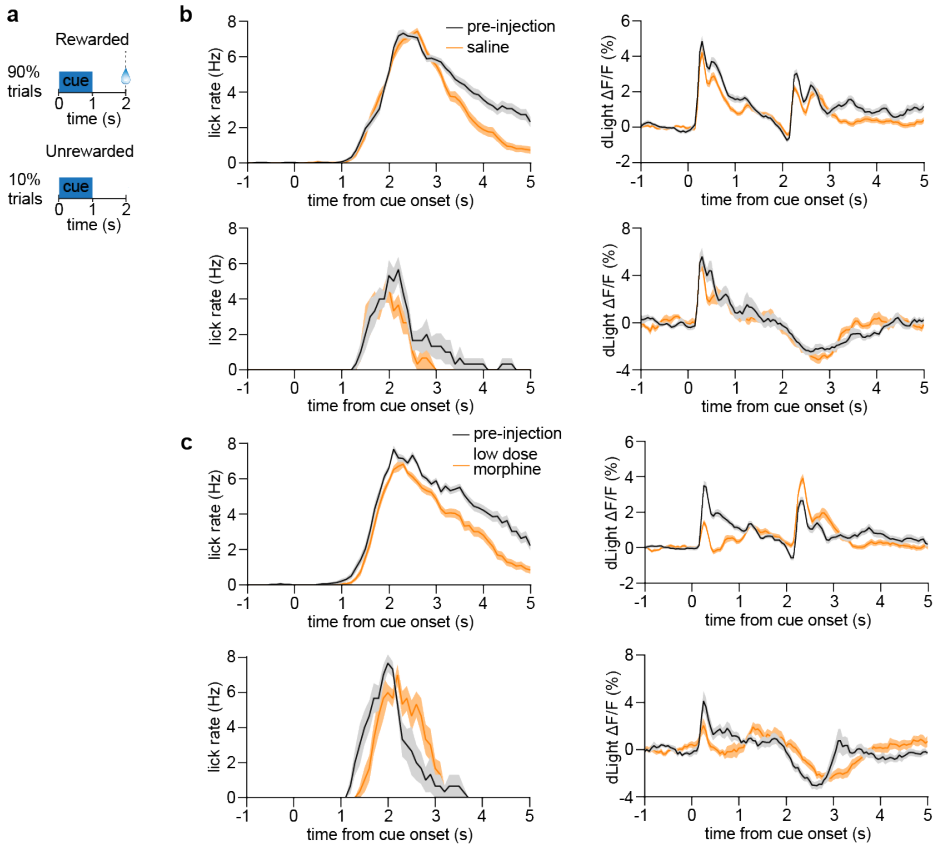
b. dLight fractional fluorescence change signal in response to rewarding (left), neutral (center), and aversive (right) stimuli from a female mouse before and after high dose morphine injection.



Extended Data Fig. 14: Sensitization of effects of morphine administration on DA responses over consecutive daily injections.

a. dLight fractional fluorescence change signal in response to rewarding (left column), neutral (center column), and aversive (right column) stimuli from a single mouse before and after daily high dose morphine injection over the course of 7 days.

b. Mean air puff-evoked dLight fluorescence across 7 days of daily high dose morphine injection (n = 4 mice, mixed-effects model, drug effect: $F_{1,6}$ = 10, P = 0.02, time effect: $F_{1.9,11.4}$ = 2.3, P = 0.15, time x drug effect: $F_{2,12}$ = 10, P = 0.001).

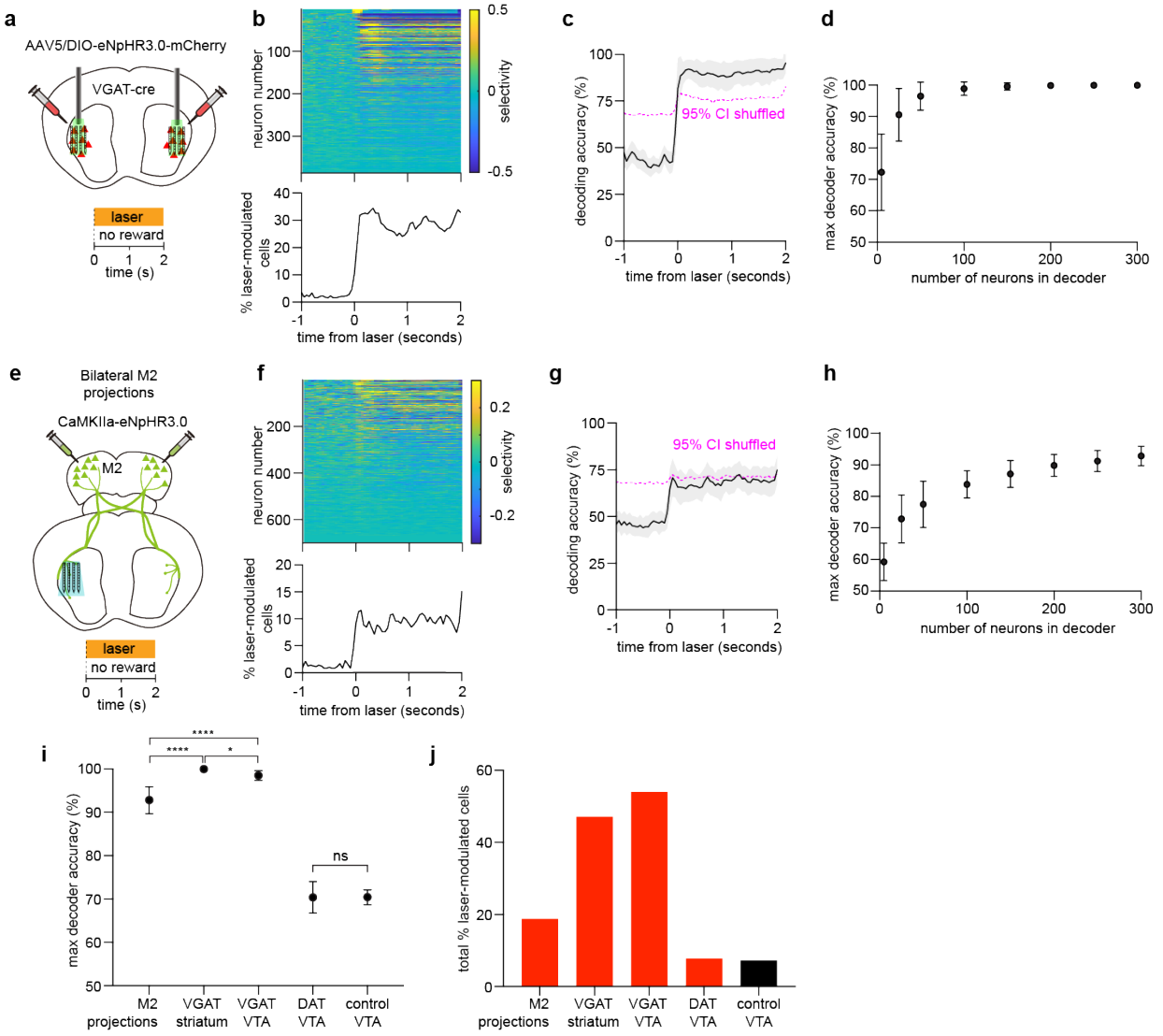


Extended Data Fig. 15: Sensitization of effects of morphine administration on DA responses over consecutive daily injections.

a. Schematic of Pavlovian task with rewarded and unrewarded trials.

b. Licking responses (left) and dLight fractional fluorescence change (right) for trials that were rewarded (top) and unrewarded (bottom) before and after saline injection.

c. Licking responses (left) and dLight fractional fluorescence change (right) for trials that were rewarded (top) and unrewarded (bottom) before and after low dose morphine injection.



Extended Data Fig. 16: Large effect of inhibiting striatal GABA interneurons and bilateral M2 projections.

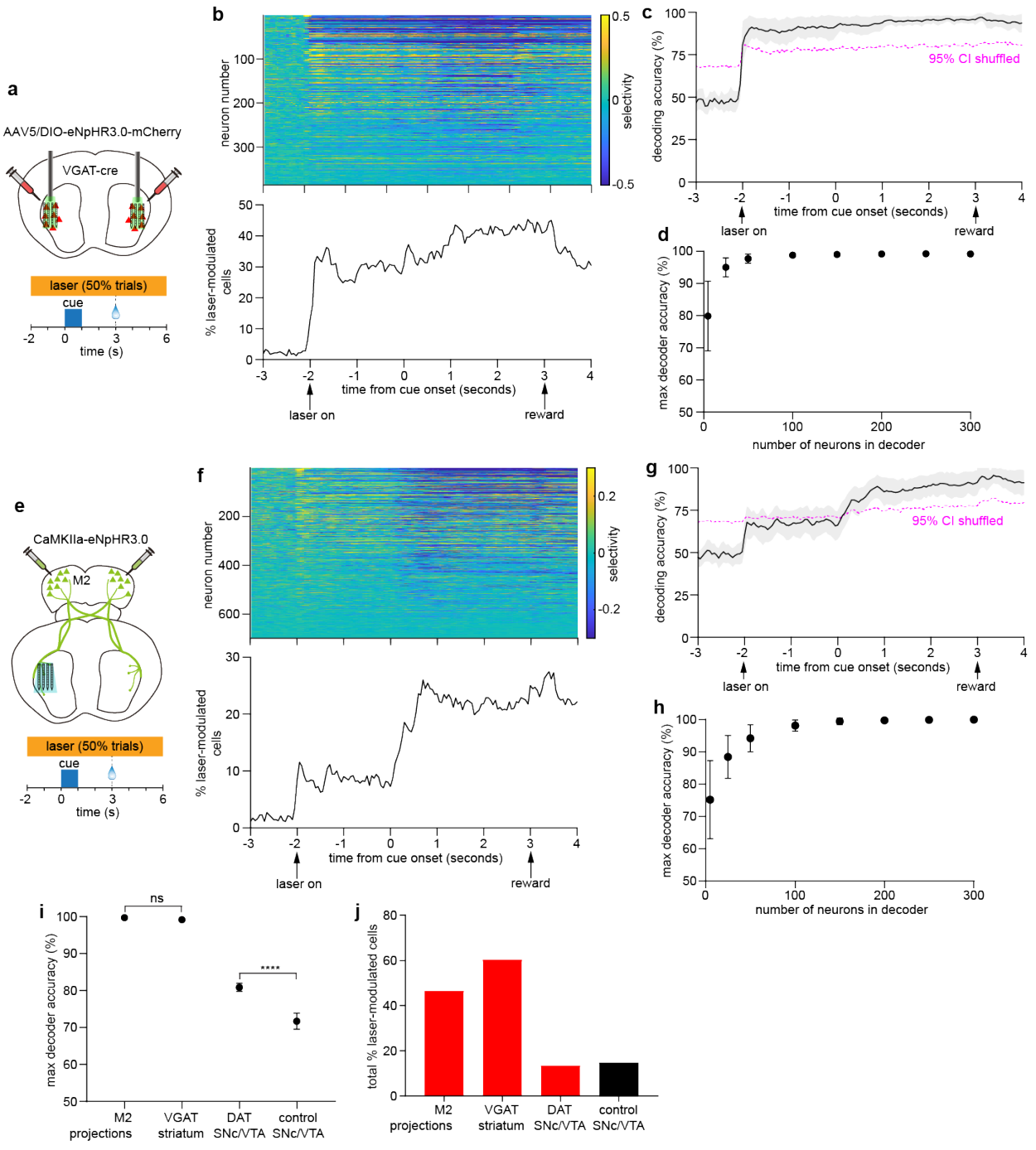
a. Top: Experimental approach for simultaneous local striatal GABA neuron inhibition and electrophysiological recording. Bottom: task schematic showing 2 seconds of continuous optogenetic inhibition in the absence of other stimuli.

b. Top: Selectivity index of 386 neurons pooled across 7 separate recordings in VGAT-cre mice. Bottom: Percentage of neurons that were significantly modulated by the laser relative to baseline, as a function of time.

c. Mean accuracy of an SVM decoder trained using 50 neurons to discriminate laser-evoked from baseline activity. Shaded area represents the SD across 50 random drawings of neurons. Magenta line indicates the 95 % confidence interval of decoder performance trained on trial-shuffled data.

d. Maximum decoder accuracy as a function of neuron number (one-way ANOVA, neuron number effect: $F_{7,392} = 150.3$, $P < 0.0001$). Post-hoc Sidak's test: $P < 0.0001$. Data represent the mean and SD across 50 random drawings of neurons.

e. Top: Experimental approach for simultaneous inhibition of bilateral M2 inputs to striatum and electrophysiological recording. Bottom: task schematic showing 2 seconds of continuous optogenetic inhibition in the absence of other stimuli.
f. Top: Selectivity index of 699 neurons pooled across 9 separate recordings. Bottom: Percentage of neurons that were significantly modulated by the laser relative to baseline, as a function of time.
g. Mean accuracy of an SVM decoder trained using 50 neurons to discriminate laser-evoked from baseline activity. Shaded area represents the SD across 50 random drawings of neurons. Magenta line indicates the 95 % confidence interval of decoder performance trained on trial-shuffled data.
h. Maximum decoder accuracy as a function of neuron number (one-way ANOVA, neuron number effect: $F_{7,392} = 241.6$, $P < 0.0001$). Post-hoc Sidak's test: $P < 0.0001$. Data represent the mean and SD across 50 random drawings of neurons.
i. Maximum decoder accuracy across experiments involving optogenetic manipulations of different cell populations in the absence of other stimuli (one-way ANOVA, experimental group effect: $F_{4,245} = 2059$, $P < 0.0001$). Post-hoc Sidak's test: $P = 0.0137$ for VGAT-Striatum vs. VGAT-VTA, $P = 0.999$ for DAT-VTA vs. control-VTA, all other comparisons $P < 0.0001$. Data represent the mean and SD across 50 random drawings of neurons.
j. Total percentage of neurons that were laser-modulated relative to baseline across experiments involving optogenetic manipulations of different cell populations in the absence of other stimuli.



Extended Data Fig. 17: Large effect of inhibiting striatal GABA interneurons and bilateral M2 projections during Pavlovian conditioning task.

a. Top: Experimental approach for simultaneous local striatal GABA neuron inhibition and electrophysiological recording. Bottom: task schematic in which half of trials include optogenetic inhibition, which begins 2 seconds before cue onset.

b. Top: Selectivity index of 386 neurons pooled across 7 separate recordings in VGAT-cre mice. Bottom: Percentage of neurons that were significantly modulated by the laser relative to trials without laser, as a function of time.

c. Mean accuracy of an SVM decoder trained using 50 neurons to discriminate laser-paired trials and trials without laser. Magenta dashed line indicates the 95 % confidence interval of decoder performance trained on trial-shuffled data. Shaded area represents the SD across 50 random drawings of neurons.
d. Maximum decoder accuracy within 2 seconds after cue onset as a function of neuron number (one-way ANOVA, neuron number effect: $F_{7,392} = 138.4$, $P < 0.0001$). Post-hoc Sidak's test: $P < 0.0001$. Data represent the mean and SD across 50 random drawings of neurons.
e. Top: Experimental approach for simultaneous inhibition of bilateral M2 inputs to striatum and electrophysiological recording. Bottom: task schematic in which half of trials include optogenetic inhibition, which begins 2 seconds before cue onset.
f. Top: Selectivity index of 699 neurons pooled across 9 separate recordings. Bottom: Percentage of neurons that were significantly modulated by the laser relative to trials without laser, as a function of time.
g. Mean accuracy of an SVM decoder trained using 50 neurons to discriminate laser-paired trials and trials without laser. Magenta dashed line indicates the 95 % confidence interval of decoder performance trained on trial-shuffled data. Shaded area represents the SD across 50 random drawings of neurons.
h. Maximum decoder accuracy within 2 seconds after cue onset as a function of neuron number (one-way ANOVA, neuron number effect: $F_{7,392} = 143.5$, $P < 0.0001$). Post-hoc Sidak's test: $P < 0.0001$. Data represent the mean and SD across 50 random drawings of neurons.
i. Maximum decoder accuracy (with 200 neurons used in the decoder) within 2 seconds after cue onset across experiments involving optogenetic manipulations of different cell populations during a Pavlovian conditioning task (one-way ANOVA, experimental group effect: $F_{3,196} = 6344$, $P < 0.0001$). Post-hoc Sidak's test: $P = 0.1134$ for M2 Projections vs. VGAT-striatum, all other comparisons $P < 0.0001$. Data represent the mean and SD across 50 random drawings of neurons.
j. Total percentage of neurons that were laser-modulated (within 2 seconds of cue onset) relative to trials without laser across experiments involving optogenetic manipulations of different cell populations during a Pavlovian conditioning task.

References:

1. Long, C. et al. Physiological constraints on the rapid dopaminergic modulation of striatal reward activity. *bioRxiv*, 2022.2009.2016.508310 (2022).
2. Nicola, S.M., Surmeier, J. & Malenka, R.C. Dopaminergic modulation of neuronal excitability in the striatum and nucleus accumbens. *Annu Rev Neurosci* **23**, 185-215 (2000).
3. Albin, R.L., Young, A.B. & Penney, J.B. The functional anatomy of basal ganglia disorders. *Trends in neurosciences* **12**, 366-375 (1989).
4. Tye, K.M. et al. Dopamine neurons modulate neural encoding and expression of depression-related behaviour. *Nature* (2012).
5. Berke, J.D. What does dopamine mean? *Nat Neurosci* **21**, 787-793 (2018).
6. Schultz, W., Dayan, P. & Montague, P.R. A neural substrate of prediction and reward. *Science* **275**, 1593-1599 (1997).
7. Koob, G.F. Dopamine, addiction and reward. *Seminars in Neuroscience* **4**, 139-148 (1992).
8. Wise, R.A. & Hoffman, D.C. Localization of drug reward mechanisms by intracranial injections. *Synapse* **10**, 247-263 (1992).
9. Di Chiara, G. The role of dopamine in drug abuse viewed from the perspective of its role in motivation. *Drug Alcohol Depend* **38**, 95-137 (1995).
10. Wise, R.A. & Rompre, P.P. Brain dopamine and reward. *Annu Rev Psychol* **40**, 191-225 (1989).
11. Le Moal, M. & Simon, H. Mesocorticolimbic dopaminergic network: functional and regulatory roles. *Physiol Rev* **71**, 155-234 (1991).
12. Steinberg, E.E. et al. A causal link between prediction errors, dopamine neurons and learning. *Nat Neurosci* **16**, 966-973 (2013).

13. Waelti, P., Dickinson, A. & Schultz, W. Dopamine responses comply with basic assumptions of formal learning theory. *Nature* **412**, 43-48 (2001).

14. Yagishita, S. et al. A critical time window for dopamine actions on the structural plasticity of dendritic spines. *Science* **345**, 1616-1620 (2014).

15. Tanaka, J. et al. Protein synthesis and neurotrophin-dependent structural plasticity of single dendritic spines. *Science* **319**, 1683-1687 (2008).

16. Dan, Y. & Poo, M.M. Spike timing-dependent plasticity of neural circuits. *Neuron* **44**, 23-30 (2004).

17. Reynolds, J.N., Hyland, B.I. & Wickens, J.R. A cellular mechanism of reward-related learning. *Nature* **413**, 67-70 (2001).

18. Roitman, M.F., Stuber, G.D., Phillips, P.E., Wightman, R.M. & Carelli, R.M. Dopamine operates as a subsecond modulator of food seeking. *J Neurosci* **24**, 1265-1271 (2004).

19. Berridge, K.C., Robinson, T.E. & Aldridge, J.W. Dissecting components of reward: 'liking', 'wanting', and learning. *Curr Opin Pharmacol* **9**, 65-73 (2009).

20. da Silva, J.A., Tecuapetla, F., Paixao, V. & Costa, R.M. Dopamine neuron activity before action initiation gates and invigorates future movements. *Nature* **554**, 244-248 (2018).

21. Armstrong, M.J. & Okun, M.S. Diagnosis and Treatment of Parkinson Disease: A Review. *JAMA* **323**, 548-560 (2020).

22. Schultz, W. Multiple dopamine functions at different time courses. *Annu Rev Neurosci* **30**, 259-288 (2007).

23. Tecuapetla, F., Jin, X., Lima, S.Q. & Costa, R.M. Complementary Contributions of Striatal Projection Pathways to Action Initiation and Execution. *Cell* **166**, 703-715 (2016).

24. Kravitz, A.V. et al. Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry. *Nature* **466**, 622-626 (2010).

25. Howe, M.W. & Dombeck, D.A. Rapid signalling in distinct dopaminergic axons during locomotion and reward. *Nature* **535**, 505-510 (2016).

26. Hughes, R.N. et al. Ventral Tegmental Dopamine Neurons Control the Impulse Vector during Motivated Behavior. *Curr Biol* **30**, 2681-2694 e2685 (2020).
27. Soares, S., Atallah, B.V. & Paton, J.J. Midbrain dopamine neurons control judgment of time. *Science* **354**, 1273-1277 (2016).
28. Hamilos, A.E. et al. Slowly evolving dopaminergic activity modulates the moment-to-moment probability of reward-related self-timed movements. *Elife* **10** (2021).
29. Howard, C.D., Li, H., Geddes, C.E. & Jin, X. Dynamic Nigrostriatal Dopamine Biases Action Selection. *Neuron* **93**, 1436-1450 e1438 (2017).
30. Zhou, S., Masmanidis, S.C. & Buonomano, D.V. Neural Sequences as an Optimal Dynamical Regime for the Readout of Time. *Neuron* **108**, 651-658 e655 (2020).
31. Hamid, A.A. et al. Mesolimbic dopamine signals the value of work. *Nat Neurosci* **19**, 117-126 (2016).
32. Liu, C. et al. An action potential initiation mechanism in distal axons for the control of dopamine release. *Science* **375**, 1378-1385 (2022).
33. Mohebi, A. et al. Dissociable dopamine dynamics for learning and motivation. *Nature* **570**, 65-70 (2019).
34. Engelhard, B. et al. Specialized coding of sensory, motor and cognitive variables in VTA dopamine neurons. *Nature* (2019).
35. Sippy, T. & Tritsch, N.X. Unraveling the dynamics of dopamine release and its actions on target cells. *Trends in neurosciences* **46**, 228-239 (2023).
36. Lahiri, A.K. & Bevan, M.D. Dopaminergic Transmission Rapidly and Persistently Enhances Excitability of D1 Receptor-Expressing Striatal Projection Neurons. *Neuron* **106**, 277-290 e276 (2020).
37. Coddington, L.T. & Dudman, J.T. The timing of action determines reward prediction signals in identified midbrain dopamine neurons. *Nat Neurosci* **21**, 1563-1573 (2018).

38. Lee, K. et al. Temporally restricted dopaminergic control of reward-conditioned movements. *Nat Neurosci* **23**, 209-216 (2020).

39. Pan, W.X., Coddington, L.T. & Dudman, J.T. Dissociable contributions of phasic dopamine activity to reward and prediction. *Cell Rep* **36**, 109684 (2021).

40. Markowitz, J.E. et al. Spontaneous behaviour is structured by reinforcement without explicit reward. *Nature* **614**, 108-117 (2023).

41. Coddington, L.T., Lindo, S.E. & Dudman, J.T. Mesolimbic dopamine adapts the rate of learning from action. *Nature* **614**, 294-302 (2023).

42. Xiao, C. et al. Cholinergic Mesopontine Signals Govern Locomotion and Reward through Dissociable Midbrain Pathways. *Neuron* **90**, 333-347 (2016).

43. Parker, N.F. et al. Reward and choice encoding in terminals of midbrain dopamine neurons depends on striatal target. *Nat Neurosci* **19**, 845-854 (2016).

44. Patriarchi, T. et al. An expanded palette of dopamine sensors for multiplex imaging in vivo. *Nat Methods* **17**, 1147-+ (2020).

45. Collins, A.L. & Saunders, B.T. Heterogeneity in striatal dopamine circuits: Form and function in dynamic reward seeking. *J Neurosci Res* **98**, 1046-1069 (2020).

46. 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).

47. Vu, M.T. et al. Targeted micro-fiber arrays for measuring and manipulating localized multi-scale neural dynamics over large, deep brain volumes during behavior. *Neuron* **112**, 909-923 e909 (2024).

48. Yin, H.H. et al. Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. *Nat Neurosci* **12**, 333-341 (2009).

49. Klaus, A., Alves da Silva, J. & Costa, R.M. What, If, and When to Move: Basal Ganglia Circuits and Self-Paced Action Initiation. *Annu Rev Neurosci* **42**, 459-483 (2019).

50. Volkow, N.D., Fowler, J.S., Wang, G.J., Swanson, J.M. & Telang, F. Dopamine in drug abuse and addiction: results of imaging studies and treatment implications. *Arch Neurol* **64**, 1575-1579 (2007).

51. Pierce, R.C. & Kumaresan, V. The mesolimbic dopamine system: the final common pathway for the reinforcing effect of drugs of abuse? *Neurosci Biobehav Rev* **30**, 215-238 (2006).

52. Laviolette, S.R., Gallegos, R.A., Henriksen, S.J. & van der Kooy, D. Opiate state controls bi-directional reward signaling via GABAA receptors in the ventral tegmental area. *Nat Neurosci* **7**, 160-169 (2004).

53. Wise, R.A. & Jordan, C.J. Dopamine, behavior, and addiction. *J Biomed Sci* **28**, 83 (2021).

54. Volkow, N.D., Michaelides, M. & Baler, R. The Neuroscience of Drug Reward and Addiction. *Physiol Rev* **99**, 2115-2140 (2019).

55. Blanco, C., Wall, M.M. & Olfson, M. Expanding Current Approaches to Solve the Opioid Crisis. *JAMA Psychiatry* **79**, 5-6 (2022).

56. Di Chiara, G. & Imperato, A. Drugs abused by humans preferentially increase synaptic dopamine concentrations in the mesolimbic system of freely moving rats. *Proc Natl Acad Sci U S A* **85**, 5274-5278 (1988).

57. Pothos, E., Rada, P., Mark, G.P. & Hoebel, B.G. Dopamine microdialysis in the nucleus accumbens during acute and chronic morphine, naloxone-precipitated withdrawal and clonidine treatment. *Brain Res* **566**, 348-350 (1991).

58. Vander Weele, C.M. et al. Rapid dopamine transmission within the nucleus accumbens: dramatic difference between morphine and oxycodone delivery. *Eur J Neurosci* **40**, 3041-3054 (2014).

59. Xi, Z.X., Fuller, S.A. & Stein, E.A. Dopamine release in the nucleus accumbens during heroin self-administration is modulated by kappa opioid receptors: an in vivo fast-cyclic

voltammetry study. *The Journal of pharmacology and experimental therapeutics* **284**, 151-161 (1998).

60. Gomez, A.A. et al. Local mu-Opioid Receptor Antagonism Blunts Evoked Phasic Dopamine Release in the Nucleus Accumbens of Rats. *ACS Chem Neurosci* **10**, 1935-1940 (2019).

61. Lefevre, E.M. et al. Interruption of continuous opioid exposure exacerbates drug-evoked adaptations in the mesolimbic dopamine system. *Neuropsychopharmacology* **45**, 1781-1792 (2020).

62. Fields, H.L. & Margolis, E.B. Understanding opioid reward. *Trends Neurosci* **38**, 217-225 (2015).

63. Johnson, S.W. & North, R.A. Opioids excite dopamine neurons by hyperpolarization of local interneurons. *J Neurosci* **12**, 483-488 (1992).

64. Amo, R., Uchida, N. & Watabe-Uchida, M. Glutamate inputs send prediction error of reward, but not negative value of aversive stimuli, to dopamine neurons. *Neuron* **112**, 1001-1019 e1006 (2024).

65. Bloom, F.E., Costa, E. & Salmoiraghi, G.C. Anesthesia and the responsiveness of individual neurons of the caudate nucleus of the cat to acetylcholine, norepinephrine and dopamine administered by microelectrophoresis. *J Pharmacol Exp Ther* **150**, 244-252 (1965).

66. Kitai, S.T., Sugimori, M. & Kocsis, J.D. Excitatory nature of dopamine in the nigro-caudate pathway. *Exp Brain Res* **24**, 351-363 (1976).

67. Yim, C.Y. & Mogenson, G.J. Response of nucleus accumbens neurons to amygdala stimulation and its modification by dopamine. *Brain research* **239**, 401-415 (1982).

68. Connor, J.D. Caudate unit responses to nigral stimuli: evidence for a possible nigro-neostriatal pathway. *Science* **160**, 899-900 (1968).

69. Gonon, F. Prolonged and extrasynaptic excitatory action of dopamine mediated by D1 receptors in the rat striatum in vivo. *J Neurosci* **17**, 5972-5978 (1997).

70. Cheer, J.F., Heien, M.L., Garris, P.A., Carelli, R.M. & Wightman, R.M. Simultaneous dopamine and single-unit recordings reveal accumbens GABAergic responses: implications for intracranial self-stimulation. *Proc Natl Acad Sci U S A* **102**, 19150-19155 (2005).

71. Cheer, J.F. et al. Coordinated accumbal dopamine release and neural activity drive goal-directed behavior. *Neuron* **54**, 237-244 (2007).

72. du Hoffmann, J. & Nicola, S.M. Dopamine invigorates reward seeking by promoting cue-evoked excitation in the nucleus accumbens. *J Neurosci* **34**, 14349-14364 (2014).

73. Ketzef, M. et al. Dopamine Depletion Impairs Bilateral Sensory Processing in the Striatum in a Pathway-Dependent Manner. *Neuron* **94**, 855-865 e855 (2017).

74. Maltese, M., March, J.R., Bashaw, A.G. & Tritsch, N.X. Dopamine differentially modulates the size of projection neuron ensembles in the intact and dopamine-depleted striatum. *Elife* **10** (2021).

75. Kim, D.S., Szczypka, M.S. & Palmiter, R.D. Dopamine-deficient mice are hypersensitive to dopamine receptor agonists. *J Neurosci* **20**, 4405-4413 (2000).

76. Wang, D.V. et al. Disrupting Glutamate Co-transmission Does Not Affect Acquisition of Conditioned Behavior Reinforced by Dopamine Neuron Activation. *Cell Rep* **18**, 2584-2591 (2017).

77. Ferenczi, E.A. et al. Prefrontal cortical regulation of brainwide circuit dynamics and reward-related behavior. *Science* **351** (2016).

78. Gerfen, C.R. & Surmeier, D.J. Modulation of striatal projection systems by dopamine. *Annu Rev Neurosci* **34**, 441-466 (2011).

79. Tritsch, N.X. & Sabatini, B.L. Dopaminergic modulation of synaptic transmission in cortex and striatum. *Neuron* **76**, 33-50 (2012).

80. Kyrozis, A. & Reichling, D.B. Perforated-patch recording with gramicidin avoids artifactual changes in intracellular chloride concentration. *J Neurosci Methods* **57**, 27-35 (1995).

81. Lindroos, R. et al. Basal Ganglia Neuromodulation Over Multiple Temporal and Structural Scales-Simulations of Direct Pathway MSNs Investigate the Fast Onset of Dopaminergic Effects and Predict the Role of Kv4.2. *Front Neural Circuits* **12**, 3 (2018).

82. McGregor, M.M. & Nelson, A.B. Directly to the Point: Dopamine Persistently Enhances Excitability of Direct Pathway Striatal Neurons. *Neuron* **106**, 201-203 (2020).

83. Dorst, M.C. et al. Polysynaptic inhibition between striatal cholinergic interneurons shapes their network activity patterns in a dopamine-dependent manner. *Nat Commun* **11**, 5113 (2020).

84. Morgenstern, N.A., Isidro, A.F., Israely, I. & Costa, R.M. Pyramidal tract neurons drive amplification of excitatory inputs to striatum through cholinergic interneurons. *Sci Adv* **8**, eabh4315 (2022).

85. Brown, H.D., McCutcheon, J.E., Cone, J.J., Ragozzino, M.E. & Roitman, M.F. Primary food reward and reward-predictive stimuli evoke different patterns of phasic dopamine signaling throughout the striatum. *Eur J Neurosci* **34**, 1997-2006 (2011).

86. Roitman, M.F., Wheeler, R.A. & Carelli, R.M. Nucleus accumbens neurons are innately tuned for rewarding and aversive taste stimuli, encode their predictors, and are linked to motor output. *Neuron* **45**, 587-597 (2005).

87. Lee, K. et al. Gain Modulation by Corticostriatal and Thalamostriatal Input Signals during Reward-Conditioned Behavior. *Cell Rep* **29**, 2438-2449 e2434 (2019).

88. Li, N. & Jasanoff, A. Local and global consequences of reward-evoked striatal dopamine release. *Nature* **580**, 239-244 (2020).

89. Brown, M.T. et al. Ventral tegmental area GABA projections pause accumbal cholinergic interneurons to enhance associative learning. *Nature* **492**, 452-456 (2012).

90. Eshel, N. et al. Arithmetic and local circuitry underlying dopamine prediction errors. *Nature* (2015).

91. Chuhma, N., Mingote, S., Moore, H. & Rayport, S. Dopamine neurons control striatal cholinergic neurons via regionally heterogeneous dopamine and glutamate signaling. *Neuron* **81**, 901-912 (2014).

92. Berridge, K.C. The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology (Berl)* **191**, 391-431 (2007).

93. Dodson, P.D. et al. Representation of spontaneous movement by dopaminergic neurons is cell-type selective and disrupted in parkinsonism. *Proc Natl Acad Sci U S A* **113**, E2180-2188 (2016).

94. Dobbs, L.K. et al. Dopamine Regulation of Lateral Inhibition between Striatal Neurons Gates the Stimulant Actions of Cocaine. *Neuron* **90**, 1100-1113 (2016).

95. Stuber, G.D., Hnasko, T.S., Britt, J.P., Edwards, R.H. & Bonci, A. Dopaminergic terminals in the nucleus accumbens but not the dorsal striatum corelease glutamate. *J Neurosci* **30**, 8229-8233 (2010).

96. Tritsch, N.X., Ding, J.B. & Sabatini, B.L. Dopaminergic neurons inhibit striatal output through non-canonical release of GABA. *Nature* **490**, 262-266 (2012).

97. Liu, H. et al. A permissive role for dopamine in the production of vigorous movements. *bioRxiv*, 2022.2011.2003.514328 (2022).

98. Goldman, M.S. & Fee, M.S. Computational training for the next generation of neuroscientists. *Curr Opin Neurobiol* **46**, 25-30 (2017).

99. Grisham, W., Lom, B., Lanyon, L. & Ramos, R.L. Proposed Training to Meet Challenges of Large-Scale Data in Neuroscience. *Front Neuroinform* **10**, 28 (2016).

100. Juavinett, A.L. The next generation of neuroscientists needs to learn how to code, and we need new ways to teach them. *Neuron* **110**, 576-578 (2022).

101. Pisner, D.A. & Schnyer, D.M. in Machine Learning. (eds. A. Mechelli & S. Vieira) 101-121 (Academic Press, 2020).

102. Noble, W.S. What is a support vector machine? *Nat Biotechnol* **24**, 1565-1567 (2006).

103. Schultz, W. Predictive reward signal of dopamine neurons. *J Neurophysiol* **80**, 1-27 (1998).

104. Jaskir, A. & Frank, M.J. On the normative advantages of dopamine and striatal opponency for learning and choice. *Elife* **12** (2023).

105. Nicola, S.M. & Malenka, R.C. Modulation of synaptic transmission by dopamine and norepinephrine in ventral but not dorsal striatum. *J Neurophysiol* **79**, 1768-1776 (1998).

106. Lohani, S., Martig, A.K., Deisseroth, K., Witten, I.B. & Moghaddam, B. Dopamine Modulation of Prefrontal Cortex Activity Is Manifold and Operates at Multiple Temporal and Spatial Scales. *Cell Rep* **27**, 99-114 e116 (2019).

107. Vander Weele, C.M. et al. Dopamine enhances signal-to-noise ratio in cortical-brainstem encoding of aversive stimuli. *Nature* **563**, 397-401 (2018).

108. Kutlu, M.G. et al. Dopamine release in the nucleus accumbens core signals perceived saliency. *Curr Biol* **31**, 4748-4761.e4748 (2021).

109. Diaz-Hernandez, E. et al. The Thalamostriatal Projections Contribute to the Initiation and Execution of a Sequence of Movements. *Neuron* **100**, 739-752 e735 (2018).

110. Yang, L., Lee, K., Villagracia, J. & Masmanidis, S.C. Open source silicon microprobes for high throughput neural recording. *J Neural Eng* **17**, 016036 (2020).

111. Pachitariu, M., Steinmetz, N., Kadir, S., Carandini, M. & Harris, K.D. Kilosort: realtime spike-sorting for extracellular electrophysiology with hundreds of channels. *bioRxiv* (2016).

112. Bakhurin, K.I., Mac, V., Golshani, P. & Masmanidis, S.C. Temporal correlations among functionally specialized striatal neural ensembles in reward-conditioned mice. *J Neurophysiol* **115**, 1521-1532 (2016).

113. Dong, Z. et al. Minian, an open-source miniscope analysis pipeline. *Elife* **11** (2022).

114. Chang, C.C. & Lin, C.J. LIBSVM: A Library for Support Vector Machines. *Acm T Intel Syst Tec* **2** (2011).

115. Yang, L. & Masmanidis, S.C. Differential encoding of action selection by orbitofrontal and striatal population dynamics. *J Neurophysiol* **124**, 634-644 (2020).

Grant support: SCM was supported by NIH grants NS100050, NS125877, DA042739, DA005010, and NSF NeuroNex Award 1707408. CL was supported by NIH grant 5T32DA024635.