Constraint of Motivation and Value by Limbic Opioidergic Systems

Christian E. Pedersen

A dissertation
submitted in partial fulfillment of the
requirements of the degree of

Doctor of Philosophy

University of Washington
2021

Reading Committee:
Michael R. Bruchas, Chair
Amy L. Orsborn
Samuel A. Golden

Program Authorized to Offer Degree:
Bioengineering

University of Washington

Abstract

Constraint of Motivation and Value by Limbic Opioidergic Systems

Christian E. Pedersen

Chair of the Supervisory Committee:
Michael R. Bruchas
Department of Bioengineering

The mesolimbic pathway connects the ventral tegmental area (VTA) to the nucleus accumbens (NAc) and has been extensively studied for its role in processing rewards, mediating reinforcement learning and promoting motivated behavior. The role of small molecule neurotransmitters such as dopamine (DA), glutamate and GABA has been extensively studied in the mesolimbic pathway within the context of reward and motivation. However, there are many opioid peptide systems that directly modulate neuronal populations within the mesolimbic pathway that have yet to be thoroughly investigated for their contribution to reward behavior. DA neurons in VTA highly express nociceptin opioid peptide receptors (NOPR) that are endogenously activated by the nociceptin ligand. Downstream of VTA DA neurons in NAc, there are two principal spiny projection neuron (SPN) cell types that express either excitatory dopamine receptors (D1R) or inhibitory dopamine receptors (D2R). D1R-SPNs also express and release the endogenous opioid dynorphin, while D2R-SPNs express another critical endogenous opioid, enkephalin. The presented studies investigate how these opioid-expressing neuronal populations are involved in facilitating reward behavior and motivation.

TABLE OF CONTENTS

Acknowledgements

Kyle Parker and Adrian Gomez were primary contributors to the content of Chapter 2 of this dissertation. Together, we were co-first authors on the related publication (Parker 2019) and in many cases they independently performed the experiments presented in the dissertation. Specifically, Kyle Parker designed and conducted the optogenetics and chemogenetics manipulation experiments in **Fig. 4-6 and Fig. S4-6**. Adrian Gomez conducted most of the initial Pnoc-cre line validation, tracing and anatomy in **Fig. 1 and Fig. S1**. Both Kyle and Adrian also contributed substantially to the writing of the manuscript (Parker 2019), from which the dissertation text in Chapter 2 is adapted.

Daniel C. Castro contributed heavily to experimental design for the content of Chapter 3. Specifically, he helped to conceive of the variable sucrose access behavior paradigm that is used in **Fig. 7, 10.** Daniel also independently designed and conducted retrotracing experiments presented in **Fig. S13.**

Michael R. Bruchas contributed heavily to experimental design for both Chapter 2 and 3. I would like to recognize his initial vision to investigate the nociceptin/NOPR system by having the Pnoc-cre and NOPR cKO mouse lines created. Michael was also the primary proponent for utilizing 2-photon calcium imaging to measure the activity of many genetically-identified neurons within deep brain structures.

We thank Dylan Blumenthal, Michelle Chung, Taylor Hobbs, and Carina Pizzano for animal colony maintenance. We thank the Bruchas lab, Stuber lab and NAPE Center for helpful discussions. **Funding:** C.E.P. was funded by NIH grant DA051124. D.C.C was funded by NIH grants NS007205, DA043999, DA049862, DA051489. M.R.B. was funded by NIH grants R37DA033396, P30DA048736, and the Mallinckrodt Endowed Professorship.

I would also like to extend a heartfelt thank you to Kyle Parker, Daniel Castro and Michael Bruchas for extensive, personalized mentorship in the lab – both in terms of fostering my intellectual growth into a mature scientist and for careful and patient bench mentorship.

Outside of Bruchas Lab mentors, I would also like to thank Tami Markovic, Jennifer Deem, Vijay Samineni and Jose Grajales-Reyes for inviting me to join them on exciting and impactful collaborations.

There were many great friends and family members along the way who kept me grounded when I was pulling long days and weekends (frequently) in the lab. In particular, I will value the friendship of Ian Berke and Raaj Gowrishankar tremendously as I reflect on my time spent in graduate school. Beyond all else, I will cherish the unconditional love of support of my girlfriend, Katie Ozgun, who was strong enough and cared enough to stick by my side as we were long-distance throughout the last 5 years.

List of Publications Contributed To

1. Arjun Ramakrishnan, Yoon Woo Byun, Kyle Rand, **Christian E. Pedersen**, Mikhail A. Lebedev, Miguel A.L. Nicolelis. Cortical neurons multiplex reward-related signals along with sensory and motor information. ***PNAS* 2017**.

2. Kyle E. Parker*, **Christian E. Pedersen***, Adrian M. Gomez*, Skylar Spangler, Bryan Copits, Shelly Feng, Dionnet Bhatti, Marie Walicki, Ream Al-Hasani, Thomas Jhou, Thomas J. Kash, Garret D. Stuber, Michael R. Bruchas. A paranigral VTA nociceptin circuit that constrains motivation for reward. ***Cell* 2019**.

3. Jennifer Deem, Chelsea L Faber, **Christian E. Pedersen**, Bao Anh Phan, Sarah A Larsen, Kayoko Ogimoto, Jarrell T Nelson, Vincent Damian, Megan A Tran, Richard D Palmiter, Karl J Kaiyala, Jarrad M Scarlett, Michael R. Bruchas, Michael W Schwartz, Gregory J Morton. Cold-induced hyperphagia requires AgRP-neuron activation in mice. ***eLife* 2020**.

4. Andrew T Luskin, Dionnet L Bhatti, Bernard Mulvey, **Christian E Pedersen**, Kasey S Girven, Hannah Oden-Brunson, Kate Kimbell, Taylor Blackburn, Abbie Sawyer, Robert W Gereau, Joseph D Dougherty, Michael R Bruchas. Extended amgydala-parabrachial circuits alter threat assessment and regulate feeding. ***Science Advances* 2021**.

5. Ream Al-Hasani*, Raajaram Gowrishankar*, Gavin Schmidt*, **Christian E. Pedersen**, Julia Lemos, Sofia Shirley, Abi Elerding, Sophie Renaud, Veronica Alvarez, Michael R. Bruchas. A long range GABAergic projection from VTA to NAc shell signals reward reinforcement. (Accepted, ***Nature Neuroscience***).

6. Daniel C Castro, Corinna S Oswell, Eric T Zhang, **Christian E Pedersen**, Sean C Piantadosi, Mark A Rossi, Avery Hunker, Anthony Guglin, Jose A Morón, Larry S Zweifel, Garret D Stuber, Michael R Bruchas. An endogenous opioid circuit determines state-dependent appetitive behavior. (In revision, ***Nature***).

7. Vijay Samineni, Bryan Copits, **Christian E. Pedersen**, Jose Grajales-Reyes, Jordan McCall, Michael R. Bruchas, Robert Gereau. Role of central amygdala gabaergic neurons in modulating itch behavior. (Accepted, ***eLife***).

8. Tamara Markovic, **Christian E. Pedersen**, Nicolas Massaly, Yvan Vachez, Kavitha Abiraman, Brian Ruyle, Jung Hoon Shin, Jennifer Garcia, Veronica Alvarez, Michael R. Bruchas, Meaghan Creed, Jose A. Morón. Pain induces somatic adaptations in ventral tegmental area dopamine neurons to drive anhedonia-like behavior. (In revision, ***Nature Neuroscience***).

9. Jose Grajales-Reyes, Vijay Samineni, Bryan Copits, **Christian E. Pedersen**, Michael R. Bruchas, Rob Gereau. Role of glutamatergic and gabaergic periaqueductal gray neuronal populations in persistent inflammatory pain. (In preparation).

10. **Christian E. Pedersen**, Daniel C. Castro, Madelyn M. Gray, Zhe C. Zhou, Shane Kan, Patrick Murphy, Patrick R. O'neill, Michael R. Bruchas. Medial accumbens shell spiny projection neurons encode relative reward value. (Impending submission).

CHAPTER 1: BACKGROUND

1.1 VTA DA neurons are activated during Reward and Cue-Reward Learning

As early as the 1950s, it was demonstrated that rodents would repeatedly press levers that would cause electrical stimulation of neurons in the midbrain or basal forebrain via implanted electrodes. It was found that animals would electrically self-stimulate their midbrains repeatedly if given the opportunity, even to the extent that they would ignore food rewards or mates in their environments (Olds 1954). It was also observed that after self-stimulation sessions, always would obsessively approach the lever and repeat the same operant lever pressing behavior, even when they were no longer receiving the stimulation (Olds 1958, Lovaas 1987). These early studies established that stimulation of the midbrain was inherently highly valued by the animals. They also demonstrated that stimulation imbued the associated cues and objects with their own incentive value that persisted long after the stimulation had stopped.

Later studies identified that this obsessive reward behavior was largely mediated by dopamine (DA) neurons in the ventral tegmental area (VTA) or neurons downstream of VTA DA neurons in the nucleus accumbens (NAc) (Fibiger 1987, Fiorino 1993). This large projection of DA neurons from VTA to NAc is known as the mesolimbic pathway and it has been extensively studied for its role in invigorating motivated reward seeking and in driving primary reinforcement from reward consumption. It was also demonstrated that pharmacological blockade of DA receptors caused extinction of intracranial self-

stimulation behavior in the midbrain, further proving that DA neurotransmission was a major component of how the mesolimbic pathway mediates behavioral reinforcement (Fouriezos 1976).

In the late 1950s, neuroscientists began recording the neural activity of individual neurons in living animals via single unit electrophysiology (Hubel and Wiesel 1959). This technique was applied to neurons in the VTA and revealed that VTA DA neurons were recruited when animals consumed primary rewards such as food or sucrose solution. It was also discovered that VTA dopamine neurons initially were not activated during the presentation of neutral cues but became increasingly activated to cues as animals learned that they predicted subsequent primary rewards (Schultz 1997). These neural recordings, combined with earlier self-stimulation experiments, gave rise to canonical models of how VTA DA neurons are activated during the expectation of rewards and are endogenously recruited to drive associative learning between cues and rewards (Schultz 2002, Berridge 1998, Flagel 2011, Berke 2018).

1.2 Role of Mesolimbic Dopamine in Motivation

DA release in NAc has been shown to occur during reward expectation and reward consumption, but in the natural environment reward-predictive cues are complex and continuous, rather than being a discrete time-locked event in time. In the real world, animals must engage in reward seeking behavior over sustained periods to attain reward payout. Examples of sustained reward seeking include foraging for food, pursuit of prey

or soliciting of mates. For this reason, mesolimbic DA has also been studied for its role in mediating motivated reward seeking behavior.

It has been found that DA is consistently released in NAc as animals are in pursuit of rewards throughout an operant task and that the magnitude of this tonic release increases as rewards become more certain and closer in space and time (Phillips 2003, Hamid 2016, Engelhard 2019). Suppressing the amount of DA transmission has also been shown to make animals lethargic and to reduce the amount of motivated reward seeking (Fouriezos 1976, Gerber 1981). Altogether, these many studies demonstrate that DA transmission in NAc is required for reward seeking behavior and is a key component in driving animal motivation.

Although DA has a critical role in driving motivated seeking behavior, it remains unresolved how neurons upstream of the VTA influence the activity of DA neurons to tune the appropriate amount of motivation. Efforts to address this enduring question have largely been focused on glutamatergic and GABAergic innervation from regions such as lateral habenula, rostromedial tegmental nucleus, lateral hypothalamus and bed nucleus of the stria terminalis (Geisler 2007, Jhou 2009, Stamatakis 2012, Jennings 2013). VTA DA neurons and local GABAergic interneurons in the VTA are also highly enriched in opioid receptors in addition to GABA and glutamate receptors. Specifically, VTA DA neurons express the nociceptin opioid peptide receptor (NOPR) (Norton 2002) and midbrain GABA neurons express the mu opioid receptor (MOR) (Johnson 1992, Margolis 2014). The persistent and robust neuromodulatory effects of these inhibitory opioid receptors could potentially have a major role in modulating tonic DA neuron activity in the VTA during motivated behavior. This is of particular interest given that changes in

motivation have been shown to result from shifts in the persistent and tonic ramp-up of DA neurons during reward seeking behavior.

1.3 Spiny Projection Neurons in Nucleus Accumbens

During primary reward and reward expectation, mesolimbic DA neurons release DA into the NAc. There are two principal spiny projection neuron (SPN) cell types in the NAc that are primarily defined by the expression of either excitatory dopamine receptors (D1R) or inhibitory dopamine receptors (D2R) (Castro 2019). Canonically, D1R-SPNs are excited by DA release in NAc while D2R-SPNs have their activity inhibited (Beaulieu 2011). D1R-SPNs also co-express and release dynorphin, the endogenous ligand for the kappa opioid receptor (Anderson 1990, Al-Hasani 2015). D2R-SPNs also co-express enkephalin, which is a critical endogenous ligand for MOR (Schiffmann 1993). Single unit electrophysiology recordings have also been conducted in NAc and have revealed that SPNs can either be excited or inhibited by primary rewards and in response to reward-predictive cues (Taha 2005, Roitman 2005). However, these in vivo neural recordings could not resolve the molecular identity of the recorded neurons or if the degree of excitation or inhibition during reward aligned with DA receptor expression. It remains unresolved whether DA release in NAc is a major factor in whether NAc SPNs are reward-excited or reward-inhibited.

It has also been shown that subregions of NAc may be functionally distinct in their role in reward behavior. SPNs in NAc core have a higher tendency to be responsive to reward predictive cues than in medial NAc shell (mNAcSh) (Day 2006). There is also

recent evidence that NAc core and mNAcSh have substantial differences in their dopaminergic innervation from VTA. Axon terminals from VTA DA neurons are not recruited in mNAcSh to reward-predictive cues (de Jong 2018). It has also been shown that optogenetic stimulation of VTA DA axons in mNAcSh during cue presentation does not imbue cues with incentive value, while it does in NAc core (Saunders 2018).

Altogether, recent evidence suggests that NAc SPNs likely have a complex and critical role in driving reward behavior that goes beyond mesolimbic dopamine transmission. Understanding how NAc SPNs diverge in their endogenous activity from canonical models of mesolimbic dopamine transmission during cues and rewards will likely provide new insight into how the larger limbic system processes reward information and invigorates motivated behavior.

CHAPTER 2: PARANIGRAL VTA PNOC+ NEURONS CONSTRAIN MOTIVATION

2.1 Chapter Introduction

Neural circuits that underlie reward seeking are regulated by homeostatic states and adapt to reward availability. Animals are less motivated to seek rewards in environments where resources are scarce (Salamone 2009, Kobayashi 2008). Persistence to seek uncertain rewards can be disadvantageous due to risky exposure to predators or from energy expenditure. To prevent aberrant seeking behavior, neural mechanisms constrain motivation to seek rewards. In humans, deficits within these regulatory processes can manifest into behavioral dysfunctions including depression, addiction, and feeding disorders (Russo 2013, Volkow 2011). In particular, disruption in

dopaminergic activity is implicated in the dysregulation of reward seeking and consumption (Morales 2017, Stauffer 2016, Salamone 2009). Evidence suggests heterogeneity of neuronal subtypes and anatomical localization in the VTA, including transmitter and neuropeptide systems that can modulate dopaminergic outputs (Morales 2017, van Zessen 2012, Tan 2012, Jhou 2009). However, how specific neuropeptides regulate motivation through the VTA remains an underexplored area.

Efforts to understand the anatomical basis and specificity of reward regulation by neuropeptides has provided converging evidence that presents the nociceptin (N/OFQ) / nociceptin opioid peptide receptor (NOPR) system as a reward-related neuropeptide system (Castro 2019, Der-Avakian 2017, Toll 2016, Ozawa 2015, Norton 2002, Zheng 2002, Mollereau 2000, Mollereau 1994). This opioid system has been proposed to have an endogenous role in regulating motivational state; however, how this regulation may occur remains unresolved. NOPR is broadly expressed in tyrosine hydroxylase (TH)-positive neurons (Norton 2002) and NOPR activation negatively regulates DA neuron activity (Zheng 2002), yet the behavioral consequences of this neuropeptide system's impact in the VTA are unknown. Anatomical studies highlight this system spans multiple reward-related brain nuclei including nucleus accumbens (NAc), striatum, bed nucleus of stria terminalis (BNST), central amygdala (CeA), hypothalamus, hippocampus, and the VTA (Maidment 2002, Ciccocioppo 2000, Mollereau 2000, Darland 1998, Anton 1996). Pharmacological evidence has suggested that NOPR stimulation can reduce reward-related behaviors, including food consumption and preference for drugs of abuse (Witkin 2014, Zaveri 2011, Zhao 2003, Kotlinska 2002, Ciccocioppo 2000). Thus, understanding the mechanisms by which this neuropeptide system modulates reward behavior provides

essential circuit-based mechanisms for endogenous VTA regulation, and how opioid peptides in this critical structure impact motivation.

To address a possible link between endogenous nociceptin systems and regulation of reward and motivation, we determined if VTA-projecting neurons expressing the precursor gene for nociceptin, prepronociceptin *(Pnoc*), are able to gate reward-related behaviors. First, we generated a prepronociceptin (*Pnoc*)-IRES-Cre knock-in mouse line, implemented retrograde tracing, and identified an enriched *Pnoc*+ neuronal nucleus within the border of the ventral VTA and interpeduncular nucleus (pnVTA) that projects into lateral portions of the VTA. Neural activity recordings revealed that VTA nociceptin neuron activity increases as mice reduce their effort to obtain rewards. This led to the hypothesis that pnVTA^{Pnoc} neurons may regulate motivation for rewards, and that the pnVTA^{Pnoc} neurons could bi-directionally alter motivation to seek rewards. Deletion and inhibition of pnVTA^{Pnoc} neurons dramatically increased breakpoints to obtain reward whereas optogenetic and chemogenetic stimulation of VTA *Pnoc* inputs decreased the breakpoints for rewards. We also report that reward seeking is reduced by nociceptin release and NOPR's action specifically on VTA DA neurons, while the deletion of VTA NOPR or DA neuron NOPR increased reward seeking. Collectively, our findings demonstrate a previously unknown population of pnVTA^{Pnoc} neurons positioned to constrain motivation for rewards via specific neuropeptide signaling on DA neurons in the VTA. The identification and interrogation of this pnVTA nociceptin system provides critical insight into the cell types that limit and regulate motivated behavior.

2.2 VTA Pnoc+ neurons synapse locally onto DA neurons

To virally target nociceptin-containing neurons that project into the VTA, we generated a *prepronociceptin*-IRES-Cre mouse (*Pnoc*-Cre). A targeted knock-in was used to generate a bi-cistronic transcript with an IRES to conserve *Pnoc* expression while providing Cre expression under control of the endogenous locus (see Taniguchi 2011) (**Figure S1A**). We crossed *Pnoc*-Cre mice to the tdTomato flox-stop reporter line (Ai9 and Ai14) (Madisen 2010) as well as the channelrhodopsin2-eYFP reporter line (Ai32) (**Figures 1A-B, see Methods**). These *Pnoc*-Cre^{tdTomato} and *Pnoc*-Cre$^{ChR2/eYFP}$ lines provided a map for *Pnoc*-expressing cells and axonal projections. Using these new mouse lines, we observed an abundance of *Pnoc*-containing cells within the VTA, BNST, CeA, septum, and DG (**Figures 1C-E, S1C-D**), demonstrating concordance with prior ISH data. We next validated the fidelity of the *Pnoc*-Cre mouse line by fluorescent *in situ* hybridization (FISH) co-labeling of *Pnoc* and Cre in known *Pnoc*-enriched regions. We observed strong Cre-fidelity in the *Pnoc*-Cre mouse line, where FISH shows that *Pnoc* and Cre substantial overlap in *Pnoc*+ mice in multiple brain areas (BNST, CeA, and VTA), whereas Cre is absent in wildtype littermate mice (**Figure S1C**). While the majority of these brain regions project to the VTA, it was not known if these projections are specifically *Pnoc*-expressing.

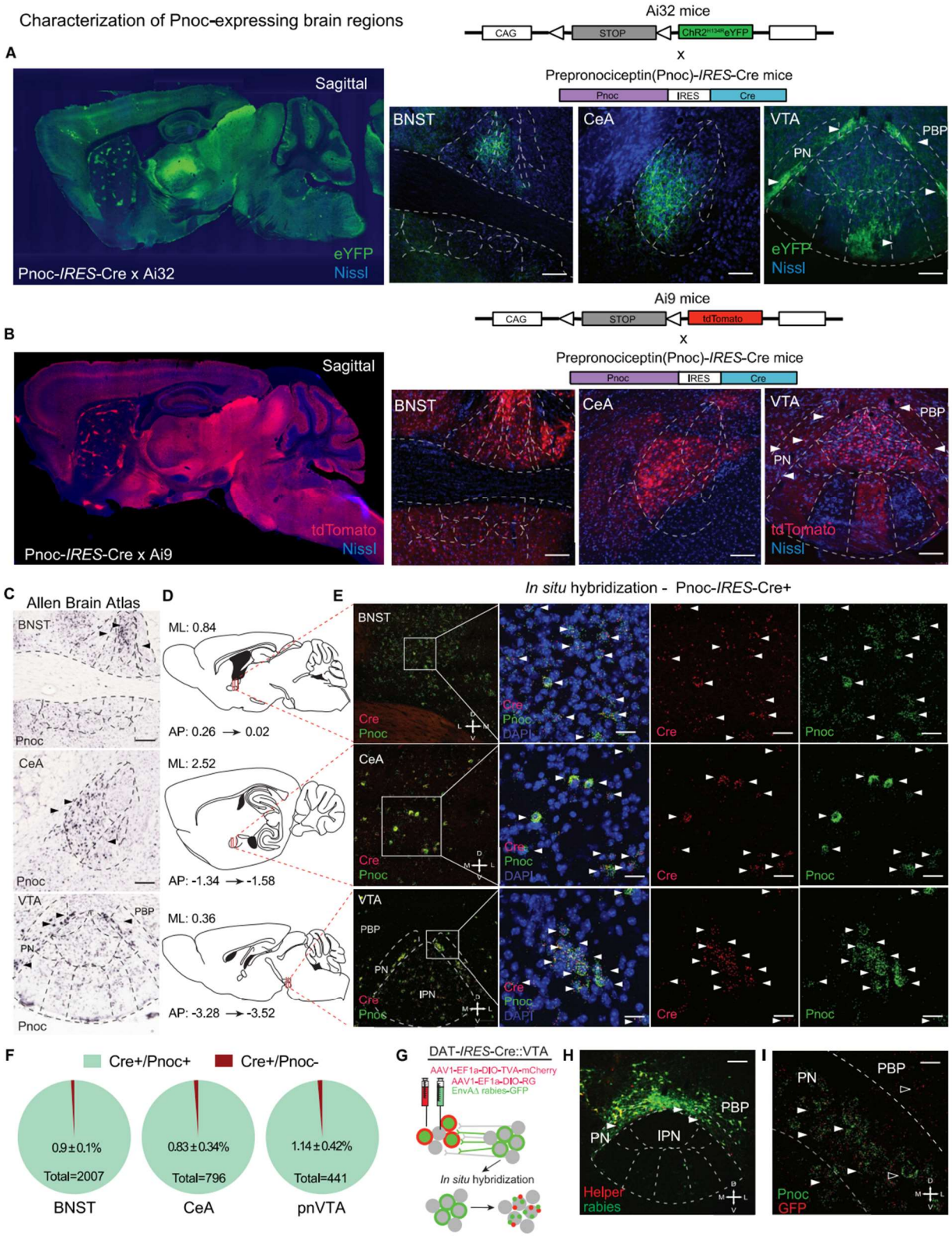

Characterization of Pnoc-expressing brain regions
A
Ai32 mice
CAG
STOP
ChR2H134REYFP
x
Prepronociceptin(Pnoc)-IRES-Cre mice
Pnoc
IRES
Cre
Sagittal
eYFP
Nissl
Pnoc-IRES-Cre x Ai32
BNST
CeA
VTA
PBP
PN
eYFP
Nissl
B
Ai9 mice
CAG
STOP
tdTomato
x
Prepronociceptin(Pnoc)-IRES-Cre mice
Pnoc
IRES
Cre
Sagittal
tdTomato
Nissl
Pnoc-IRES-Cre x Ai9
BNST
CeA
VTA
PBP
PN
tdTomato
Nissl
C
Allen Brain Atlas
BNST
Pnoc
CeA
Pnoc
VTA
PBP
PN
Pnoc
D
ML: 0.84
AP: 0.26 → 0.02
ML: 2.52
AP: -1.34 → -1.58
ML: 0.36
AP: -3.28 → -3.52
E
In situ hybridization - Pnoc-IRES-Cre+
BNST
Cre
Pnoc
DAPI
CeA
Cre
Pnoc
DAPI
VTA
PBP
PN
IPN
Cre
Pnoc
DAPI
F
Cre+/Pnoc+
Cre+/Pnoc-
0.9±0.1%
Total=2007
0.83±0.34%
Total=796
1.14±0.42%
Total=441
BNST
CeA
pnVTA
G
DAT-IRES-Cre::VTA
AAV1-EF1a-DIO-TVA-mCherry
AAV1-EF1a-DIO-RG
EnvAΔ rabies-GFP
In situ hybridization
H
PBP
PN
IPN
Helper
rabies
I
PN
PBP
Pnoc
GFP

Figure 1: Anatomical Identification of Endogenous *Pnoc*-expressing VTA Inputs

(A-B) Generation of the *Pnoc*-Cre^{tdTomato} and *Pnoc*-Cre$^{ChR2/eYFP}$ mouse lines from the cross between the *Pnoc*-Cre x Ai9-tdTomato or *Pnoc*-Cre x Ai32-ChR2/eYFP. Sagittal images of *Pnoc* labeling in *Pnoc*-Cre x Ai9 or Ai32 with coronal images depicting bed nucleus of stria terminalis (BNST), central amygdala (CeA), and paranigral ventral tegmental area (pnVTA). Images show tdTomato (red) or ChR2/eYFP (green) and Nissl (blue) staining. Scale bars are 100 μm. (C) Representative coronal images for *Pnoc in situ* hybridization from Allen Brain Institute of the BNST, CeA, VTA/IPN regions. All scale bars are 100 μm. (D) Sagittal atlas images depicting the location of BNST, CeA, and VTA regions used for quantification. (E) *Pnoc* and Cre expression patterns in *Pnoc*-Cre+ mice via *in situ* hybridization of *Pnoc* (green), Cre (red), and DAPI (blue) for coronal images of BNST, CeA, and pnVTA regions corresponding to panel D. Scale bars are 25 μm. (F) Quantification of *Pnoc* expression within Cre-expressing cells in the BNST, CeA, and pnVTA (n=4, 4 mice, 3 slices each). (G) Schematic depicting *in situ* hybridization for GFP (red) and Pnoc (green) mRNA following recombinant AAV helper viral and rabies viral injections into the VTA of DAT-Cre mice. (H) Coronal image depicting cre-dependent helper (red) and monosynaptic rabies (green) viral expression in the VTA / IPN regions of *DAT*-Cre mice. PBP-parabrachial pigmented nucleus, PN – paranigral VTA. Scale bars are 200 μm. (I) Coronal image depicting fluorescent *in situ* hybridization (FISH) of *Pnoc* (green) and GFP (red) colocalization in the pnVTA of *DAT*-Cre mice following recombinant AAV helper viral and rabies viral injections into the VTA. Scale bars are 25 μm.

To identify putative *Pnoc*+ VTA-projecting brain regions we utilized two retrograde tracing approaches. First, to survey potential projections interacting with DA neurons, we injected a mixture of AAV1-EF1α-DIO-TVA-mCherry: AAV1-EF1α-DIO-RG into the lateral VTA of *DAT*-Cre mice followed by an injection SADΔG-GFP (EnvA, rabies) 20 days later (Tian 2016, Watabe-Uchida 2012). Here, we identified multiple sources of monosynaptic inputs onto VTA DA neurons. Of particular interest was the dense labeling in ventral portions of the VTA, particularly in the paraintrafasciular (PIF) and paranigral nucleus (PN) (**Figure 1F, S1F**). This area showed robust GFP labeling without helper virus (mCherry), suggesting that local VTA *Pnoc* cells may connect directly onto VTA DA neurons as we observed robust *Pnoc* expression patterns within the paranigral VTA (**Figure 1A-C).** To confirm this observation, we performed *in situ* hybridization for *Pnoc*

and GFP following this rabies tracing approach (**Figure 1F)**. Here we visualized GFP and *Pnoc* co-labeling in the pnVTA (**Figure 1E, S1G**) suggesting *Pnoc*-expressing neurons have synaptic inputs onto DA (*DAT*-Cre+) cells. To further examine connectivity of this nucleus with the VTA, we injected a retroAAV2-EF1α-DIO-hChR2-(H134R)-eYFP (Tervo 2016), into the lateral VTA of *Pnoc*-Cre mice and visualized eYFP-labeled ChR2+ expressing neurons within the pnVTA. Additionally, we performed in situ hybridization for *Pnoc* and DAT and found that the majority of pnVTA^{Pnoc} neurons do not express *DAT* (**Figure S1J-K**). Together, these data support the conclusion that *Pnoc*-expressing neurons in the pnVTA synapse onto VTA DA neurons.

2.3 VTA Pnoc+ neurons are active during demotivation

To determine if VTA^{Pnoc} neurons are active during reward seeking, we recorded the intracellular calcium dynamics (Pologruto 2004) of VTA^{Pnoc} neurons and their terminals in the VTA. We used fiber photometry (Gunaydin 2014) while mice performed an operant task to receive sucrose reward (**Figure 2B, 2C)**. Mice were injected with 400 nl of AAV-DJ-EF1α-DIO-GCaMP6s unilaterally into the pnVTA and implanted with a 400mm optical fiber in the VTA (**Figure 2A**). Mice were food restricted and underwent Pavlovian conditioning to associate a house light (CS) with access to a sucrose sipper (US). Mice were trained on fixed ratio (FR) operant schedules to nosepoke for 10% sucrose under an FR1 and then under a FR3 schedule. Following FR3 testing, mice were put through a progressive ratio (PR) test, a task that directly measures the breakpoint at which an animal is unwilling to further work for reward (Hodos, 1961, Richardson and Roberts, 1996). Here, nosepoke criterion exponentially increases for each subsequent

reward during the task. Neural activity was recorded during the final FR3 day and the PR test day for two training cycles. Typical baseline drift was observed over the hour-long recording sessions and corrected using standard methods (Parker 2016, Gunaydin 2014) (**Figure S2A-C**). We found that *Pnoc*+ neural activity in the VTA increased phasically when mice were nosepoking in both FR3 and PR tasks (**Figure 2E**). Mice nosepoke at a consistent rate throughout the FR3 task, yet in the PR task, mice initially poke at a higher rate, and progressively reduce their poke rate as rewards became more difficult to attain (**Figure 2F**). Further, evidence shows that VTA DA cells are engaged during operant reward seeking and DA cells are particularly active when effort is high-yield and rewards are attained at a high rate (Gan, Walton, and Phillips, 2010, Hamid 2016, Oliva and Wanat, 2019).

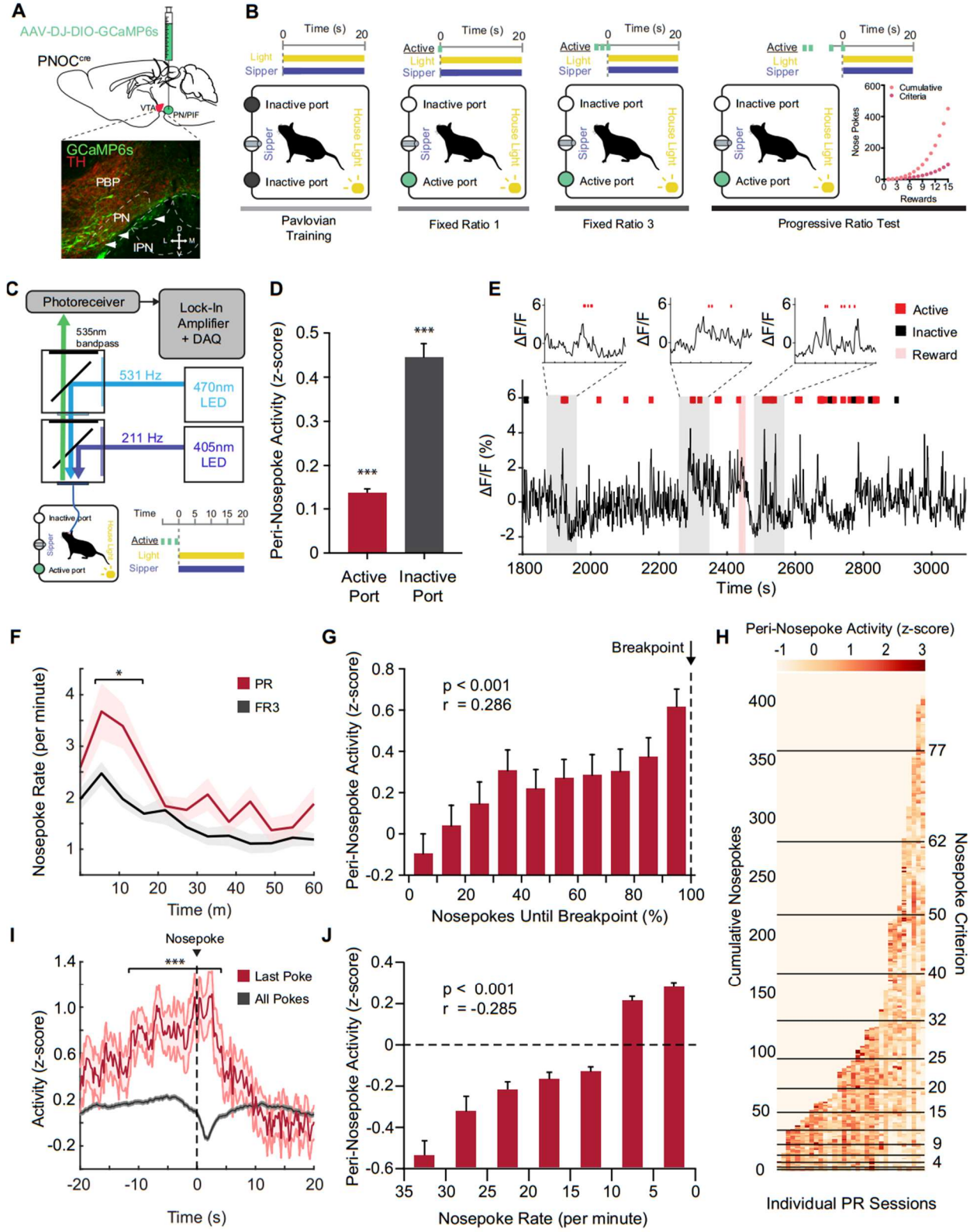

A
AAV-DJ-DIO-GCaMP6s
PNOC^cre
VTA
PN/PIF
GCaMP6s
TH
PBP
PN
IPN
B
Time (s)
Light
Sipper
Inactive port
Active port
House Light
Pavlovian Training
Fixed Ratio 1
Fixed Ratio 3
Progressive Ratio Test
Cumulative Criteria
Nose Pokes
Rewards
C
Photoreceiver
Lock-In Amplifier + DAQ
535nm bandpass
531 Hz
470nm LED
211 Hz
405nm LED
D
Peri-Nosepoke Activity (z-score)
Active Port
Inactive Port
E
ΔF/F
ΔF/F (%)
Active
Inactive
Reward
Time (s)
F
Nosepoke Rate (per minute)
PR
FR3
Time (m)
G
Breakpoint
p < 0.001
r = 0.286
Peri-Nosepoke Activity (z-score)
Nosepokes Until Breakpoint (%)
H
Peri-Nosepoke Activity (z-score)
Cumulative Nosepokes
Nosepoke Criterion
Individual PR Sessions
I
Nosepoke
Last Poke
All Pokes
Activity (z-score)
Time (s)
J
p < 0.001
r = -0.285
Peri-Nosepoke Activity (z-score)
Nosepoke Rate (per minute)

Figure 2: *Pnoc*+ VTA Inputs are Engaged During Low-Yield Reward Seeking
(A) Sagittal brain cartoon (top panel) of viral injection of GCaMP6s into the pnVTA of *Pnoc*-Cre+ mice and representative coronal image (bottom panel) showing immunohistochemistry for GCaMP6s and tyrosine hydroxylase (TH) staining. Images show GCaMP6s (green), and TH (red). Scale bars are 100 µm. (B) Cartoon schematic depicting training regimen for Pavlovian conditioning and operant task within operant box. (C) Cartoon schematic of fiber photometry setup for recording from mouse within operant box. (D) Neural activity during 2 second time bins centered around active or inactive nosepoke events (n = 18 mice, 65 sessions: One sample t-test, ****p<0.001*). (E) Representative recorded activity during PR operant task (red ticks = active nosepoke, black ticks = inactive nosepoke, red highlight = 20s reward period). (F) Mean nosepoke rate (active nosepokes per minute) over entire hour-long FR3 sessions and PR sessions (n = 18 mice, 65 sessions: Two sample t-test, **p=0.031*). (G) Mean peri-nosepoke activity relative to proportion of nosepokes until breakpoint for each PR session (n = 18 mice, 32 sessions: Pearson's correlation on uncategorized data points, ****p<0.001, r = 0.286*). (H) Individual 2s time bins of mean neural activity centered around active nosepoke events; shown for every active nosepoke during PR sessions (n = 18 mice, 32 sessions). Horizontal black lines indicate attainment of reward for given PR level. (I) Mean neural activity aligned to nosepoke times for all active nosepokes and the last active nosepoke (breakpoint) for PR sessions (n = 18 mice, 32 sessions: Two sample t-test, ****p<0.001*). (J) Mean peri-nosepoke activity relative to nosepoke rate (active nosepokes per minute) (n = 18 mice, 32 sessions: Pearson's correlation on uncategorized data points, ****p<0.001, r = -0.285*).

To determine how pnVTA^{Pnoc} neurons are active during low and high-yield reward seeking, we analyzed *Pnoc*+ neural activity during nosepoke epochs for different states of PR task performance. To quantify nosepoke-aligned *Pnoc*+ neural activity, mean fluorescence was assessed in 2-second time bins centered on nosepoke events. All peri-nosepoke activity is shown in **Figure 2H** for individual PR sessions, where each colored pixel represents mean neural activity for an individual active nosepoke. Peri-nosepoke activity was consistently above baseline for nosepokes into the active and inactive ports for the FR3 and PR tasks (**Figure 2D**). Assessing nosepoke-aligned activity during PR task performance revealed that activity was initially low during nosepokes early in the PR task when rewards were readily attainable. However, activity increased later in the PR session as more nosepokes were required for each subsequent reward. Interestingly,

activity was highest during animals' final nosepokes before reaching breakpoint and ceasing reward seeking (**Figure 2G, 2I, S2F**). Activity was also examined relative to the rate of nosepoking. Nosepoke rate was calculated as the total number of active nosepokes occurring within a 1-minute window centered around a given nosepoke event. Peri-nosepoke activity was at its minimum during high rates of nosepoking when mice were seeking rewards vigorously and increased as the poke rate diminished throughout the PR session (**Figure 2J**). Altogether, these findings indicate that *Pnoc*+ VTA inputs are engaged during reward seeking, and this activity is negatively correlated to the value of effort. Notably, *Pnoc*+ neural activity returned to baseline several seconds after an animal ceases its reward-seeking behavior (**Figure 2I**). These data support the conclusion that VTA *Pnoc*+ neurons are engaged during low-yield reward seeking, rather than active in the absence of reward anticipation.

We also analyzed the activity of VTA^{Pnoc} neurons during other behavioral epochs to offer a comprehensive evaluation of these unique neurons' responsivity. We found a divergence in VTA *Pnoc*+ neural activity during reward consumption between two distinct groups of mice (**Figure 3D and 3G**). Analysis of optical fiber placement with matched recordings revealed anatomically distinct neural activity during consumption. Mice with anterior optical fibers implanted over pnVTA (anterior to -3.5 AP) had increased activity during reward consumption in FR3 and PR tasks (**Figure 3E and 3F)** while mice implanted in posterior VTA (posterior to -3.5 AP) had decreased neural activity during reward consumption in FR3 and PR tasks (**Figure 3H and 3I**). However, analysis revealed that *Pnoc*+ VTA inputs for both groups of mice had increased activity as mice

reduced their rate of reward seeking (**Figure 2G and 2J**) and that there was no significant effect of fiber placement location and PR task progress in determining peri-nosepoke activity (**Figure S2E)**. This suggests VTA^{Pnoc} neurons and their fiber terminals in the anterior and posterior VTA are engaged similarly during nosepoking for reward but diverge in their neural activity during reward consumption.

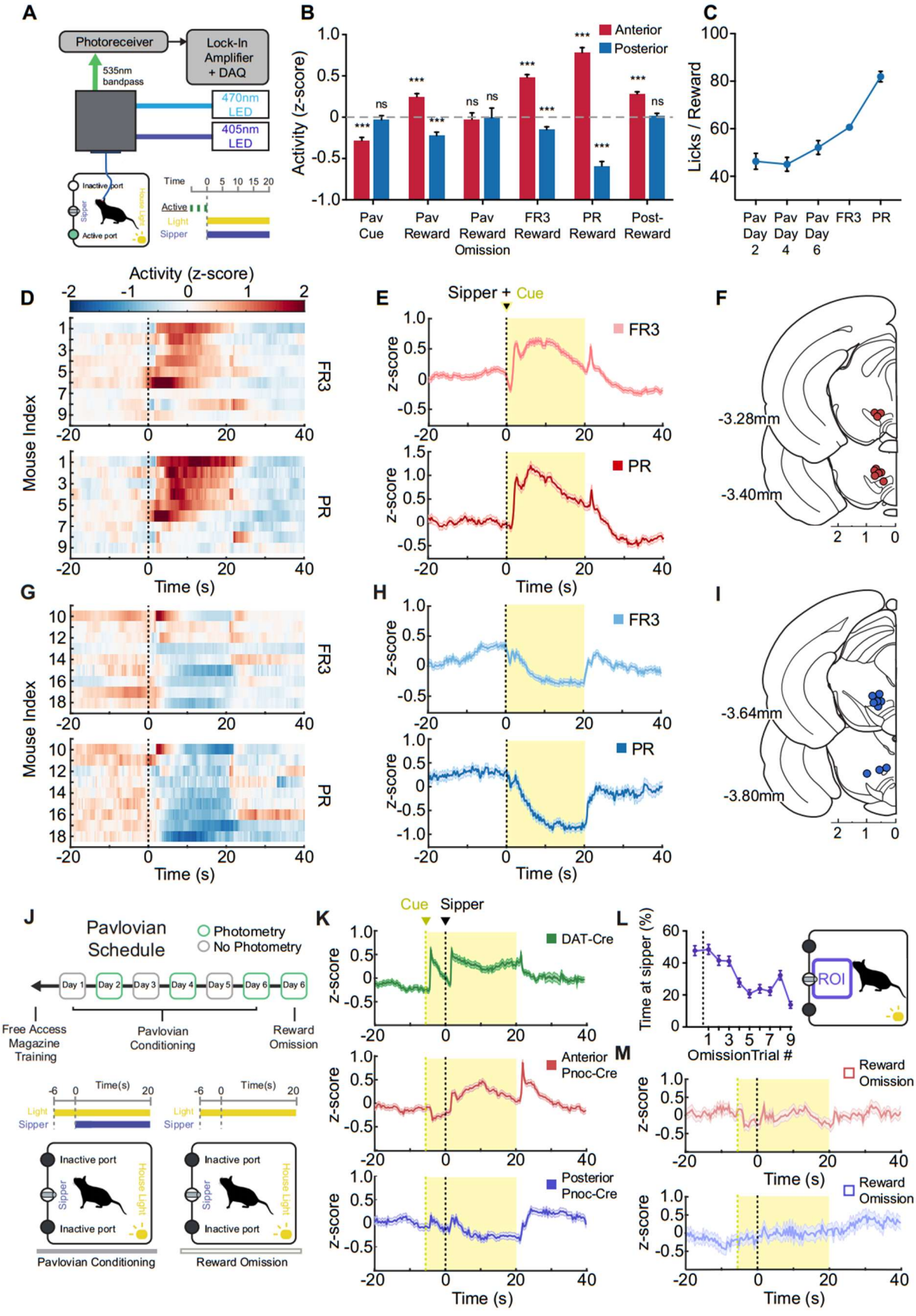

A
Photoreceiver
Lock-In Amplifier + DAQ
535nm bandpass
470nm LED
405nm LED
Inactive port
Sipper
Active port
House Light
Time
0 5 10 15 20
Active
Light
Sipper
B
Activity (z-score)
Anterior
Posterior
Pav Cue
Pav Reward
Pav Reward Omission
FR3 Reward
PR Reward
Post-Reward
C
Licks / Reward
Pav Day 2
Pav Day 4
Pav Day 6
FR3
PR
D
Activity (z-score)
Mouse Index
FR3
PR
Time (s)
E
Sipper + Cue
z-score
FR3
PR
Time (s)
F
-3.28mm
-3.40mm
G
Mouse Index
FR3
PR
Time (s)
H
z-score
FR3
PR
Time (s)
I
-3.64mm
-3.80mm
J
Pavlovian Schedule
Photometry
No Photometry
Day 1
Day 2
Day 3
Day 4
Day 5
Day 6
Day 6
Free Access Magazine Training
Pavlovian Conditioning
Reward Omission
Time(s)
Light
Sipper
Inactive port
Sipper
Inactive port
House Light
Pavlovian Conditioning
Reward Omission
K
Cue
Sipper
z-score
DAT-Cre
Anterior Pnoc-Cre
Posterior Pnoc-Cre
Time (s)
L
Time at sipper (%)
Omission Trial #
ROI
M
Reward Omission
z-score
Time (s)

Figure 3: Anatomically Distinct Subdivisions of pnVTA^{Pnoc} Neurons are Engaged During Reward Anticipation and Consumption

(A) Cartoon schematic of fiber photometry setup. (B) Summary of mean neural activity across various time intervals (intervals specified in Figure S3B) associated with reward anticipation and consumption during Pavlovian conditioning (n = 10 mice, 30 sessions) and operant task performance (n = 18 mice, 65 sessions; data presented as mean ± SEM; one sample t-tests: ****p<0.001*). (C) Number of licks per reward period during Pavlovian conditioning (n = 10 mice, 30 sessions) and operant task performance (n = 18 mice, 65 sessions; data presented as mean ± SEM). (D) Heatmap of neural activity averaged across every reward period during FR3 (top panel) and PR (bottom panel) tasks. Each row corresponds to an individual mouse with anterior VTA fiber placement (n = 9 mice, 17 FR3 sessions, 17 PR sessions). (E) Trace of neural activity averaged across every reward period during FR3 (top panel) and PR (bottom panel) tasks. Generated from all mice with anterior VTA fiber placement (n = 9 mice, 17 FR3 sessions, 17 PR sessions). (F) Coronal section showing anatomical location of 400μm optic fiber implant placement for mice with anterior VTA fiber placements. (G) Heatmap of neural activity averaged across every reward period during FR3 (top panel) and PR (bottom panel) tasks. Each row corresponds to an individual mouse with posterior VTA fiber placement (n = 9 mice, 16 FR3 sessions, 15 PR sessions). (H) Trace of neural activity averaged across every reward period during FR3 (top panel) and PR (bottom panel) tasks. Generated from all mice with posterior VTA fiber placement (n = 9 mice, 16 FR3 sessions, 15 PR sessions). (I) Coronal section showing anatomical location of 400μm optic fiber implant placement for mice with posterior VTA fiber placements. (J) Pavlovian conditioning schedule and cartoon depicting time course of house light (CS) and sipper access (US) during Pavlovian conditioning. (K) Trace of neural activity averaged across every reward period during Pavlovian conditioning. Generated from *Pnoc*-Cre mice with anterior VTA fiber placement (middle panel) (n = 5 mice, 15 sessions) and posterior fiber

placement (bottom panel) (n = 5 mice, 15 sessions). Trace of neural activity also shown for *DAT*-cre mice (top panel) (n = 7 mice, 14 sessions) for comparison of DA cell and *Pnoc* cell activity over identical Pavlovian conditioning paradigm. (L) Proportion of time spent near sipper (ROI) during reward omission periods. Data point left of dashed line represents the last rewarded trial on day 6 of Pavlovian conditioning (n = 10 mice, 10 sessions). (M) Trace of neural activity averaged across every reward omission period during day 6 of Pavlovian conditioning. Mice with anterior VTA fiber placement (top panel) (n = 5 mice, 5 sessions) and posterior fiber placement (bottom panel) (n = 5 mice, 5 sessions).

Further, VTA^{Pnoc} activity was also recorded in a Pavlovian conditioning paradigm (**Figure 3J**). The cue was presented 6 seconds before the sipper to allow for dissociation of *Pnoc*+ neural activity during reward anticipation from neural activity during reward consumption. A parallel group of *DAT*-Cre mice (n=7) were also recorded from in an identical Pavlovian paradigm to allow for direct comparison of DA cell activity and *Pnoc*+ cell activity within the VTA in this task. Anterior pnVTA^{Pnoc} cells increased their activity during reward consumption while posterior VTA^{Pnoc} inputs decreased their activity during reward consumption (**Figure 3K**). Additionally, anterior pnVTA^{Pnoc} cells had decreased activity in response to a reward-predictive cue while posterior VTA^{Pnoc} inputs showed no significant change from baseline during the cue period. Anterior pnVTA^{Pnoc} cells also had sharply increased activity after the end of the 20 second reward period for Pavlovian and operant sessions (**Figure 3E and 3K**). Reward omission trials were also tested with CS presentation, but no US access (**Figure 3J**). The consumption-related activity of anterior and posterior VTA^{Pnoc} inputs was absent during the omission trials (**Figure 3M**) despite mice remaining near the sipper port during extinction (**Figure 3L**). Lick rate during reward

was lowest during Pavlovian conditioning, higher for the FR3 task, and highest for the PR task (**Figure 3C**). For both anatomically and functionally distinct *Pnoc*+ VTA inputs, the intensity of fluorescence during reward corresponds to lick rate (Pearson's correlation: r = 0.849, p* = 0.033) and is negligible when mice are not consuming reward during omission trials (**Figure 3B**). Although anterior pnVTA^{Pnoc} cells decrease activity during reward anticipation, they did not exhibit any signal that parallels reward prediction error signals associated with reward omission (Schultz 1997, 2007, Hart 2014). However, these neurons are indeed recruited when access to sucrose reward is taken away during the post-reward period (**Figure 3B**). Altogether, these findings suggest that pnVTA^{Pnoc} cells act to limit seeking behavior, but do not likely represent error in reward anticipation.

2.4 VTA Pnoc+ neurons causally reduce motivation for reward

To determine the functional necessity of the pnVTA^{Pnoc} neurons in regulating natural reward-seeking behavior, we utilized three separate, but parallel approaches to assess if deletion or inhibition of pnVTA^{Pnoc} neurons impacts motivation for rewards. First, we injected an AAV5-FLEX-taCasp3-TEVp into the pnVTA of *Pnoc*-Cre+ and *Pnoc*-Cre- mice to induce apoptosis in Cre-positive pnVTA^{Pnoc} neurons. This caused a significant loss in pnVTA^{Pnoc} neurons of *Pnoc*+ mice (**Figure 4C-D, S4A**). We observed that pnVTA^{Pnoc} neuron ablation increased responding under the PR task showing increases in total nosepokes, rewards received, as well as robust increases in operant responding across the entire PR test (**Figure 4B,E-F, S4B-E**). These manipulations increased breakpoint and seeking for reward (FR3) but did not impact reward consumption (i.e. lick rate) (**Figure S4F**). Taken together, these data demonstrate that selective removal of

pnVTA^{Pnoc} neurons enhances PR responding for sucrose, implicating a necessary role for these unique pnVTA neurons in limiting motivation for rewards.

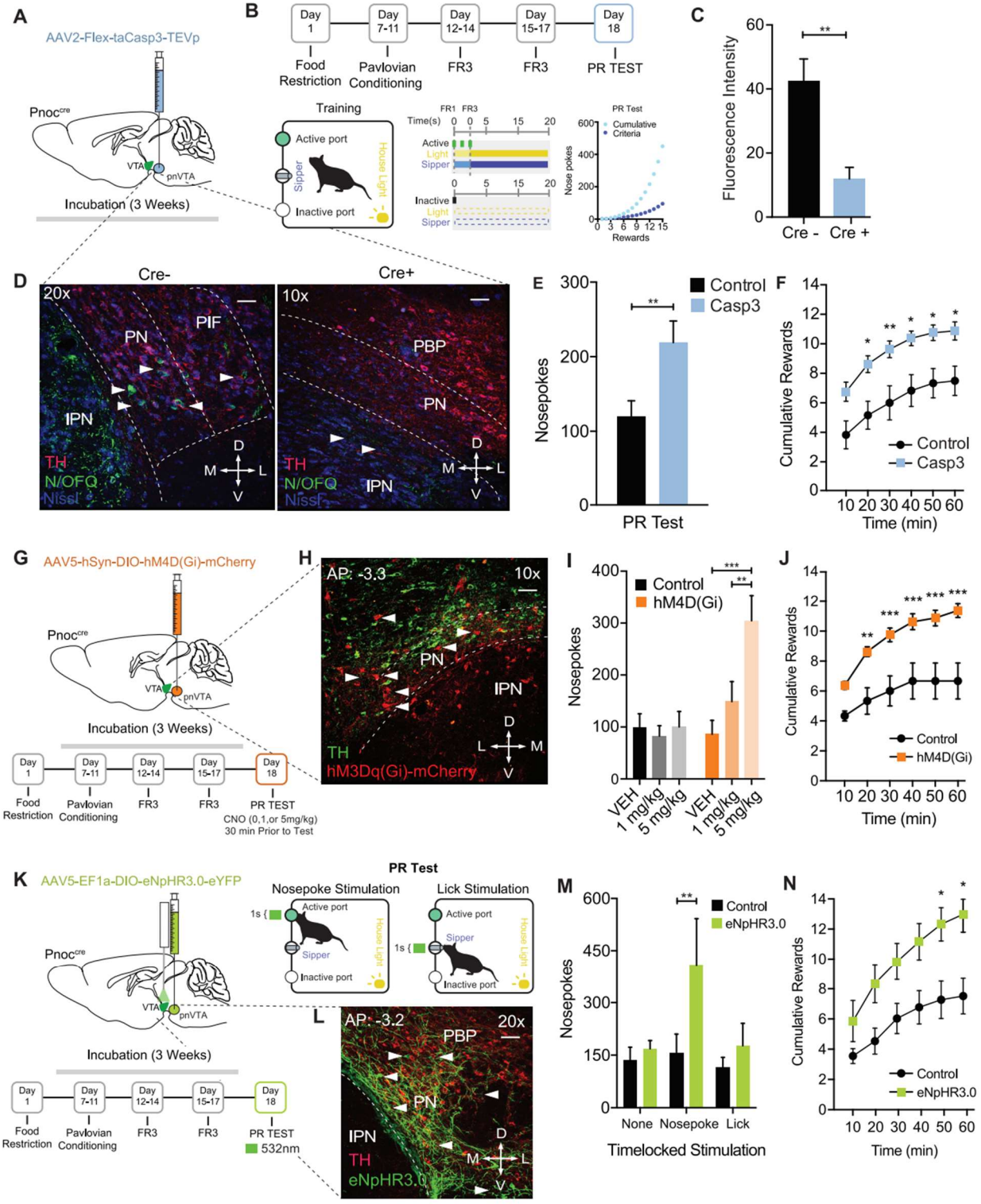

A
AAV2-Flex-taCasp3-TEVp
Pnoc^cre
VTA
pnVTA
Incubation (3 Weeks)
B
Day 1
Food Restriction
Day 7-11
Pavlovian Conditioning
Day 12-14
FR3
Day 15-17
FR3
Day 18
PR TEST
Training
Active port
Sipper
House Light
Inactive port
FR1 FR3
Time(s)
Active
Light
Sipper
Inactive
PR Test
Cumulative
Criteria
Nose pokes
Rewards
C
Fluorescence Intensity
Cre -
Cre +
D
Cre-
Cre+
20x
10x
PN
PIF
IPN
PBP
TH
N/OFQ
Nissl
D
M
L
V
E
Control
Casp3
Nosepokes
PR Test
F
Cumulative Rewards
Time (min)
G
AAV5-hSyn-DIO-hM4D(Gi)-mCherry
Pnoc^cre
VTA
pnVTA
Incubation (3 Weeks)
Day 1
Food Restriction
Day 7-11
Pavlovian Conditioning
Day 12-14
FR3
Day 15-17
FR3
Day 18
PR TEST
CNO (0,1,or 5mg/kg)
30 min Prior to Test
H
AP: -3.3
10x
PN
IPN
TH
hM3Dq(Gi)-mCherry
I
Control
hM4D(Gi)
Nosepokes
VEH
1 mg/kg
5 mg/kg
J
Cumulative Rewards
Time (min)
K
AAV5-EF1a-DIO-eNpHR3.0-eYFP
PR Test
Nosepoke Stimulation
Lick Stimulation
Pnoc^cre
VTA
pnVTA
Incubation (3 Weeks)
Day 1
Food Restriction
Day 7-11
Pavlovian Conditioning
Day 12-14
FR3
Day 15-17
FR3
Day 18
PR TEST
532nm
L
AP: -3.2
20x
PBP
PN
IPN
TH
eNpHR3.0
M
Control
eNpHR3.0
Nosepokes
None
Nosepoke
Lick
Timelocked Stimulation
N
Cumulative Rewards
Time (min)

Figure 4*:* Selective Ablation and Inhibition of pnVTA^{Pnoc} Neurons Enhances Operant Responding for Natural Rewards

(A) Cartoon depicting injection of AAV2-FLEX-taCasp3-TEVp into the pnVTA of *Pnoc*-Cre+ and *Pnoc*-Cre- mice. (B) Schematic and timeline for the operant training schedule, FR, and PR test days. (C) Quantification of fluorescence intensity from immunohistochemistry analysis for N/OFQ fluorescence in *Pnoc*-Cre+ and *Pnoc*-Cre- mice following cell ablation). (D) Representative coronal images of the VTA and IPN showing immunohistochemistry for nociceptin (N/OFQ) and tyrosine hydroxylase (TH) staining in *Pnoc*-Cre- (left panel) and *Pnoc*-Cre+ (right panel) mice. Images show N/OFQ (green), DAPI (blue) and TH (red). Scale bars are 50 and 100μm, respectively. (E) Nosepokes performed during PR test. *Pnoc*-Cre^{taCasp3} mice have significantly increased nosepokes and rewards, compared to controls (n = 11 to 13: two-way repeated-measures ANOVA, Bonferroni post hoc; Nosepokes for control versus Casp3 during PR test ***p < 0.01*. (F) Cumulative rewards received in Control and *Pnoc*-Cre^{taCasp3} mice. *Pnoc*-Cre^{taCasp3} mice show significantly increased number of rewards received, compared to controls (n = 11 to 13: two-way repeated-measures ANOVA, Bonferroni post hoc; Nosepokes for Control versus Casp3 during PR test **p < 0.05, **p<0.01*. (G) Calendar for injection of AAV5-hSyn-DIO-hM4D(Gi)-mCherry into the pnVTA of *Pnoc*-Cre+ and *Pnoc*-Cre- mice. Schematic and timeline for the operant training schedule and FR and PR test days. (H) Coronal images of the pnVTA and IPN showing immunofluorescence staining tyrosine hydroxylase (TH) and hM4D(Gi)-mCherry following viral injection in *Pnoc*-Cre+ mice. Images show TH (green) and hM4D(Gi)-mCherry (red). (I) Number of nosepokes performed during PR test. *Pnoc*-Cre$^{hM4D(Gi)}$ mice show significantly increased active nosepokes compared to Controls (n = 8: two-way repeated-measures ANOVA, Bonferroni post hoc; Nosepokes for Vehicle versus 5mg/kg, ***p < 0.001,* 1mg/kg versus 5mg/kg **p < 0.01.* (J) Cumulative rewards received in Control and *Pnoc*-Cre$^{hM4D(Gi)}$ mice. *Pnoc*-Cre$^{hM4D(Gi)}$ mice receive significantly more rewards over session, compared to controls (n = 8: two-way repeated-measures ANOVA, Bonferroni post hoc; Nosepokes for control versus hM4D(Gi) during PR test, **p < 0.01, ***p < 0.001*. (K) Calendar for injection of AAV5- EF1□-DIO-eNpHR3.0-eYFP into the pnVTA of *Pnoc*-Cre+ and *Pnoc*-Cre- mice. Schematic and timeline for the operant training schedule, FR and PR test days. (L) Coronal images of pnVTA immunofluorescence staining of tyrosine hydroxylase (TH) and eNpHR3.0-eYFP following viral injection in *Pnoc*-Cre+ mice. Images show TH (red) and eNpHR3.0-eYFP (green). (M) Number of nosepokes performed during PR test. *Pnoc*-Cre$^{eNpHR3.0}$ mice show significantly increased active nosepokes compared to Controls following nosepoke-paired inhibition (n = 6: two-way repeated-measures ANOVA, Bonferroni post hoc; **p<0.01*). (N) Cumulative rewards received in Control and *Pnoc*-Cre$^{eNpHR3.0}$ mice. *Pnoc*-Cre$^{eNpHR3.0}$ mice receive significantly more rewards over session compared to control mice following nosepoke-paired inhibition (n = 6: two-way repeated-measures ANOVA, Bonferroni post hoc; Nosepoke-paired inhibition Rewards for Control versus *Pnoc*-Cre^{eNpHR3} during PR test, **p<0.05*.

We injected virus encoding the inhibitory DREADD receptor (AAV5-hSyn-DIO-hM4Di-mCherry into the pnVTA of *Pnoc*-Cre+ and *Pnoc*-Cre- mice (**Figure 4G and 4H**) and tested mice in a PR task following the administration of two CNO doses: 1mg/kg (i.p.) and 5mg/kg (i.p.) (**Figure 4G**). We found that chemogenetic inhibition of pnVTA^{Pnoc} neurons significantly increased nosepoke responding following CNO (5 mg/kg) and increased the number rewards received following both CNO doses (1 and 5 mg/kg) (**Figure 4I, 4H, S4G and S4H**). However, the above approaches do not assess how inhibiting pnVTA^{Pnoc} neurons during specific epochs of the behavioral task might modulate the motivation to obtain natural rewards. Therefore, in a parallel series of experiments, we injected an AAV5-EF1α-DIO-eNpHR3.0-eYFP, and chronically implanted bilateral fiber optics into their terminal region of the lateral VTA of *Pnoc*-Cre^{+} and *Pnoc*-Cre^{-} mice (**Figure 4K-L**). We then assessed whether motivation to obtain sucrose rewards was altered by time-locked photoinhibition of pnVTA^{Pnoc} neurons during nosepoking or licking. Photoinhibition of pnVTA^{Pnoc} neurons was found to increase the total number of nosepokes and the number of total rewards during the PR task, but only when inhibition was paired to nosepoking behavior (**Figure 4M-N, S4I-J**). These data suggest that pnVTA^{Pnoc} neurons are necessary for operant responding for rewards, but not for consumption of a natural reward. Taken together, these data demonstrate that selective removal and inhibition of pnVTA^{Pnoc} neurons robustly enhances PR responding, implicating a key role for these neurons in regulating low-yield reward-seeking behavior.

To determine the sufficiency of pnVTA^{Pnoc} neurons in modulating sucrose reward seeking, we injected AAV5-EF1□-DIO-ChR2-eYFP, and chronically implanted a fiber optic into their terminal region of the lateral VTA of *Pnoc*-Cre+ and *Pnoc*-Cre- mice

(**Figure 5A-B, S5A-B**). During the PR task, mice received 60 minutes of 5Hz photo-stimulation or had the photo-stimulation time-locked (10Hz, 1 sec) to either their nosepoking or licking. We found that 5Hz stimulation of *Pnoc*+ VTA neurons and their terminals significantly decreased total nosepokes and rewards during FR3 and PR tests (**Figure 5C-F, Figure S5E-F**). We also determined whether the selective NOPR antagonist, J-113397, (3mg/kg, i.p.) would block this reduction in behavior. Indeed, antagonist administration prevented the photo-stimulated pnVTA^{Pnoc}-mediated suppression of sucrose seeking (**Figure 5G**), suggesting that tonic stimulation of pnVTA^{Pnoc} neurons releases NOPR agonist nociceptin. Time-locked stimulation did not affect sucrose seeking behaviors (**Figure S5C**), yet we found that nosepoke-paired stimulation led to either diminished (n = 7) or enhanced (n = 4) sucrose seeking (**Figure S5D**). These data suggest that while continuous, tonic unpaired pnVTA^{Pnoc} stimulation reduces PR performance, behavior-paired acute stimulation results in a divergent responding during the task, underscoring the additional complex features of this unique population within the pnVTA.

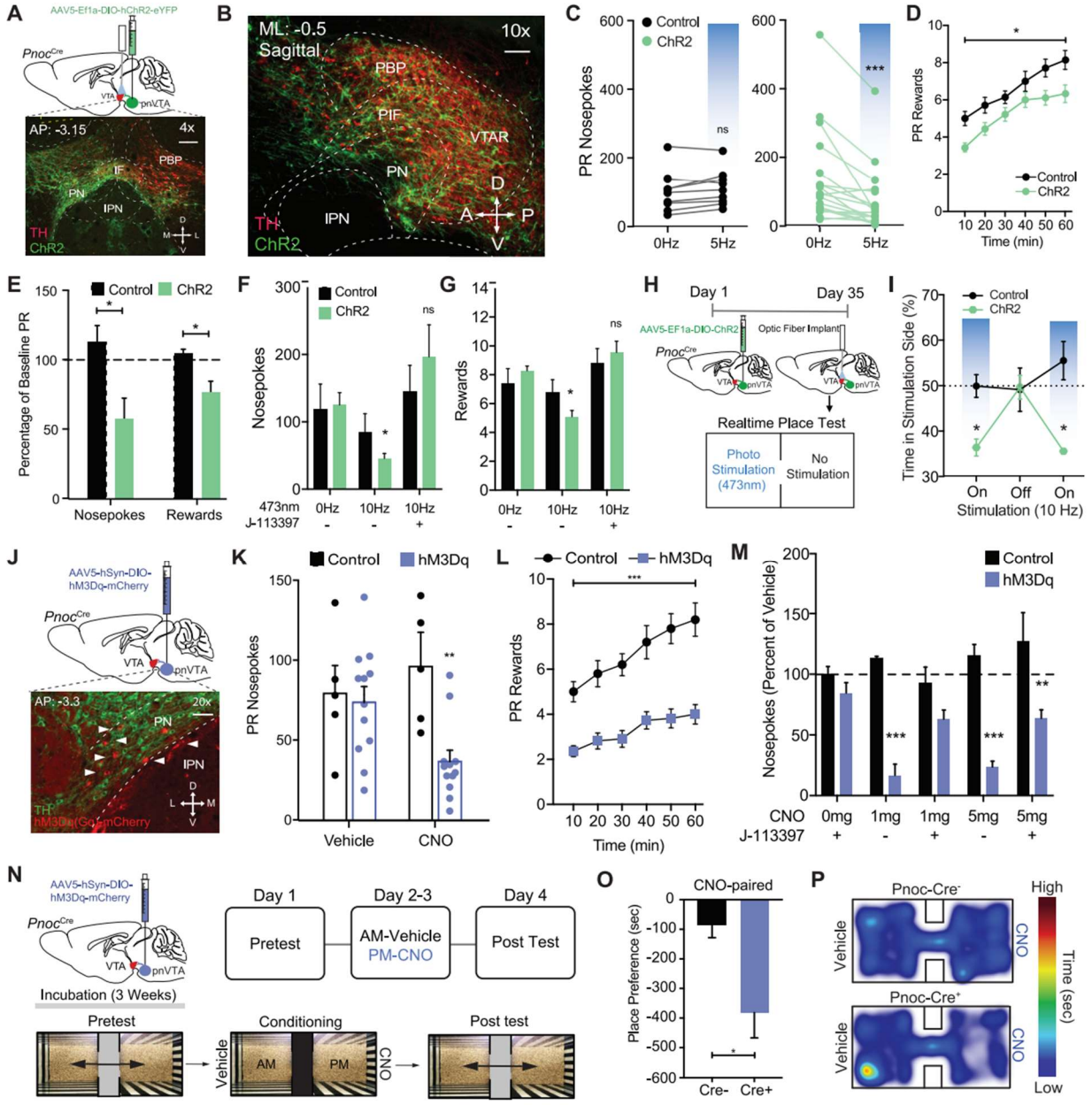


Figure 5: Optogenetic and Chemogenetic Stimulation of pnVTA^{Pnoc} Neurons Decreases Effort to Receive a Natural Reward and Promotes Aversion

(A) Sagittal brain cartoon (top) of viral injection of ChR2 into the pnVTA of *Pnoc*-Cre+ mice and representative sagittal image (bottom) showing immunohistochemistry for ChR2 (green) and tyrosine hydroxylase (TH) (red). (B) Calendar outlining timeline and stimulation parameters for operant training schedule during FR1, FR3, and PR tasks. (C) Number of nosepokes during PR test. *Pnoc*-Cre^{ChR2} mice decrease responding during continuous 5 Hz (10ms) blue light stimulation compared to controls (n=12: two-way repeated-measures ANOVA, Bonferroni post hoc;

Nosepokes for 0Hz versus 5Hz during PR test, ****p<0.001*). (D) Number of rewards received over time during PR test. *Pnoc*-Cre^{ChR2} mice show a significant reduction in the number of rewards received during the PR test received following 5Hz stimulation compared to controls (n=12: Student's t-test, **p<0.05*). (E) Percentage of baseline nosepokes and rewards during PR test. *Pnoc*-Cre^{ChR2} mice show a significant decrease in active nosepokes and number of rewards received during 5 Hz stimulation compared to baseline and controls (n=12: One way ANOVA, Bonferroni post hoc; Nosepokes for baseline versus 5Hz during PR test **p<0.05*, Rewards for baseline versus 5Hz during PR test **p<0.05*). (F) Number of nosepokes and (G) rewards during PR test in *Pnoc*-Cre^{ChR2} mice that received photostimulation (10Hz, 10ms) and the NOPR antagonist (J-113397, 3 mg/kg). *Pnoc*-Cre^{ChR2} mice show a significant decrease in active nosepokes and rewards compared to controls during photostimulation that was rescued by treatment with J-113397 (n=12: two-way repeated-measures ANOVA, Bonferroni post hoc; *Pnoc*-Cre^{ChR2} 0Hz versus 10 Hz, **p<0.05*, 0Hz versus 10 Hz + J-113397, *ns).* (H) Calendar outlining timeline for viral injection of ChR2 and VTA fiber optic implantation, followed by real time place testing. (I) Time-lock (ON-OFF-ON) patterns of photostimulation-induced avoidance. Percentage of total time spent receiving 10Hz stimulation in control and *Pnoc*-Cre^{ChR2} mice (n=9 to 12: two-way repeated-measures ANOVA, Bonferroni post hoc; percentage of time spent in stimulation side for control versus *Pnoc*-Cre^{ChR2} at 10Hz **p < 0.05*). (J) Sagittal brain cartoon (top panel) of viral injection of hM3D(Gq)-mCherry into the pnVTA of *Pnoc*-Cre+ mice and representative coronal image (bottom panel) showing immunohistochemistry for hM3D(Gq)-mCherry (red) and tyrosine hydroxylase (TH) (green). (K) Number of nosepokes during PR test. *Pnoc*-Cre$^{hM3D(Gq)}$ mice show a significant decrease in active nosepokes CNO (5mg/kg) administration compared to controls (data presented as mean ± SEM, n=5 to 12: two-way repeated-measures ANOVA, Bonferroni post hoc; Nosepokes for baseline versus CNO (5mg/kg) during PR test ***p<0.01*) Control mice show no effect in nosepokes or number of rewards received following CNO (5mg/kg) administration. (L) Number of rewards received over time during PR test. *Pnoc*-Cre$^{hM3D(Gq)}$ mice show a significant reduction in the number of rewards received during PR test received following CNO (5mg/kg) administration compared to controls (n=5 to 12: Student's t-test, ****p<0.001*). (M) Number of nosepokes (as a % of vehicle treatment) during PR test in *Pnoc*-Cre$^{hM3D(Gq)}$ mice that received CNO (1 and 5 mg/kg) and J-113397 (3 mg/kg). CNO treated mice show a significant decrease in nosepokes compared to controls that was partially rescued by treatment with J-113397 (n=5 to 12: two-way repeated-measures ANOVA, Bonferroni post hoc; Nosepokes for baseline versus CNO (5mg/kg) during PR test **p < 0.01*, Rewards for baseline versus CNO (5mg/kg) during PR test **p < 0.01*). (N) Timeline for pnVTA viral injection and detailed calendar outlining CPP experimental paradigm. (O) Amount of time spent on CNO-treated side during conditioned place post-test between *Pnoc*-Cre- and *Pnoc*-Cre+ mice. Post test reveals *Pnoc*-Cre$^{hM3D(Gq)}$ mice spend significantly less time in the side previously paired with CNO (5mg/kg). (n=5, 12 sessions: Student's t-test, Tukeys post-hoc, *****p<0.01*). (P) Representative heat map of relative time spent in CPP chambers for a *Pnoc*-Cre- and *Pnoc*-Cre+ mouse during CPP post-test following 5 mg/kg CNO treatment.

We next determined whether selective stimulation of pnVTA^{Pnoc} neurons causes a particular valence response (Namburi 2016). We injected AAV5-DIO-EF1□-ChR2-eYFP into the pnVTA of *Pnoc*-Cre+ or *Pnoc*-Cre- mice and used a real-time place test (RTPT) paradigm (**Figure 5H**) to determine whether mice prefer or avoid neutral contexts paired with various photostimulation frequencies (0-60hz) (Al-Hasani 2015). We found that photostimulation of pnVTA^{Pnoc} neurons (5Hz and 10Hz) produced significant real-time avoidance behavior (**Figure S5H-I**). This real time avoidance behavior is reversible, conserved following multiple test days, as well as within test (ON-OFF-ON), as demonstrated in **Figure 5I.**

In a parallel series of experiments, we determined the effects of stimulating pnVTA^{Pnoc} neurons via excitatory (hM3Gq) chemogenetic manipulation. We injected AAV5-hSyn-DIO-hM3Dq-mCherry into the pnVTA region of *Pnoc*-Cre+ and *Pnoc*-Cre- mice (**Figure 5J**) and tested mice in the PR task following CNO administration (1 and 5mg/kg, i.p.) We found that chemogenetic stimulation (5mg/kg CNO) of pnVTA^{Pnoc} neurons decreased operant responding and the number of rewards received during the PR task (**Figure 5K-L, Figure S5K-L**). In fact, we also found that CNO administration dose-dependently (1mg/kg and 5mg/kg) decreased the total number of nosepokes and the total number of sucrose rewards received in *Pnoc*-Cre$^{hM3D(Gq)}$ expressing mice (**Figure 5M, Figure S5M**). Both 1mg and 5mg CNO doses reduced operant responding and subsequent reward received only in *Pnoc*-Cre$^{hM3D(Gq)}$ expressing mice (**Figure 5G-6H**). Nosepoke responses and rewards for *Pnoc*-Cre$^{hM3D(Gq)}$ mice treated with CNO + the NOPR antagonist J-113397 returned to similar levels as compared to mice following vehicle administration (**Figure 5M, Figure S5M**). These results demonstrate a reversal

of CNO-induced reduction in PR responding and further indicate that stimulating pnVTA^{Pnoc} neurons releases nociceptin to act on NOP receptors.

We next determined whether stimulation of pnVTA^{Pnoc} neurons integrates salient negative affective information with associated environmental cues. Here we injected AAV5-hSyn-DIO-hM3Dq-mCherry in *Pnoc*-Cre+ and *Pnoc*-Cre- mice and measured conditioned place preference (CPP) for the context paired with DREADD agonist CNO (**Figure 5N**). We observed that repeated administration of CNO (5 mg/kg, i.p.) in a paired context induced a robust conditioned place aversion (CPA) only in *Pnoc*-Cre$^{hM3D(Gq)}$ mice (**Figure 5O-P**), without impacting locomotor activity, (**Figure S5N**). These data indicate that pnVTA^{Pnoc} neuron activation drives aversion in neutral environments and integrates negative valence with relevant contextual stimuli.

2.5 NOPR expression on DA neurons is required to constrain motivation

Our results above demonstrated involvement of both the neuropeptide nociceptin and NOPR in regulating sucrose seeking behavior; however, how and where these NOPR-mediated effects act is unknown. We next determined whether VTA NOPR expression is necessary and sufficient to coordinate this behavior. Using wildtype and NOPR knockout (KO) mice, we examined how selective NOPR agonist SCH-221510 administration impacts reward seeking within the PR task (**Figure 6A**). Here, we found that SCH-221510 administration (10mg/kg, i.p.) significantly reduced the number of nosepokes and rewards compared to baseline PR responding of wildtype mice, while having no significant effect in NOPR$^{-/-}$ mice (**Figure 6B**), complimenting previous experimental pnVTA^{Pnoc} manipulation results in **Figure 5**. Additionally, the administration NOPR agonist had no effect on the average lick rate during sucrose consumption (**Figure**

6C). This data indicates that exogenous NOPR stimulation is sufficient to reduce breakpoints for reward.

Next, to determine how and where NOPR mediates these effects on reward-seeking behavior, we recorded calcium transients in DA neurons in the VTA during pavlovian behavior following selective NOPR manipulation. In this experiment, *DAT*-Cre mice were injected with AAV-DJ-EF1α-DIO-GCaMP6s unilaterally into the VTA and implanted with optical fiber in the medial-lateral VTA (**Figure 6D**). Mice were trained in a Pavlovian conditioning task to associate house light presentation with sucrose reward. Mice were either injected with vehicle, SCH-221015 (10 mg/kg, i.p.), or J-113,397 (3 mg/kg, i.p.) prior to *in vivo* photometry recording of DA neuron activity during responding to cue and sucrose rewards (**Figure 6A, S6C-E**). NOPR activation significantly reduced overall baseline DA neural activity, while NOPR blockade increased baseline activity as compared to vehicle treated groups. This bidirectional effect was observed during consumption, but not during cue presentation (**Figure 6E-G**). Despite the effects on DA cell activity, there were no differences in lick rate or consumption of sucrose across treatments (**Figure S6D**). These data suggest that NOPR impacts DA neuron excitability during reward behavior.

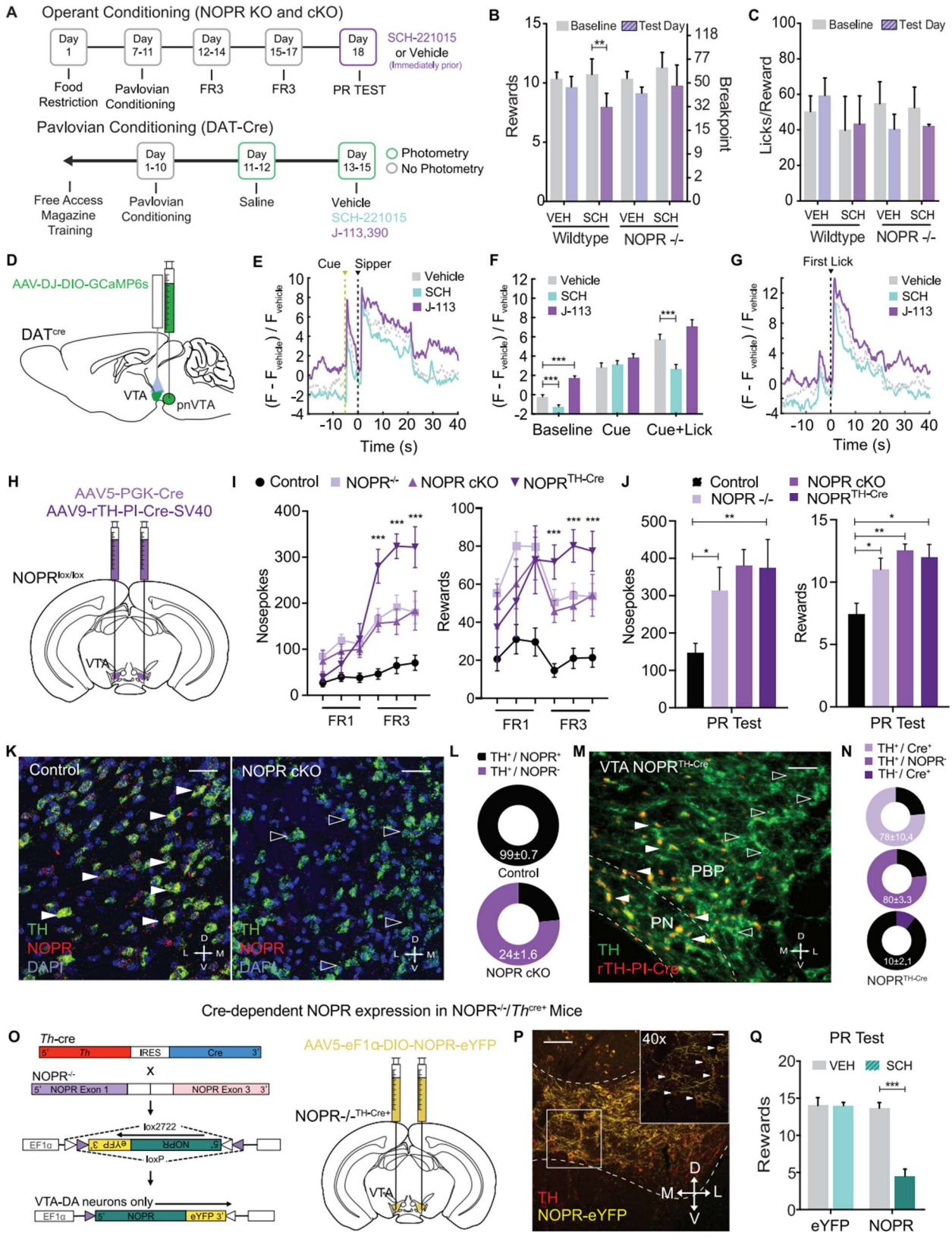

A
Operant Conditioning (NOPR KO and cKO)
Day 1
Day 7-11
Day 12-14
Day 15-17
Day 18
SCH-221015 or Vehicle
(Immediately prior)
Food Restriction
Pavlovian Conditioning
FR3
FR3
PR TEST
Pavlovian Conditioning (DAT-Cre)
Day 1-10
Day 11-12
Day 13-15
Photometry
No Photometry
Free Access Magazine Training
Pavlovian Conditioning
Saline
Vehicle
SCH-221015
J-113,390
B
Baseline
Test Day
Rewards
Breakpoint
VEH
SCH
Wildtype
NOPR -/-
C
Licks/Reward
D
AAV-DJ-DIO-GCaMP6s
DATcre
VTA
pnVTA
E
Cue
Sipper
Vehicle
SCH
J-113
Time (s)
F
Baseline
Cue
Cue+Lick
G
First Lick
H
AAV5-PGK-Cre
AAV9-rTH-PI-Cre-SV40
NOPRlox/lox
VTA
I
Control
NOPR-/-
NOPR cKO
NOPRTH-Cre
Nosepokes
Rewards
FR1
FR3
J
PR Test
K
Control
NOPR cKO
TH
NOPR
DAPI
L
TH+ / NOPR+
TH+ / NOPR-
99±0.7
24±1.6
M
VTA NOPRTH-Cre
PBP
PN
rTH-PI-Cre
N
TH+ / Cre+
TH+ / NOPR-
TH- / Cre+
78±10.4
80±3.3
10±2.1
Cre-dependent NOPR expression in NOPR-/-/Thcre+ Mice
O
Th-cre
IRES
Cre
NOPR-/-
NOPR Exon 1
NOPR Exon 3
lox2722
EF1α
eYFP
NOPR
loxP
VTA-DA neurons only
AAV5-eF1α-DIO-NOPR-eYFP
NOPR-/-TH-Cre+
P
40x
TH
NOPR-eYFP
Q
PR Test
eYFP
NOPR

Figure 6: VTA Dopamine Neuron NOPR Expression is Necessary and Sufficient for Regulating Motivation for Sucrose

(A) Calendar outlining timeline for operant training schedule in NOPR KO and cKO experiments or pavlovian training schedule for in vivo photometry recording of *DAT*-Cre experiments. (B) Number of rewards and licks per reward during PR test following vehicle and SCH221510 (10mg/kg) administration. Wildtype (WT) mice show a significant decrease in the number of rewards received following SCH221510 (10mg/kg) administration compared to $NOPR^{-/-}$ mice (n=6 to 8: two-way repeated-measures ANOVA, Bonferroni post hoc; Rewards for Baseline versus SCH221510 (10mg/kg) PR test $^{**}p < 0.01$). (C) Data depicting lick rate between vehicle and SCH221510 (10 mg/kg) treated WT and NOPR KO mice during PR testing. (n=6 to 8: two-way repeated-measures ANOVA, ns). (D) Cartoon representation of viral injection of GCaMP6s in the VTA of *DAT*-cre animals. (E) Comparison of mean DA cell dynamics across reward-predictive cue and sipper presentation in Pavlovian conditioning paradigm with systemic activation or blockade of NOPR (n=7 animals, 7 sessions per drug treatment). (F) Activity of DA cells during Pavlovian conditioning paradigm organized into 2-s time bins based on behavioral state of animal (n=7, 7 sessions per drug treatment; data represented as mean ± SEM; two sample t-tests: ***p<0.001). (G) Comparison of mean DA cell dynamics across first lick event within reward period of Pavlovian conditioning paradigm with systemic activation or blockade of NOPR (n = 7 , 7 sessions per drug treatment). (H) Cartoon depicting viral conditional knock out of NOPR in the VTA of $NOPR^{loxP/YFP}$ mice after bilateral injection of AAV5-PGK-Cre or AAV9-rTH-PI-Cre-SV40 into the VTA. (I) Number of nosepokes and rewards during FR1 and FR3 test sessions between control, NOPR KO, NOPR cKO, and $NOPR^{TH\text{-}Cre}$ mice. NOPR KO, NOPR cKO, and $NOPR^{TH\text{-}Cre}$ mice show a significant increase in the number of rewards received and total nosepokes in comparison to WT mice (n=6 to 8: two-way repeated-measures ANOVA, Bonferroni post hoc; Nosepokes and Rewards for WT versus NOPR KO, NOPR cKO, and $NOPR^{TH\text{-}Cre}$ FR1 and FR3 tests $^{***}p<0.001$). (J) Number of nosepokes and rewards during PR test sessions between control, NOPR KO, NOPR cKO, and $NOPR^{TH\text{-}Cre}$ mice. NOPR KO, NOPR cKO, and $NOPR^{TH\text{-}Cre}$ mice show a significant increase in the number of rewards received and total nosepokes in comparison to WT mice (n=6 to 8: two-way repeated-measures ANOVA, Bonferroni post hoc; Nosepokes and Rewards for WT versus NOPR KO and NOPR cKO PR tests, $^{***}p<0.001$). (K) Representative 40X coronal images for *in situ* hybridization showing colabeling of NOPR (red) and TH (green) mRNA in $NOPR^{lox/lox-}$ (left panel) and $NOPR^{lox/lox+}$ (right panel) mice that received injections of AAV5-PGK-Cre. (L) Pie chart graph depicting %Co-expression of TH and NOPR mRNA in the VTA of $NOPR^{lox/lox-}$ (top panel) and $NOPR^{lox/lox+}$ (bottom panel) mice that received injections of AAV5-PGK-Cre (n=4, 4 slices per n). (M) Representative 20X coronal images for immunofluorescence showing colabeling of Cre (red) and TH (green) in $NOPR^{lox/lox+}$ mice that received injections of AAV9-rTH-PI-Cre-SV40. (L) Pie chart graph depicting %Co-expression of TH and Cre (top panel) and %Co-expression of TH and NOPR (bottom panel) in the VTA of $NOPR^{lox/lox+}$ mice that received injections of AAV9-rTH-PI-Cre-SV40 (n=4, 4 slices per n). (O) Cartoon for viral rescue of NOPR in VTA dopamine cells of $NOPR^{-/-}$ X TH-Cre mice. (P) Representative image (10x) showing NOPR-eYFP infected VTA TH^{+}-cells. Images show TH (red) and NOPR-eYFP (yellow). Inset image depicts 40x image. Arrows indicate TH and NOPR-eYFP co-labeling. (Q) Number of rewards during PR test sessions following vehicle and SCH221510 (10mg/kg) administration. TH-Cre x $NOPR^{-/-}$ “rescue” mice show a significant decrease in the

number of rewards received following SCH221510 (10mg/kg) administration compared to NOPR$^{-/-}$ mice (n=8 to 10: two-way repeated-measures ANOVA, Bonferroni post hoc; Training: Rewards for NOPR-rescue baseline PR test versus SCH221510 (10mg/kg) PR test ****p<0.001*. No Training: Rewards for NOPR-rescue mice with Vehicle PR test versus SCH221510 (10mg/kg) PR test versus CNO (5mg/kg) during PR test ****p<0.001)*.

To determine the necessity of NOPR expression specifically within the VTA on motivation and reward-seeking behavior, we further examined these behaviors following VTA-specific conditional deletion of NOPR. We first generated an original conditional NOPR knockout mouse (**Figure S6A-B**) in order to selectively remove NOPR from regions and cell-types of interest. Here, we bilaterally injected 350nl of AAV5-PGK-Cre or AAV9-rTH-PI-Cre-SV40 virus (for selective DA neuron expression) into the VTA of NOPR$^{lox/lox}$ (labeled NOPR cKO and NOPR$^{TH\text{-}Cre}$, respectively) and control mice (**Figure 6H**). Using the same operant training and testing paradigms, we found that NOPR KO, VTA^{NOPR} cKO, and VTA NOPR$^{TH\text{-}Cre}$ cKO mice had significantly higher nosepokes and rewards received during operant responding as compared to littermate controls (**Figure 6I**). Similar to mice that have had pnVTA^{Pnoc} ablation or inhibition of pnVTA neurons (**Figure 4**), these NOPR KO, VTA^{NOPR} cKO, and VTA NOPR$^{TH\text{-}Cre}$ cKO mice showed a significant increase in responding for sucrose during the PR task (**Figure 6J**). Although, VTA NOPR cKO mice consumed more sucrose during the PR task, there was no difference in the average lick rate during reward consumption as a result of this manipulation (**Figure S6F-G**). Furthermore, to characterize this regional KO of NOPR, we used *in situ* hybridization for NOPR in VTA slices and found a significant reduction of NOPR mRNA expression in DA neurons (24±1.6%) in Cre-injected NOPR$^{lox/lox}$ mice, while leaving NOPR expression intact (99±0.7%) in control mice (**Figure 6K-L**). Additionally,

VTA NOPR$^{TH\text{-}Cre}$ cKO mice show a significant reduction in NOPR expression as revealed by immunofluorescence of Cre in DA neurons (97±0.6%) and NOPR expression in DA neurons (20.3±3.3%) in rTH-PI-Cre-injected NOPR$^{lox/lox}$ mice (**Figure 6M-N**).

To determine the sufficiency of NOPR in VTA DA neurons to produce these effects on reward seeking, we generated a mouse line of constitutive NOPR KO mice crossed to a *Th*-Cre mouse line (**Figure 6O**). We then constructed an AAV5-EF1□-DIO-NOPR-eYFP virus and injected it into the VTA of NOPR$^{-/-}$ x TH-Cre+ mice to express NOPR only within VTA DA neurons (**Figure 6O-P**). We then trained these mice as in previous experiments and examined how the selective NOPR agonist SCH-221510 (10 mg/kg, i.p.) impacted PR responding for sucrose in mice that express NOPR only in VTA DA neurons. Here we found that SCH-221510 significantly reduced responding in NOPR$^{-/-::NOPR\text{-}VTA}$ mice, while leaving control (AAV5-EF1α-DIO-eYFP) animal responding was unaffected (**Figure 6Q, S6I**). This selective VTA DA NOPR expression and agonist stimulation did not alter lick rate or impact locomotor activity (**Figure S6J-K).** These results indicate a critical role for endogenous nociceptin and NOPR activity on DA neurons. Together these data strongly support the conclusion that VTA NOPR activation constrains motivation to seek rewards.

2.6 Chapter Summary

We generated a *Pnoc*-Cre mouse line to specifically target and isolate *Pnoc*-expressing inputs into the VTA and map endogenous sources of nociceptin. Following examination of *Pnoc* expression patterns in two new *Pnoc* reporter lines (**Figure 1**) we identified numerous brain regions rich in *Pnoc*-expressing neurons. Considering our data

indicating monosynaptic inputs onto VTA DA neurons, we investigated multiple potential neurobiological sources of nociceptin inputs into the VTA, and found that the pnVTA not only expressed a distinct pattern of nociceptin neuron labeling, but was also comprised of likely dense inputs onto VTA DA neurons. Considering that the VTA and other ventral midbrain nuclei contain heterogeneous neuronal populations, no studies to date have demonstrated the function of this unique *Pnoc*-containing neuron population within the pnVTA region.

When examining the activity of pnVTA^{Pnoc} neurons during reward seeking conditions, we found they have relatively low activity when rewards are readily accessible and effort requirements are low during the PR task or during reward anticipation in Pavlovian conditioning tasks. These neurons are dynamically engaged during higher PR levels when mice are nosepoking infrequently for reward and not achieving sipper access (**Figure 2G**). In fact, we find that neural activity is highest when mice are reaching their breakpoint and terminate their reward-seeking behavior. Distinct activation of pnVTA^{Pnoc} neurons was also observed directly after the reward period when the sucrose sipper retracted and mice could no longer consume the reward. Although pnVTA^{Pnoc} neurons are very active when current reward access is removed, they surprisingly did not exhibit any significant signal associated with reward prediction errors that have been canonically observed in DA cells within the VTA. This is a surprising finding as DA cells phasically increase firing in response to reward predictive cues in anticipation of reward and are also dramatically suppressed when expected rewards are not received (Schultz 1997, Hart 2014). Together, the endogenous activity of pnVTA^{Pnoc} neurons indicates that they do not signal errors in reward prediction, but rather signal to limit the motivation for reward.

Indeed, mesolimbic DA is necessary for potentiating reward-seeking behavior when the effort required for reward is greater and the rate of reward attainment is lower, while DA activity is lower when future rewards are uncertain and the effort to reward payout ratio is greater (Gan, Walton, and Phillips, 2010, Hamid 2016). Here we report that the pnVTA^{Pnoc} neurons are increasingly active as nosepoke rate declines and likely inhibit VTA DA neurons during low-yield reward seeking via release of nociception and subsequent NOPR activation. It is possible that pnVTA^{Pnoc} neurons increasingly suppress DA neuron tonic activity to diminish and regulate seeking for less attainable rewards.

Given endogenous pnVTA^{Pnoc} activity during reward seeking, we hypothesized that these neurons contribute inhibitory tone during sucrose seeking to limit motivation. Indeed, we observed increased PR performance following chronic ablation or inhibition of pnVTA^{Pnoc} neurons. Interestingly, PR performance also increased when inhibition of pnVTA^{Pnoc} neurons was time-locked only to nosepoke events (**Figure 4K-N**). These data suggest that pnVTA^{Pnoc} neurons terminate seeking behavior particularly when the effort to obtain the reward increases. Accordingly, continuous pnVTA^{Pnoc} neuron stimulation reduced PR performance indicating that enhanced pnVTA^{Pnoc} activity is sufficient to diminish the pursuit of natural rewards, while NOPR antagonist administration reversed this effect. It was uncertain whether this loss in motivation is due to the promotion of a negative affective state or the enhancement of an existing inhibitory tone that coordinates the intersection of reward value and effort. pnVTA^{Pnoc} neuron stimulation resulted in active avoidance during real-time stimulation and mice avoided stimulation-paired environmental stimuli within the CPP assay. These data are indicative of a neuronal population that also promotes a negative affective state. Interestingly, this data parallels

other studies that demonstrate that mice avoid VTA DA neuron inhibition (Tan 2012, Danjo 2014). pnVTA^{Pnoc} neurons are likely positioned to receive input related to affective stimuli that subsequently drive inhibitory tone onto local VTA DA neuron populations, which would translate into reduced motivation in reward contexts and a negative affective state in non-reward contexts.

Recent studies have demonstrated that monoamine nuclei including the VTA, dorsal raphe, and locus coeruleus have a diverse host of neuronal populations based on their anatomical localization and projection target (Matthews 2016, Seo 2016, Beier 2015). However, prior studies (Morales 2017) have yet to determine the connectivity and/or function of critical neuropeptide producing neurons in the VTA. Further investigation of VTA neuropeptide and neurotransmitter heterogeneity using intersectional genetic methods will elucidate how the VTA coordinates complex inputs to shape motivated behavior.

Given the concomitant intersection of results that indicate a unique pnVTA^{Pnoc} population regulates reward seeking by nociceptin release, we further determined whether there is a critical, necessary role for NOPR within the VTA in maintaining normal reward seeking. Global NOPR KO and conditional NOPR deletion from VTA DA cells dramatically enhanced reward-seeking behaviors in a manner similar to pnVTA^{Pnoc} neuron ablation and inhibition (**Figure 4**). This data indicates that, along with our electrophysiological and tracing studies, the action of pnVTA^{Pnoc} neurons on reward seeking is mediated through NOPR activation on dopaminergic VTA neurons. Interestingly, fiber photometry recordings of VTA DA activity during Pavlovian conditioning revealed that agonist activation of NOPR suppressed baseline DA neural

activity. Phasic DA activity in response to the reward-predictive cue was left largely intact during NOPR activation while persistent DA activity during reward consumption was suppressed. This suggests that VTA NOPR activation may selectively limit PR performance by dramatically shifting dopaminergic tone. This finding is further supported by pnVTA^{Pnoc} manipulation affecting effort for reward, but not the rate of reward consumption. Lastly, we demonstrated that the conditional rescue and stimulation of NOPR specifically in VTA DA neurons greatly diminishes breakpoint responding. These data demonstrate the effectiveness of NOPR stimulation in inhibiting seeking for sucrose when it is exclusively expressed and stimulated in VTA DA neurons. Collectively, these data support a unique role of intra-VTA nociceptin release and subsequent NOPR activation acting to dynamically constrain DA neuron activity during reward-seeking behavior. In our case, an absent VTA nociceptin or NOPR system allows higher breakpoints for obtaining rewards in which a normally functioning NOPR system would typically engage and regulate appropriate reward seeking under various environmental conditions through DA neuron inhibition.

Our findings present a previously unknown pnVTA cell population of nociceptin-containing neurons that are positioned to suppress reward seeking via localized peptidergic action on NOPR-expressing VTA dopamine neurons. Understanding how this discrete pnVTA nucleus is critical to the regulation of motivation provides essential insights into the diverse behavioral functions of the VTA. How local VTA regions and neuropeptides are functionally defined at the circuit, cellular, and molecular level reveals insights into how the VTA coordinates homeostatic motivational balance during dysregulated motivational states like depression and addiction.

2.7 Supplementary Information

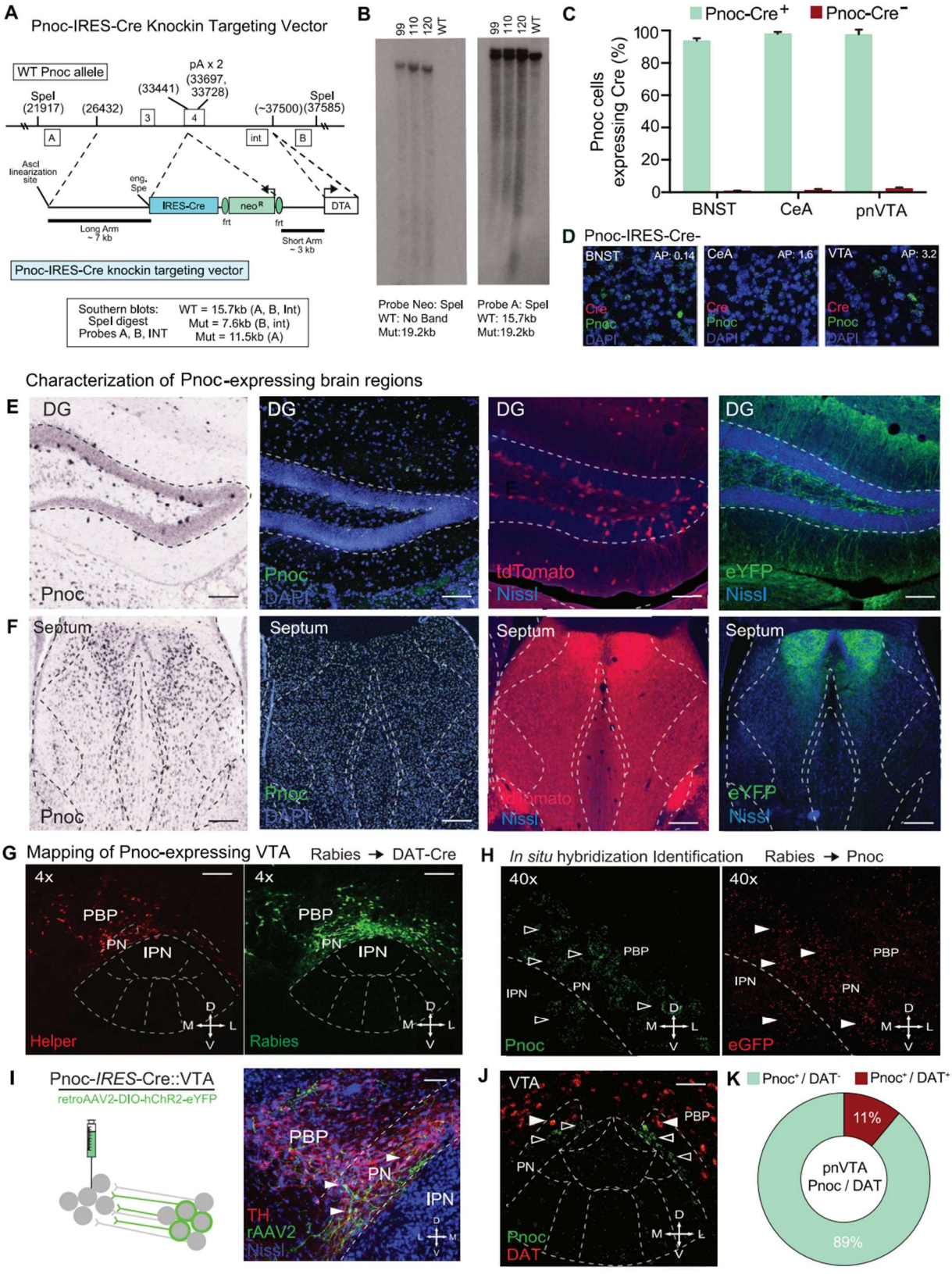

A
Pnoc-IRES-Cre Knockin Targeting Vector
WT Pnoc allele
SpeI (21917)
(26432)
(33441)
pA x 2 (33697, 33728)
SpeI (~37500)(37585)
3
4
int
A
B
AscI linearization site
eng. Spe
IRES-Cre
neoR
DTA
frt
frt
Long Arm ~ 7 kb
Short Arm ~ 3 kb
Pnoc-IRES-Cre knockin targeting vector
Southern blots: SpeI digest Probes A, B, INT
WT = 15.7kb (A, B, Int)
Mut = 7.6kb (B, int)
Mut = 11.5kb (A)
B
99
110
120
WT
Probe Neo: SpeI
WT: No Band
Mut:19.2kb
Probe A: SpeI
WT: 15.7kb
Mut:19.2kb
C
Pnoc-Cre+
Pnoc-Cre−
Pnoc cells expressing Cre (%)
100
80
60
40
20
0
BNST
CeA
pnVTA
D
Pnoc-IRES-Cre-
BNST
AP: 0.14
CeA
AP: 1.6
VTA
AP: 3.2
Cre
Pnoc
DAPI
Characterization of Pnoc-expressing brain regions
E
DG
Pnoc
DAPI
tdTomato
Nissl
eYFP
F
Septum
G
Mapping of Pnoc-expressing VTA
Rabies → DAT-Cre
4x
PBP
PN
IPN
Helper
Rabies
H
In situ hybridization Identification
Rabies → Pnoc
40x
eGFP
I
Pnoc-IRES-Cre::VTA
retroAAV2-DIO-hChR2-eYFP
TH
rAAV2
Nissl
J
VTA
DAT
K
Pnoc+ / DAT-
Pnoc+ / DAT+
11%
pnVTA
Pnoc / DAT
89%

Figure S1. Related to Figure 1: Anatomical Identification of Endogenous *Pnoc*-expressing VTA Inputs

(A) Schematic of targeting vector for generation of *Pnoc*-Cre mouse line. (B) Representative southern blot of *Pnoc*-Cre knock-in. (C) Quantification of Cre expression within *Pnoc*-expressing cells in the BNST (1988/2122 cells), CeA (786/805 cells), and pnVTA (435/450 cells). (n=4, 4 mice, 3 slices each). (D) Confirmation and validation of *Pnoc*-Cre mouse line depicting coronal images for *Pnoc* in *Pnoc*-Cre- mice for *in situ* hybridization colocalization of *Pnoc* (green) and Cre (red) for coronal images of BNST, CeA, and IPN/VTA regions. (E) Confirmation of *Pnoc*-labeling within coronal images of the dentate gyrus (DG) of the hippocampus from left to right of: *Pnoc* mRNA labeling from mouse Allen Brain Atlas, RNAscope in situ hybridization of DAPI (blue) and *Pnoc* mRNA (green), *Pnoc* (red) labeling in *Pnoc*-Cre x Ai9-tdTomato against Nissl (blue) immunostaining, and *Pnoc* (green) labeling in *Pnoc*-Cre x Ai32-ChR2/eYFP against Nissl (blue). (F) Confirmation of *Pnoc*-labeling within coronal images of the septum from left to right of: *Pnoc* mRNA labeling from mouse Allen Brain Atlas, RNAscope in situ hybridization of DAPI (blue) and *Pnoc* mRNA (green), *Pnoc* (red) labeling in *Pnoc*-Cre x Ai9-tdTomato against Nissl (blue) immunostaining, and *Pnoc* (green) labeling in *Pnoc*-Cre x Ai32-ChR2/eYFP against Nissl (blue). (G) Coronal images depicting cre-dependent helper (red channel, left panel) and monosynaptic rabies (green channel, right panel) viral expression in the VTA / IPN regions of *DAT*-Cre mice corresponding to Figure 1F. PBP- parabrachial pigmented nucleus. Representative image from n=3. (H) Coronal images that correspond to Figure 1H showing GFP (red, right panel) and *Pnoc* (green, left panel) mRNA following recombinant AAV helper viral and rabies viral injections into the VTA of *DAT*-Cre mice PN-paranigral, PBP- parabrachial pigmented nucleus. Representative images from n=3. (I) Cartoon depicting recombinant retrograde AAV2 virus injection into the VTA of Pnoc-Cre mice. Coronal image depicts GFP (green) Pnoc+ cells in the paranigral (PN) region of the VTA against DAPI (blue) and tyrosine hydroxylase (red) immunostaining. PBP- parabrachial pigmented nucleus. Representative images taken from n=3. (J) *In situ* hybridization colocalization of *Pnoc* (green) and *DAT* (red) for coronal images of zoomed VTA regions. Closed arrows represent colocalization of *Pnoc* and *DAT*, and empty arrows represent Pnoc labeled cells that are absent of *DAT*. (K) Pie chart depicting quantification of *Pnoc* and *DAT* colocalization within the paranigral VTA (pnVTA) (n=3).

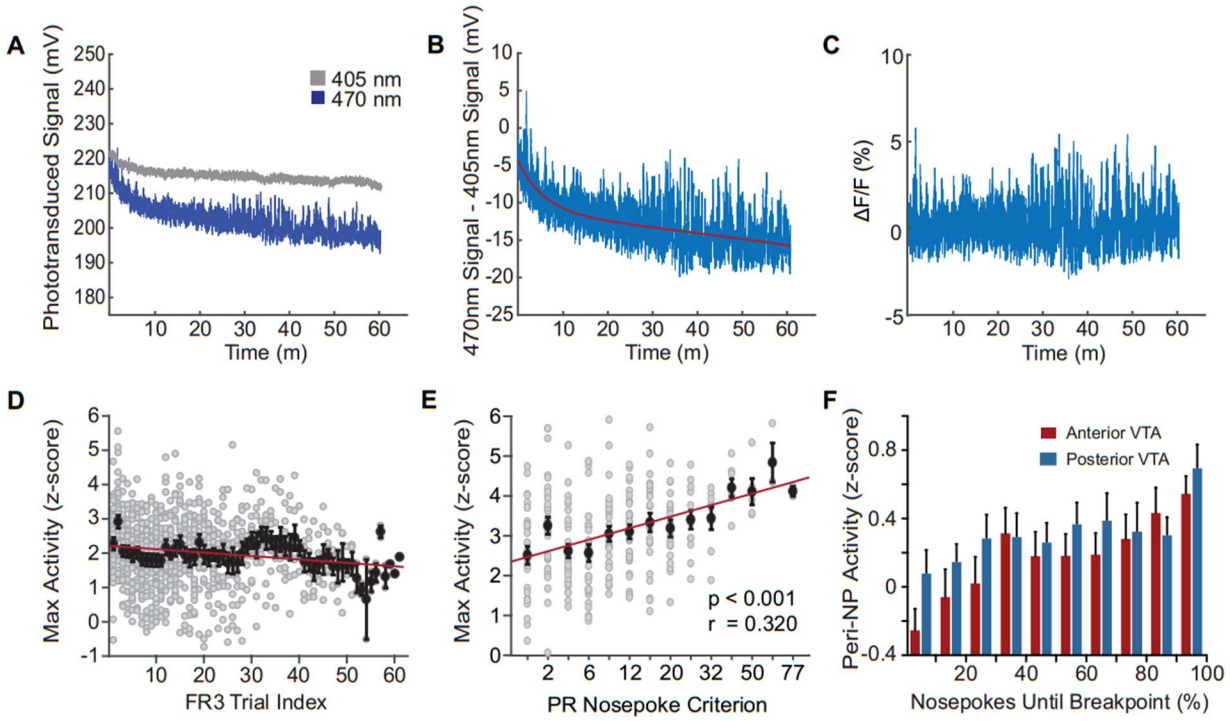


Figure S2. Related to Figure 2: *Pnoc*+ VTA Inputs are Engaged During Low-Yield Reward-Seeking

(A) Raw transduced photon emissions from excitation with 405 nm light and 470 nm light. (B) Double exponential curve (red line) fit to signal after subtraction of 405 nm channel from 470 nm channel. (C) Resulting baseline-corrected signal after subtraction of double exponential curve. (D) Maximum activity (maximum 0.5 seconds of activity prior to reward epoch) for each FR3 level (n = 18 mice, 33 sessions: Pearson's correlation, *p = 0.033, r = -0.073). (E) Maximum activity (maximum 0.5 seconds of activity prior to reward epoch) relative to nosepoke criterion for each PR level (n = 18 mice, 32 sessions: Pearson's correlation, ***p<0.001, r = 0.320). (F) Mean neural activity during active nosepoking (see Figure 2H) relative to proportion of nosepokes until breakpoint for each PR session. Same data as Figure 2G, but split based on anatomical location of fiber implant within VTA (see Figure 3). (n = 18 mice, 32 sessions; Two-way repeated measures ANOVA: interaction, F = 0.49, p = 0.883).

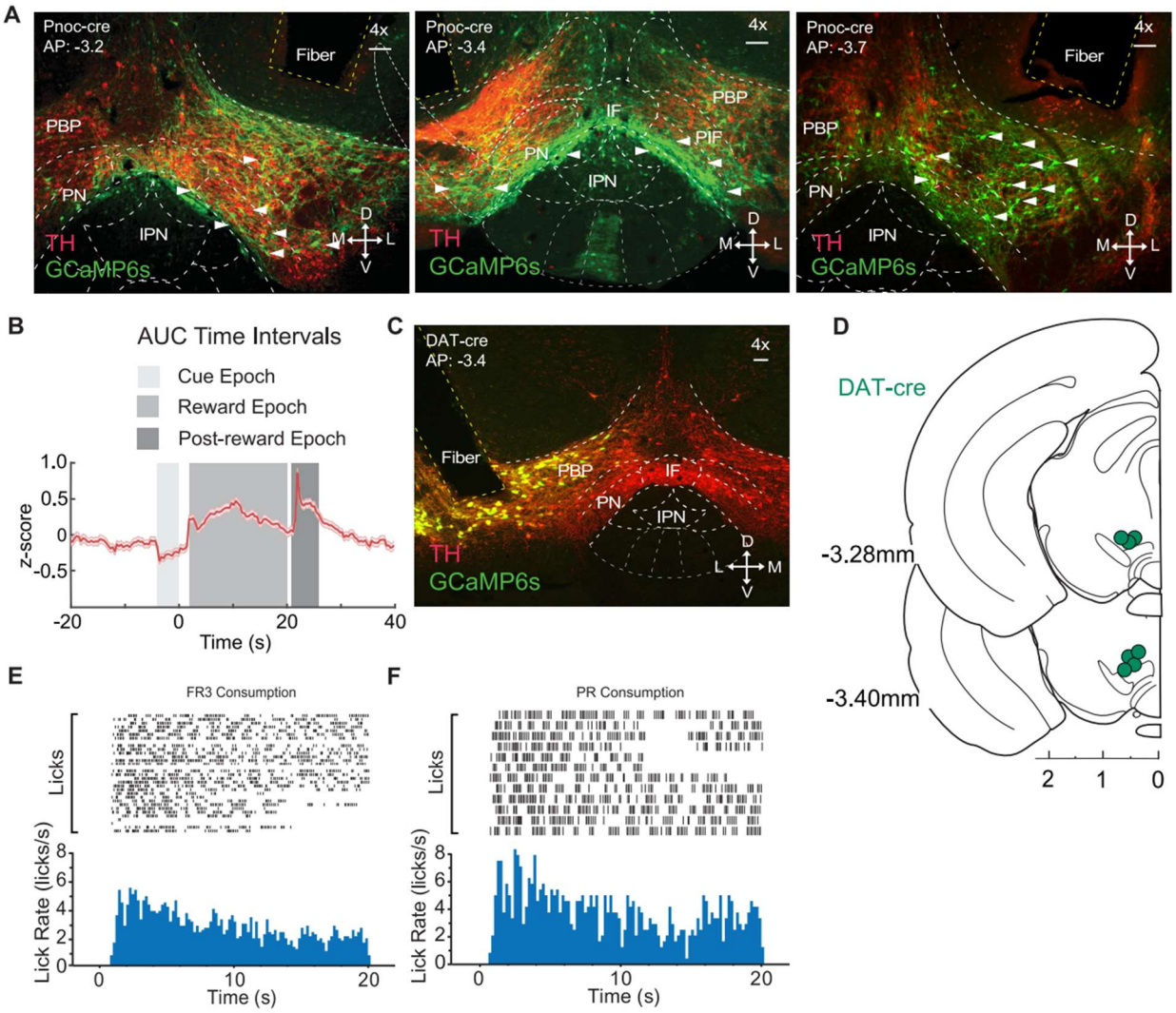


Figure S3. Related to Figure 2: Anatomically Distinct Subdivisions of VTA *Pnoc*+ Neurons are Engaged Differentially During Reward Anticipation and Consumption

(A) Immunohistochemistry for GCaMP6s and tyrosine hydroxylase (TH) staining in *Pnoc*-Cre mice. Coronal sections show GCaMP6s (green), TH (red) and fiber implant tract (yellow) at various anterior/posterior coordinates throughout VTA. (B) Time intervals relative to reward period that were used to calculate mean neural activity in Figure 3B. Trace of mean neural activity (red) taken from Figure 3K (middle panel) for reference. (C) Immunohistochemistry for GCaMP6s and tyrosine hydroxylase (TH) staining in *DAT-Cre* mice. Coronal sections show GCaMP6s (green), TH (red) and fiber implant tract (yellow) at representative implant site. (D) Coronal section cartoon showing anatomical location of 400µm optic fiber implant placement for *DAT-Cre* mice. (E) Raster plot depicting reward consumption (lick rate) during a representative FR3 task session with simultaneous fiber photometry recording. (F) Raster plot depicting reward consumption (lick rate) during a representative PR task session with simultaneous fiber photometry recording.

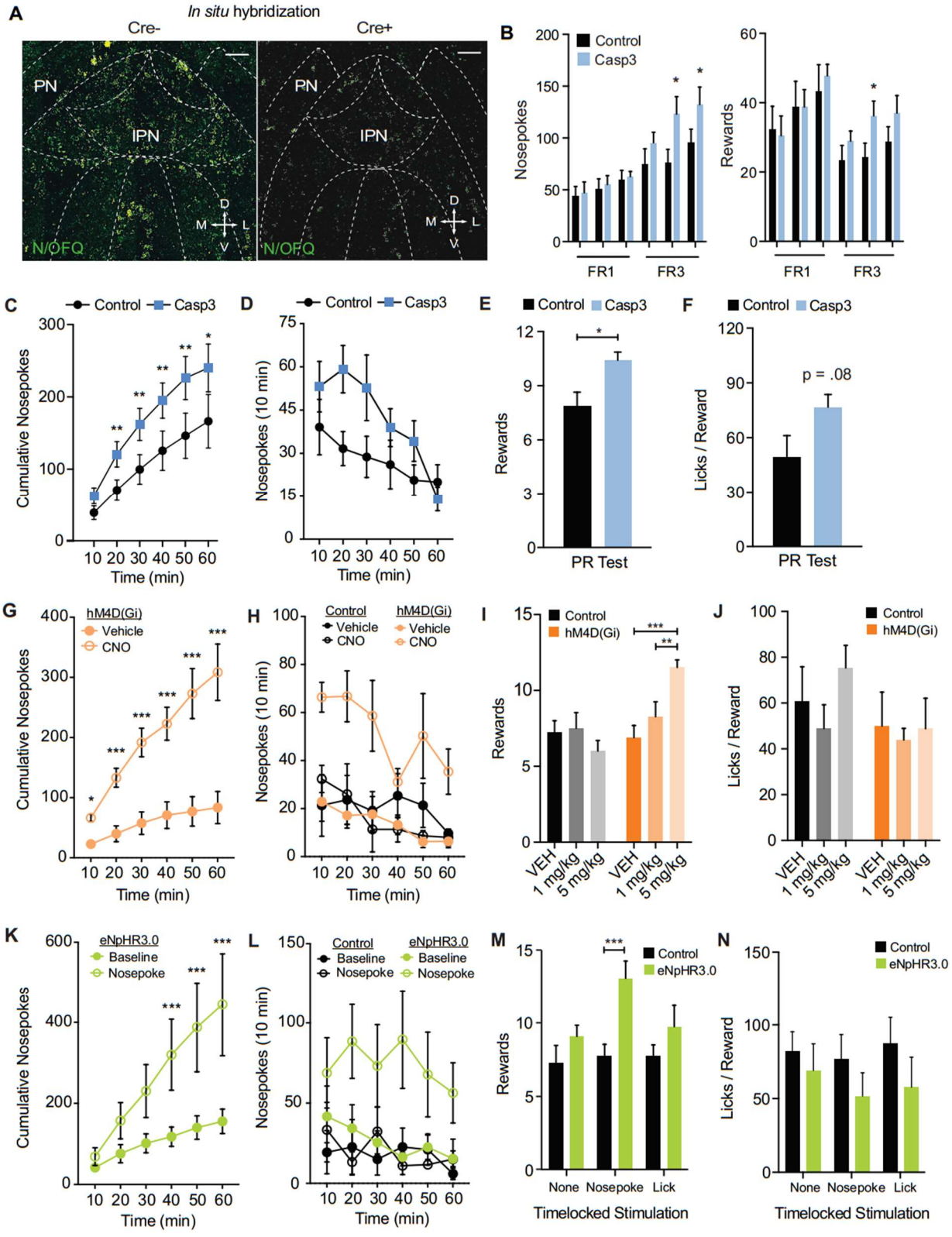

A
In situ hybridization
Cre-
Cre+
PN
IPN
N/OFQ
D
M
L
V
B
Control
Casp3
Nosepokes
Rewards
FR1
FR3
C
Cumulative Nosepokes
Time (min)
D
Nosepokes (10 min)
E
PR Test
F
Licks / Reward
p = .08
G
hM4D(Gi)
Vehicle
CNO
H
I
VEH
1 mg/kg
5 mg/kg
J
K
eNpHR3.0
Baseline
Nosepoke
L
M
None
Nosepoke
Lick
Timelocked Stimulation
N

Figure S4. Related to Figure 3*:* Selective Ablation and Inhibition of pnVTA^{Pnoc} Neurons Enhances Operant Responding for Natural Rewards

(A) Representative coronal images of the pnVTA and IPN showing *in situ* hybridization for nociceptin (N/OFQ) mRNA following viral injection in *Pnoc*-Cre- (left panel) and *Pnoc*-Cre+ (right panel) mice. Images show N/OFQ (green). (B) Number of nosepokes (left panel) and rewards (right panel) across 3 days of FR1 and 3 days of FR3 sucrose self-administration. *Pnoc*-Cre^{taCasp3} mice show significantly increased nosepokes and rewards received, compared to controls (n = 11 to 13: two-way repeated-measures ANOVA, Bonferroni post hoc; Nosepokes for Control versus Casp3 on day 2 and 3 of FR3 training $*p < 0.05$, Rewards for control versus Casp3 on day 2 and 3 of FR3 training $*p < 0.05$). (C) Cumulative nosepokes across 60 min session *Pnoc*-Cre+ and *Pnoc*-Cre- mice that received injections of AAV2-FLEX-taCasp3-TEVp. (n = 11 to 13: two-way repeated-measures ANOVA, Bonferroni post hoc; Nosepokes for Control versus Casp3 $**p < 0.05$, $**p < 0.01$). (D) Total nosepokes for each 10 min epoch across 60 min session in *Pnoc*-Cre+ and *Pnoc*-Cre- mice that received injections of AAV2-FLEX-taCasp3-TEVp. (n = 11 to 13: two-way repeated-measures ANOVA, *ns*). (E) Average number of rewards received during PR test. *Pnoc*-Cre^{taCasp3} mice show significantly increased number of rewards received, compared to controls (n = 11 to 13: unpaired t-test, Rewards for Control versus Casp3 during PR test $*p < 0.05$). (F) Average number of licks per reward performed during PR test. *Pnoc*-Cre^{taCasp3} mice mice show increased number of licks per reward received compared to Controls (n = 11 to 13: unpaired t-test, licks per reward for Control versus Casp3 during PR test, $p = 0.08$). (G) Cumulative nosepokes across 60 min session of *Pnoc*-Cre$^{hM4D(Gi)}$ given vehicle or CNO (5mg/kg). *Pnoc*-Cre$^{hM4D(Gi)}$ mice show significantly increased number of cumulative nosepokes following CNO administration, compared to Vehicle. (n = 8: two-way repeated-measures ANOVA, Bonferroni post hoc; Nosepokes for Vehicle versus CNO, $* p < 0.05$, $*** p < 0.001$). (H) Total nosepokes for each 10 min epoch across 60 min session of *Pnoc*-Cre$^{hM4D(Gi)}$ mice given vehicle or CNO (5mg/kg). (n = 8: two-way repeated-measures ANOVA, *ns*). (I) Average number of rewards received during PR test. *Pnoc*-Cre$^{hM4D(Gi)}$ mice show significantly increased number of rewards received following CNO (5mg/kg administration), compared to vehicle treatment. (n = 8: two-way repeated-measures ANOVA, Bonferroni post hoc; Rewards for Vehicle versus 5mg/kg, $***p < 0.001$, 1mg/kg versus 5mg/kg $**p < 0.01$). (J) Average number of licks per reward performed during PR test. There is no significant difference in lick rate between Control and *Pnoc*-Cre$^{hM4D(Gi)}$ mice following administration of all doses of CNO (n=8: two-way repeated-measures ANOVA, *ns*). (K) Cumulative nosepokes across 60 min session of *Pnoc*-Cre$^{eNpHR3.0}$ mice given either no inhibition or nosepoke-paired photoinhibition. *Pnoc*-Cre$^{eNpHR3.0}$ mice show significantly increased number of cumulative nosepokes following nosepoke-paired photoinhibition, compared to Baseline. (n=6: two-way repeated-measures ANOVA, Bonferroni post hoc; Cumulative nosepokes for Baseline versus Nosepoke, $*p < 0.05$, $**p < 0.01$). (L) Total nosepokes for each 10 min epoch across 60 min session for *Pnoc*-Cre$^{eNpHR3.0}$ mice given either no inhibition or nosepoke-paired photoinhibition. (n = 6: two-way repeated-measures ANOVA, *ns*). (M) Average number of rewards received during PR test. *Pnoc*-Cre$^{eNpHR3.0}$ mice show significantly increased number of rewards received following nosepoke-paired photoinhibition, compared to Controls. (n=6: two-way repeated-measures ANOVA, Bonferroni post hoc; Nosepoke-paired inhibition Rewards for Control versus *Pnoc*-Cre$^{eNpHR3.0}$, $***p < 0.001$). (N) Average number of licks per reward performed

during PR test. There is no significant difference in lick rate between Control and *Pnoc*-Cre$^{eNpHR3.0}$ mice following all stimulation parameters. (n = 6: two-way repeated-measures ANOVA, *ns*).

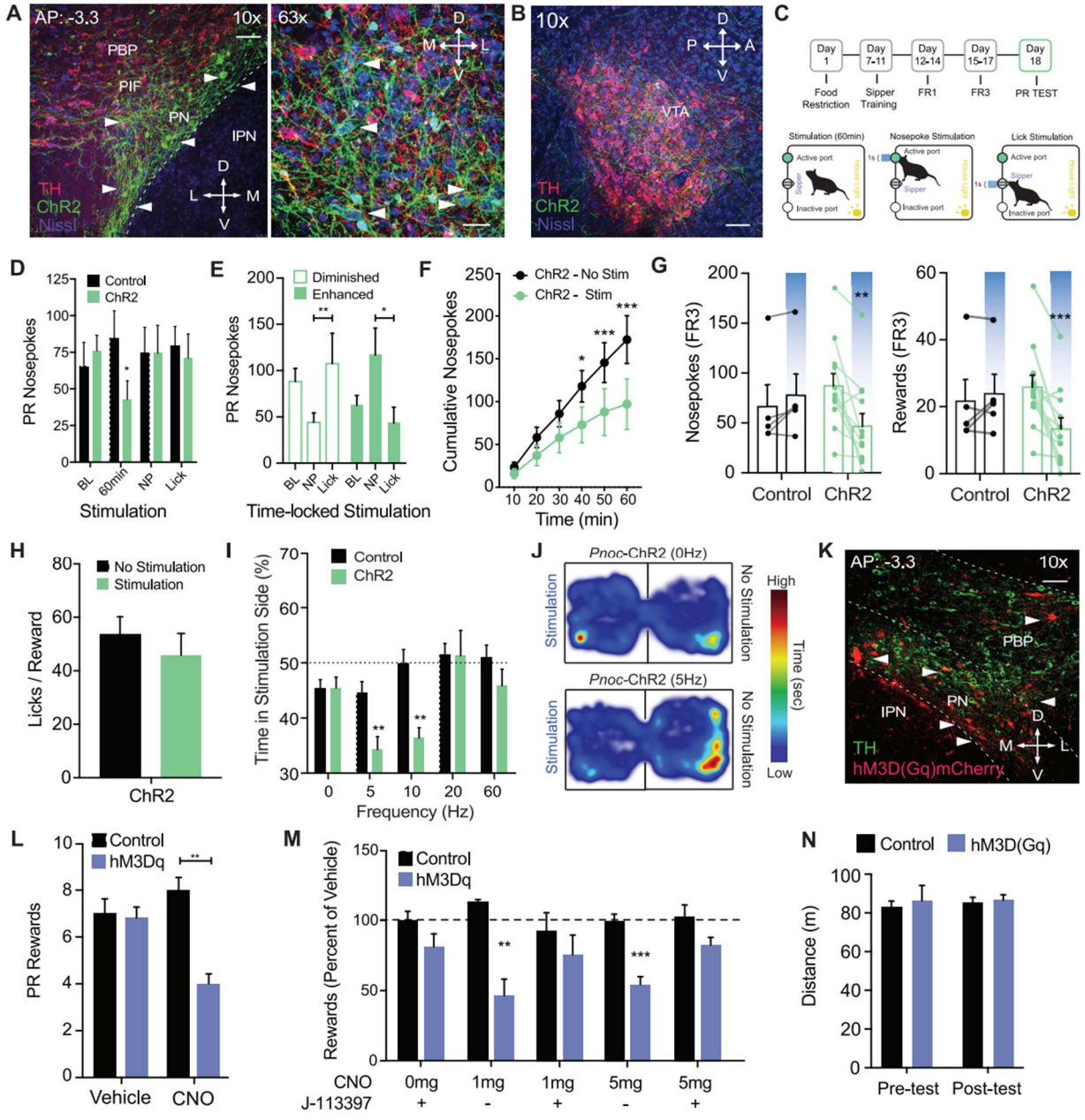


Figure S5. Related to Figure 4: Optogenetic and Chemogenetic Stimulation of pnVTA^{Pnoc} Neurons Decreases Effort to Receive a Natural Reward

(A) Coronal brain images depicting representative ChR2-eYFP expression following viral injection into the pnVTA of *Pnoc*-Cre+ mice (AP:-3.3) Image shows 10x (left panel) and 63X (right panel)). Nissl (Blue), Tyrosine hydroxylase (red) and ChR2-eYFP (green). (B) Sagittal brain image depicting representative ChR2-eYFP expression following viral injection into the pnVTA of *Pnoc*-Cre+ mice. Nissl (Blue), Tyrosine hydroxylase (red) and ChR2-eYFP (green). (C) Number of nosepokes performed during PR test during photostimulation time-locked to either nosepokes or

licks. *Pnoc*-Cre$^{ChR2/eYFP}$ mice show significantly decreased active nosepokes compared to Controls following 60-minute stimulation, but are not different following nosepoke or lick time-locked stimulation. (n = 8: two-way repeated-measures ANOVA, Bonferroni post hoc; $*p < 0.05$). (D) Performance-based separation of *Pnoc*-Cre$^{ChR2/eYFP}$ mice showing average number of nosepokes performed during PR test and during Baseline (BL), Nosepoke-paired (NP) or Lick-paired (Lick) photostimulation. "Diminished" responding *Pnoc*-Cre$^{ChR2/eYFP}$ mice (n=8) show significantly decreased nosepokes compared "Enhanced" responding *Pnoc*-Cre$^{ChR2/eYFP}$ (n=4) during nosepoke-paired stimulation. (n = 4 - 8: two-way repeated-measures ANOVA, Diminished NP vs. Lick Stimulation $**p < 0.01$; Enhanced NP vs. Lick Stimulation $*p < 0.05$). (E) Average cumulative nosepokes during PR test for sucrose in *Pnoc*-Cre$^{ChR2/eYFP}$ mice (n = 12: two-way repeated-measures ANOVA, Nosepokes for *Pnoc*-Cre$^{ChR2/eYFP}$ mice, No Stimulation versus Stimulation, $*p < 0.05$, $***p < 0.001$). (F) Number of nosepokes (left panel) and rewards (right panel) during FR3 operant conditioning. *Pnoc*-Cre$^{ChR2/eYFP}$ mice show reduction in nosepokes and number of rewards received during 10 Hz stimulation compared to baseline and controls (n = 8 to 12, two-way repeated-measures ANOVA, Bonferroni post hoc; Nosepokes for *Pnoc*-Cre$^{ChR2/eYFP}$ mice, baseline versus stimulation, $**p < 0.01$; Rewards for *Pnoc*-Cre$^{ChR2/eYFP}$ mice, baseline versus stimulation $***p < 0.001$. (G) Average number of licks per reward performed during (PR) test. There is no significant difference in lick rate in *Pnoc*-Cre$^{ChR2/eYFP}$ mice between no stimulation and stimulation. (n = 12: Student's t-test, *ns*). (H) Percentage of time spent in the stimulation side for different frequencies (5, 10, 20, 60Hz) in control and *Pnoc*-Cre$^{ChR2/eYFP}$ mice (n = 9 to 12: two way repeated-measures ANOVA, Bonferroni post hoc; percentage of time spent in stimulation side for control versus ChR2 at (5Hz) $*p < 0.05$, percentage of time spent in stimulation side for control versus ChR2 at (10Hz) $*p < 0.05$). (I) Representative heatmap of relative time spent during real-time place testing for a *Pnoc*-Cre$^{ChR2/eYFP}$ mouse that received 5 Hz photostimulation in the stimulation chamber (top) and the same mouse that received no photostimulation in the stimulation chamber (bottom). (J) Coronal image depicting representative Gq-mCherry expression following viral injection in *Pnoc*-Cre+ mice. Image shows Tyrosine hydroxylase (green) and Gq-mCherry (red). (K) Number of rewards for baseline versus CNO (5mg/kg) during PR test $**p < 0.01$). Control mice show no effect in number of rewards received following CNO (5mg/kg) administration. (L) Number of rewards during PR test. *Pnoc*-Cre$^{hM3D(Gq)}$ mice show a significant decrease in the number of rewards received following CNO (5mg/kg) administration compared to controls (data presented as mean ± SEM, n = 8 to 11: two way repeated-measures ANOVA, Bonferroni post hoc; Rewards for baseline versus CNO (5mg/kg) during PR test $**p < 0.01$). Control mice show no effect in number of rewards received following CNO (5mg/kg) administration. (M) Number of rewards (as % of vehicle treatment) during PR test in *Pnoc*-Cre$^{hM3D(Gq)}$ mice that received CNO (1 mg/kg or 5 mg/kg) and the NOPR antagonist (J-113397, 10 mg/kg). CNO treated mice show a significant decrease in rewards compared to controls that was partially rescued by treatment with J-113397. (data presented as mean ± SEM, n = 8 to 11: two way repeated-measures ANOVA, Bonferroni post hoc; Nosepokes for baseline versus CNO (5mg/kg) during PR test $**p < 0.01$, Rewards for baseline versus CNO (5mg/kg) during PR test $**p < 0.01$). (N) Distance traveled during Pre-test and Post-test following conditioning paradigm. There is no difference in the distance traveled between control and *Pnoc*-Cre$^{hM3D(Gq)}$ mice.

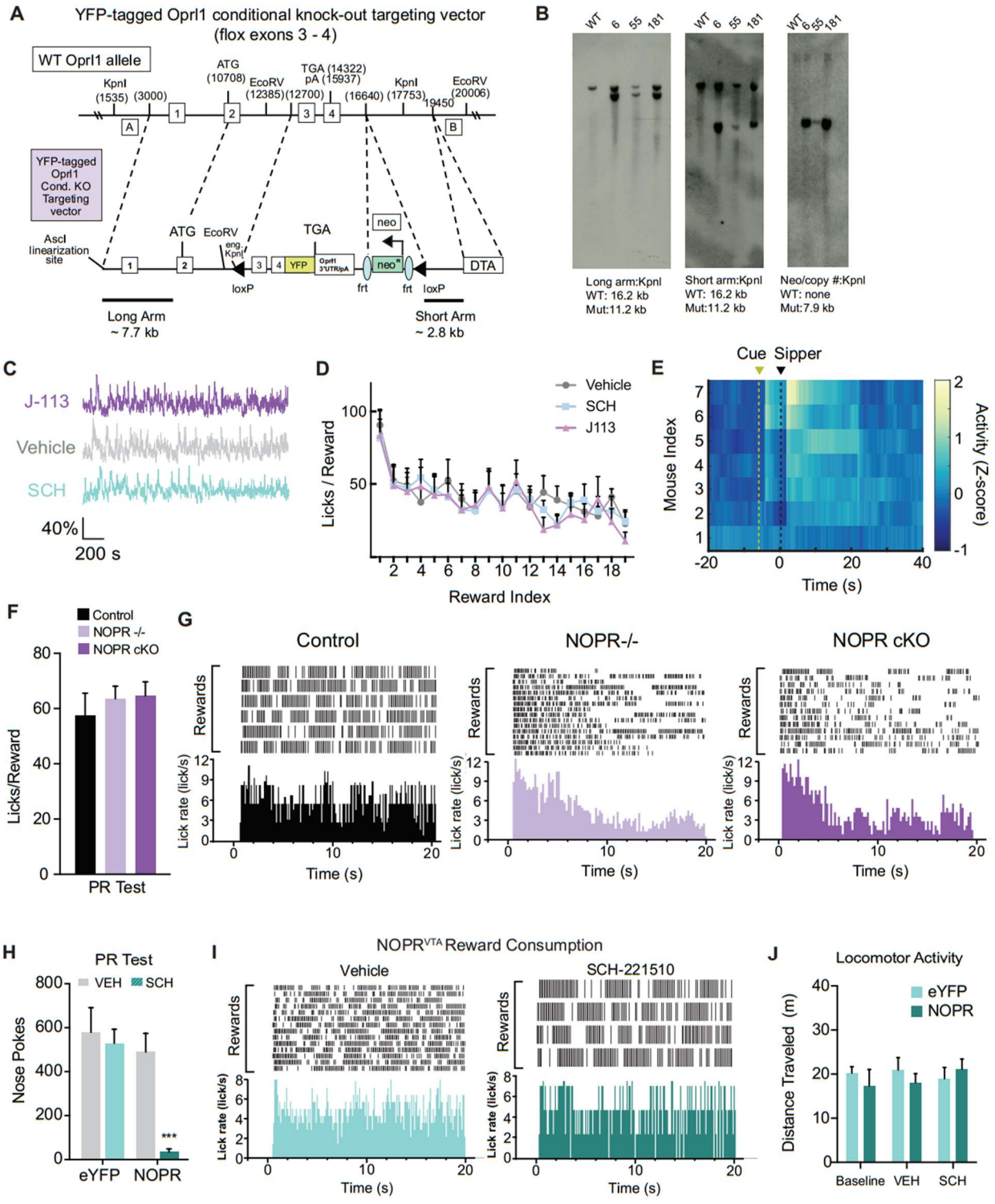


Figure S6. Related to Figure 5: VTA NOPR Expression is Necessary and Sufficient for Regulating Motivation for Sucrose

(A) Schematic of YFP-tagged NOPR conditional knock-out targeting vector. (B) Representative southern blots for Oprl1 conditional KO (KpnI digest). (C) Representative traces for J-113397

(top), Vehicle (middle) and SCH-221510 (bottom). (D) Average lick rate for *DAT*-Cre mice during Pavlovian conditioning over rewards following vehicle, SCH-221510, or J-113397 administration. (E) Representative heat map displaying average neural activity for each mouse during Pavlovian conditioning. (F) Average lick rate for control, NOPR KO and NOPR cKO mice during PR test (Two-way repeated measures ANOVA, *ns*). (G) Representative lick rate rasters for control, NOPR KO and NOPR cKO mice (H) Pie chart graph depicting %Co-expression Cre and TH (top panel) in the VTA of $NOPR^{lox/lox+}$ mice that received injections of AAV9-rTH-PI-Cre-SV40 (n=4, 4 slices per n). (I) Number of nosepokes during PR test session following vehicle and SCH221510 (10mg/kg) administration. TH-Cre x $NOPR^{-/-}$ "rescue" mice show a significant decrease in the number of rewards received following SCH221510 (10mg/kg) administration compared to $NOPR^{-/-}$ mice (n = 8 to 10: two way repeated-measures ANOVA, Bonferroni post hoc; Nosepokes for NOPR-rescue mice Vehicle versus SCH221510 (10mg/kg) PR test $***p < 0.001$. (J) Representative lick rate raster for NOPR-rescue mice given either vehicle or SCH-221510 prior to PR test. (K) Total distanced traveled in an open field assay at baseline, vehicle, and SCH-221510 (10mg/kg) treated NOPR KO X TH-Cre mice that were injected with eYFP or NOPR.

CHAPTER 3: MEDIAL ACCUMBENS SHELL NEURONS ENCODE REWARD VALUE

3.1 Chapter Introduction

Rewards are stimuli that have intrinsic incentive value to an animal. Rewards with greater incentive value will be sought out over alternatives of lesser value. The brain must compute and represent the relative value of rewards to facilitate goal directed behavior to optimize reward attainment. Nucleus accumbens (NAc) is a critical brain region for driving motivated behavior (Manella 2013, Castro 2019). NAc integrates many incoming reward-related information streams such as glutamatergic innervation from areas including basolateral amygdala, prefrontal cortex, ventral pallidum (VP) and ventral hippocampus and dopaminergic innervation from ventral tegmental area (VTA) (Britt 2012). These tributary NAc afferents have been shown to be reward modulated, and in many cases, even reflect the quantity and value of rewards (Stuber 2011, Otis 2017, Ottenheimer 2018, Richard 2016, Vachez 2021, Schultz 2016).

The two principal types of spiny projection neurons (SPN) in NAc are predominantly defined by their expression of either D1R or D2R type dopamine receptors. D1R-SPNs also largely co-express the endogenous opioid dynorphin, while D2R-SPNs primarily express enkephalin (Besson 1990, Castro 2021). Previous studies have shown NAc SPNs to be modulated by rewards and reward-predictive cues but have not resolved the molecular identity of the recorded neurons (Taha 2005, Roitman 2005). Previous studies have also predominantly focused on the reward-responsivity of neurons in NAc core or made no distinction between recordings in NAc core and medial accumbens shell (mNAcSh). mNAcSh has recently been shown to be very distinct from NAc core in its

connectivity, function and dopaminergic innervation (Day 2006, Britt 2012, Ambroggi 2011, Saunders 2018, de Jong 2018). Even within mNAcSh, there have been functional distinctions discovered across anterior-posterior and dorsal-ventral axes (Al-Hasani 2015, Castro 2019). However, these subregion differences are likely due to a continuum of functionally distinct individual neurons, rather than resulting from absolute boundaries.

Given the unique connectivity and function of mNAcSh, we sought to characterize the responses of mNAcSh SPNs to rewards of different value and across cue-reward learning with proper consideration of known differences in SPN function due to molecular identity and relative anatomical position. Recent experimental advances have allowed for the simultaneous recording of dozens of genetically defined neurons in deep brain structures using 2-photon calcium imaging through endoscopic lenses (Resendez 2016, Rossi 2020). The low noise-floor of 2-photon microscopy also allows for the resolution of individual cell morphology and permits high fidelity tracking of individual neurons across multiple days (Namboodiri 2019). In the current study, we leverage this technique to conduct longitudinal studies on how mNAcSh SPNs respond to multiple reward value conditions and change across cue-reward learning. This revealed a distinct subpopulation of enkephalinergic SPNs (Penk-SPN) that are selectively activated during the consumption of lower value reward. It was also found that an Penk-SPN circuit from mNAcSh→VP is also selectively activated during lower value reward and causally drives a low, positive reward value. These results provide foundational insight into how reward value information is integrated and separated in mNAcSh and causally contributes to driving appropriate consummatory behavior.

3.2 mNAcSh SPNs encode reward value

In order to characterize any potential differences in how rewards of different value are processed in mNAcSh, we presented mice with 4 distinct reward conditions that had differing motivational value but similar taste: 0% (water), 3%, 10% and 30% sucrose. To test the relative value of the 4 sucrose reward conditions to water-restricted mice, 2-bottle choice experiments were conducted in which the animal was given access to 2 of the 4 conditions at a time and had to choose the preferred reward (**Fig. S7A**). Previous studies have consistently found that mice and rats prefer sucrose concentrations between 10-15% over higher and lower concentrations, and that the distribution of sucrose preferences is shifted toward lower concentrations during water deprivation (Pothion 2004, Wallace 2008, Ottenheimer 2020). Mice also tend to drink smaller volumes of higher concentration sucrose (20%+) due to satiety (Davis 1973).

To test the relative value of the 4 sucrose reward conditions while minimizing effects from satiety, we tested mice in an adapted 2-bottle choice assay where the animals were only given brief access (20 s) to sucrose sippers with a variable intertrial interval between access periods. In the first phase, animals were given access to either sipper in isolation so that they could sample them independently. In the second phase, animals were simultaneously presented with both sippers side by side and had to choose where to allocate their sipper access time (**Fig. S7B**). It was found that water-restricted C57 mice preferred 10% sucrose over all other conditions and 30% sucrose less than all other conditions (**Fig. S7E**). The 6 possible pairings of the 4 sucrose reward conditions were presented in a pseudorandomized and counterbalanced sequence to minimize effects from condition history and right versus left sipper location bias. Preference test

days were also buffered by 10% versus 10% sucrose days to neutralize any biases that arose from test days (**Fig. S7C-D**).

To measure the activity of mNAcSh SPNs during the consumption of varying reward value conditions, we imaged the calcium dynamics of individual SPNs using 2-photon fluorescence microscopy through endoscopic GRIN lenses. Preprodynorphin-cre (Pdyn-cre) and proenkephalin-cre (Penk-cre) animals were stereotaxically injected with 500 nl of AAVdj-EF1a-DIO-GCaMP6s virus and implanted with a 600 um diameter cylindrical GRIN lens in mNAcSh (**Fig. 6A**). After 4-6 weeks of recovery from surgery, animals were water restricted and trained to sip 10% sucrose reward at variable intervals while reversibly immobilized and head-restrained (**Fig. 6B**). All animals displayed consistent licking behavior while immobilized after 4 days of lick training and were then tested in 30 minute, 20 trial calcium imaging sessions each day for 4 consecutive days. In each session, animals consumed 1 of 4 possible sucrose reward conditions (0%, 3%, 10% or 30% sucrose in water), with the order of condition days pseudorandomized and counterbalanced across animals (**Fig. 6C**). Animals immediately licked the presented sucrose reward, with an average licking bout duration of 7-8 seconds (**Fig. 6D**). Given the small 10ul reward volumes, there was no significant difference between reward conditions in the average amount of licking for each reward (**Fig. 6E**). This uniformity in lick behavior allowed for accurate comparison of SPN activity in response to the content of the different reward identities with minimal variation in behavioral output.

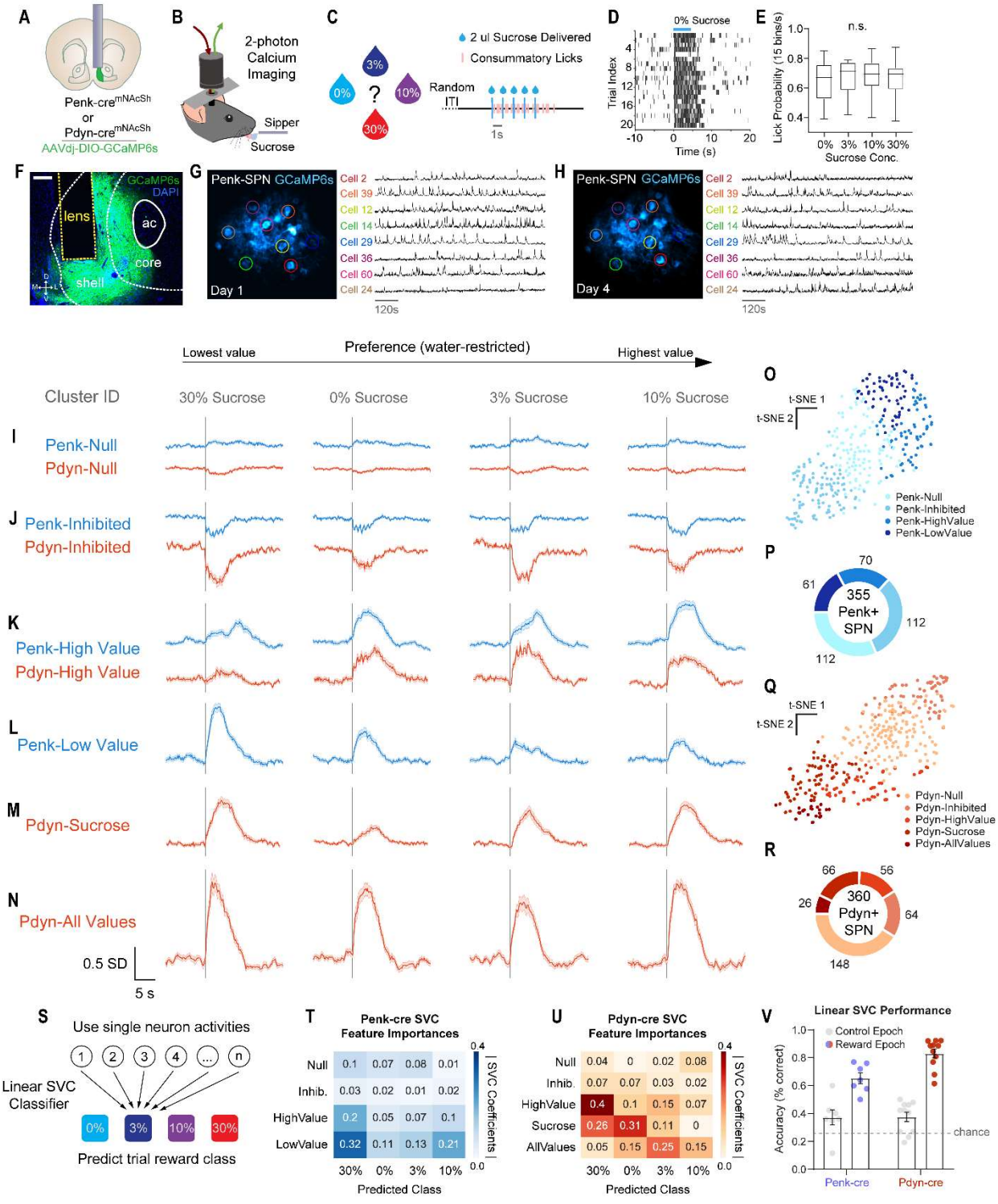

A
Penk-cre mNAcSh
or
Pdyn-cre mNAcSh
AAVdj-DIO-GCaMP6s
B
2-photon
Calcium
Imaging
Sipper
Sucrose
C
2 ul Sucrose Delivered
Consummatory Licks
3%
0%
?
10%
30%
Random
ITI
1s
D
0% Sucrose
Trial Index
Time (s)
E
n.s.
Lick Probability (15 bins/s)
0% 3% 10% 30%
Sucrose Conc.
F
GCaMP6s
DAPI
lens
ac
core
shell
G
Penk-SPN GCaMP6s
Day 1
Cell 2
Cell 39
Cell 12
Cell 14
Cell 29
Cell 36
Cell 60
Cell 24
120s
H
Penk-SPN GCaMP6s
Day 4
120s
Lowest value
Preference (water-restricted)
Highest value
Cluster ID
30% Sucrose
0% Sucrose
3% Sucrose
10% Sucrose
I
Penk-Null
Pdyn-Null
J
Penk-Inhibited
Pdyn-Inhibited
K
Penk-High Value
Pdyn-High Value
L
Penk-Low Value
M
Pdyn-Sucrose
N
Pdyn-All Values
0.5 SD
5 s
O
t-SNE 1
t-SNE 2
Penk-Null
Penk-Inhibited
Penk-HighValue
Penk-LowValue
P
70
61
355
Penk+
SPN
112
112
Q
t-SNE 1
t-SNE 2
Pdyn-Null
Pdyn-Inhibited
Pdyn-HighValue
Pdyn-Sucrose
Pdyn-AllValues
R
66
56
26
360
Pdyn+
SPN
64
148
S
Use single neuron activities
1 2 3 4 ... n
Linear SVC
Classifier
0% 3% 10% 30%
Predict trial reward class
T
Penk-cre SVC
Feature Importances
Null 0.1 0.07 0.08 0.01
Inhib. 0.03 0.02 0.01 0.02
HighValue 0.2 0.05 0.07 0.1
LowValue 0.32 0.11 0.13 0.21
30% 0% 3% 10%
Predicted Class
SVC Coefficients
0.4
0.0
U
Pdyn-cre SVC
Feature Importances
Null 0.04 0 0.02 0.08
Inhib. 0.07 0.07 0.03 0.02
HighValue 0.4 0.1 0.15 0.07
Sucrose 0.26 0.31 0.11 0
AllValues 0.05 0.15 0.25 0.15
30% 0% 3% 10%
Predicted Class
SVC Coefficients
0.4
0.0
V
Linear SVC Performance
Control Epoch
Reward Epoch
Accuracy (% correct)
chance
Penk-cre
Pdyn-cre

Figure 7: Distinct subpopulations of mNAcSh SPNs are recruited in response to rewards of different value. (A) Coronal section cartoon displaying target brain region, cell-types and lens implant location. (B) Schematic of 2-photon calcium imaging through endoscopic lens during consumption of variable sucrose reward. (C) Time-course of sucrose reward delivery for a given trial. Droplets were either 0%, 3%, 10% or 30% sucrose solution in water. (D) Representative lick raster depicting individual licks that were detected by contact licometer during behavior and imaging session. (E) Mean lick probability during reward epoch (0-8 seconds from reward onset) for each sucrose reward condition (n = 18 animals; one-way ANOVA: p = 0.9275). (F) Representative coronal section showing enk-SPNs expressing GCaMP6s underneath the implanted endoscopic GRIN lens in mNAcSh (scale bar = 200 um). (G and H) Fluorescence time series of individual SPNs that were tracked from (G) day 1 session through (H) day 4 session. (I) SPN clusters that were not consistently recruited during any of the reward conditions (averaged response to 20 trials of each reward condition). (J) SPN clusters that were consistently transiently inhibited during all reward conditions. (K) SPN clusters that were consistently transiently excited during the 0%, 3% and 10% conditions, but less for the 30% condition. (L) enk-SPN cluster that was consistently transiently excited during 30% condition, moderately for the 0% condition and less for the 3% and 10% condition. (M) dyn-SPN cluster that was consistently transiently excited during the 3%, 10% and 30% sucrose conditions but less for the 0% (water) condition. (N) dyn-SPN cluster that was consistently transiently excited for all reward conditions. (O) t-sne plot of variance in activity of enk-SPN clusters during all reward conditions. Each dot represents an individual tracked neuron that composes the cluster. (P) Fraction of all tracked enk-SPN neurons that belong to each functional cluster. (Q) t-sne plot of variance in activity of dyn-SPN clusters during all reward conditions. Each dot represents an individual tracked neuron that composes the cluster. (R) fraction of all tracked dyn-SPN neurons that belong to each functional cluster. (S) Cartoon depicting multiclass linear Support Vector Classifier implementation (features = neurons, observations = mean activity during given trial, target = class of reward condition). (T) Mean of SVC coefficients for all enk-SPNs (SVC features) of each functional cluster. (U) Mean of SVC coefficients for all dyn-SPNs (SVC features) of each functional cluster. (V) Linear SVC classification accuracy for every animal using mean SPN activity during reward epoch (0-8s from reward onset) and during control epoch (12-4s before reward onset) (n = 7 enk-cre and 11 dyn-cre animals).

The resolution and low noise floor of 2-photon calcium imaging allowed for visualization of individual cell morphology and enabled high fidelity tracking of over 700 individual SPNs across multiple days and conditions (**Fig. 7G-H, S8A**). Individual tracked Pdyn-SPNs and Penk-SPNs displayed reward responses that were consistent for a given sucrose concentration, but that varied between the different sucrose concentrations (**Fig. S8B-C**). Groups of individual SPNs were also correlated in their selectivity for specific

sucrose concentrations. In order to group individual neurons together that had similar response profiles across the 4 sucrose conditions, k-means clustering was performed across all tracked SPNs. This unbiased approach to grouping neurons on the basis of similar activity revealed that there were 4 distinct clusters of Penk-SPNs and 5 distinct clusters of Pdyn-SPNs (**Fig. 7O-R, S8D-G**). Notably, every cluster was composed of neurons from nearly all, if not all, of the mice that were tested, demonstrating that these functionally-defined clusters are generalizable across animals (**Fig. S8B,C**). 32% of tracked Penk-SPNs and 41% of tracked Pdyn-SPNs were not consistently modulated by any of the reward conditions (**Fig. 7I**). A greater proportion of Penk-SPNs (32%) were inhibited in response to all 4 sucrose conditions than for Pdyn-SPNs (18%) (**Fig. 7J**). A large cluster of Penk-SPNs and Pdyn-SPNs were found to be active in response to the higher value 0%, 3% and 10% conditions, but were relatively quiescent in response to the lower value (see **Fig. S7**) 30% sucrose condition (**Fig. 7K**). Interesting, nearly half of all reward-excited Penk-SPNs were selectively active far more for the lower value 30% sucrose condition (**Fig. 7L**). The presence of 3 robust and complimentary SPN clusters that are differently activated at the 30% condition demonstrates a dramatic shift of SPN recruitment during low value rewards. A large 30% sucrose selective population of Penk-SPNs is recruited only for the lowest value condition and this occurs simultaneously as otherwise reward-excited SPNs are systematically silent (**Fig. 7K-L**).

A distinct cluster of Pdyn-SPNs was also found to be active to all sucrose reward conditions (3%, 10% and 30%) but much less active to water (0% sucrose) (**Fig. 7M**). A small minority of Pdyn-SPNs (7%) were also found to be highly excited by every reward condition (**Fig. 7N**). To further quantify the specificity of mNAcSh SPN activity to the 4

reward value conditions, a linear Support Vector Classifier (SVC) model was created. This SVC model was trained to predict which of the 4 sucrose conditions an animal consumed on a given trial only by considering the mean activity from each SPN during the reward epoch (**Fig. 7S**). The SVC performed with high accuracy and specificity for all 4 reward conditions for both Pdyn-SPNs and Penk-SPNs (**Fig. 7V**). The SVC model coefficients were greater for SPNs in the HighValue and LowValue clusters for discriminating between 30% sucrose and the other reward conditions (**Fig. 7T-U**). This reaffirms that the primary differences in recruitment of reward-excited SPNs were occurring between the low value and high value reward conditions. Altogether, these mNAcSh calcium imaging experiments across multiple reward value conditions have revealed substantial heterogeneity in the reward-evoked activity of individual SPNs. There are mNAcSh SPN populations that are selective for higher value reward conditions and mNAcSh SPN populations that are selective for lower value reward conditions. Consistent recruitment of distinct mNAcSh SPN subpopulations to rewards of different value suggests that these neurons are encoding specific reward value information.

3.3 mNAcSh SPNs do not signal reward anticipation

As animals learn the association between valued rewards and previously un-valued cues, the reward-predictive cues gradually become imbued with motivational value as well (Flagel 2011, Cartoni 2016). Since we observed heterogeneity in the activity of mNAcSh SPNs to rewards of different value, we then sought to characterize potential differences in how individual mNAcSh SPNs respond to reward-predictive cues as a function of cue-reward associative learning. The same animals that were tested in the variable sucrose 2-photon imaging paradigm (see **Fig. 7**) were also trained to associate

an auditory cue to subsequent delivery of 10% sucrose reward across 7 days of Pavlovian conditioning (**Fig. 8A-C**). Animals again displayed immediate licking in response to sucrose reward presentation, but initially did not show robust anticipatory licking to the cue (**Fig. 8D,H**). After multiple cue-reward Pavlovian conditioning sessions, animals began to display consistent anticipatory licking, indicating that the cue-reward association had been learned and the cue itself had been imbued with motivational value (**Fig. 8E,H**).

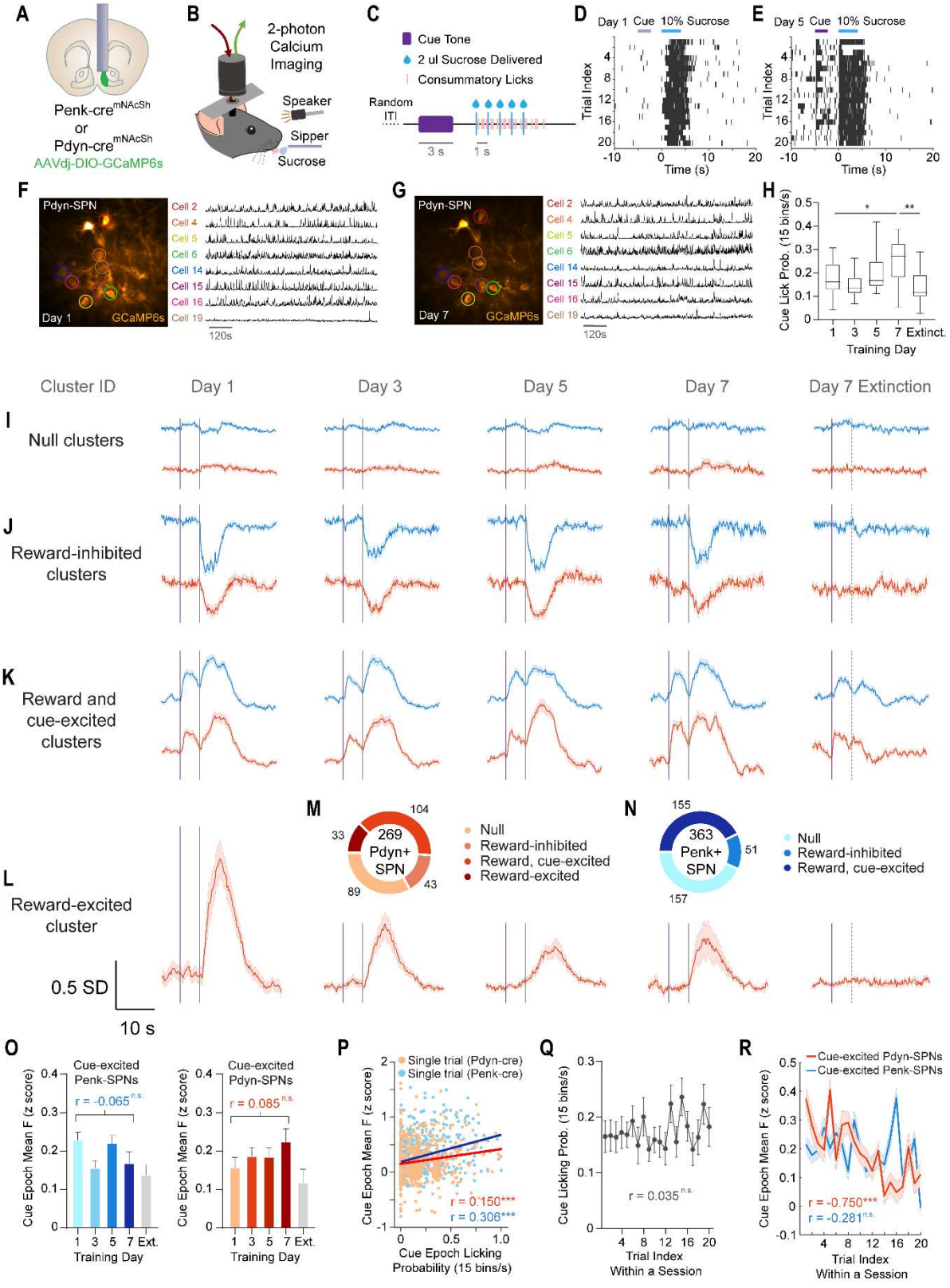

A
Penk-cre mNAcSh
or
Pdyn-cre mNAcSh
AAVdj-DIO-GCaMP6s
B
2-photon
Calcium
Imaging
Speaker
Sipper
Sucrose
C
Cue Tone
2 ul Sucrose Delivered
Consummatory Licks
Random
ITI
3 s
1 s
D
Day 1 Cue 10% Sucrose
Trial Index
Time (s)
E
Day 5 Cue 10% Sucrose
Trial Index
Time (s)
F
Pdyn-SPN
Day 1
GCaMP6s
Cell 2
Cell 4
Cell 5
Cell 6
Cell 14
Cell 15
Cell 16
Cell 19
120s
G
Pdyn-SPN
Day 7
GCaMP6s
120s
H
Cue Lick Prob. (15 bins/s)
Training Day
Extinct.
Cluster ID
Day 1
Day 3
Day 5
Day 7
Day 7 Extinction
I
Null clusters
J
Reward-inhibited clusters
K
Reward and cue-excited clusters
L
Reward-excited cluster
0.5 SD
10 s
M
269
Pdyn+
SPN
104
33
89
43
Null
Reward-inhibited
Reward, cue-excited
Reward-excited
N
363
Penk+
SPN
155
51
157
Null
Reward-inhibited
Reward, cue-excited
O
Cue-excited
Penk-SPNs
r = -0.065 n.s.
Cue-excited
Pdyn-SPNs
r = 0.085 n.s.
Cue Epoch Mean F (z score)
Training Day
P
Single trial (Pdyn-cre)
Single trial (Penk-cre)
r = 0.150***
r = 0.306***
Cue Epoch Mean F (z score)
Cue Epoch Licking Probability (15 bins/s)
Q
Cue Licking Prob. (15 bins/s)
r = 0.035 n.s.
Trial Index
Within a Session
R
Cue-excited Pdyn-SPNs
Cue-excited Penk-SPNs
Cue Epoch Mean F (z score)
r = -0.750***
r = -0.281 n.s.
Trial Index
Within a Session

Figure 8: Most reward-excited mNAcSh SPNS are cue-excited, but do not reflect reward expectation (A) Coronal section cartoon displaying target brain region, cell-types and lens implant location. (B) Schematic of 2-photon calcium imaging through endoscopic lens during cue-reward Pavlovian conditioning. (C) Time-course of auditory cue and subsequent sucrose reward delivery for a given trial. (D and E) Representative lick raster depicting individual licks that were detected by contact licometer during cue and reward when animals were cue-naïve (D) and after conditioning (E). (F and G) Fluorescence time series of individual SPNs that were tracked from Pavlovian (F) day 1 session through (G) day 7 session. (H) Mean lick probability during cue epoch (0-5 seconds from cue-onset) for each day of Pavlovian conditioning and extinction (n = 16 animals; One-way ANOVA: Day 1 vs Day 7, *p = 0.020, Day 7 vs Extinction, **p = 0.002). (I) SPN clusters that were not consistently recruited during cue or reward throughout 7 days of Pavlovian conditioning (averaged response to 20 trials of each reward condition). (J) SPN clusters that were consistently transiently inhibited during reward throughout 7 days of Pavlovian conditioning. (K) SPN clusters that were consistently transiently excited during cue and reward throughout 7 days of Pavlovian conditioning and extinction. (L) dyn-SPN cluster that was consistently transiently excited during reward but not cue throughout 7 days of Pavlovian conditioning. (M) Fraction of all tracked dyn-SPN neurons that belong to each Pavlovian functional cluster. (N) Fraction of all tracked enk-SPN neurons that belong to each Pavlovian functional cluster. (O) Mean cue response (0-5 seconds from cue-onset) of all cue-excited SPNs during each day of Pavlovian conditioning and extinction (n = 7 enk-cre animals, 155 neurons; Pearson's correlation: p = 0.105; n = 9 dyn-cre animals, 104 neurons; Pearson's correlation: p = 0.083). (P) Correlation between mean cue response of all cue-excited cluster SPNs and mean cue epoch lick probability for every Pavlovian conditioning trial. Blue and red lines capture linear fit between cue response and cue licking for each cell-type. (Pearson's correlation: ***p < 0.001). (Q) Mean lick probability during cue epoch (0-5 seconds from cue-onset) for each trial within a session (Days 1, 3, 5 combined) (n = 16 animals; Pearson's correlation: p = 0.288). (R) Mean cue response of all cue-excited cluster SPNs during each trial within a session (Days 1, 3, 5 combined) (n = 7 enk-cre animals; Pearson's correlation: p = 0.254; n = 9 dyn-cre animals; Pearson's correlation: ***p < 0.001).

2-photon calcium imaging of mNAcSh SPNs expressing GCaMP6s occurred simultaneously to Pavlovian conditioning on training days 1, 3, 5 and 7, and 632 individual SPNs were successfully tracked across all Pavlovian imaging sessions (**Fig. 8F-G, S8A**). K-means clustering was performed on all 632 SPNs to stratify groups of neurons on the basis of similar cue and reward responses. It was revealed that there were 3 distinct clusters of Penk-SPNs and 4 distinct clusters of Pdyn-SPNs within the context of the Pavlovian conditioning paradigm (**Fig. 8M-N, S8B-G**). Not all reward-responsive mNAcSh

SPNs (61% of all SPNs) were found to also be cue-responsive (41% of all SPNs) (**Fig. 8M-N**). There were also no cue-inhibited SPNs observed and all reward-inhibited SPNs had no consistent cue-response (**Fig. 8J**). There was also a distinct cluster of Pdyn-SPNs that was reward-excited but not cue-excited (**Fig. 8L**), while all reward-excited Penk-SPNs were also cue-excited (**Fig. 8K**). The Pdyn-SPN cluster that was only reward-excited also showed smaller reward-excitation to the 10% sucrose reward in later Pavlovian conditioning days, suggesting that the expectation of reward diminished the magnitude of reward-excitation (**Fig. 8L**). To test for the persistence of cue and reward-responses during the absence of expected reward delivery, the latter half (last 10 out of 20 trials) of trials on Pavlovian day 7 were reward omission trials, where the cue is presented as usual but the expected sucrose reward is not delivered. The SPNs that were found to be exclusively reward-responsive but not cue-responsive showed no response during Pavlovian extinction trials when reward was omitted (**Fig. 8J,L**), suggesting that the reward-responses were truly specific to the content of the reward itself. However, cue and reward-excited SPNs displayed a persisting cue and reward-excitation during the 10 extinction trials (**Fig. 8K**).

Surprisingly, the magnitude of cue-excitation of the cue-excited SPN clusters did not increase across Pavlovian training days (**Fig. 8O**) as is canonically observed in VTA DA neurons during cue-reward associative learning (Schultz 1997, Berridge 1998). We also found that the magnitude of cue-excitation was only weakly correlated with amount of anticipatory licking per trial (**Fig. 8P**). Taken together, these results demonstrate that mNAcSh SPN cue-response does not robustly change as a function of cue-reward associative learning and does not reflect the motivational value of the reward-predictive

cues. Additionally, Pdyn-SPNs decreased the magnitude of their cue-excitation across trials within an imaging session, suggesting that the cue-response of Pdyn-SPNs may reflect the novelty and salience of the cue tone rather than the motivational value (**Fig. 8Q-R**). Altogether, the presence of reward-responsive SPNs that are not cue-responsive and the decoupling of cue-responses from reward expectation in mNAcSh suggest that this region diverges from the canonical reward-expectation based models of dopamine release in NAc as some recent studies have proposed it may (de Jong 2019, Saunders 2018).

3.4 Functionally distinct mNAcSh SPN populations are anatomically separable

Previous studies have demonstrated stark differences in the molecular identity and function of SPNs, even within the mNAcSh (Castro 2019, Al-Hasani 2015). Accordingly, the functionally defined SPN clusters that were identified in the variable sucrose and Pavlovian conditioning behavioral paradigms (see **Fig. 7-8**) were analyzed to relate the functional identity of SPNs to their relative position within the focal plane during in vivo 2-photon calcium imaging. The implanted endoscopic lenses were 600 um in diameter and thus, segregation of functional SPN clusters within the imaging focal plane represents hundreds of microns of anatomical separation within mNAcSh. Visualizing the spatial distribution of functionally defined SPN clusters revealed that reward-inhibited SPNs were specific to anterior portions of the focal plane, for the variable sucrose and Pavlovian imaging sessions (**Fig. 9A-J**). The observation that reward-inhibited SPNs are anterior to other cue and reward-excited SPN clusters is consistent with recent studies showing that incoming glutamatergic innervation from basolateral amygdala, ventral hippocampus and medial thalamus to anterior NAc is decreased during reward consumption and

glutamatergic innervation to posterior NAc is increased during reward consumption (Reed 2018). This finding strongly suggests that differences in glutamatergic innervation in NAc along the anterior-posterior axis result in differences in SPN reward processing and that the anterior-posterior division is a continuum of individual neurons rather than a strict boundary.

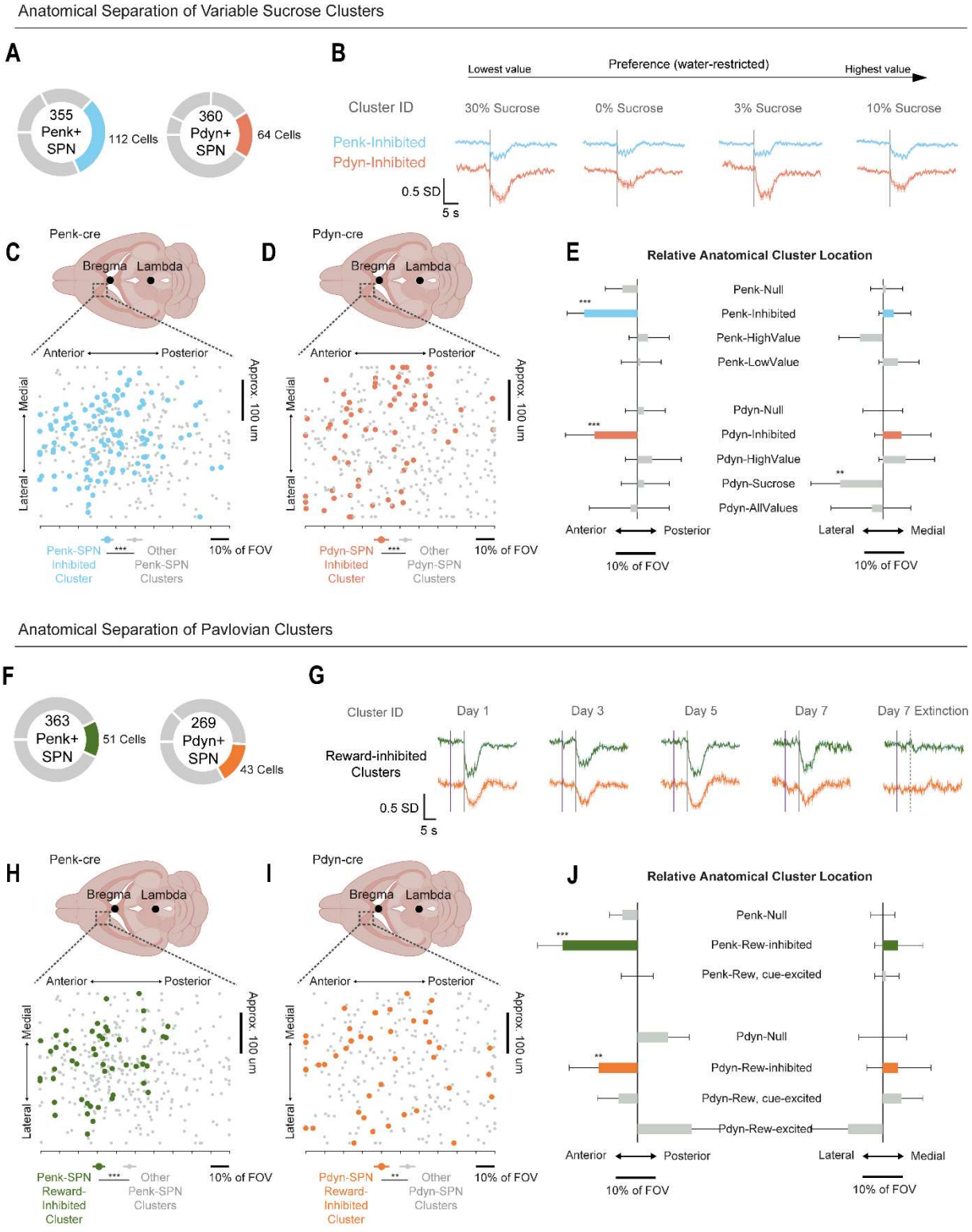

Anatomical Separation of Variable Sucrose Clusters
A
355 Penk+ SPN
112 Cells
360 Pdyn+ SPN
64 Cells
B
Lowest value
Preference (water-restricted)
Highest value
Cluster ID
30% Sucrose
0% Sucrose
3% Sucrose
10% Sucrose
Penk-Inhibited
Pdyn-Inhibited
0.5 SD
5 s
C
Penk-cre
Bregma
Lambda
Anterior
Posterior
Medial
Lateral
Approx. 100 um
Penk-SPN Inhibited Cluster
***
Other Penk-SPN Clusters
10% of FOV
D
Pdyn-cre
Bregma
Lambda
Anterior
Posterior
Medial
Lateral
Approx. 100 um
Pdyn-SPN Inhibited Cluster
***
Other Pdyn-SPN Clusters
10% of FOV
E
Relative Anatomical Cluster Location
Penk-Null
Penk-Inhibited
Penk-HighValue
Penk-LowValue
Pdyn-Null
Pdyn-Inhibited
Pdyn-HighValue
Pdyn-Sucrose
Pdyn-AllValues
***
***
**
Anterior
Posterior
10% of FOV
Lateral
Medial
10% of FOV
Anatomical Separation of Pavlovian Clusters
F
363 Penk+ SPN
51 Cells
269 Pdyn+ SPN
43 Cells
G
Cluster ID
Day 1
Day 3
Day 5
Day 7
Day 7 Extinction
Reward-inhibited Clusters
0.5 SD
5 s
H
Penk-cre
Bregma
Lambda
Anterior
Posterior
Medial
Lateral
Approx. 100 um
Penk-SPN Reward-Inhibited Cluster
***
Other Penk-SPN Clusters
10% of FOV
I
Pdyn-cre
Bregma
Lambda
Anterior
Posterior
Medial
Lateral
Approx. 100 um
Pdyn-SPN Reward-Inhibited Cluster
**
Other Pdyn-SPN Clusters
10% of FOV
J
Relative Anatomical Cluster Location
Penk-Null
Penk-Rew-inhibited
Penk-Rew, cue-excited
Pdyn-Null
Pdyn-Rew-inhibited
Pdyn-Rew, cue-excited
Pdyn-Rew-excited
***
**
Anterior
Posterior
10% of FOV
Lateral
Medial
10% of FOV

Figure 9: Reward-inhibited SPN clusters are anatomically anterior to other SPN clusters

(A) Fraction of all tracked enk-SPNs and dyn-SPNs that are part of reward-inhibited clusters. (B) Mean activity of SPN clusters that were consistently transiently inhibited during all variable sucrose reward conditions. (C) Position of individual enk-SPNs in focal plane during in vivo 2-photon calcium imaging sessions; arranged by variable sucrose response cluster (n = 4 clusters, 7 animals; one-way ANOVA: ***p < 0.001). (D) Position of individual dyn-SPNs in focal plane during in vivo 2-photon calcium imaging sessions; arranged by variable sucrose response cluster (n = 5 clusters, 11 animals; one-way ANOVA: ***p < 0.001). (E) Relative anatomical location of SPN variable sucrose response clusters. (F) Fraction of all tracked enk-SPNs and dyn-SPNs that are part of Pavlovian reward-inhibited clusters. (G) Mean activity of SPN clusters that were consistently transiently inhibited during reward throughout 7 days of Pavlovian conditioning. (H) Position of individual enk-SPNs in focal plane during in vivo 2-photon calcium imaging sessions; arranged by Pavlovian conditioning response cluster (n = 3 clusters, 7 animals; one-way ANOVA: ***p < 0.001). (I) Position of individual dyn-SPNs in focal plane during in vivo 2-photon calcium imaging sessions; arranged by Pavlovian conditioning response cluster (n = 4 clusters, 9 animals; one-way ANOVA: **p = 0.004). (J) Relative anatomical location of SPN Pavlovian conditioning response clusters.

3.5 Low-value selective mNAcSh SPNs project to VP

Distinct subpopulations of Pdyn-SPNs and Penk-SPNs are recruited in response to rewards of different value (see **Fig. 7**). This suggests that different reward values are encoded by distinct neuronal information streams upstream of mNAcSh SPNs. However, we do not know if the stratification of reward value information in mNAcSh is relayed to downstream brain regions via major mNAcSh efferent projections. To explore this possibility, we first retrotraced the major output projections from mNAcSh Penk-SPNs and Pdyn-SPNs using AAVretro viral techniques, retrobeads and fluorescent in situ hybridization (FISH). In congruence with previous literature (Creed 2016), Penk-SPNs were found to densely project to ventral pallidum (VP) and somewhat to lateral hypothalamus (LH) and these circuits had minimal collateralization (**Fig. S13B**). It was also found that Pdyn-SPNs densely project to LH and sparsely to VTA with minimal

collateralization (**Fig. S13A**), as has been previously reported of D1R expressing SPNs (Thoeni 2020).

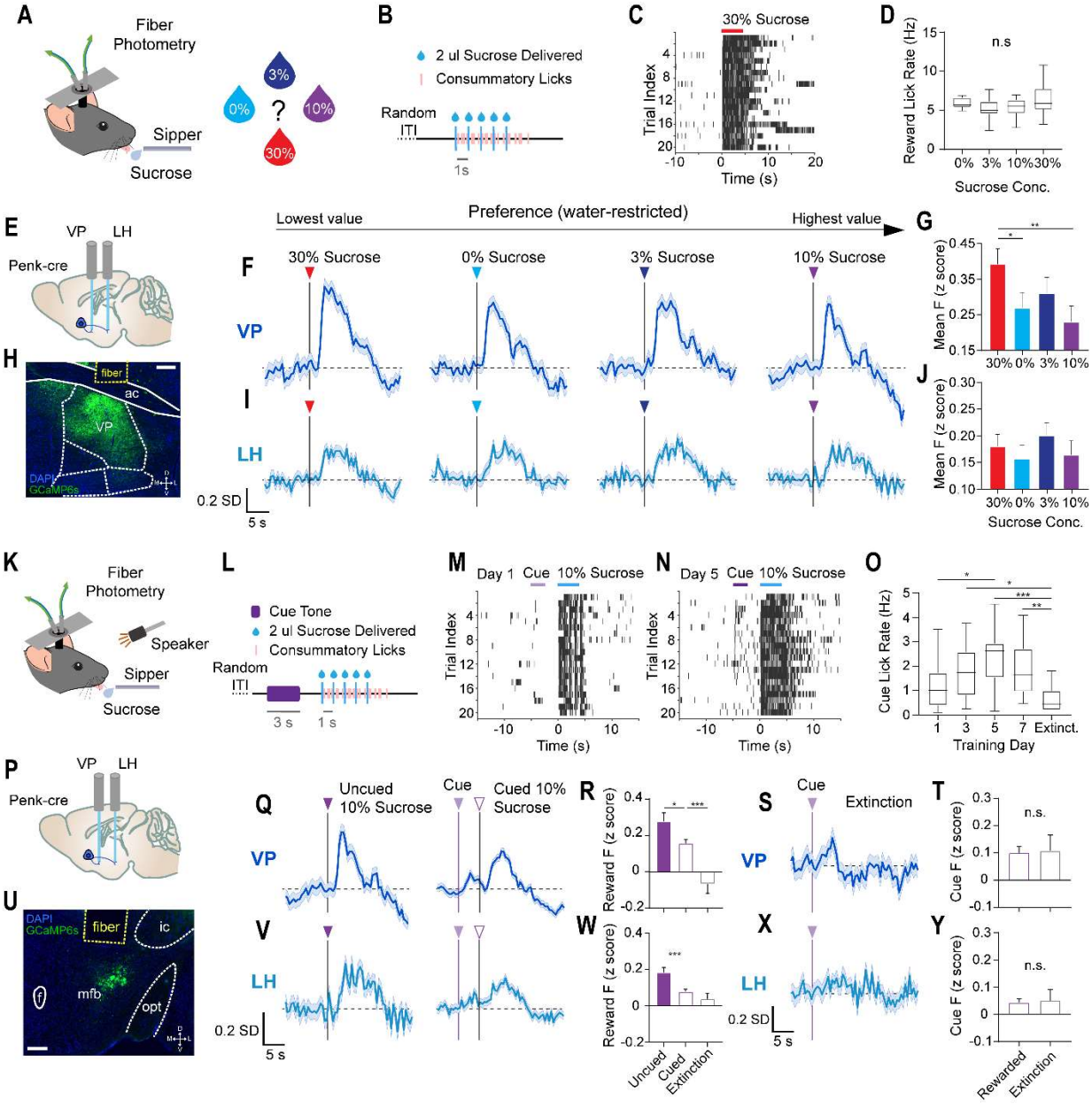


Figure 10: mNAcSh-VP^{enk} parallels low-value reward selectivity and cue-excitation of enk-SPN cluster (A) Schematic of simultaneous two-site fiber photometry during consumption of variable sucrose reward. (B) Time-course of sucrose reward delivery for a given trial. Droplets were either 0%, 3%, 10% or 30% sucrose in water. (C) Representative lick raster depicting individual licks that were detected by contact licometer during behavior and photometry session. (D) Mean lick rate during reward epoch (0-8 seconds from reward onset) for each sucrose reward condition (n = 15 animals; one-way ANOVA: p = 0.280). (E) Sagittal section cartoon displaying target injection region, cell-type and fiber implant locations. (F) Mean mNAcSh-VP^{enk} reward-response for every trial of each variable sucrose reward condition. (G) Mean mNAcSh-VP^{enk} activity over reward response epoch (2-12 seconds from reward-onset) for each condition (n = 9 animals, 190+ trials; one-way ANOVA: 30% vs 0%, *p = 0.044, 30% vs 10%, **p = 0.009). (H) Representative coronal section showing enk-SPN terminals in VP expressing GCaMP6s underneath the fiber implant (scale bar = 200 um). (I) Mean mNAcSh-LH^{enk} reward-response for every trial of each variable sucrose reward condition. (J) Mean mNAcSh-LH^{enk} activity over reward response epoch (2-12 seconds from reward-onset) for each condition (n = 6 animals, 127+ trials; one-way ANOVA: p = 0.6167). (K) Schematic of simultaneous two-site fiber photometry during Pavlovian cue-reward conditioning. (L) Time-course of auditory cue and subsequent sucrose reward delivery for a given trial. (M and N) Representative lick raster depicting individual licks that were detected by contact licometer during cue and reward when animals were cue-naïve (M) and after conditioning (N). (O) Mean lick rate during cue epoch (0-5 seconds from cue-onset) for each day of Pavlovian conditioning and extinction (n = 15 animals; one-way ANOVA: Day 1 vs Day 5, *p = 0.041, Day 3 vs Ext., *p = 0.011, Day 5 vs Ext., ***p < 0.001, Day 7 vs Ext., **p = 0.005). (P) Sagittal section cartoon displaying target injection region, cell-types and fiber implant locations. (Q) Mean mNAcSh-VP^{enk} reward-response for every uncued 10% sucrose trial and every cued 10% sucrose trial. (R) Mean mNAcSh-VP^{enk} activity over reward response epoch (2-12 seconds from reward-onset) for uncued 10% sucrose, cued 10% sucrose and extinction trials (n = 9 animals, 96+ trials; one-way ANOVA: Uncued vs Cued, *p = 0.014, Cued vs Ext., ***p < 0.001). (S) Mean mNAcSh-VP^{enk} reward-response for every trial of day 7 Pavlovian extinction. (T) Mean mNAcSh-VP^{enk} activity over cue response epoch (2-5 seconds from cue-onset) for cued 10% sucrose and extinction trials (n = 9 animals, 96+ trials; two sample t-test: p = 0.909). (U) Representative coronal section showing enk-SPN terminals in LH expressing GCaMP6s underneath the fiber implant (scale bar = 200 um). (V) Mean mNAcSh-LH^{enk} reward-response for every uncued 10% sucrose trial and every cued 10% sucrose trial. (W) Mean mNAcSh-LH^{enk} activity over reward response epoch (2-12 seconds from reward-onset) for uncued 10% sucrose, cued 10% sucrose and extinction trials (n = 6 animals, 64+ trials; one-way ANOVA: Uncued vs Cued, ***p < 0.001). (X) Mean mNAcSh-LH^{enk} reward-response for every trial of day 7 Pavlovian extinction. (Y) Mean mNAcSh-LH^{enk} activity over cue response epoch (2-5 seconds from cue-onset) for cued 10% sucrose and extinction trials (n = 6 animals, 64+ trials; two sample t-test: p = 0.851).

To determine if these major SPN efferent circuits were selectively recruited to particular reward value conditions and not others, we conducted simultaneous dual-site fiber photometry recordings while animals consumed the 4 reward value conditions (0%, 3%, 10% or 30% sucrose). Penk-cre mice were first injected with 300 nl of AAVdj-EF1a-DIO-GCaMP6s in the mNAcSh and then implanted 400 um diameter fiber optic cannulas into the VP and LH (**Fig. 10E**). Simultaneous recording of calcium dynamics at the axon terminals of two mNAcSh efferents allowed us to make direct comparisons of reward value-related activity within-subject. Photometry animals were tested in a head-restrained behavior paradigm identical to the 2-photon calcium imaging experiments (**Fig. 10A-B**). Notably, fiber photometry animals also showed no significant differences in their licking for the 4 different reward conditions given the small 10 ul reward quantities per trial (**Fig. 10D**), allowing for comparison of reward-evoked activity during uniform behavior between conditions.

For the Penk-cre mice, both the mNAcSh-VP^{enk} and mNAcSh-LH^{enk} circuits had reward-excitation to all of the 4 sucrose value conditions. However, the magnitude of mNAcSh-VP^{enk} reward-excitation was significantly greater for the lower-value 30% sucrose reward condition than the higher-value reward conditions (**Fig. 10F-G**). This greater excitation to the low-value condition directly parallels the low-value reward selective Penk-SPN cluster (see **Fig. 7L**).

The dual site fiber photometry animals were also tested in a 7 day Pavlovian conditioning paradigm identical to the head-restrained paradigm used during 2-photon calcium imaging (see **Fig. 8B-C**). mNAcSh-VP^{enk} was found to be cue-excited throughout Pavlovian conditioning in addition to having consistent 10% sucrose reward-excitation

(**Fig. 10Q**). Furthermore, the cue-excitation of mNAcSh-VP^{enk} persisted even during the 10 extinction trials where the reward was omitted (**Fig. 10S-T**). The magnitude of mNAcSh-VP^{enk} reward-excitation was also significantly lower when animals were expecting the cued 10% sucrose reward than when they receive uncued 10% sucrose reward (**Fig. 10R**). This mNAcSh-VP^{enk} response profile within the Pavlovian conditioning paradigm parallels the cue and reward-excitation that was observed in Penk-SPNs cue and reward-excited clusters (see **Fig. 8K**). Given our extensive calcium imaging-based characterization of 350+ Penk-SPNs in mNAcSh, this suggests that the mNAcSh-VP^{enk} circuit is likely largely composed of the low-value reward selective Penk-SPN subpopulation in mNAcSh. Thus, the separation of distinct reward value information streams in mNAcSh is continued through the mNAcSh-VP^{enk} circuit, which preferentially encodes low-value rewards.

3.6 mNAcSh-VP circuit causally drives low reward value

mNAcSh-VP^{enk} was found to have greater reward-excitation to the lower value 30% sucrose reward condition than the higher value conditions. However, it is still unclear whether mNAcSh-VP^{enk} reward-related activity causally contributes to consummatory behavior. To test the role of mNAcSh-VP^{enk} in casually driving consummatory behavior, we injected Penk-cre mice with 300 nl of AAV5-DIO-parapinopsin(PPO) (Copits 2021) or AAV5-DIO-ChR2 bilaterally into mNAcSh and implanted 200 um diameter fiber optic cannulas bilaterally into VP (**Fig. 11E,I**). After recovering from surgery, animals were water-restricted and trained to lick 10% sucrose reward while reversibly immobilized and head-restrained (**Fig. 11A**). After 4 days of lick training, all animals displayed consistent

lick behavior in response to variable interval sucrose presentation and were then put through 2 consecutive test days.

Each test day involved one 30 minute, 20 trial session in which the animals were presented with either 10% or 30% sucrose reward. The order of sucrose condition days was pseudorandomized and counterbalanced across animals. For consistency with the imaging and photometry behavior paradigms, animals were presented with 5 droplets of sucrose reward per trial with a total aggregate volume of 10 ul and 1 second between droplet delivery (**Fig. 11B**). Animals immediately licked the presented sucrose reward, with an average licking bout durations of 7-8 seconds (**Fig. 11C**). On a pseudorandom 50% of the trials within a session, time-locked optogenetic manipulation of mNAcSh-VP^{enk} activity would be paired with reward delivery, via 4 seconds of 10 Hz blue light delivery (**Fig. 11B**). 10Hz optogenetic stimulation frequency was chosen because it is physiologically relevant level of neural activity for SPNs during phasic reward-related activity (Taha 2005, Roitman 2005). Any differences in consummatory licking between no laser and laser trials were recorded via a contact licometer with millisecond precision (**Fig. 11D**) and all differences were evaluated within subject to account for the baseline variability in lick rate between animals. To control for possible behavioral effects from the delivery of the blue light itself, cre- littermate controls were also injected, implanted, and tested in identical conditions.

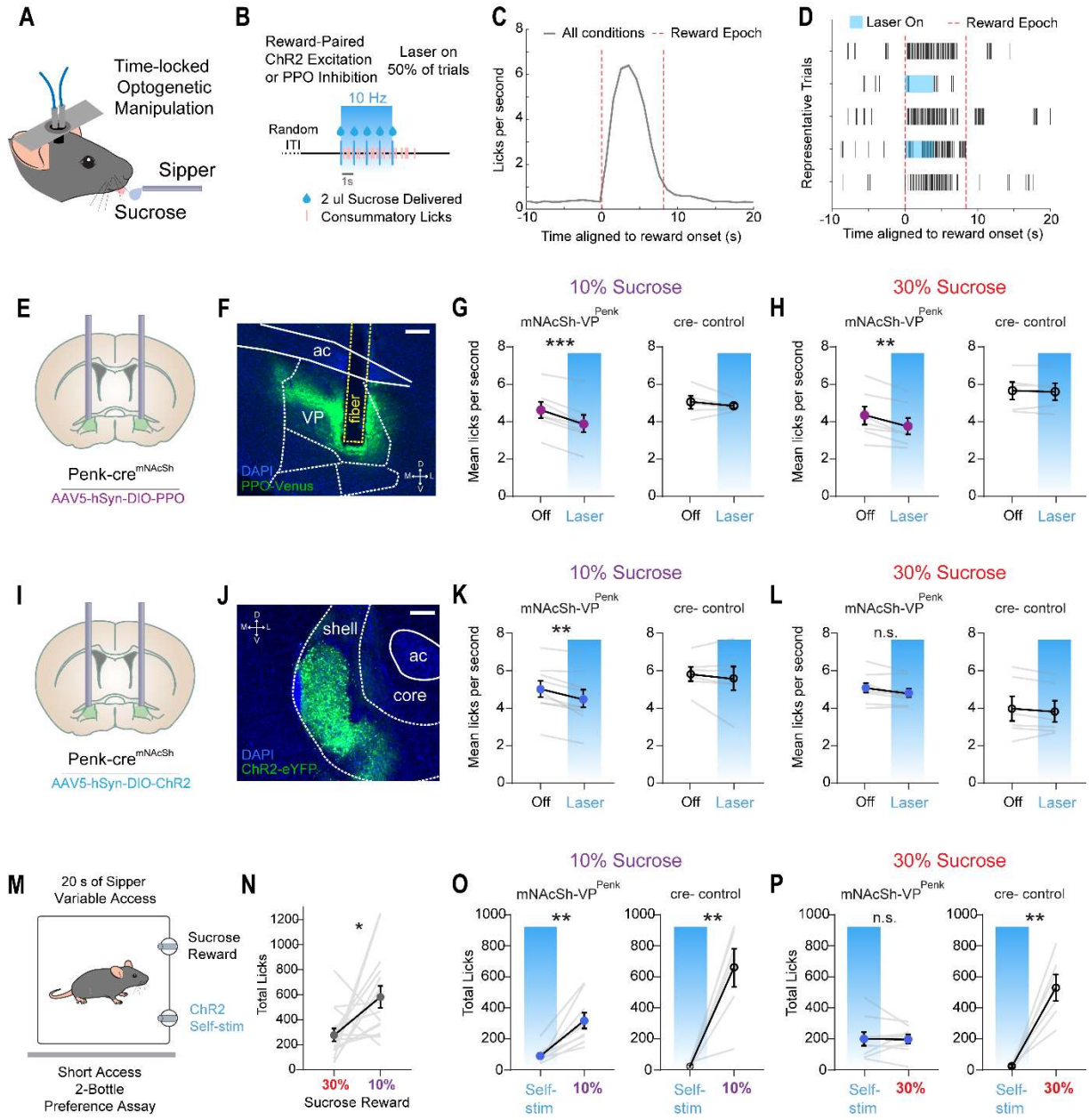

A
Time-locked
Optogenetic
Manipulation
Sipper
Sucrose
B
Reward-Paired
ChR2 Excitation
or PPO Inhibition
Laser on
50% of trials
10 Hz
Random
ITI
1s
2 ul Sucrose Delivered
Consummatory Licks
C
All conditions
Reward Epoch
Licks per second
Time aligned to reward onset (s)
D
Laser On
Reward Epoch
Representative Trials
Time aligned to reward onset (s)
E
Penk-cre mNAcSh
AAV5-hSyn-DIO-PPO
F
ac
VP
fiber
DAPI
PPO-Venus
G
10% Sucrose
mNAcSh-VP Penk
cre- control
Mean licks per second
***
Off
Laser
H
30% Sucrose
mNAcSh-VP Penk
cre- control
**
I
Penk-cre mNAcSh
AAV5-hSyn-DIO-ChR2
J
shell
ac
core
DAPI
ChR2-eYFP
K
10% Sucrose
mNAcSh-VP Penk
cre- control
**
L
30% Sucrose
mNAcSh-VP Penk
cre- control
n.s.
M
20 s of Sipper
Variable Access
Sucrose
Reward
ChR2
Self-stim
Short Access
2-Bottle
Preference Assay
N
Total Licks
*
30%
10%
Sucrose Reward
O
10% Sucrose
mNAcSh-VP Penk
cre- control
Total Licks
**
Self-
stim
10%
P
30% Sucrose
mNAcSh-VP Penk
cre- control
n.s.
**
Self-
stim
30%

Figure 11: mNAcSh-VP^{enk} causally drives low-positive reward value

(A) Schematic of timelocked-optogenetic neural activity manipulation during consumption of variable sucrose reward. (B) Time-course of sucrose reward delivery for a given trial. 50% of trials also had 4 seconds of 10 Hz optogenetic stimulation concurrent with reward delivery. (C) Lick rate distribution relative to reward-onset for all conditions. Comparison of lick rate between conditions was restricted to the indicated reward epoch (0-8 seconds from reward-onset). (D) Representative lick raster displaying individual measured licks during trials without stimulation and trials with stimulation. (E) Coronal section cartoon displaying target injection region, cell-type and bilateral fiber implant location. (F) Representative coronal section showing enk-SPN terminals in VP expressing PPO underneath optic fiber implant (scale bar = 200 um). (G) Within-session comparison of 10% sucrose consumption between trials with and without reward-paired optogenetic inhibition (n = 7 treatment, 5 control animals; pairwise t-test: ***p < 0.001 , p = 0.418). (H) Within-session comparison of 30% sucrose consumption between trials with and without reward-paired optogenetic inhibition (n = 7 treatment, 5 control animals; pairwise t-test: **p = 0.003 , p = 0.619). (I) Coronal section cartoon displaying target injection region, cell-type and bilateral fiber implant location. (J) Representative coronal section showing enk-SPN terminals in VP expressing ChR2 underneath optic fiber implant (scale bar = 200 um). (K) Within-session comparison of 10% sucrose consumption between trials with and without reward-paired optogenetic excitation (n = 9 treatment, 6 control animals; pairwise t-test: **p = 0.003 , p = 0.437). (L) Within-session comparison of 30% sucrose consumption between trials with and without reward-paired optogenetic excitation (n = 9 treatment, 6 control animals; pairwise t-test: p = 0.057 , p = 0.327). (M) Schematic of brief access 2-bottle choice task: animal makes choice between sucrose reward sipper and lick-triggered 2 second, 10 Hz optogenetic self-stimulation. (N) Preference between 10% sucrose reward sipper and 30% sucrose reward sipper (n = 15 animals; pairwise t-test: *p = 0.021). (O) Preference between 10% sucrose reward and optogenetic self-stimulation (n = 9 treatment, 6 control animals; pairwise t-test: **p = 0.001 , **p = 0.004). (P) Preference between 30% sucrose reward and optogenetic self-stimulation (n = 9 treatment, 6 control animals; pairwise t-test: p = 0.956 , **p = 0.002).

Time-locked photoinhibition of mNAcSh-VP^{enk} during consumption significantly reduced the reward epoch lick rate regardless of whether animals were consuming 10% or 30% sucrose reward (**Fig. 11G-H**). This demonstrates that mNAcSh-VP^{enk} reward-excitation is required for normal consummatory licking. Since mNAcSh Penk-SPNs specifically encode reward value information and mNAcSh-VP^{enk} is reward-excited, the reduction in consummatory licking during mNAcSh-VP^{enk} photoinhibition may be due to complete disruption of reward-value information in this circuit.

Interestingly, time-locked photoexcitation of mNAcSh-VP^{enk} during consumption had an effect on consummatory licking that was reward value-dependent. Time-locked photoexcitation of mNAcSh-VP^{enk} reduced the consummatory licking of higher-value 10% sucrose reward (**Fig. 11K**) but did not significantly change the consumption of lower-value 30% sucrose reward (**Fig. 11L**). The fact that mNAcSh-VP^{enk} already has greater endogenous activity during the consumption of 30% sucrose reward (see **Fig. 10**) suggests that exogenous photoexcitation is not changing the activity of this circuit as much during consumption of low value reward. While in the higher value reward condition, mNAcSh-VP^{enk} has less endogenous activity and 10 Hz exogenous photoexcitation would more profoundly enhance the activity of the circuit. Perhaps the exogenous photoexcitation is driving this circuitry into a low-value reward state abruptly and overriding the higher-value reward encoding, leading to a lower perceived reward value and a measurable reduction in licking.

Notably, mNAcSh-VP^{enk} exogenous photoexcitation is not disrupting the motor action of licking itself, as no disruption of licking behavior was observed during the consumption of 30% sucrose reward (**Fig. 11L**). To further rule out any possible effects from motor action impairment, a set of additional experiments were designed in which mNAcSh-VP^{enk} was photoexcited or photoinhibited as animals freely moved through a behavioral arena in red light conditions (**Fig. S16H-I**). Movement velocity was not changed by mNAcSh-VP^{enk} photoinhibition (**Fig. S16K**), nor by mNAcSh-VP^{enk} photoexcitation (**Fig. S16M**), further suggesting that manipulation of this circuit during headfixed licking behavior was influencing the processing of reward information rather than disrupting the motor execution of licking.

If mNAcSh-VP^{enk} is causally driving a low positive reward value, then photoexcitation of this circuit alone should be inherently rewarding. To test this hypothesis, we again tested mice in our brief access 2-bottle choice paradigm (see **Fig. S7**). The mNAcSh-VP$^{enk:ChR2}$ optogenetics animals were trained to lick 10% sucrose reward from two sucrose spouts within a behavioral chamber (**Fig. 11M**). After training the mice to approach and lick the sucrose spouts for 5 days, we put the animals through 6 consecutive test days. Test days 1, 3, 5 were 'reset' days to neutralize any left versus right biases that resulted from other test conditions as described before (see **Fig. S7**). On test day 2, animals were given brief access to both a 10% and a 30% sucrose sipper. In accordance with our earlier results (see **Fig. S7**), mNAcSh-VP$^{enk:ChR2}$ optogenetics animals showed a preference for 10% sucrose over 30% sucrose (**Fig. 11N**). On test days 4 and 6, the animals were again tested for their preference between two sippers. However, one of the two sippers contained no liquid and licking it triggered 2 seconds of 10 Hz closed loop optogenetic photoexcitation (**Fig. 11M**). This stimulation duration and frequency was chosen to reasonably recapitulate mNAcSh-VP^{enk} activity during 30% sucrose reward licking. The other sipper that the animals could choose instead of the self-stimulation sipper contained either 10% or 30% sucrose as before. The order of test days in which animals got the 10% or 30% alternative sipper and the sipper location were pseudorandomized and counterbalanced across animals. This experiment was designed to test if animals find mNAcSh-VP^{enk} photoexcitation inherently rewarding and how the inherent motivational value of mNAcSh-VP^{enk} photoexcitation compared to the motivational value of the 10% and 30% sucrose conditions. It was found that mNAcSh-VP$^{enk:ChR2}$ animals licked the empty self-stimulation sipper significantly more than cre-

controls, demonstrating that mNAcSh-VP^{enk} photoexcitation has an inherent positive value. Furthermore, cre+ animals still licked the 10% sucrose sipper significantly more than the self-stimulation sipper (**Fig. 11O**) but had no significant preference between the 30% sucrose and the self-stimulation sippers (**Fig. 11P**). This allows us to assess the relative positive value of exogenous mNAcSh-VP^{enk} excitation and to observe that it is comparable to the positive value of 30% sucrose reward but measurably lower than the value of the 10% sucrose reward.

Overall, these complementary optogenetics experiments demonstrate that mNAcSh-VP^{enk} causally drives a low positive reward value. This implies that not only is mNAcSh-VP^{enk} endogenously recruited during the consumption of lower value rewards, but that this endogenous activity during reward causally contributes to driving perceived reward value and reward consumption.

3.7 Chapter Summary

Here we demonstrate that distinct mNAcSh SPN ensembles encode rewards of different value. Separate subpopulations of mNAcSh SPN are recruited during the consumption of lower value rewards and higher value rewards. We find that this stratification of reward value information in mNAcSh is then propagated through efferent projections, with lower value reward signals relayed to VP. Optogenetic activation of the mNAcSh-VP^{enk} circuit was found to cause no change in consummatory behavior when the animals were consuming lower value reward but caused reductions in the consumption of higher value reward. Self-stimulation experiments further revealed that mNAcSh-VP^{enk} activation itself was inherently rewarding and had an incentive value approximately equivalent to the lower value natural reward. These findings present a

detailed mechanism for how mNAcSh integrates incoming reward-related information, represents reward information on the basis of incentive value and coordinates the appropriately motivated consummatory behavior.

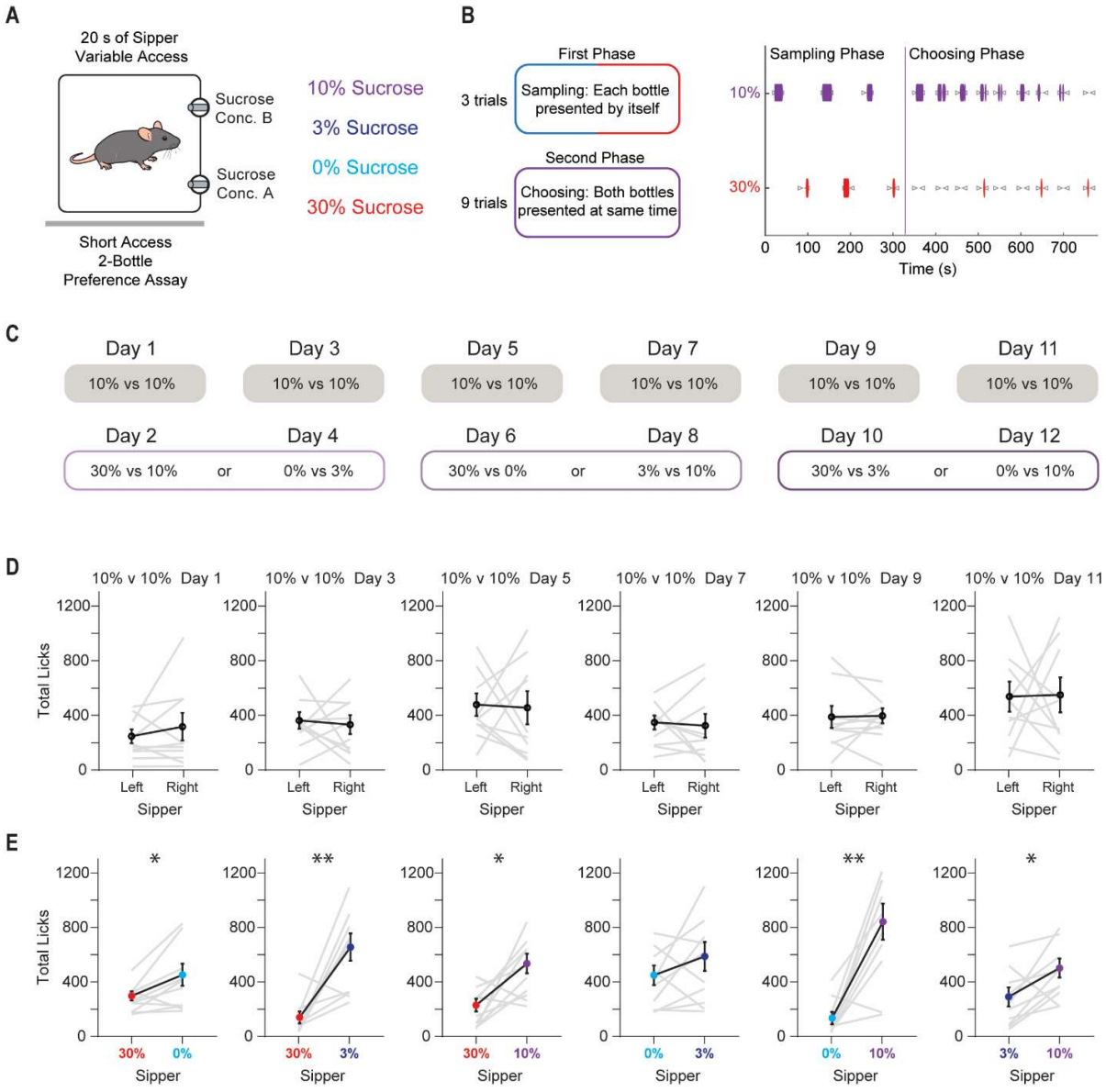
It remains unresolved how much separation of reward value information is present in brain regions upstream of accumbens and which afferents are supplying this information. There is recent evidence that neurons in VP encode reward value and project to NAc (Richard 2016, Vachez 2021). Furthermore, reward value signals in VP may even be present earlier than in NAc (Ottenheimer 2018). These previous findings and our current results suggest that some reward value information may by originally transmitted from VP to mNAcSh, where it is further amplified and propagated.

Reward-related signaling in NAc core has been found to be largely mediated by mesolimbic dopamine release (Phillips 2003, Saunders 2018). However, it remains unclear how the differences in dopaminergic innervation of NAc core and mNAcSh contribute to differences in reward processing. Our current results demonstrate that mNAcSh SPNs are cue-excited but do not encode reward expectation as is canonically observed in NAc core (Roitman 2005, Day 2006). DA is also likely not released as robustly to primary rewards or reward-predictive cues in mNAcSh as it is in NAc core (Saddoris 2015, de Jong 2018). We also find that many Penk-SPNs, that largely co-express inhibitory D2Rs, (Bertran 2008, Castro 2019) are excited during reward and reward-predictive cues (see **Fig. 7-8**). This may be attributable to unique dopamine receptor expression in mNAcSh, where D3-type excitatory dopamine receptors are highly enriched (Pich 2015, Berke 2018).

Mu opioid receptor (MOR) activation in the basal forebrain has been thoroughly characterized for its localized robust induction of reward 'liking' and 'disliking'. In particular, MOR activation in anterior VP suppresses reward 'liking' while MOR activation in posterior VP enhances reward 'liking' (Smith 2007, Castro 2015). Penk-SPNs in mNAcSh likely release enkephalin, which is a high affinity endogenous ligand for MOR (Castro 2021, Jordan 2000). Here we find that mNAcSh Penk-SPNs densely project to VP and are activated by lower value rewards. Enkephalinergic activation of MOR from Penk-SPNs may be a potential mechanism by which reward value signals are transduced in VP. This potential opioidergic mechanism remains unresolved by the present study but presents an interesting avenue for further investigation.

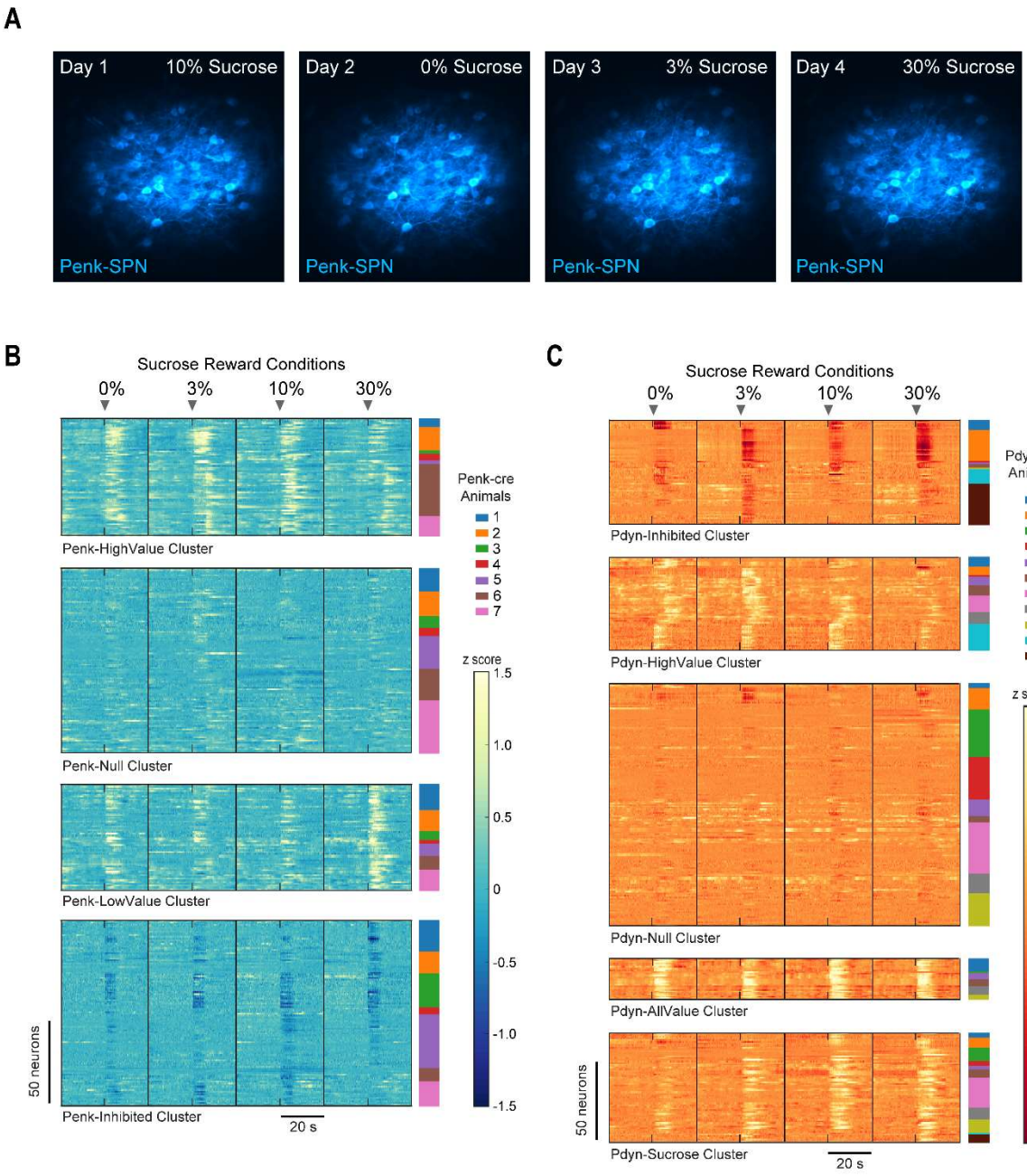
Although further work remains to understand the neurotransmitter and circuit-level mechanisms underlying reward value computation in the NAc-VP reward system, this works provides critical insight that mNAcSh robustly represents distinct reward values and transmits this information through a pallidal efferent. This critical neural encoding mechanism may play a central role in how NAc integrates incoming reward value information and uses it to drive motivated and consummatory behavior.

3.8 Supplementary Information



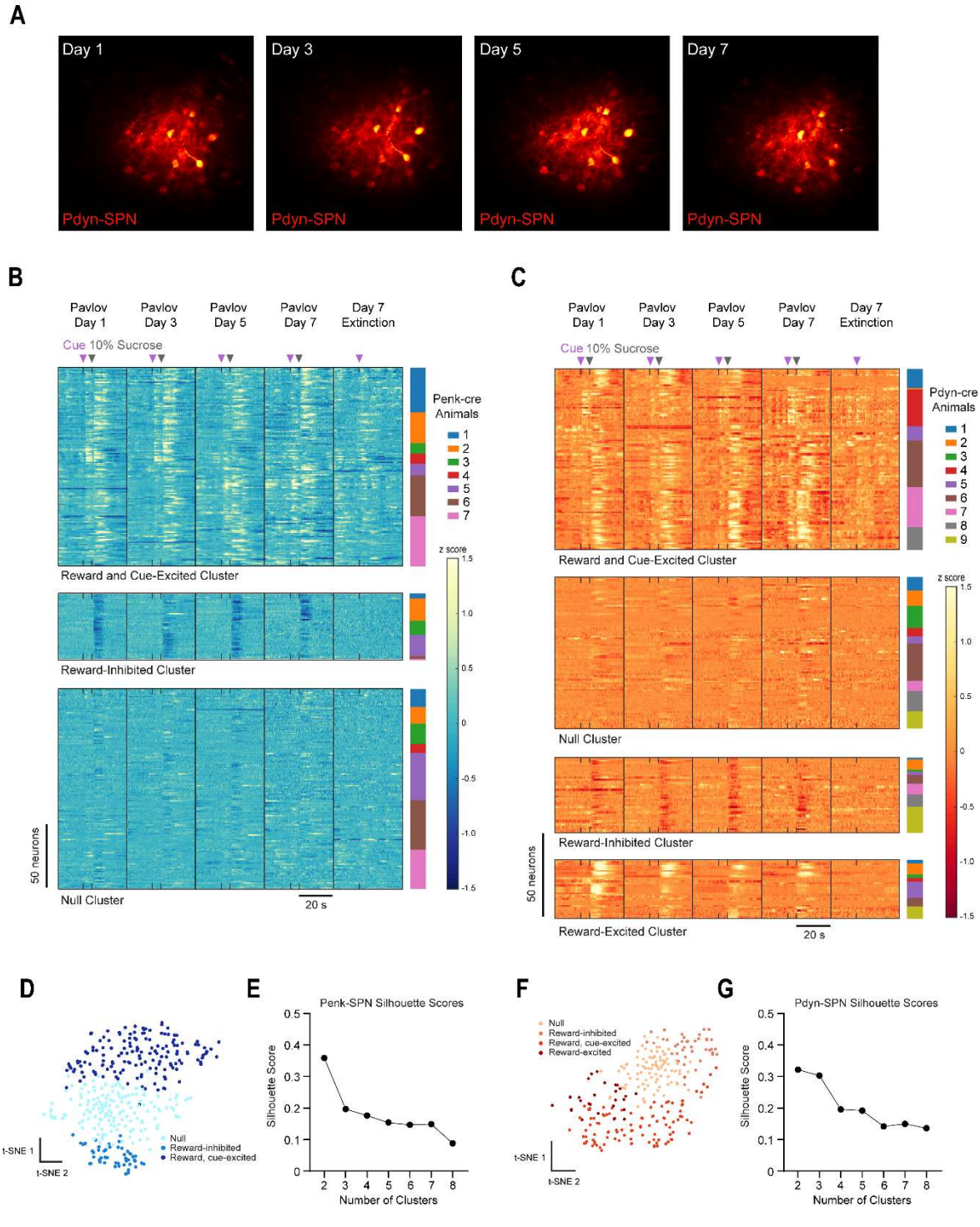
Supp. Figure S7: Sucrose concentration preference of C57 mice under water restriction

(A) Schematic of brief access 2-bottle choice task: animal makes choice between two sippers containing different concentrations of sucrose reward. (B) Time-course of task and representative lick data throughout session. (C) Brief access 2-bottle choice task schedule: Alternating between pseudorandomized test days and 10% versus 10% days to neutralize spatial biases that arise during test days. (D) Preference between left and right sippers containing identical rewards on spatial bias neutralization days (n = 10 animals, pairwise t-test: p = 0.411, p = 0.760, p = 0.885, p = 0.784, p = 0.911, p = 0.949). (E) Preference between each pairing of the 4 sucrose concentration conditions; graphs rearranged in order of preference (n = 10 animals, pairwise t-test: *p = 0.031, **p = 0.003, *p = 0.012, p = 0.227, **p = 0.001, p = *0.026).


A
Day 1
10% Sucrose
Penk-SPN
Day 2
0% Sucrose
Penk-SPN
Day 3
3% Sucrose
Penk-SPN
Day 4
30% Sucrose
Penk-SPN
B
Sucrose Reward Conditions
0%
3%
10%
30%
Penk-HighValue Cluster
Penk-Null Cluster
Penk-LowValue Cluster
Penk-Inhibited Cluster
50 neurons
20 s
Penk-cre Animals
1
2
3
4
5
6
7
z score
1.5
1.0
0.5
0
-0.5
-1.0
-1.5
C
Sucrose Reward Conditions
0%
3%
10%
30%
Pdyn-Inhibited Cluster
Pdyn-HighValue Cluster
Pdyn-Null Cluster
Pdyn-AllValue Cluster
Pdyn-Sucrose Cluster
50 neurons
20 s
Pdyn-cre Animals
1
2
3
4
5
6
7
8
9
10
11
z score
1.5
1.0
0.5
0
-0.5
-1.0
-1.5
D
Penk-HighValue
Penk-LowValue
Penk-Null
Penk-Inhibited
0
0.2
0.4
0.6
0.8
1
Silhouette Coefficient Values
E
Penk-SPN Silhouette Scores
Silhouette Score
0.5
0.4
0.3
0.2
0.1
0
2
3
4
5
6
7
8
Number of Clusters
F
Pdyn-AllValue
Pdyn-HighValue
Pdyn-Sucrose
Pdyn-Null
Pdyn-Inhibited
0
0.2
0.4
0.6
0.8
1
Silhouette Coefficient Values
G
Pdyn-SPN Silhouette Scores
Silhouette Score
0.5
0.4
0.3
0.2
0.1
0
2
3
4
5
6
7
8
Number of Clusters

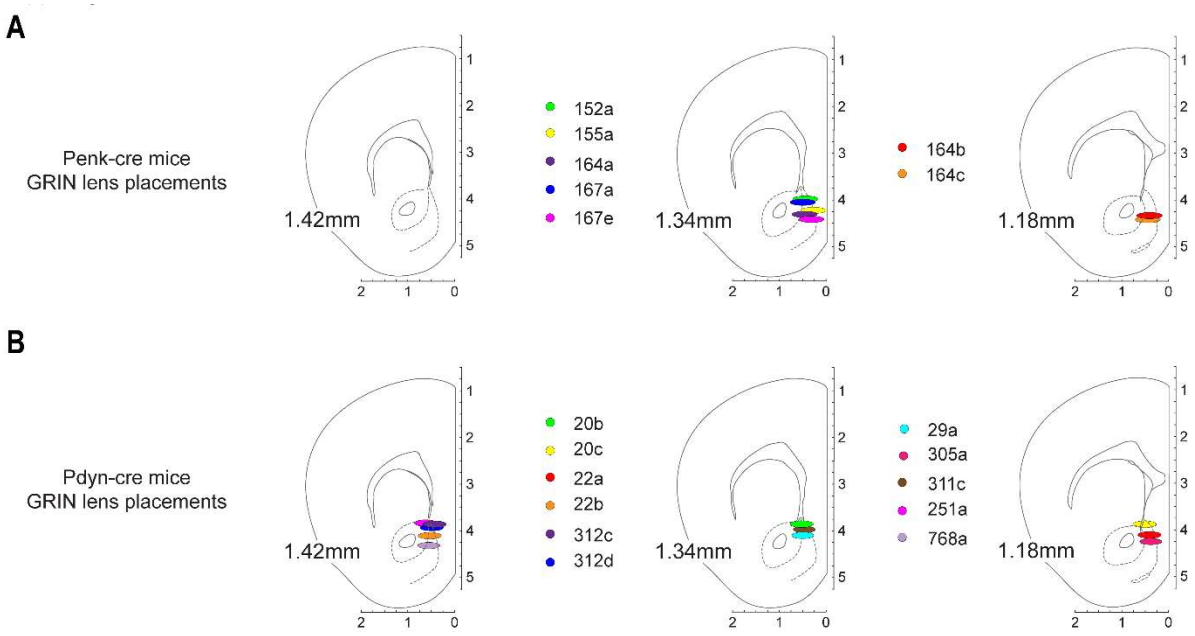
Supp. Figure S8: Individual SPN tracking and clustering based on response to variable sucrose conditions

(A) Mean pixel values across entire in vivo 2-photon calcium imaging session for a representative enk-cre animal to demonstrate the ability to resolve individual cell morphology and track neurons reliably across multiple days and conditions. (B) Mean response (20 trials per condition) of individual enk-SPNs to all sucrose concentration conditions. Individual tracked cells are arranged by their cluster ID and animal ID to show representation of different animals within the neuronal clusters. (C) Mean response (20 trials per condition) of individual dyn-SPNs to all sucrose concentration conditions. Individual tracked cells are arranged by their cluster ID and animal ID to show representation of different animals within the neuronal clusters. (D) Silhouette plot shown for optimal k-means clustering of enk-SPNs. (E) Silhouette scores for different enk-SPN cluster counts to determine the optimal number of clusters. (F) Silhouette plot shown for optimal k-means clustering of dyn-SPNs. (G) Silhouette scores for different dyn-SPN cluster counts to determine the optimal number of clusters.


A
Day 1
Day 3
Day 5
Day 7
Pdyn-SPN
B
Pavlov Day 1
Pavlov Day 3
Pavlov Day 5
Pavlov Day 7
Day 7 Extinction
Cue 10% Sucrose
Penk-cre Animals
Reward and Cue-Excited Cluster
Reward-Inhibited Cluster
Null Cluster
50 neurons
20 s
z score
C
Pdyn-cre Animals
Reward and Cue-Excited Cluster
Null Cluster
Reward-Inhibited Cluster
Reward-Excited Cluster
D
t-SNE 1
t-SNE 2
Null
Reward-inhibited
Reward, cue-excited
E
Penk-SPN Silhouette Scores
Silhouette Score
Number of Clusters
F
Reward-excited
G
Pdyn-SPN Silhouette Scores

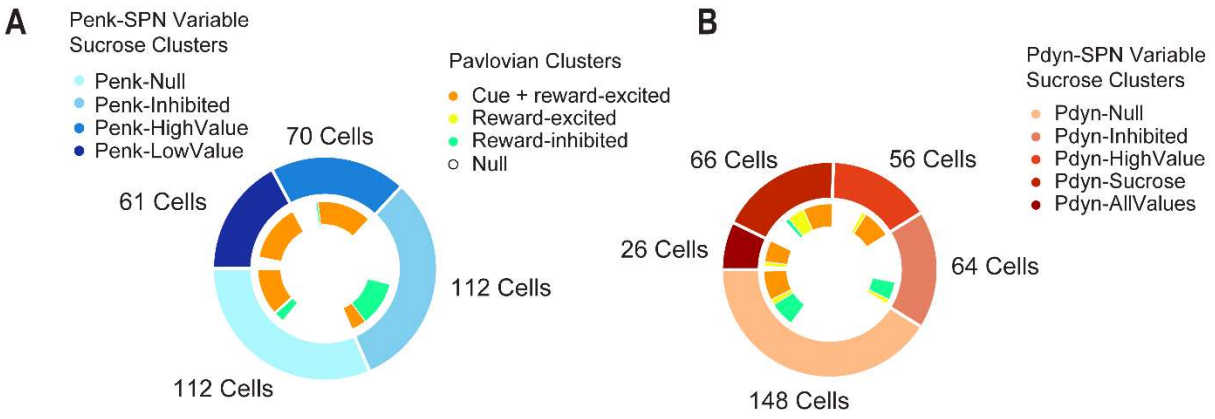
Supp. Figure S9: Individual SPN tracking and clustering based on response to Pavlovian conditioning

(A) Mean pixel values across entire in vivo 2-photon calcium imaging session for a representative dyn-cre animal to demonstrate the ability to resolve individual cell morphology and track neurons reliably across multiple days and conditions. (B) Mean response (20 trials per condition) of individual enk-SPNs to cue and reward for all Pavlovian conditioning days. Individual tracked cells are arranged by their cluster ID and animal ID to show representation of different animals within the neuronal clusters. (C) Mean response (20 trials per condition) of individual dyn-SPNs to cue and reward for all Pavlovian conditioning days. Individual tracked cells are arranged by their cluster ID and animal ID to show representation of different animals within the neuronal clusters. (D) t-sne plot of variance in activity of enk-SPN clusters during cue and reward for all Pavlovian conditioning days. Each dot represents an individual tracked neuron that composes the cluster. (E) Silhouette scores for different enk-SPN Pavlovian response cluster counts to determine the optimal number of clusters. (F) t-sne plot of variance in activity of dyn-SPN clusters during cue and reward for all Pavlovian conditioning days. Each dot represents an individual tracked neuron that composes the cluster. (G) Silhouette scores for different dyn-SPN Pavlovian response cluster counts to determine the optimal number of clusters.



Supp. Figure S10: GRIN lens placements

(A) Surgical placements of endoscopic GRIN lenses in NAc of enk-cre animals. (B) Surgical placements of endoscopic GRIN lenses in NAc of dyn-cre animals.



C

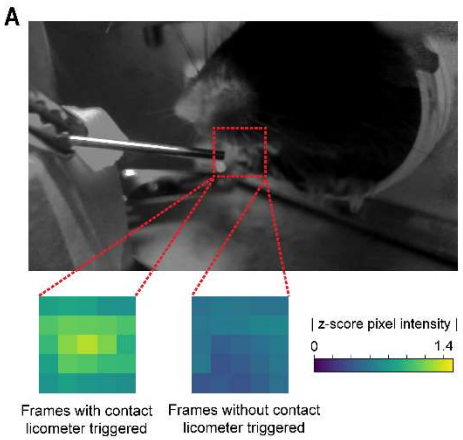
8 neurons not tracked	Penk-Null	Penk-Inhibited	Penk-HighValue	Penk-LowValue
Null	62	60	19	10
Reward-inhibited	9	39	2	1
Cue + reward-excited	41	13	49	50

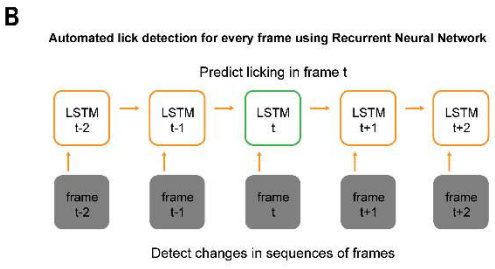
D

112 neurons not tracked	Pdyn-Null	Pdyn-Inhibited	Pdyn-HighValue	Pdyn-Sucrose	Pdyn-AllValues
Null	49	11	4	10	3
Reward-inhibited	22	16	0	4	0
Reward-excited	5	4	4	15	4
Cue + reward-excited	26	1	25	26	19

Supp. Figure S11: Overlap of variable sucrose clusters and Pavlovian conditioning clusters

(A) Fraction of enk-SPNs in each variable sucrose cluster (outer donut) that were found to be cue or reward-responsive (inner donut) within the context of the Pavlovian conditioning paradigm. (B) Fraction of dyn-SPNs in each variable sucrose cluster (outer donut) that were found to be cue or reward-responsive (inner donut) within the context of the Pavlovian conditioning paradigm. (C) Number of enk-SPNs that were tracked across entirety of variable sucrose and Pavlovian conditioning paradigms, organized by the intersection of their variable sucrose cluster and Pavlovian conditioning cluster. (D) Number of dyn-SPNs that were tracked across entirety of variable sucrose and Pavlovian conditioning paradigms, organized by the intersection of their variable sucrose cluster and Pavlovian conditioning cluster.

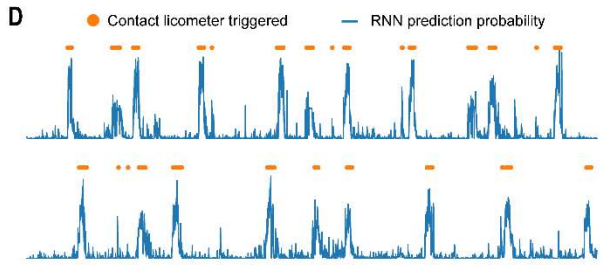


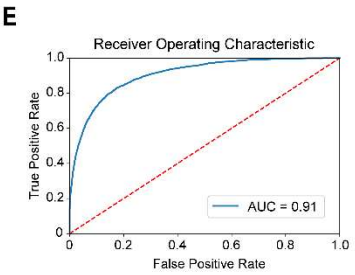


C

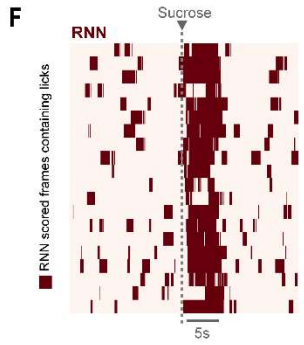
		Behavior Video	Simultaneous Contact Licometer	Manual Video Scoring
Supervised Training	48 videos	X	X	
Predicting	104 videos	X		
Manual Validation	104 videos	X		X

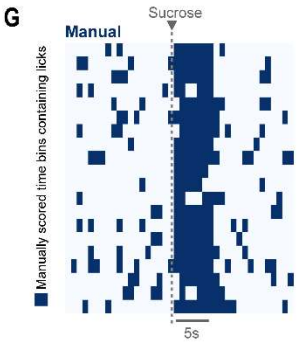
Supervised Training

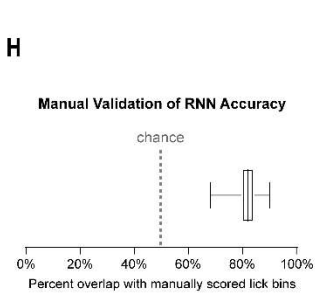




Manual Validation

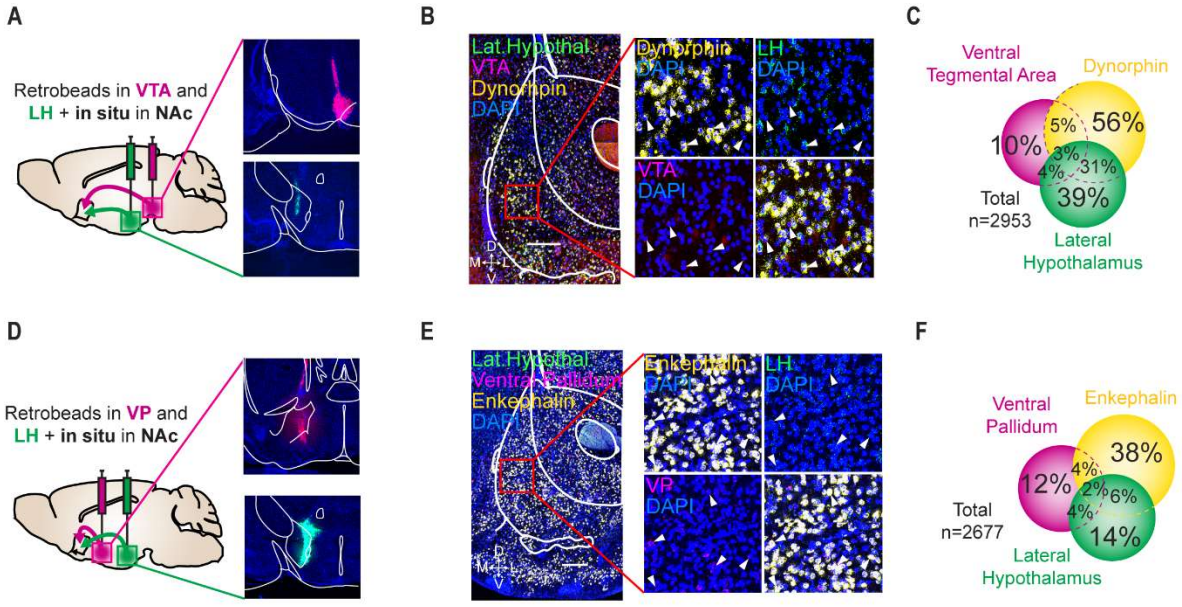






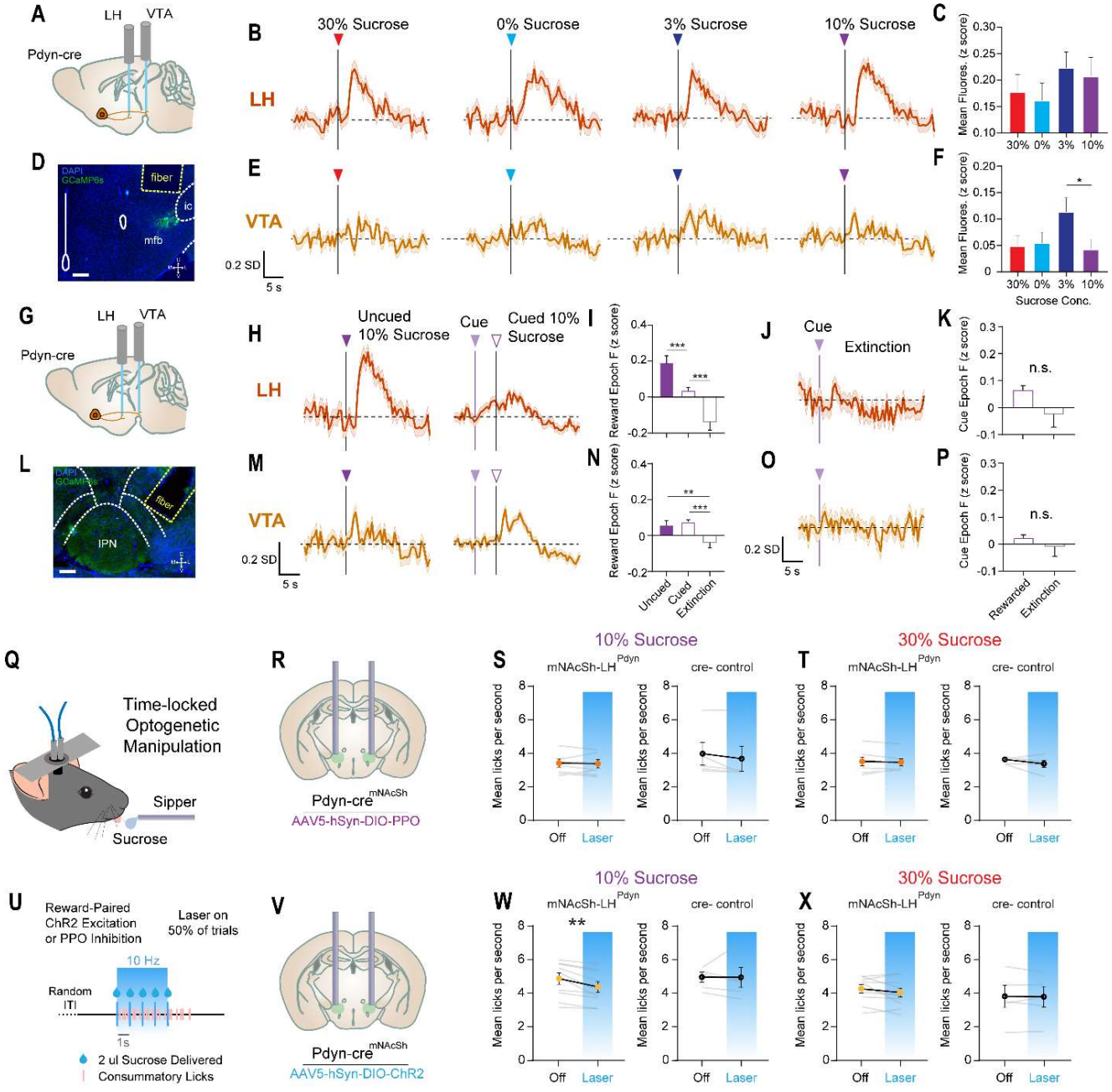
Supp. Figure S12: Training and Validation of Recurrent Neural Network for Lick Detection

(A) Representative frame from head-fixed behavior video. ROI near animals mouth cropped from video, spatially downsampled and normalized prior to frame classification. (B) Cartoon depicting organization of RNN that uses changes in previous and future frames to predict licking in the current frame. (C) Table depicting which data sets were used to initially train the RNN model using contact licometer information as frame labels and used to manually validate RNN model accuracy. (D) Representative sequential RNN prediction probabilities for when animal is licking (blue) overlaid with timestamps of animal licking from contact licometer (orange). (E) ROC curve showing specificity of RNN binary classification on entire 48 video training set. (F and G) Representative session frame classification by RNN model (F) and manual video scoring (G). (H) Overlap between manually scored video and RNN model video classification for all 104 behavioral videos.



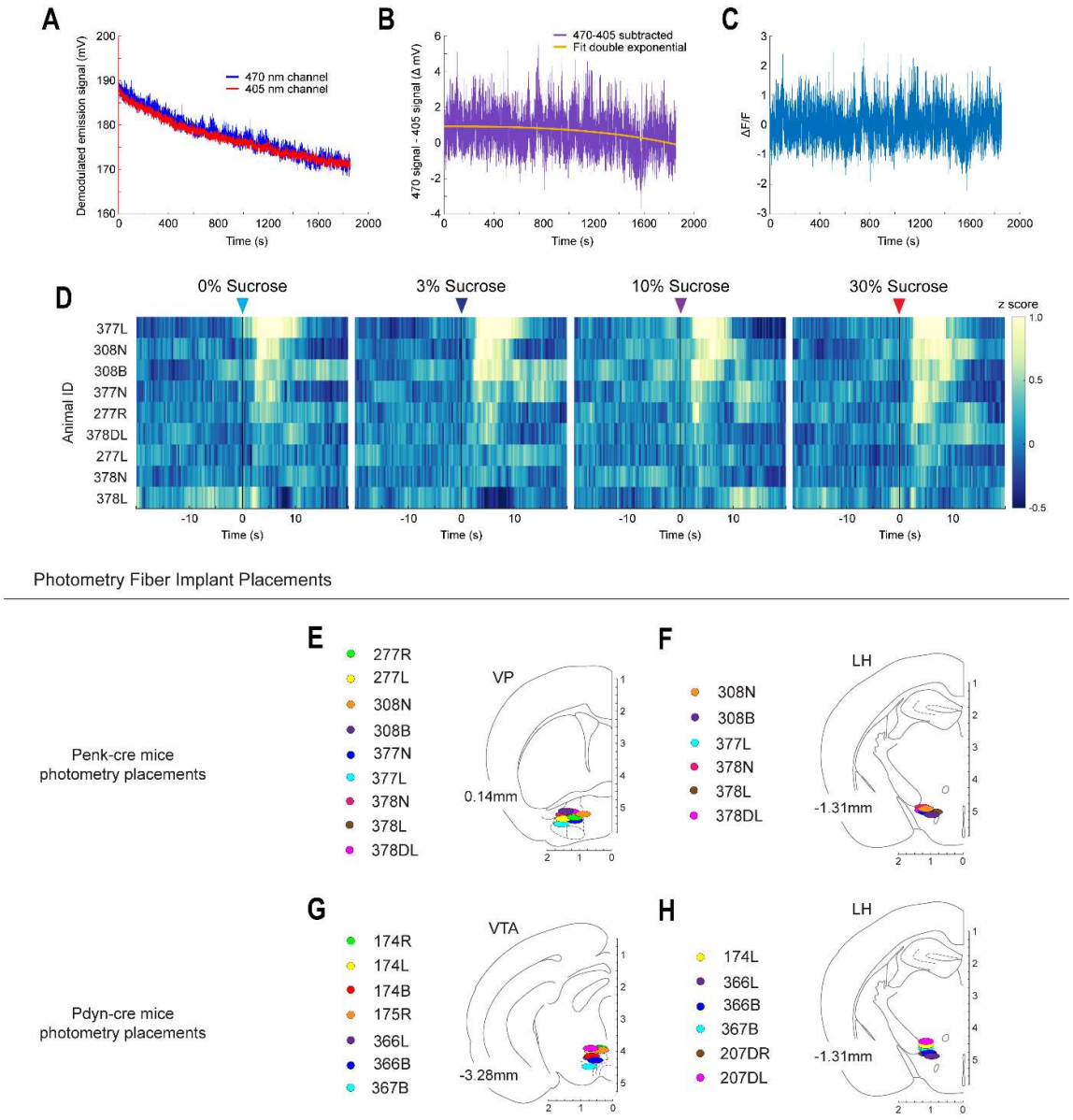
Supp. Figure S13: There is limited collateralization of enk-SPN and dyn-SPN efferent circuits

(A) Schematic (left) and representative images (right) of ventral tegmental area (top) or lateral hypothalamus (bottom) of retrobead injection sites. (B and C) In situ hybridization (B) and quantification (C) of prodynorphin, red, and green retrobeads in mNAcSh (scale bar = 200um). Percentage of DAPI-labelled cells that are also labelled with indicated probe. (D) Schematic (left) and representative images (right) of ventral pallidum (top) or lateral hypothalamus (bottom) of retrobead injection sites. (E and F) In situ hybridization (E) and quantification (F) of proenkephalin, red, and green retrobeads in mNAcSh (scale bar = 200um). Percentage of DAPI-labelled cells that are also labelled with indicated probe.


A
LH
VTA
Pdyn-cre
B
30% Sucrose
0% Sucrose
3% Sucrose
10% Sucrose
LH
C
Mean Fluores. (z score)
30% 0% 3% 10%
D
DAPI
GCaMP6s
fiber
ic
mfb
E
VTA
0.2 SD
5 s
F
Mean Fluores. (z score)
*
30% 0% 3% 10%
Sucrose Conc.
G
LH
VTA
Pdyn-cre
H
Uncued
10% Sucrose
Cue
Cued 10%
Sucrose
LH
I
Reward Epoch F (z score)
***
***
J
Cue
Extinction
K
Cue Epoch F (z score)
n.s.
L
DAPI
GCaMP6s
fiber
IPN
M
VTA
0.2 SD
5 s
N
Reward Epoch F (z score)
**
***
Uncued
Cued
Extinction
O
0.2 SD
5 s
P
Cue Epoch F (z score)
n.s.
Rewarded
Extinction
Q
Time-locked
Optogenetic
Manipulation
Sipper
Sucrose
R
Pdyn-cre mNAcSh
AAV5-hSyn-DIO-PPO
S
10% Sucrose
mNAcSh-LH Pdyn
cre- control
Mean licks per second
Off
Laser
T
30% Sucrose
mNAcSh-LH Pdyn
cre- control
Mean licks per second
Off
Laser
U
Reward-Paired
ChR2 Excitation
or PPO Inhibition
Laser on
50% of trials
10 Hz
Random
ITI
1s
2 ul Sucrose Delivered
Consummatory Licks
V
Pdyn-cre mNAcSh
AAV5-hSyn-DIO-ChR2
W
10% Sucrose
mNAcSh-LH Pdyn
**
cre- control
Mean licks per second
Off
Laser
X
30% Sucrose
mNAcSh-LH Pdyn
cre- control
Mean licks per second
Off
Laser

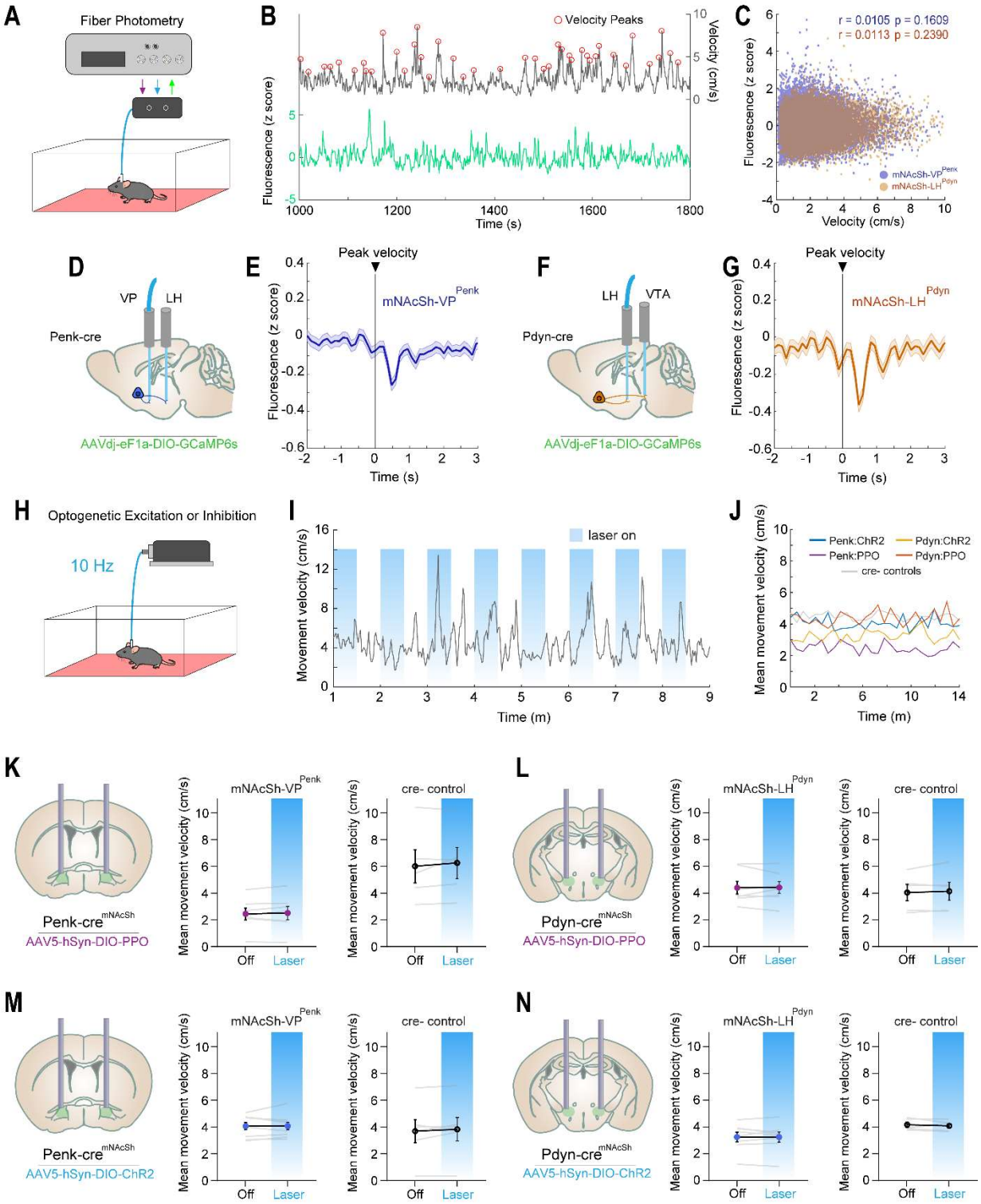
Supp. Figure S14: mNAcSh-LH^{dyn} is reward-responsive but may not causally drive sucrose consumption

(A) Sagittal section cartoon displaying target injection region, cell-types and fiber implant locations. (B) Mean mNAcSh-LH^{dyn} reward-response for every trial of each variable sucrose condition. (C) Mean mNAcSh-LH^{dyn} activity over reward response epoch (2-12 seconds from reward-onset) for each condition (n = 6 animals, 123+ trials, one-way ANOVA: p = 0.550). (D) Representative coronal section showing dyn-SPN terminals in LH expressing GCaMP6s underneath the fiber implant (scale bar = 150 um). (E) Mean mNAcSh-VTA^{dyn} reward-response for every trial of each variable sucrose condition. (F) Mean mNAcSh-VTA^{dyn} activity over reward response epoch (2-12 seconds from reward-onset) for each condition (n = 7 animals, 143+ trials, one-way ANOVA: 3% vs 10%, *p = 0.031). (G) Sagittal section cartoon displaying target injection region, cell-types and fiber implant locations. (H) Mean mNAcSh-LH^{dyn} reward-response for every uncued 10% sucrose trial and every cued 10% sucrose trial. (I) Mean mNAcSh-LH^{dyn} activity over reward response epoch (2-12 seconds from reward-onset) for uncued 10% sucrose, cued 10% sucrose and extinction trials (n = 6 animals, 63+ trials, one-way ANOVA: Uncued vs Cued, ***p < 0.001, Cued vs Ext., ***p < 0.001). (J) Mean mNAcSh-LH^{dyn} reward-response for every trial of day 7 Pavlovian extinction. (K) Mean mNAcSh-LH^{dyn} activity over cue response epoch (2-5 seconds from cue-onset) for cued 10% sucrose and extinction trials (n = 6 animals, 63+ trials, two sample t-test: p = 0.073). (L) Representative coronal section showing dyn-SPN terminals in VTA expressing GCaMP6s underneath the fiber implant (scale bar = 200 um). (M) Mean mNAcSh-VTA^{dyn} reward-response for every uncued 10% sucrose trial and every cued 10% sucrose trial. (N) Mean mNAcSh-VTA^{dyn} activity over reward response epoch (2-12 seconds from reward-onset) for uncued 10% sucrose, cued 10% sucrose and extinction trials (n = 7 animals, 71+ trials, one-way ANOVA: Uncued vs Ext., **p = 0.009, Cued vs Ext., ***p < 0.001). (O) Mean mNAcSh-VTA^{dyn} reward-response for every trial of day 7 Pavlovian extinction. (P) Mean mNAcSh-VTA^{dyn} activity over cue response epoch (2-5 seconds from cue-onset) for cued 10% sucrose and extinction trials (n = 7 animals, 71+ trials, two sample t-test: p = 0.387). (Q) Schematic of time-locked optogenetic neural activity manipulation during consumption of variable sucrose reward. (R) Coronal section cartoon displaying target injection region, cell-type and bilateral fiber implant location. (S) Within-session comparison of 10% sucrose consumption between trials with and without reward-paired optogenetic inhibition (n = 7 treatment, 5 control animals; pairwise t-test: p = 0.776, p = 0.187). (T) Within-session comparison of 30% sucrose consumption between trials with and without reward-paired optogenetic inhibition (n = 7 treatment, 5 control animals; pairwise t-test: p = 0.698, p = 0.220). (U) Time-course of sucrose reward delivery for a given trial. 50% of trials also had 4 seconds of 10 Hz optogenetic stimulation concurrent with reward delivery. (V) Coronal section cartoon displaying target injection region, cell-type and bilateral fiber implant location. (W) Within-session comparison of 10% sucrose consumption between trials with and without reward-paired optogenetic excitation (n = 10 treatment, 5 control animals; pairwise t-test: **p = 0.002, p = 0.973). (X) Within-session comparison of 30% sucrose consumption between trials with and without reward-paired optogenetic excitation (n = 10 treatment, 5 control animals; pairwise t-test: p = 0.174, p = 0.857).



Supp. Figure S15: Fiber photometry signal preprocessing and implant placements

(A) Demodulated fluorescence timeseries for 405nm and 470 nm GCaMP6s excitation from same recording session. (B) Fitting a second exponential curve to the 470nm-405nm subtracted timeseries and subtracting the fitted curve to 'detrend' the slow drift artifact from the whole session timeseries. (C) Final normalized trace after calculation of ΔF/F using median of initial 470nm channel for division. (D) Mean mNAcSh-VP^{enk} reward-response for 9 animals across all variable sucrose conditions. (E) Surgical placements of 400 um diameter photometry fiber implants in VP of enk-cre animals. (F) Surgical placements of 400 um diameter photometry fiber implants in LH of enk-cre animals. (G) Surgical placements of 400 um diameter photometry fiber implants in VTA of dyn-cre animals. (H) Surgical placements of 400 um diameter photometry fiber implants in LH of dyn-cre animals.


A
Fiber Photometry
B
Velocity Peaks
Fluorescence (z score)
Velocity (cm/s)
Time (s)
C
r = 0.0105 p = 0.1609
r = 0.0113 p = 0.2390
Fluorescence (z score)
Velocity (cm/s)
mNAcSh-VP Penk
mNAcSh-LH Pdyn
D
VP
LH
Penk-cre
AAVdj-eF1a-DIO-GCaMP6s
E
Peak velocity
mNAcSh-VP Penk
Fluorescence (z score)
Time (s)
F
LH
VTA
Pdyn-cre
AAVdj-eF1a-DIO-GCaMP6s
G
Peak velocity
mNAcSh-LH Pdyn
Fluorescence (z score)
Time (s)
H
Optogenetic Excitation or Inhibition
10 Hz
I
laser on
Movement velocity (cm/s)
Time (m)
J
Penk:ChR2
Pdyn:ChR2
Penk:PPO
Pdyn:PPO
cre- controls
Mean movement velocity (cm/s)
Time (m)
K
Penk-cre mNAcSh
AAV5-hSyn-DIO-PPO
mNAcSh-VP Penk
cre- control
Mean movement velocity (cm/s)
Off
Laser
L
Pdyn-cre mNAcSh
AAV5-hSyn-DIO-PPO
mNAcSh-LH Pdyn
cre- control
Mean movement velocity (cm/s)
Off
Laser
M
Penk-cre mNAcSh
AAV5-hSyn-DIO-ChR2
mNAcSh-VP Penk
cre- control
Mean movement velocity (cm/s)
Off
Laser
N
Pdyn-cre mNAcSh
AAV5-hSyn-DIO-ChR2
mNAcSh-LH Pdyn
cre- control
Mean movement velocity (cm/s)
Off
Laser

Supp. Figure S16: mNAcSh-VP^{enk} and mNAcSh-LH^{dyn} are not recruited during locomotion and cannot disrupt locomotion

(A) Schematic of fiber photometry recording during locomotion in a behavioral arena in red light conditions. (B) Simultaneous recording of fiber photometry signal and tracking of animal movement velocity. Detection of instantaneous velocity peaks was performed post-hoc to identify movement bouts. (C) Correlation between 1 second time bins of movement velocity and 1 second time bins of GCaMP6s fluorescence (Pearson's correlation; n = 10 enk-cre mice, 6 dyn-cre mice). (D) Sagittal section cartoon displaying target injection region, cell-types and photometry recording location. (E) Mean mNAcSh-VP^{enk} activity during every movement velocity peak. (F) Sagittal section cartoon displaying target injection region, cell-types and photometry recording location. (G) Mean mNAcSh-LH^{dyn} activity during every movement velocity peak. (H) Schematic of optogenetic stimulation (30 seconds on, 30 seconds off, repeating) during locomotion in a behavioral arena in red light conditions. (I) Representative timeseries of movement velocity throughout behavioral session, with alternating stimulation schedule displayed. (J) Mean movement velocity throughout duration of session for all stimulation and control conditions. (K) Within session comparison of movement velocity during mNAcSh-VP^{enk} photo-inhibition via PPO (n = 7 treatment, 5 control animals; pairwise t-test: p = 0.497, p = 0.378). (L) Within session comparison of movement velocity during mNAcSh-LH^{dyn} photo-inhibition via PPO (n = 7 treatment, 5 control animals; pairwise t-test: p = 0.903, p = 0.489). (M) Within session comparison of movement velocity during mNAcSh-VP^{enk} photo-excitation via ChR2 (n = 9 treatment, 6 control animals; pairwise t-test: p = 0.910, p = 0.440). (N) Within session comparison of movement velocity during mNAcSh-LH^{dyn} photo-excitation via ChR2 (n = 8 treatment, 5 control animals; pairwise t-test: p = 0.923, p = 0.480).

4. Discussion

As neuroscientists, we tend to think of rewards as discrete and binary events; reward occurred at this particular instant or it did not. Even more, we tend to always assume rewards are what comes at the end of a sequence of behavior. It is the ultimate payout at the end of a trial and signifies the end of a behavioral epoch.

In the real world, rewards are rarely so absolute and discrete. Animals are not always compelled to engage in the same redundant behavioral cycles to get the same quick and quantized rewards of great value. In the natural environment, rewards can be much more modest such as less appetitive food or food in sparse quantities. In the real world, animals must always be assessing their environment and choosing between a wide range of behavioral policies that have only slightly differentiated outcomes. There are fine distinctions that must be evaluated in the brain to make decisions between multiple paths of relatively equivalent predicted value. Accordingly, computing reward value information is critical and knowing when not to pursue rewards is just as important.

Neuroscientists have extensively studied how highly valued rewards cause robust changes in neural activity within particular brain regions and how these signals contribute to behavior and learning. However, we still have little idea of how these reward signals in the brain are tempered in the more common and ambiguous cases of reward where reward value is unclear and modest. We know how reward states in the brain are accelerated but we still know relatively little about the brakes.

The presented studies seek to address this underappreciated, yet environmentally common, category of reward behavior. We asked at which point animals decide it is no longer worth it to engage in reward seeking and which physiological mechanisms shift the

brain away from a state of motivation and engagement. We demonstrated that the neuropeptide nociceptin within the mesolimbic system is uniquely positioned to be a regulators that exercises constraint on the brain's mechanisms for reward and motivation.

The current studies also detail how neurons in a brain region critical for reward and motivation are engaged when animals consume rewards of lower value versus rewards of higher value. NAc representing differently valued rewards with dramatic differences in neuronal recruitment suggests that the brain may be designed to robustly separate reward information on the basis of value. Furthermore, we demonstrate that differences in reward value encoding are also consequential for consummatory behavior and low value encoding neurons causally reduce the consumption of high value rewards.

In the modern world, humans have engineered a vast toolkit for the exploitation of our evolutionarily-derived reward and motivation systems. Among the most directly damaging of these tools are drugs of abuse that directly chemically modulate reward pathways in the brain. The social and economic costs of illicit drug taking remain immense in our society, robbing millions of their lives and also their capacity to make their own decisions in their own long-term interest. According to the CDC, 130 Americans die every day from an opioid overdose on average. The number of opioid overdose deaths in 2017 was 6 times higher than in 1999 (Hedegaard 2018).

A health crisis that is not as obvious but is much more prevalent is our growing obesity epidemic. Food has been artfully refined for centuries to optimally activate the reward pathways in our brains and to cause us to overindulge. Obesity rates among American adults have increased from 30% in 2000 to 42% in 2018 (Hales 2020).

Other extremely prevalent modern interfaces with the reward pathways of our brains are those that are engineered to keep us engaged and locked in perpetually on our phones and other electronics. I personally can be enthralled in scrolling social media or playing video games for multiple hours a day, often closing applications only to lose focus and immediately reopen them again. Recent studies also detail the psychological harm of chronic social media and videogame addiction. The CDC reports around a 30% increase in suicide rates from the year 1999, with young people ages 10-34 as one of the groups most affected (Hedegaard 2020). For girls, a high level of social media use in early adolescence followed by a marked increase over time was most predictive of suicide risk in emerging adulthood (Coyne 2021). Young men have also suffered tremendously due to excessive virtual engagement. The number of men in their 20s without a college degree who had not worked a single day in the past year has doubled since the year 2000. Time-use data show that younger men distinctly shifted their leisure time to video gaming over this period (Aguiar 2017). Entire segments of the population, young and old find themselves perpetually locked into their cell phones and video games at the cost of their long term happiness and ability to grow socially and financially.

It can be argued that some of these modern creations, although addictive, can add utility to our lives. However, the danger is clear when we are increasingly engaging with them in a way that feels increasingly involuntary. For these reasons, it is paramount to understand how we can constrain our exploited reward seeking systems in the brain. We must study how nature applies the brakes to these reward seeking processes so that we understand when our neural systems for reward and motivation constraint have become maladapted. Furthermore, the fact that these mechanisms for the limitation of reward and

motivation are potentially largely mediated by opioid systems suggests that they could be pharmacologically actionable. Understanding how opioid systems in the brain constrain aberrant reward seeking may be a promising new avenue for the treatment of maladapted motivated behavior.

5. Methods

Animals

Adult (20–35g) male and female prepronociceptin-IRES-Cre (*Pnoc*-Cre), NOPR KO, NOPR cKO mice were group housed, given access to food pellets and water ad libitum, and maintained on a 12 hr:12 hr light:dark cycle (lights on at 7:00 AM). All mice were kept in a sound-attenuated, isolated holding facility in the lab one week prior to surgery, post-surgery, and throughout the duration of the behavioral assays to minimize stress. For ablation, chemogenetic, and optogenetic experiments, we used Cre- cage and littermate controls. For NOPR KO and NOPR cKO behavioral experiments, experimental mice were compared to age-matched wildtype controls. Unless otherwise noted, animals had ad libitum access to food and water. The mice were bred at Washington University in Saint Louis by crossing the *Pnoc*-Cre mice, NOPR KO, and NOPR cKO with C57BL/6 wild-type mice and backcrossed for seven generations. Additionally, where needed, *Pnoc*-Cre mice were then crossed to Ai9-tdTomato and Ai32-ChR2-eYFP mice on C57BL/6 background, bred, and backcrossed for seven generations. For the generation of NOPRlox/lox mice, Exon 4 of NOPR was flanked on both sides by loxP sites. All animals were drug and test naïve, individually assigned to specific experiments as described, and not involved with other experimental procedures. Student's t-tests did not detect any statistically significant difference when comparing animals grouped by sex, therefore male and female mice were combined to complete final group sizes. All animals were monitored for health status daily and before experimentation for the entirety of the study. All procedures were approved by the Animal Care and Use Committee of Washington University in St. Louis and performed in accordance with the National Institute of Health's guide for the care and use of laboratory animals.

Generation of *Pnoc*-IRES-Cre mouse line

129 BAC genomic clones containing the Pnoc genes was used to target a cassette containing the Cre recombinase gene preceded by an internal ribosomal entry sequence (IRES-Cre) downstream of the *Pnoc* stop codon, so as to have endogenous *Pnoc* genes drive Cre recombinase expression. The *Pnoc*-IRES-Cre targeting constructs were electroporated into W4 ES cells and screened using an FRT-flanked neo-cassette. Targeted clones were injected into blastocysts. Chimeras were obtained and bred for germline transmission of the *Pnoc*-IRES-Cre allele. Following germline transmission, *Pnoc*-IRES-cre mice were crossed to Rosa-FLPR mice to remove the neo cassette at the 3′ end of the IRES-cre sequence.

Tissue processing

Unless otherwise stated, animals were transcardially perfused with 0.1 M phosphate-buffered saline (PBS) and then 40 mL 4% paraformaldehyde (PFA). Brains were dissected and post-fixed in 4% PFA overnight and then transferred to 30% sucrose solution for cryoprotection. Brains were sectioned at 30 μM on a microtome and stored in a 0.01M phosphate buffer at 4°C prior to immunohistochemistry and tracing experiments. For behavioral cohorts, viral expression and optical fiber placements were confirmed before inclusion in the presented datasets.

Immunohistochemistry

Immunohistochemistry was performed as previously described by (Al-Hasani 2013, Kim 2013, McCall 2015). In brief, mice were intracardially perfused with 4% paraformaldehyde, and then brains were sectioned (30μm) and placed in 0.1 M PB until immunohistochemistry. Free-floating sections were washed in 0.1 M PBS for 3 x 10 minutes intervals. Sections were then placed in blocking buffer (0.5% Triton X-100 and 5% natural goat serum in 0.1 M PBS) for 1 hr at room temperature. After blocking buffer, sections were placed in primary antibody (chicken anti-tyrosine hydroxylase,1:2000, Aves Labs, Inc.; rabbit anti-nociceptin, 1:500, abcam) overnight at

room temperature. After 3 x 10-minute 0.1 M PBS washes, sections were incubated in secondary antibody (AlexaFluor 488 goat anti-rabbit; AlexaFluor 594 or 633 goat anti-chicken, Life Technologies) for 2 hours at room temperature, followed by subsequent washes (3 X 10 minute in 0.1 M PBS). Later, sections were incubated in NeuroTrace (435/455 blue fluorescent Nissl stain, ThermoFisher Scientific) for 1 hour, followed by 3 x 10 minute 0.1 MPBS then 3 x 10-minute 0.1 M PB washes. After immunostaining, sections were mounted and coverslipped with Vectashield HardSet mounting medium (Vector Laboratories) and imaged on a Leica TCS SPE confocal microscope.

RNAscope Fluorescent In Situ Hybridization (FISH)

Following rapid decapitation of mice, brains were flash frozen in -50°C 2-methylbutane and stored at -80°C for further processing. Coronal sections containing the BNST, CeA, and VTA/IPN regions, corresponding to the injection plane used in the behavioral experiments, were cut at 20μM at -20°C and thaw-mounted onto Super Frost Plus slides (Fisher). Slides were stored at -80°C until further processing. FISH was performed according to the RNAScope® 2.0 Fluorescent Multiplex Kit for Fresh Frozen Tissue (Advanced Cell Diagnostics, Inc.) as described by Wang 2012. Slides containing the specified coronal brain sections were fixed in 4% paraformaldehyde, dehydrated, and pretreated with protease IV solution for 30 mins. Sections were then incubated for target probes for mouse *Pnoc*, (accession number NM_010932.2, probe region 325-1263), VGlut (*slc17a6*, accession number NM_080853.3, probe region 1986-2998), VGat (*slc32a1*, accession number NM_009508.2, probe region 894-2037), Cre (accession number KC845567.1, probe region 1058-2032), NOPR (accession number NM_011012.5, probe region 988-1937, TH (accession number NM_009377.1, probe region 483 – 1603), DAT (*slc6a3*, accession number NM_010020.3, probe region 1486-2525) and GFP (accession number AF275953.1, probe region 12-686) for 2 hrs. All target probes consisted of 20 ZZ oliogonucleotides and were obtained from Advanced Cell Diagnostics. Following probe

hybridization, sections underwent a series of probe signal amplification steps (AMP1-4) including a final incubation of fluorescently labeled probes (Alexa 488, Atto 550, Atto 647), designed to target the specified channel (C1-C3 depending on assay) associated with the probes. Slides were counterstained with DAPI and coverslips were mounted with Vectashield Hard Set mounting medium (Vector Laboratories). Images were obtained on a Leica TCS SPE confocal microscope (Leica), and Application Suite Advanced Fluorescence (LAS X) and ImageJ/Fiji software were used for analyses. Images for *Pnoc* Cre+ and *Pnoc* Cre- mice for quantification of *Pnoc* / Cre coexpression were carried out under the same standards (i.e. intensity, threshold, exposure time for each slide). To analyze the images, each image was opened in ImageJ software using the "Color Threshold" function, selecting the default threshold method and RGB color space. Within the threshold window, the default the settings were kept consistent for each slide. After adjusting for fluorescence threshold, we counted the total pixels of the fluorescent signal, assuming that each pixel represents a single molecule of RNA as per RNAscope's methodology. A positive cell consisted of an area within the radius of a DAPI nuclear staining that measured at least 10 total positive pixels and an integrated density of >5. Three separate slices from the BNST, CeA, and pnVTA were used for each animal and that total is reflected in the data (n = 4).

Stereotaxic viral / optical fiber implant surgeries

After the mice were acclimated to the holding facility for at least seven days, the mice were anaesthetized in an induction chamber (1-4% isoflurane) and placed into a stereotaxic frame (Kopf Instruments, Model 1900) where they were maintained at 1-2% isoflurane. Mice were then injected unilaterally or bilaterally, depending on experimental paradigm, using a blunt needle (86200, Hamilton Company) at a rate of 100 nL / min. The type of virus, injection volume, and stereotaxic coordinates for each experiment are listed in Supplementary Table 1. Mice were allowed to recover for five weeks prior to behavioral testing, permitting optimal expression of the virus. For optogenetic experiments, one week prior to behavioral testing, intracranial optic fiber

implants were directed above the VTA (AP -3.15, ML ± 0.5, DV -4.25). Implants were secured using at least two bone screws and a dental cement headcap (Lang Dental).

Generation of TRAP virus

The TRAP construct (eGFP/Rpl10a), a generous gift from N. Heintz lab, was PCR cloned into a commercial Gateway compatible entry vector (pENTR-3C), sequenced to confirm coding sequence, then transferred into an AAV-EF1α-DIO-Gateway destination vector using LR-clonase (Invitrogen) and following the manufacture's protocol (Heiman 2008). AAV-EF1α-DIO-Gateway plasmid was previously generated by replacing Chr2-eYfp from AAV-EF1α-DIO-ChR2-eYFP with an ATTR flanked Gateway acceptor cassette. The resulting AAV-EF1α-DIO-TRAP Plasmid was maxiprepped, sequence confirmed, then packaged by UNC Vector Core using standard protocols. Cre dependence was confirmed anatomically in independent mice and TRAP construct showed cytoplasmic and nucleolar localization within cells consistent with incorporation into ribosomes.

TRAP RNAseq Assay

TRAP was carried out as described in Sakers 2017. Briefly, a 1mm coronal slice was dissected 6mm caudal to the base of the olfactory bulbs. Five slices (across 5 mice) were pooled and homogenized in cold lysis buffer (10 mM HEPES [pH 7.4], 150 mM KCl, 5 mM MgCl2, 0.5 mM dithiothreitol, 100 µg/ml cycloheximide, protease inhibitors, and recombinant RNase inhibitors) using a Teflon-glass homogenizer and a power drill. Homogenates were spun at 2,000xg for 10 minutes, and the resulting supernatant was lysed with 30mM DHPC and 1% NP40 followed by incubation on ice for 15 minutes. The lysate was then cleared at 20,000xg for 20 minutes, and the resulting supernatant was applied to EGFP-coated magnetic beads. 10% of this input was kept for a 'Pre-IP' sample. Immunoprecipitations were carried out for 4 hours at 4C with constant rotation. Four 1mL washes were carried out with wash buffer (10 mM HEPES [pH 7.4], 350 mM KCl, 5 mM MgCl2, 1% NP-40, 0.5mM dithiothreitol, 100 µg/ml cycloheximide and

recombinant RNase inhibitors). RNA extraction using Trizol LS was carried out followed by Zymo RNA clean and concentrator-5 kit, with DNAse treatment, according to manufacturer's guidelines.

Library Prep and Sequencing

cDNA was made and amplified from 1ng of RNA using the Ovation RNA Amplification System V2 from NuGEN. cDNA was sheared to ~200bp using a Covaris sonicator. RNA-sequencing libraries were made from 600ng of sheared cDNA with the NEBNext Ultra DNA Library Prep Kit for Illumina. ~90-95M 50bp reads were obtained, per sample, on an Illumina HiSeq 2500 and analyzed using an in-house pipeline as described in Sakers 2017. Briefly, reads were trimmed using Trimmomatic (v0.33), aligned to the genome using STAR (Mouse Ensembl 77) and counted with HTseq. Counts per million analyses were performed in R using the edgeR package (R Foundation for Statistical Computing, Vienna, Austria).

Gene ontology

Gene ontology analysis was performed using the BiNGO application from Cytoscape (v3.4.0). Enriched genes were identified with a Benjamini-Hochberg corrected Hypergeometric test, at a significance level of $p < 0.05$. Pre-IP genes with counts above 2 were used as a reference set for enrichment.

Patch-clamp electrophysiology

Mice were anesthetized with pentobarbital (50 mg/kg) before transcardial perfusion with ice-cold sucrose cutting solution containing the following (in mM): 225 sucrose, 119 NaCl, 1.0 NaH_2P0_4, 4.9 $MgCl_2$, 0.1 $CaCl_2$, 26.2 $NaHCO_3$, 1.25 glucose, 305 mOsm. Brains were then rapidly removed, and coronal sections 300μm thick were taken using a vibratome (Leica, VT 1200). Sections were then incubated in aCSF (32°C) containing the following (in mM): 119 NaCl, 2.5 KCl, 1.0 NaH_2P0_4, 1.3 MgCl, 2.5 $CaCl_2$, 26.2 $NaHCO_3$, 15 glucose, 305 mOsm. After an hour of recovery, slices were constantly perfused with aCSF (32°C) and visualized using differential

interference contrast through a 40x water-immersion objective mounted on an upright microscope (Olympus BX51WI). Whole-cell recordings were obtained using borosilicate pipettes (3–5 MΩ) back-filled with internal solution containing the following (in mM): 130 K gluconate, 10 KCl, 10 HEPES, 10 EGTA, 2 $MgCl_2$, 2 ATP, 0.2 GTP (pH 7.35, 270-285 mOsm). Current-clamp recordings were obtained from eYFP-expressing neurons to identify spike fidelity in optogenetic stimulation. During recordings, a train of 10 action potentials were evoked through presentation of a blue LED (488nm, 1mW, 1 and 10ms pulses) at frequencies of 1, 5, 10, 20, and 60Hz. Spike fidelity was quantified as the percentage of light presentations that resulted in action potential induction. Data acquisition occurred at 10 kHz sampling rate through a MultiClamp 700B amplifier connected to a Digidata 1440A digitizer (Molecular Devices) and were analyzed using threshold analysis in Clampfit 10.3 (Molecular Devices).

Behavior

All behaviors were performed within a sound-attenuated room maintained at 23°C at least one week after habituation to the holding room and the final surgery. For all experiments, mice were brought into the experimental behavior space and allowed to acclimate for at least 30 minutes. For all experiments, experimenters were blinded to mouse genotype or experimental manipulation. For chemogenetic and optogenetic behavior experiments, CNO dose and frequency of light stimulation were randomly assigned using block randomization, unless otherwise indicated. All pharmacological interventions (e.g. agonists and antagonists) were randomized and counterbalanced via block randomization prior to experimentation. Replication of operant responding behavior (e.g. PR test) was only used to compare experimental and control treatments with treatments counterbalanced between all mice. Sample-size estimation was determined via power analysis (G*Power 3, Faul 2007) with a power of 0.85 and a standard deviation of 15-20% of the mean. Lighting was stabilized at ~1,500 lux for aversion behaviors, ~250 lux for anxiety-like behaviors. Movements were video recorded via CCD camera and

analyzed using Ethovision XT 11 (Noldus Information Technologies). At the end of each study, mice were perfused with 4% paraformaldehyde followed by anatomical analysis to confirm injection sites and cell-type-specific expression.

Operant Progressive-Ratio Test

Food restricted (~90% of free feeding body weight) mice were placed in a Med-Associates operant conditioning box and subjected to magazine training during which a retractable lickometer containing a 10% sucrose solution extended for 20s at 20 random intervals for 60 minutes. Each sipper presentation was accompanied by illumination of a houselight. Following magazine training, mice were trained to nosepoke for lickometer access on a fixed ratio 1 (FR1) for 3 days (60 minutes/session). This required the mouse to nosepoke in the active nosepoke port one time to receive 20 seconds of access to the sipper. Following FR1 training, the ratio was increased to an FR3 schedule requiring the mouse to perform 3 active port nosepokes to receive access to the sipper. This schedule was used for 3 days (60 minutes/session). Finally, mice were given a progressive ratio schedule (PR) task following the geometric progression, $nj = 5e^{j/5} - 5$, in which the criteria for rewards increased in an exponential manner (1, 2, 4, 6, 9, 12…) over the training period (Richardson and Roberts, 1996). For photostimulation and photoinhibition, mice received 5 Hz or 10 Hz stimulation (473 nm, 10 ms pulse width, ~10 mW light power) and (532 nm, ~10 mW light power), respectively upon placement into the operant chamber. During all optogenetic experiments, mice were counterbalanced to receive the appropriate photostimulation or photoinhibition parameters. In photostimulation experiments, animals received 5Hz stimulation during the entire PR session, 10Hz stimulation time-locked to nosepoking during the PR session, or 10Hz stimulation time-locked to licking of sucrose reward during PR session. In photoinhibition experiments, animals received either 1 sec of illumination time-locked to nosepoking during the PR session or 1 sec of illumination time-locked to licking of sucrose reward during PR session. For chemogenetic experiments, mice received an injection of vehicle or CNO (1 mg/kg; 5 mg/kg,

i.p.) 30 minutes prior to being placed in the operant chamber on PR test day. For antagonist experiments, NOPR antagonist, J-113397 was administered (3 mg/kg, i.p.) 15 minutes prior to behavioral test.

Pavlovian Conditioning Paradigm

One week prior to Pavlovian conditioning, mice were food restricted down to ~90% of free feeding body weight. A subset of the *Pnoc*-Cre fiber photometry mice (n = 10) and *DAT-Cre* fiber photometry mice (n = 7) were trained to associate illumination of a house light (CS) with access to a sipper with 10% sucrose solution (US) within a Med-Associates operant box (ENV-307A). The house light would illuminate 6 s prior to sipper presentation. The sipper remained accessible for 20 seconds before it retracted and the house light shut off. A randomized intertrial interval of between 30-90 seconds separated consecutive trials. Pavlovian conditioning sessions lasted for 30 minutes, over which an average of 19-20 rewards were presented. For reward omission sessions, the first 10 rewards were cued and subsequently delivered as normal and every following trial only involved house light illumination with no sipper presentation. Simultaneous fiber photometry recordings were made during Pavlovian conditioning sessions as indicated in Results section. NOPR antagonist, J-113397 (3 mg/kg, i.p.) or SCH-221510 (10mg/kg, i.p.) were administered 15 minutes prior to Pavlovian conditioning.

In Vivo Fiber Photometry

Fiber photometry recordings were made throughout the entirety of 30-minute Pavlovian conditioning sessions and 1-hour FR3 and PR sessions. Prior to recording during Pavlovian or operant behavior sessions, an optic fiber was attached to the implanted fiber using a ferrule sleeve (Doric, ZR_2.5). Two LEDs were used to excite GCaMP6s. A 531-Hz sinusoidal LED light (Thorlabs, LED light: M470F3; LED driver: DC4104) was bandpass filtered (470 ± 20 nm, Doric, FMC4) to excite GCaMP6s and evoke Ca^{2+}-dependent emission. A 211-Hz sinusoidal LED light

(Thorlabs, LED light: M405FP1; LED driver: DC4104) was bandpass filtered (405 ± 10 nm, Doric, FMC4) to excite GCaMP6s and evoke Ca^{2+}-independent isosbestic control emission. Prior to recording, a 120 second period of GCaMP6s excitation with 405 nm and 470 nm light was used to remove the majority of baseline drift. Laser intensity for the 470 nm and 405 nm wavelength bands were measured at the tip of the optic fiber and adjusted to ~50 µW before each day of recording. GCaMP6s fluorescence traveled through the same optic fiber before being bandpass filtered (525 ± 25 nm, Doric, FMC4), transduced by a femtowatt silicon photoreceiver (Newport, 2151) and recorded by a real-time processor (TDT, RZ5P). The envelopes of the 531-Hz and 211-Hz signals were extracted in real-time by the TDT program Synapse at a sampling rate of 1017.25 Hz.

Photometry Analysis

Custom MatLab scripts were developed for analyzing fiber photometry data in context of mouse behavior. The isosbestic 405 nm excitation control signal was subtracted from the 470 nm excitation signal to remove movement artifacts from intracellular Ca^{2+}-dependent GCaMP6s fluorescence (see **Figure S2A**). Baseline drift was evident in the signal due to slow photobleaching artifacts, particularly during the first several minutes of each hour-long recording session. A double exponential curve was fit to the raw trace and subtracted (see **Figure S2B**) to correct for baseline drift. After baseline correction, the photometry trace was z-scored relative to the mean and standard deviation of the hour-long session. The post-processed fiber photometry signal was analyzed in the context of animal behavior during Pavlovian conditioning and operant task performance. Pearson correlations, one sample t-tests, two sample t-tests and two-way ANOVAs were performed using standard MATLAB functions “corr”, “ttest”, “ttest2” and “anovan”, respectively. ΔF/F was calculated as individual fluorescence intensity measurements relative to median fluorescence of entire session for 470nm channel. ΔF/F for DA cell NOPR manipulation experiments (see **Figure 7**) was calculated as individual fluorescence intensity measurements

relative to median fluorescence of the entire vehicle treatment session. Recorded calcium dynamics were averaged into 2-second time bins and functionally categorized based on behavioral state of the animal: baseline activity during inter-trial interval, reward anticipation activity during cue presentation, or consumption activity when cue is present and animal is actively licking. A minority of scheduled FR3 and PR recording sessions were either excluded from analysis due to animal licking while sipper was retracted or not recorded due to fiber optic implant detachment before behavioral schedule was completed. 65/65 Pavlovian conditioning sessions, 33/36 planned FR3 sessions and 32/36 planned PR sessions were recorded and retained.

Real-Time Place Testing

We used custom-made unbiased, balanced two-compartment conditioning apparatus (52.5 x 25.5 x 25.5 cm) as described previously (Jennings 2013, Stamatakis 2012). Each compartment was assigned either constant laser stimulation or no stimulation in a counterbalanced manner. Each animal was connected to a 473 nm laser via 200μm fiber optic patch cable then placed into the non-stimulation compartment at the onset of the experiment. Mice were allowed to freely roam between the two compartments for 30 minutes. Entry into the stimulation-paired compartment triggered constant photostimulation at either 5Hz, 10Hz, 20Hz or 60Hz (473 nm, 10 ms pulse width, ~10 mW light power) while the mouse remained in the stimulation-paired compartment. Entry into the non-stimulation compartment ended the photostimulation. The compartment paired with photostimulation was counterbalanced across all mice. Time spent in each compartment and total distance traveled for the entire 30-minute trial was recorder via CCD camera and measured using Ethovision 10 (Noldus Information Technologies, Leesburg, VA). Real-time preference was calculated by dividing the time spent in the stimulation-paired side by the total time in the apparatus.

Conditioned Place Preference

We used a modified three-chamber CPP apparatus consisting of two square boxes (27 cm × 27 cm) that served as the conditioning chambers separated by a small center area that served as the passageway (5 cm wide × 8 cm long) between boxes. Boxes had 2.5 cm black-and-white vertical stripes or horizontal stripes and floors were covered with 500 ml of bedding on each side. The floor of the center area was smooth Plexiglas. Mice were transported to the CPP behavior testing room and handled once per day for at least 7 days before behavioral testing. Mice were then conditioned using a semi-biased and counterbalanced CPP. On day one, mice were pretested for an initial bias to the conditioning apparatus and allowed to explore all three regions of the box drug-free for 20 minutes. Mice were paired in a counterbalanced fashion. Mice that spent >300 s on one side during the pretest were excluded due to potential side bias. Conditioning occurred over the following 2 days in which mice received a subcutaneous injection of saline in the morning and immediately confined to one side of the CPP box for 20 min. In the afternoon, at least 4 hours after the morning conditioning session, mice received an injection of CNO (5 mg/kg, i.p.) and were immediately confined to the opposite side for 20 min. This dose of CNO has been shown to be effective in activating DREADDS. Control (vehicle-injected) mice received saline injections during both A.M. and P.M. conditioning sessions. The following day, mice were tested for a preference for the CNO-paired side following the same procedure as the pretest. Preference scores were calculated by subtracting time spent in the CNO-paired side during the pretest from time spent in the CNO-paired side during the posttest.

Animals

Adult C57BL/6 (18–35 g) male and female wildtype, prodynorphin-IRES-Cre (*dyn*-cre) and preproenkephalin-IRES-Cre (*enk*-cre) mice were group housed, given access to food pellets and water ad libitum, and maintained on a 12 hr:12 hr light:dark cycle. All mice were kept in a sound-

attenuated, isolated holding facility one week prior to surgery, post-surgery, and throughout the duration of the behavioral assays to minimize stress. For cell-type specific optogenetic experiments, we used Cre- cage and littermate controls. Unless otherwise noted, animals had *ad libitum* access to food and water. The mice were bred at Washington University in Saint Louis or the University of Washington. All animals were initially test naive, individually assigned to specific experiments as described, and not involved with other experimental procedures. Statistical comparisons did not detect any significant differences between male and female mice and were therefore combined to complete final group sizes. All animals were monitored for health status daily and before experimentation for the entirety of the study. All procedures were approved by the Animal Care and Use Committee of Washington University, Animal Care and Use Committee of the University of Washington and conformed to US National Institutes of Health guidelines.

Tissue collection after behavioral experiments

Unless otherwise stated, animals were transcardially perfused with 0.1 M phosphate-buffered saline (PBS) and then 40 mL 4% paraformaldehyde (PFA). Brains were dissected and post-fixed in 4% PFA overnight and then transferred to 30% sucrose solution for cryoprotection. Brains were sectioned at 30 mM on a microtome and stored in a 0.01M phosphate buffer at 4ºC prior to immunohistochemistry and tracing experiments. For behavioral cohorts, viral expression and optical fiber placements were confirmed before inclusion in the presented datasets.

RNAscope Fluorescent In Situ Hybridization

Following rapid decapitation of wildtype mice, brains were rapidly frozen in 100mL -50°C isopentane and stored at -80°C. Coronal sections corresponding to the site of interest or injection plane used in the behavioral experiments were cut at 20uM at -20°C and thaw-mounted onto

SuperFrost Plus slides (Fisher). Slides were stored at -80°C until further processing. Fluorescent in situ hybridization was performed according to the RNAscope 2.0 Fluorescent Multiple Kit User Manual for Fresh Frozen Tissue (Advanced Cell Diagnostics, Inc.) as described by Wang et al. (2012). Briefly, sections were fixed in 4% PFA, dehydrated, and treated with pretreatment 4 protease solution. Sections were then incubated for target probes for mouse proenkephalin (*Penk*, accession number NM_001002927.2, probe region 106 – 1332) or prodynorphin (*Pdyn*, accession number NM_018863.3, probe region 33 - 700). All target probes consisted of 20 ZZ oligonucleotides and were obtained from Advanced Cell Diagnostics. Following probe hybridization, sections underwent a series of probe signal amplification steps followed by incubation of fluorescently labeled robes designed to target the specific channel associated with the probes. Slides were counterstained with DAPI, and coverslips were mounted with Vectashield Hard Set mounting medium (Vector Laboratories). Images were obtained on an Olympus Fluoview 3000 confocal microscope and analyzed with HALO software. To analyze the images, each image was opened in the HALO software. DAPI positive cells were then registered and used as markers for individual cells. An observer blind to the brain tissue origin and probes used then counted the total number of probe-labeled cells for each channel separately. A positive cell consisted of an area within the radius of a DAPI nuclear staining that measured at least 10 total positive pixels for neurotransmitter probes. Two - three separate slices from the NAc or were used for each animal and that total is presented in the data.

Stereotaxic Surgery

After mice were acclimated to the holding facility for at least seven days, the mice were anaesthetized in an induction chamber (1%-4% isoflurane) and placed into a stereotaxic frame (Kopf Instruments, model 1900) where they were mainlined at 1%-2% isoflurane. A small hole was drilled in the skull above the target brain region. For mice receiving viral injections, a needle

syringe (Hamilton Company, 65458) was used to deliver the vector at a rate of 100 μL/min. Animals used for 2-photon calcium imaging experiments were injected unilaterally with 500 uL of AAVdj-ef1a-DIO-GCaMP6s into the mNAcSh (coordinates: AP +1.3, ML +0.5, DV -4.4 from bregma), implanted with a GRIN lens (Inscopix, 7.3 mm length, 0.6 mm diameter) and had a 1 cm diameter headring horizontally adhered to the top of the skull. Animals used for fiber photometry experiments were injected unilaterally with 400 uL of AAVdj-ef1a-DIO-GCaMP6s into the mNAcSh (coordinates: AP +1.3, ML +0.5, DV -4.4 from bregma), implanted ipsilaterally with fiber optic cannulas (Doric Lenses, 400 um diameter, 6 mm length, 0.48 NA) in VP (coordinates: AP +0.3, ML +1.2, DV -4.7 from bregma), LH (coordinates: AP -1.3, ML +1.1, DV -4.7 from bregma) or VTA (coordinates: AP -3.3, ML +0.4, DV -4.4 from bregma), and had a 1 cm diameter headring horizontally adhered to the top of the skull. Animals used for retrotracing and FISH were injected unilaterally with 200 uL of green or red retrobeads (Lumafluor, Retrobeads IX) in VP (coordinates: AP +0.3, ML +1.2, DV -4.7 from bregma), LH (coordinates: AP -1.3, ML +1.1, DV -4.7 from bregma) or VTA (coordinates: AP -3.3, ML +0.4, DV -4.4 from bregma). Animals used for optogenetics experiments were injected bilaterally with 200 uL of AAV5-DIO-ChR2-EYFP or AAV5-DIO-parapinopsin-Venus into the mNAcSh (coordinates: AP +1.3, ML +/-0.5, DV -4.4 from bregma), implanted bilaterally with custom-built fiber optic cannulas (200 um diameter) in VP (coordinates: AP +0.3, ML +1.2, DV -4.7 from bregma) or LH (coordinates: AP -1.3, ML +1.1, DV -4.7 from bregma), and had a 1 cm diameter headring horizontally adhered to the top of the skull. All implants and headrings were secured to the skull using dental cement (Parkell, C&B Metabond). Animals were injected with slow-release Carprofen (5 mg/kg) prior to surgery to act as a post-surgery analgesic. Animals were monitored daily for 7 days following surgery.

Water Access Restriction prior to behavioral testing

Mice were allowed to recover from surgery for a minimum of 4 weeks prior to water restriction or behavioral testing. To adequately motivate consistent sucrose consumption in behavioral subjects, animals were water-restricted by giving them access to water only once per day as previously described (Otis 2017, Rossi 2019, Namboodiri 2019). Starting 5 days prior to behavioral testing animals were given 1.0 ml of water in a 2x2cm paper tray in their home cage each day at approximately 5-6pm, after behavioral testing. Animals readily consumed the water while supervised by the researcher to ensure they consumed most of the presented amount. Animals were also weighed routinely to ensure they stayed at 85-90% of their original body weight while under water-restriction. No health problems related to dehydration arose throughout the water-restriction period of up to ~16 days. After the conclusion of a behavioral testing schedule, water-restriction was immediately stopped and animals were given ad-libitum access to water in their home cages again.

Brief Access Two Bottle Choice

To test the relative value of the 4 variable sucrose reward conditions (0%, 3%, 10%, 30% sucrose dissolved in water) while minimizing effects from satiety, we tested mice in an adapted 2-bottle choice assay where the animals were only given brief access (20 s) to sucrose sippers with a variable intertrial interval (20-40s) between access periods. Brief access 2-bottle choice assays were conducted within enclosed operant chambers, that had two retractable sucrose sippers side-by-side (MedAssociates, MED-307A, ENV-352). Discrete licks on the metallic sippers were recorded by a contact licometer system (MedAssociates, ENV-250). Prior to 2-bottle choice test days, each animal was trained with free access (2 days) and variable access (2 days) to both sippers containing 10% sucrose for 30 minute sessions. Each preference test involved

presentation of 2 of the 4 different sucrose conditions and each session had two phases. In the first phase, animals were given access to either sipper in isolation so that they could sample them independently. In the second phase, animals were simultaneously presented with both sippers side-by-side and had to choose where to allocate their sipper access time. The 6 possible pairings of the 4 sucrose reward conditions were presented in a pseudorandomized and counterbalanced sequence to minimize effects from condition history and right versus left sipper location bias. Preference test days were also buffered by 10% versus 10% sucrose days to neutralize any left-versus-right biases that arose from test days.

As indicated in the Results, some sessions included a variation of this paradigm where one of the sucrose sippers contained no liquid and instead triggered closed loop optogenetic stimulation. Licking the empty sucrose spout would register a lick in the MED-PC program and immediately trigger a TTL sent from the MED-307A to an Arduino Uno. The Arduino Uno controlled a 470 nm light laser (4 mW intensity at end of patch cable, 10 Hz, 10 ms pulse width, 2 second in duration) that was optically coupled to the fiber implants on the animal's head.

Variable Sucrose Access during Reversible Head-fixation

After 4-6 weeks of recovery from surgery, animals were water restricted (see above) and trained to sip 10% sucrose reward at variable intervals while reversibly immobilized and head-restrained. Reversible head-fixation was conducted by scruffing the animals, backing them into a 50 ml canonical tube and sliding the animals and tube into place in a custom headstage as has been previously described (Rossi 2019, Namboodiri 2019). All animals displayed no signs of physical distress and consistent licking behavior while immobilized after 4 days of lick training and were then tested in 30 minute, 20 trial sessions each day for 4 consecutive days. Intertrial interval was uniformly distributed between 70 and 100 seconds. On each trial, animals were presented with 5 droplets of sucrose reward with a total aggregate volume of 10 ul, with 1 second between

droplet delivery. Reward delivery was controlled by custom code on an Arduino Uno. Animals immediately licked the presented sucrose reward, with an average licking bout duration of 7-8 seconds. In each session, animals consumed 1 of 4 possible sucrose reward conditions (0%, 3%, 10% or 30% sucrose in water), with the order of condition days pseudorandomized and counterbalanced across animals. Discrete lick events were detected and recorded with <1 ms time resolution using a contact licometer circuit as has been previously described (Rossi 2019, Namboodiri 2019). In some cases, 15-17 FPS behavioral video of the animal's face was recorded alongside the contact licometer or instead of the contact licometer. The presence of sucrose licking in a particular frame was assessed manually or through automated scoring by a recurrent neural network model (see below) that was extensively trained and validated using simultaneous contact licometer information. The behavioral testing was conducted in dark or far-red light conditions to remove contamination from visual stimuli. 2-photon calcium imaging data or photometry data was acquired concurrently with this behavioral paradigm on days 1-4.

Pavlovian conditioning during Reversible Head-fixation

The same animals that were tested in the variable sucrose during reversible head-fixation paradigm (see above) were then also trained to associate an auditory cue to subsequent delivery of 10% sucrose reward across 7 days of Pavlovian conditioning during reversible head-fixation. Animals were again tested in 30 minute, 20 trial sessions each day for 7 consecutive days. Intertrial interval was uniformly distributed between 70 and 100 seconds. On each trial, animals were presented with 3 seconds of a 4 kHz auditory cue. 2 seconds after cue termination, 10% sucrose reward was delivered in a series of 5 droplets with total aggregate volume of 10 ul, with 1 second between droplet delivery. Reward delivery and cues were controlled by custom code on an Arduino Uno. The latter half of trials (last 10 out of 20 trials) on Pavlovian day 7 were reward omission trials, where the cue was presented as usual, but the expected sucrose reward was not delivered. Licking behavior was detected and recorded with <1 ms time resolution using a contact

licometer and/or by 15-17 FPS behavioral video. The behavioral testing was conducted in dark or far-red light conditions to remove contamination from visual stimuli. 2-photon calcium imaging data or photometry data was acquired concurrently with this behavioral paradigm on Pavlovian training days 1, 3, 5 and 7.

Automated Lick Detection from Video using Recurrent Neural Network

In some reversible head-fixed behavior sessions, 15-17 FPS, 720p or 1080p behavioral video was collected concurrently with or instead of the contact licometer data for lick detection. The camera was positioned approximately 6-8 cm in front of the animal's face to capture licking behavior. Far-red or infrared light was used to illuminate the animals face during behavioral video acquisition. ROIs near the animal's mouth (160x160 pixels for 1080p video, 100x100 pixels for 720p video) were manually selected and cropped from the original video. Pixel data types were converted from 16-bit RGB to 8-bit integer grayscale. Individual pixels were normalized by taking the absolute value of the z score, using the mean and standard deviation of an individual pixel over the duration of the entire video. Normalized frames were then spatially downsampled to 5x5 pixels. 48 behavioral sessions that had simultaneous contact licometer lick detection and behavioral video were used to identify which video frames contained spout licks. 10 frame sequences of these ground-truth labelled video frames were then used to train a custom recurrent neural network (RNN) (10 long-short-term-memory node layer, followed by 10 node dense layer) using standard tools from the Python TensorFlow library. A RNN implementation was chosen because it considers changes in pixels across sequential frames, rather than only considering one frame at a time. The RNN model for binary frame classification performed with high accuracy and specificity with a ROC AUC = 0.91. The model was then used to predict licking for every frame of 104 30-minute behavioral videos. To further validate the accuracy of the automated detection approach, the predicted classifications of the model were compared to manually scored

video data where researchers scored 1 second time bins as containing animal licks or not. The model again reliably matched the compared lick scoring information, with 82% median overlap between manually scored bins and recurrent neural network frames (see **Fig. S6**). Furthermore, the overall behavioral findings of how much animals licked for different reward conditions and to reward-predictive cues across Pavlovian conditioning were consistent between contact licometer, RNN-scored video and manually-scored video approaches for lick detection. Code for RNN lick detection model is available at github.com/christianepedersen.

In vivo 2-photon Calcium Imaging Acquisition

Image acquisition was performed simultaneously with reversible head-fixed behavior (see above) using the Olympus Fluoview FVMPE-RS 2-photon microscope. A 20x magnification air objective lens (Olympus, LCPLN20XIR, 0.45 numerical aperture, 8.3mm working distance) was optically coupled to the surgically implanted endoscopic lens (Inscopix, 7.3mm length, 0.6mm diameter GRIN lens) of an animal. 920 nm laser light (SpectraPhysics MaiTai Laser, ~100-femtosecond pulse width) excited GCaMP6s fluorophores during resonant scanning (30-Hz roundtrip frame rate, averaging every 6 frames in real-time). Emitted ~535 nm photons were directed to detectors (Olympus, GaAsP PMT) and attributed to the excited pixel. The same focal plane (<5 um tolerance) was optically accessed for each imaging session with the same animal so we could reliably track the activity of the same neurons across multiple days and conditions. A 200-frame, 40-second image series was initiated 20 seconds before reward delivery on each behavioral trial and was triggered by an Arduino Uno that also controlled reward delivery and cues.

In vivo 2-photon Calcium Imaging Data Processing

The Olympus OIR files collected during imaging through Olympus FluoView (FV1200) were exported as tif files. Tif files for individual behavioral trials were motion corrected using a nonrigid motion correction package (Flatiron Institute: github.com/flatironinstitute/NoRMCorre) and concatenated using custom code in MatLab. These whole-session tif files were then converted to an HDF5 format using a custom code. These HDF5 files were then motion corrected in the x–y plane using an adapted hidden Markov model (SIMA version 1.3.0: losonczylab.org/sima/1.3/api/motion.html) as previously reported (Kaifosh 2014). We had found that the imaging plane showed very little z movement in mNAcSh (<5 μm based on a random sample of sessions). Individual neurons were then identified manually and selected by drawing regions of interest (ROIs) manually using the ImageJ polygon tool on the mean projection of pixel intensity across all frames in a session. ROIs for different sessions that corresponded to the same neuron and accurately captured the dynamics of the cell were used to 'track' that neuron across days and conditions. Individual neurons that did not have consistent ROIs for all sessions were not 'tracked' and were not used for comparison of responses to different conditions. The retained ROIs were imported into SIMA and then used for signal extraction.

Custom code added neuropil correction to the SIMA signal extraction (Namboodiri 2019). GCaMP6s labelled neurons were fairly sparse in the focal plane, with minimal overlap of neurons in x, y dimensions. However, changes in the fluorescence of individual neuron ROIs were still contaminated with slight variation in back fluorescence due to neuropil. Neuropil correction was done by first calculating a neuropil signal around each ROI. This was done by calculating a weighted sum of all recorded pixels excluding those falling within a 15-pixel (~17-μm) radius of all ROIs. The weight for any pixel was calculated using a Gaussian function centered on the ROI of interest with a radius of 50 pixels (~45μm). These parameters were obtained after a systematic search of the parameter space in a small subset of sessions and visually comparing the obtained

fluorescence traces against the raw videos in the current mNAcSh dataset as well as in previous studies (Namboodiri 2019). The results in this manuscript are robust against large variation in these parameters. Once the neuropil signal was calculated for every ROI, a correction of this signal was done by subtracting 0.8 multiplied by this signal from the raw calcium trace of the ROI. The results of this previously published neuropil correction procedure (Namboodiri 2019) were also found to be highly consistent with results from another prominent background subtraction approach from the package Suite2p (Pachitariu 2017). After single cell fluorescence time series were extracted and neuropil-corrected from raw calcium image sequences, they were detrended by fitting a double exponential curve to the entire session and subtracting it. Changes in fluorescence from the now consistent baseline were z-score normalized, using the mean and standard deviation of the entire session.

In Vivo 2-photon Calcium Imaging Data Analysis

Individual neurons that were successfully tracked across days and conditions were further analyzed to understand their responsivity to different behavioral conditions. First, the average response of individual neurons to all variable sucrose and Pavlovian conditions was taken by taking the mean fluorescence of each cell across all 20 trials of a given behavioral session (or 10 rewarded trials or 10 extinction trials for Pavlovian Day 7). This trial-averaged response of individual neurons across all variable sucrose or Pavlovian conditions was then used for K-means clustering (number of clusters with lowest silhouette score was chosen, 500x bootstrapping was performed to minimize variability from initial conditions) to group together neurons that had similar responses across these different conditions. K-means clustering and silhouette score calculations were performed in Python using the Scikit-learn library. Several clusters were identified among dyn-SPN and enk-SPN neurons for each behavioral paradigm. The trial-averaged responses for

all neurons within a given cluster were then further averaged to generate the average response traces for each cluster to every condition.

To investigate whether the different reward conditions in the variable sucrose paradigm (0%, 3%, 10%, 30%) could be uniquely identified only by the simultaneous activity of mNAcSh SPNs, a multi-class linear support vector classifier (SVC) model (gamma = XX, cost = 0.8) was created. The linear SVC model was trained and tested on the data for each animal individually. The class of the reward condition and the magnitude of the reward response (0-8 s from onset of reward delivery) of every neuron for a given trial was used as input for the model (80 trials per animal, 20 trials per reward class). 80% of the trials were used to train the linear SVC and 20% were used for cross-validation. The accuracy of the model during cross-validation for each animal is reported in the Results. The linear SVC model was created in Python using the Scikit-learn library and feature coefficients were taken directly from the model and averaged by which K-means cluster the neurons belonged to. To test whether cue-excitation of mNAcSh SPNs was correlated to Pavlovian conditioning session index, trial index within a session or to anticipatory licking to the cue, Pearson's correlations were calculated relative to the magnitude of cue-response (0-5 s from onset of cue) of SPNs that belonged to clusters that were both cue-excited and reward-excited.

To investigate the relationship between SPN cue and reward-response and relative anatomical position, we calculated the relative position of tracked SPNs in the focal plane during 2-photon calcium imaging sessions. The 2-photon resonant scanner generated raw frames that were 512x512 pixels. The x and y pixel position at the center of each ROI that was manually drawn in ImageJ was considered for each successfully tracked neuron. The relative anatomical position of each neuron was calculated as the mean x and y position of its ROI centers across all imaging sessions. Medial-lateral and anterior-posterior axes were inferred based on the orientation of the animal being orthogonal to the objective lens during the imaging and behavior

session. The depth of the focal plane ventral to the GRIN lens during the imaging session introduces a complex magnification effect, so the precise ratio between pixels and anatomical distance in the brain could only be approximated based on the 600 um diameter of the lens and the width of the focal plane being approximately equivalent.

I*n Vivo* Fiber Photometry during Reversible Head-fixation

Fiber photometry recordings were made throughout the entirety of 20 trial, 30-minute head-fixed variable sucrose and Pavlovian conditioning sessions. Prior to recording, an optic fiber was attached to the implanted fiber using a ferrule sleeve (Doric, ZR_2.5). Two LEDs were used to excite GCaMP6s. A 531-Hz sinusoidal LED light (Thorlabs, LED light: M470F3; LED driver: DC4104) was bandpass filtered (470 ± 20 nm, Doric, FMC4) to excite GCaMP6s and evoke Ca^{2+}-dependent emission. A 211-Hz sinusoidal LED light (Thorlabs, LED light: M405FP1; LED driver: DC4104) was bandpass filtered (405 ± 10 nm, Doric, FMC4) to excite GCaMP6s and evoke Ca^{2+}-independent isosbestic control emission. Prior to behavior and recording, a 120 s period of GCaMP6s excitation with 405 nm and 470 nm light was used to remove the majority of baseline drift. Laser intensity for the 470 nm and 405 nm wavelength bands were measured at the tip of the optic fiber and adjusted to ~70 μW before each day of recording. GCaMP6s fluorescence traveled through the same optic fiber before being bandpass filtered (525 ± 25 nm, Doric, FMC4), transduced by a femtowatt silicon photoreceiver (Newport, 2151) and recorded by a real-time processor (TDT, RZ5P). The envelopes of the 531-Hz and 211-Hz signals were extracted in real-time by the TDT program Synapse at a sampling rate of 1017.25 Hz.

Photometry Data Analysis

Custom MATLAB scripts were developed for analyzing fiber photometry data in context of mouse behavior and can be accessed via github.com/christianepedersen. The isosbestic 405 nm excitation control signal was subtracted from the 470 nm excitation signal to remove movement artifacts from intracellular Ca^{2+}-dependent GCaMP6s fluorescence (see Figure S9). Baseline drift was evident in the signal due to slow photobleaching artifacts, particularly during the first several minutes of each 30-minute recording session. A double exponential curve was fit to the raw trace and subtracted to correct for baseline drift. After baseline correction, the photometry trace was z-scored relative to the mean and standard deviation of the entire session. The processed fluorescence timeseries was then aligned to animal behavior and the average fluorescence relative to the time of behavioral events (sucrose reward delivery) was visualized. Quantification of peri-event fluorescence was calculated by taking the average of all samples over a given time window for each trial and comparing between behavioral conditions.

Time-locked Optogenetic Stimulation during Head-fixed Sucrose Consumption

After recovering from surgery, animals were water-restricted and trained to lick 10% sucrose reward while reversibly immobilized and head-restrained. After 4 days of lick training, all animals displayed consistent lick behavior in response to variable interval sucrose presentation and were then put through 2 consecutive test days. Each test day involved one 30 minute, 20 trial session in which the animals were presented with either 10% or 30% sucrose reward. The order of sucrose condition days was pseudorandomized and counterbalanced across animals. For consistency with the imaging and photometry behavior paradigms, animals were presented with 5 droplets of sucrose reward per trial with a total aggregate volume of 10 ul and 1 second between droplet delivery. Animals immediately licked the presented sucrose reward, with average licking

bout durations of 7-8 seconds. On a pseudorandom 50% of the trials within a session, time-locked optogenetic manipulation of mNAcSh-VP^{enk} or mNAcSh-LH^{dyn} activity would be paired with reward delivery, via 4 seconds of 470 nm blue laser light delivery (4 mW for ChR2; 10 mW for PPO measured at end of patch cable, 10 Hz, 10 ms pulse width). 10Hz optogenetic stimulation frequency was chosen because it is physiologically relevant level of neural activity for SPNs during phasic reward-related activity (Taha 2005, Roitman 2005). Any differences in consummatory licking between no laser and laser trials were recorded via the contact licometer circuit with <1 ms time precision. All differences in licking between no laser and laser trials were evaluated within subject to account for the baseline variability in lick rate between animals. To control for possible behavioral effects from the delivery of the blue light itself, cre- littermate controls were also injected, implanted, and tested in identical conditions.

Open-field Locomotion Assay

To determine whether mNAcSh-VP^{enk} or mNAcSh-LH^{dyn} is naturally activated during bouts of motor activity as they are during reward consumption (see Fig. 4, S8), photometry animals were tested in a locomotion assay. Animals freely roamed a 0.5x0.5 m arena in far red light conditions in a quiet environment to measure the amount of spontaneous locomotion in a context that was not anxiogenic. Simultaneous fiber photometry recordings were conducted during the 30-minute long behavioral sessions. 15 FPS video was captured throughout the behavioral session by a camera position directly above the behavioral arena. Automated animal location tracking software (Noldus, Ethovision) was used on the video data after the session to get the x and y coordinates of the animal in the behavioral arena for every frame of the video. Changes in animal coordinates between frames were used to calculate the instantaneous velocity of the animal for each frame. Pearson's correlations were calculated between the velocity of the animal and z scored photometry fluorescence for every frame. Local maxima for movement velocity were

identified as discrete events by using the MatLab 'findpeaks()' command (peak prominence >1 SD, peak width >4 seconds). Windows of the fluorescence time series were aligned to velocity peaks to visualize peri-velocity peak averaged fluorescence. Contrary to the mNAcSh-VP^{enk} and mNAcSh-LH^{dyn} reward-excitations observed during sucrose consumption, bouts of substantial locomotion did not correlate to circuit recruitment.

To determine if mNAcSh-VP^{enk} or mNAcSh-LH^{dyn} activity manipulation could causally disrupt the execution of motor actions, we then performed a similar locomotion assay experiment with intermittent optogenetic photo-excitation or photo-inhibition. Bilaterally implanted optogenetics animals (see above) freely roamed a 0.5x0.5 m arena in far red light conditions in a quiet environment to measure the amount of spontaneous locomotion in a context that was not anxiogenic. Every minute of the 14 minute behavioral session had 30 seconds of pulsing 470 nm laser light (4 mW for ChR2; 10 mW for PPO measured at end of patch cable, 10 Hz, 10 ms pulse width) followed by 30 seconds of no laser. 15 FPS behavioral video was captured throughout the session and was again used to calculate the instantaneous velocity of the animal for every frame. Movement velocity was compared within subject and within session between laser on epochs and laser off epochs. Optogenetic photo-excitation or photo-inhibition of mNAcSh-VP^{enk} or mNAcSh-LH^{dyn} had no significant effect on animal movement velocity. There was also no cumulative effect on locomotion throughout the session as movement velocity was found to be consistent throughout the entire duration of the session.

Statistical analyses

All data collected were averaged and expressed as mean ± SEM. Statistical significance was taken as $^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$, as determined by Pearson's correlation, Student's t-test or one-way ANOVA followed by Tukey post hoc tests as appropriate. For *in situ* hybridization data, we used Student's t-test. For 2-photon imaging and photometry

experiments, we used Pearson's correlation and Student's t-tests, as appropriate. For optogenetic behavioral experiments, we used pairwise t-tests for comparison within subject and within session. All n values for each experimental group are described in the appropriate figure legend. For behavioral experiments, group size ranged from n = 6 to n = 10. For *in situ* hybridization quantification experiments, slices were collected from 2-3 mice, with data averaged from 2-3 slices per mouse. Statistical analyses were performed in GraphPad Prism 8.0 (Graphpad, La Jolla, CA), MatLab 2020 (The MathWorks, Natick, MA) or with Python SciPy library.

6. References

Aguiar, M., Bils, M., Charles, K. K. & Hurst, E. Leisure Luxuries and the Labor Supply of Young Men. w23552 http://www.nber.org/papers/w23552.pdf (2017) doi:10.3386/w23552.

Al-Hasani, R. et al. Distinct Subpopulations of Nucleus Accumbens Dynorphin Neurons Drive Aversion and Reward. Neuron 87, 1063–1077 (2015).

Ambroggi, F., Ghazizadeh, A., Nicola, S. M. & Fields, H. L. Roles of nucleus accumbens core and shell in incentive-cue responding and behavioral inhibition. J Neurosci 31, 6820–6830 (2011).

Anderson, K. D. & Reiner, A. Extensive co-occurrence of substance P and dynorphin in striatal projection neurons: an evolutionarily conserved feature of basal ganglia organization. J Comp Neurol 295, 339–369 (1990).

Anton, B., Fein J., To, T., Li, X., Silberstein, L., Evans, C.J. (1996). Immunohistochemical localization of ORL-1 in the central nervous system of the rat. J Comp Neurol. 368(2), 229-251.

Baimel, C., Bartlett, S.E., Chiou, L.C., Lawrence, A.J., Muschamp, J.W., Patkar, O., Tung, L.W., Borgland, S.L. (2015). Orexin/hypocretin role in reward: implications for opioid and other addictions. Br J Pharmacol, 172(2), 334-348.

Beaulieu, J.-M. & Gainetdinov, R. R. The physiology, signaling, and pharmacology of dopamine receptors. Pharmacol Rev 63, 182–217 (2011).

Beier, K. T., Steinberg, E.E., DeLoach, K. E., Xie, S., Miyamichi, K., Schwarz, L., Gao, X.J., E.J. Kremer, Malenka, R.C., Luo, L. (2015). Circuit Architecture of VTA Dopamine Neurons Revealed by Systematic Input–Output Mapping. Cell, 162(3), 622-634.

Berke, J. D. What does dopamine mean? Nat Neurosci 21, 787–793 (2018).

Berridge, K. C. & Robinson, T. E. What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? Brain Res Brain Res Rev 28, 309–369 (1998).

Bertran-Gonzalez, J. et al. Opposing patterns of signaling activation in dopamine D1 and D2 receptor-expressing striatal neurons in response to cocaine and haloperidol. J Neurosci 28, 5671–5685 (2008).

Besson, M. J., Graybiel, A. M. & Quinn, B. Co-expression of neuropeptides in the cat's striatum: an immunohistochemical study of substance P, dynorphin B and enkephalin. Neuroscience 39, 33–58 (1990).

Binder, E.B., Kinkead, B., Owens, M.J., Nemeroff, C.B. (2001). Neurotensin and dopamine interactions. Pharmacol Rev, 53, 453-486.

Britt, J. P. et al. Synaptic and behavioral profile of multiple glutamatergic inputs to the nucleus accumbens. Neuron 76, 790–803 (2012).

Cartoni, E., Balleine, B. & Baldassarre, G. Appetitive Pavlovian-instrumental Transfer: A review. Neurosci Biobehav Rev 71, 829–848 (2016).

Castro, D. C. & Bruchas, M. R. A Motivational and Neuropeptidergic Hub: Anatomical and Functional Diversity within the Nucleus Accumbens Shell. Neuron 102, 529–552 (2019).

Castro, D. C., Cole, S. L. & Berridge, K. C. Lateral hypothalamus, nucleus accumbens, and ventral pallidum roles in eating and hunger: interactions between homeostatic and reward circuitry. Front. Syst. Neurosci. 9, (2015).

Ciccocioppo, R., Polidori, C., Antonelli, L., Salvadori, S., Guerrini, R., Massi, M. (2002). Pharmacological characterization of the nociceptin receptor which mediates reduction of alcohol drinking in rats. Peptides, 23(1), 117-25.

Clark, J.J., Sandberg, S.G., Wanat, M.J., Gan, J.O., Horne, E.A., Hart, A.S., Akers, C.A., Parker, J.G., Willuhn, I., Martinez, V., Evans, S.B., Stella, N., Phillips, P.E. (2010). Chronic microsensors for longitudinal, subsecond dopamine detection in behaving animals. Nature Methods, 7, 126–129.

Copits, B. A. et al. A photoswitchable GPCR-based opsin for presynaptic silencing. http://biorxiv.org/lookup/doi/10.1101/2021.02.19.432008 (2021) doi:10.1101/2021.02.19.432008.

Coyne, S. M. et al. Suicide Risk in Emerging Adulthood: Associations with Screen Time over 10 years. J Youth Adolesc (2021) doi:10.1007/s10964-020-01389-6.

Creed, M., Ntamati, N. R., Chandra, R., Lobo, M. K. & Lüscher, C. Convergence of Reinforcing and Anhedonic Cocaine Effects in the Ventral Pallidum. Neuron 92, 214–226 (2016).

Danjo, T., Yoshimi, K., Funabiki, K., Yawata, S., Nakanishi, S. (2014). Aversive behavior induced by optogenetic inactivation of ventral tegmental area dopamine neurons is mediated by dopamine D2 receptors in the nucleus accumbens. Proc. Natl. Acad. Sci. USA, 111, 1-6.

Darland, T., Heinricher, M.M., Grandy, D.K. (1998). Orphanin FQ/nociceptin: a role in pain and analgesia, but so much more. Trends in Neurosciences, 21(5), 215-221.

Davis, J. D. The effectiveness of some sugars in stimulating licking behavior in the rat. Physiol Behav 11, 39–45 (1973).

Day, J. J., Wheeler, R. A., Roitman, M. F. & Carelli, R. M. Nucleus accumbens neurons encode Pavlovian approach behaviors: evidence from an autoshaping paradigm. European Journal of Neuroscience 23, 1341–1351 (2006).

de Jong, J. W. et al. A Neural Circuit Mechanism for Encoding Aversive Stimuli in the Mesolimbic Dopamine System. Neuron 101, 133-151.e7 (2019).

Der-Avakian, A., D'Souza, M.S., Potter , D.N., Chartoff E.H., Carlezon Jr., W.A., Pizzagalli, D.A., Markou, A. (2017). Social defeat disrupts reward learning and potentiates striatal nociceptin/orphanin FQ mRNA in rats. Psychopharmacology, 234, 1603-1614.

Engelhard, B. et al. Specialized coding of sensory, motor and cognitive variables in VTA dopamine neurons. Nature 570, 509–513 (2019).

Ewing, A.G. and Wightman, R.M. (1984). Monitoring the stimulated release of dopamine with in vivo voltammetry. II: clearance of released dopamine from extracellular fluid. Journal of Neurochemistry, 43, 570-577.

Fibiger, H. C., LePiane, F. G., Jakubovic, A. & Phillips, A. G. The role of dopamine in intracranial self-stimulation of the ventral tegmental area. J Neurosci 7, 3888–3896 (1987).

Fiorino, D. F., Coury, A., Fibiger, H. C. & Phillips, A. G. Electrical stimulation of reward sites in the ventral tegmental area increases dopamine transmission in the nucleus accumbens of the rat. Behav Brain Res 55, 131–141 (1993).

Flagel, S. B. et al. A selective role for dopamine in stimulus-reward learning. Nature 469, 53–57 (2011).

Fouriezos, G. & Wise, R. A. Pimozide-induced extinction of intracranial self-stimulation: response patterns rule out motor or performance deficits. Brain Res 103, 377–380 (1976).

Gan, J. O., Walton, M. E., & Phillips, P. E. (2010). Dissociable cost and benefit encoding of future rewards by mesolimbic dopamine. Nature neuroscience, 13(1), 25.

Geisler, S., Derst, C., Veh, R. W. & Zahm, D. S. Glutamatergic afferents of the ventral tegmental area in the rat. J Neurosci 27, 5730–5743 (2007).

Gerber, G. J., Sing, J. & Wise, R. A. Pimozide attenuates lever pressing for water reinforcement in rats. Pharmacol Biochem Behav 14, 201–205 (1981).

Glimcher, P.W. (2011). Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. Proc Natl Acad Sci USA, 108(3), 15647-15654.

Gunaydin, L.A., et al. (2014). Natural neural projection dynamics underlying social behavior. Cell, 157(7), 1535-1551.

Hales, C. M., Carroll, M. D., Fryar, C. D. & Ogden, C. L. Prevalence of Obesity and Severe Obesity Among Adults: United States, 2017-2018. NCHS Data Brief 1–8 (2020).

Hamid, A. A. et al. Mesolimbic dopamine signals the value of work. Nat Neurosci 19, 117–126 (2016).

Hart, A.S., Rutledge, R.B., Glimcher, P.W. & Phillips, P.E. Phasic dopamine release in the rat nucleus accumbens symmetrically encodes a reward prediction error term. J. Neurosci. 34, 698–704 (2014).

Hedegaard, H., Curtin, S. C. & Warner, M. Increase in Suicide Mortality in the United States, 1999-2018. NCHS Data Brief 1–8 (2020).

Hedegaard, H., Miniño, A. M. & Warner, M. Drug Overdose Deaths in the United States, 1999-2017. NCHS Data Brief 1–8 (2018).

Hodos, W. (1961). Progressive ratio as a measure of reward strength. Science, 134(3483), 943-944.

Hong, S., Jhou, T. C., Smith, M., Saleem, K. S., & Hikosaka, O. (2011). Negative reward signals from lateral habenula to dopamine neurons are mediated by rostromedial tegmental nucleus in primates. The Journal of Neuroscience, 31(32), 11457-11471.

Hubel, D. H. & Wiesel, T. N. Receptive fields of single neurones in the cat’s striate cortex. J Physiol 148, 574–591 (1959).

Jennings, J. H. et al. Distinct extended amygdala circuits for divergent motivational states. Nature 496, 224–228 (2013).

Jhou, T. C., Fields, H. L., Baxter, M. G., Saper, C. B. & Holland, P. C. The rostromedial tegmental nucleus (RMTg), a GABAergic afferent to midbrain dopamine neurons, encodes aversive stimuli and inhibits motor responses. Neuron 61, 786–800 (2009).

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

Jordan, B. A., Cvejic, S. & Devi, L. A. Opioids and their complicated receptor complexes. Neuropsychopharmacology 23, S5–S18 (2000).

Kobayashi S., Schultz W. (2008). Influence of reward delays on responses of dopamine neurons. J Neurosci 28:7837–7846.

Krause, M., German, P. W., Taha, S. A. & Fields, H. L. A Pause in Nucleus Accumbens Neuron Firing Is Required to Initiate and Maintain Feeding. Journal of Neuroscience 30, 4746–4756 (2010).

Lovaas, I., Newsom, C. & Hickman, C. Self-stimulatory behavior and perceptual reinforcement. J Appl Behav Anal 20, 45–68 (1987).

Lutfy, K., Do, T., Maidment, N.T. (2001). Orphanin FQ/nociceptin attenuates motor stimulation and changes in nucleus accumbens extracellular dopamine induced by cocaine in rats. Psychopharmacology (Berl), 154(1), 1-7.

Lutfy, K., Eitan S., Bryant, C.D., Yang, Y.C., Saliminejad, N., Walwyn, W., Kieffer, B.L., Takeshima, H., Carroll, F.I., Maidment, N.T., Evans, C.J. (2003). Buprenorphine-induced antinociception is mediated by mu-opioid receptors and compromised by concomitant activation of opioid receptor-like receptors. Journal of Neuroscience, 23(32), 10331-10337.

Madisen, L., Zwingman, T. A., Sunkin, S. M., Oh, S. W., Zariwala, H. A., Gu, H., Ng, L.L., Palmiter, R.D., Hawrylycz, M.J., Jones, A.R., Lein, E.S., Zeng, H. (2010). A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. Nature Neuroscience, 13(1), 133-140.

Maidment, N.T., Chen, Y., Tan, A.M., Murphy, N.P., Leslie, F.M. (2002). Rat ventral midbrain dopamine neurons express the orphanin FQ/nociceptin receptor ORL-1. Neuroreport, 13(9), 1137-40.

Mannella, F., Gurney, K. & Baldassarre, G. The nucleus accumbens as a nexus between values and goals in goal-directed behavior: a review and a new hypothesis. Frontiers in Behavioral Neuroscience 7, (2013).

Margolis, E. B., Hjelmstad, G. O., Fujita, W. & Fields, H. L. Direct bidirectional μ-opioid control of midbrain dopamine neurons. J Neurosci 34, 14707–14716 (2014).

Matthews, G. A., Nieh, E. H., Vander Weele, C. M., Halbert, S. A., Pradhan, R. V., Yosafat, A. S., Glober, G. F., Izadmehr, E. M., Thomas, R. E., Lacy, G.D., Wildes, C. P., Ungless, M. A., Tye, K. M. (2016). Dorsal Raphe Dopamine Neurons Represent the Experience of Social Isolation. Cell, 164(4), 617-631.

McCall, J.G., Al-Hasani, R., Siuda, E.R., Hong, D.Y., Norris, A.J., Ford, C.P., Bruchas, M.R. (2015). CRH engagement of the Locus Coeruleus Noradrenergic System mediates Stress-Induced anxiety, Neuron, 87, 605-620.

Mollereau, C, Mouledous, L. (2000). Tissue distribution of the opioid receptor-like (ORL1) receptor. Peptides, 21(7), 907-917.

Mollereau, C., Parmentier, M., Mailleux, P., Butour, J.L., Moisand, C., Chalon, P., Caput, D., Vassart, G., Meunier, J.C. (1994). ORL1, a novel member of the opioid receptor family. Cloning, functional expression and localization. FEBS Lett, 341, 33-38.

Morales, M., Margolis, E.B. (2017). Ventral tegmental area: cellular heterogeneity, connectivity and behaviour. Nature Reviews Neuroscience, 18, 73-85.

Mulvey, B., Bhatti, D. L., Gyawali, S., Lake, A. M., Kriaucionis, S., Ford, C. P., Bruchas, M.R., Heintz, N., Dougherty, J. D. (2018). Molecular and Functional Sex Differences of Noradrenergic Neurons in the Mouse Locus Coeruleus. Cell Reports, 23(8), 2225–2235.

Murphy, N.P., Maidment, N.T. (1999). Orphanin FQ/nociceptin modulation of mesolimbic dopamine transmission determined by microdialysis. J Neurochem 73, 179-186.

Namboodiri, V. M. K. et al. Single-cell activity tracking reveals that orbitofrontal neurons acquire and maintain a long-term memory to guide behavioral adaptation. Nat Neurosci 22, 1110–1121 (2019).

Namburi, P., Al-Hasani, R., Calhoon, G. G., Bruchas, M. R., & Tye, K. M. (2016). Architectural representation of valence in the limbic system. Neuropsychopharmacology, 41(7), 1697.

Neuhoff, H., Neu, A., Liss, B., & Roeper, J. (2002). Ih channels contribute to the different functional properties of identified dopaminergic subpopulations in the midbrain. Journal of Neuroscience, 22(4), 1290-1302.

Nishi, M., Houtani, T., Noda, Y., Mamiya, T., Sato, K., Doi, T., ... & Yamashita, T. (1997). Unrestrained nociceptive response and disregulation of hearing ability in mice lacking the nociceptin/orphaninFQ receptor. The EMBO Journal, 16(8), 1858-1864.

Norton, C.S., Neal, C.R., Kumar, S., Akil, H., Watson, S.J. (2002). Nociceptin/orphanin FQ and opioid receptor-like receptor mRNA expression in dopamine systems. J Comp Neuro, 444, 358-368.

O'Connor, E. C. et al. Accumbal D1R Neurons Projecting to Lateral Hypothalamus Authorize Feeding. Neuron 88, 553–564 (2015).

Olds, J. & Milner, P. Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. Journal of Comparative and Physiological Psychology 47, 419–427 (1954).

Olds, J. Self-stimulation of the brain; its use to study local effects of hunger, sex, and drugs. Science 127, 315–324 (1958).

Oliva, I. & Wanat, M. J. (2019) Operant Costs Modulate Dopamine Release to Self-Administered Cocaine. The Journal of Neuroscience 39, 1249–1260.

Olszewski, P.K., Billington, C.J., Levine, A.S. (2000). Fos expression in feeding-related brain areas following intracerebroventricular administration of orphanin FQ in rats. Brain Research, 855, 171-175.

Otis, J. M. et al. Prefrontal cortex output circuits guide reward seeking through divergent cue encoding. Nature 543, 103–107 (2017).

Ottenheimer, D. J. et al. Reward activity in ventral pallidum tracks satiety-sensitive preference and drives choice behavior. Sci Adv 6, (2020).

Ottenheimer, D., Richard, J. M. & Janak, P. H. Ventral pallidum encodes relative reward value earlier and more robustly than nucleus accumbens. Nat Commun 9, 4350 (2018).

Ozawa, A., Brunori, G., Mercatelli, D., Wu, J., Cippitelli, A., Zou, B., Xie, X.S., Williams, M., Zaveri, N.T., Low, S., Scherrer, G., Kieffer, B.L., Toll, L. (2015). Knock-In Mice with NOP-eGFP Receptors Identify Receptor Cellular and Regional Localization. J Neurosci, 35(33), 11682-11693.

Parker, K.E., Pedersen, C.E., Gomez, A.M., Spangler, S.M., Walicki, M.C., Feng, S.Y., Stewart, S.L., Otis, J.M., Al-Hasani, R., McCall, J.G., et al. (2019). A Paranigral VTA Nociceptin Circuit that Constrains Motivation for Reward. Cell 178, 653-671.e19.

Parker, N.F., Cameron, C.M., Taliaferro, J.P., Lee, J., Choi, J.Y., Davidson, T.J., Daw, N.D., Witten, I. (2016). Reward and choice encoding in terminals of midbrain dopamine neurons depends on striatal target. Nature Neuroscience, 19(6), 845-854.

Patriarchi, J.R. Cho, K. Merten, M.W. Howe, A. Marley, W.H. Xiong, R.W. Folk, G.J. Broussard, R. Liang, M.J. Jang, Tian, L. (2018). Ultrafast neuronal imaging of dopamine dynamics with designed genetically encoded sensor. Science, 360 eaat4422.

Phillips, P. E. M., Stuber, G. D., Heien, M. L. A. V., Wightman, R. M. & Carelli, R. M. Subsecond dopamine release promotes cocaine seeking. Nature 422, 614–618 (2003).

Pich, E. M. & Collo, G. Pharmacological targeting of dopamine D3 receptors: Possible clinical applications of selective drugs. Eur Neuropsychopharmacol 25, 1437–1447 (2015).

Pologruto, T.A., Yasuda, R., Svoboda, K. (2004). Monitoring neural activity and [Ca2+] with genetically encoded Ca2+ indicators. J Neurosci. 24(43), 9572-9579.

Pothion, S., Bizot, J.-C., Trovero, F. & Belzung, C. Strain differences in sucrose preference and in the consequences of unpredictable chronic mild stress. Behav Brain Res 155, 135–146 (2004).

Reddy, A.S., O'Brien, D., Pisat, N., Weichselbaum, C.T., Sakers, K., Lisci, M., Dalal, J.S., Dougherty, J.D. (2017). A Comprehensive Analysis of Cell Type-Specific Nuclear RNA From Neurons and Glia of the Brain. Biol Psychiatry, 81(3), 252-264.

Reed, S. J. et al. Coordinated Reductions in Excitatory Input to the Nucleus Accumbens Underlie Food Consumption. Neuron 99, 1260-1273.e4 (2018).

Resendez, S. L. et al. Visualization of cortical, subcortical and deep brain neural circuit dynamics during naturalistic mammalian behavior with head-mounted microscopes and chronically implanted lenses. Nat Protoc 11, 566–597 (2016).

Reviews in Neuroscience, 30, 259–288.

Richard, J. M., Ambroggi, F., Janak, P. H. & Fields, H. L. Ventral Pallidum Neurons Encode Incentive Value and Promote Cue-Elicited Instrumental Actions. Neuron 90, 1165–1173 (2016).

Richardson, N.R., Roberts, D.C. (1996). Progressive ratio schedules in drug self-administration studies in rats: a method to evaluate reinforcing efficacy. J Neurosci Methods, 66(1), 1-11.

Rodi, D., Polidori, C., Bregola, G., Zucchini, S., Simonato, M., Massi, M. (2002). Pro-nociceptin/orphanin FQ and NOP receptor mRNA levels in the forebrain of food deprived rats. Brain Research, 957(2), 354-361.

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).

Rossi, M. A. et al. Obesity remodels activity and transcriptional state of a lateral hypothalamic brake on feeding. Science 364, 1271–1274 (2019).

Russo, S. J., & Nestler, E. J. (2013). The Brain Reward Circuitry in Mood Disorders. Nature Reviews. Neuroscience, 14(9), 609-625.

Saddoris, M. P., Cacciapaglia, F., Wightman, R. M. & Carelli, R. M. Differential Dopamine Release Dynamics in the Nucleus Accumbens Core and Shell Reveal Complementary Signals for Error Prediction and Incentive Motivation. J Neurosci 35, 11572–11582 (2015).

Sakers, K., Lake, A.M., Khazanchi, R., Ouwenga, R., Vasek, M.J., Dani, A., Dougherty, J.D. (2017). Astrocytes locally translate transcripts in their peripheral processes. Proc Natl Acad Sci USA, 114, E3830-E3838.

Salamone, J.D., Correa, M., Farrar, A.M., Nunes, E.J. & Pardo, M. Dopamine, behavioral economics, and effort. Front. Behav. Neurosci. 3, (2009).

Saunders, B. T., Richard, J. M., Margolis, E. B. & Janak, P. H. Dopamine neurons create Pavlovian conditioned stimuli with circuit-defined motivational properties. Nat Neurosci 21, 1072–1083 (2018).

Schiffmann, S. N. & Vanderhaeghen, J. J. Adenosine A2 receptors regulate the gene expression of striatopallidal and striatonigral neurons. J Neurosci 13, 1080–1087 (1993).

Schultz, W. (2007). Multiple dopamine functions at different time courses. Annual

Schultz, W. Dopamine reward prediction-error signalling: a two-component response. Nat Rev Neurosci 17, 183–195 (2016).

Schultz, W. Getting formal with dopamine and reward. Neuron 36, 241–263 (2002).

Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. Science 275, 1593–1599 (1997).

Seo, D., Funderburk, S.C., Bhatti, D.L., Motard, L.E., Newbold, D., Girven, K.S., McCall, J.G., Krashes, M., Sparta, D.R., Bruchas, M.R. (2016). A GABAergic Projection from the Centromedial Nuclei of the Amygdala to Ventromedial Prefrontal Cortex Modulates Reward Behavior. The Journal of Neuroscience, 36(42), 10831-10842.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks, Genome Research, 13(11), 2498-2504.

Siuda, E.R., Copits, B.A., Schmidt, M.J., Baird, M.A., Al-Hasani, R., Planer, W.J., Funderburk, S.C., McCall, J.G., Gereau, R.W., Bruchas, M.R. (2015). Spatiotemporal control of opioid signaling and behavior. Neuron, 86, 923-935.

Smith, K. S. & Berridge, K. C. Opioid Limbic Circuit for Reward: Interaction between Hedonic Hotspots of Nucleus Accumbens and Ventral Pallidum. Journal of Neuroscience 27, 1594–1605 (2007).

Soares-Cunha, C., Coimbra, B., David-Pereira, A., Borges, S., Pinto, L., Costa, P., Sousa, N., Rodrigues, A.J. (2016). Activation of D2 dopamine receptor-expressing neurons in the nucleus accumbens increases motivation. Nat Commun, 7, 11829.

Stamatakis, A. M. & Stuber, G. D. Activation of lateral habenula inputs to the ventral midbrain promotes behavioral avoidance. Nat Neurosci 15, 1105–1107 (2012).

Stauffer, W,R,, Lak, A., Yang, A., Borel, M., Paulsen, O., Boyden, E.S., Schultz, W. (2016). Dopamine neuron-specific optogenetic stimulation in Rhesus macaques. Cell, 166, 1564-1571.

Stratford, T.R., Holahan, M.R., Kelley, A.E. (1997). Injections of nociceptin into nucleus accumbens shell or ventromedial hypothalamic nucleus increase food intake. Neuroreport, 8, 423-426.

Stuber, G. D. et al. Excitatory transmission from the amygdala to nucleus accumbens facilitates reward seeking. Nature 475, 377–380 (2011).

Sun, F., Zeng, J., Jing, M., Zhou, J., Feng, J., Owen, S. F., … Li, Y. (2018). A Genetically Encoded Fluorescent Sensor Enables Rapid and Specific Detection of Dopamine in Flies, Fish, and Mice. Cell, 174(2), 481-496.e19.

Taha, S. A. & Fields, H. L. Encoding of palatability and appetitive behaviors by distinct neuronal populations in the nucleus accumbens. J Neurosci 25, 1193–1202 (2005).

Tan, K.R., Yvon, C., Turiault, M., Mirzabekov, J.J., Doehner, J., Labouèbe, G., Deisseroth, K., Tye, K.M., Lüscher, C. (2012). GABA neurons of the VTA drive conditioned place aversion. Neuron, 73, 1173-1183.

Taniguchi, H., et al. (2011). A Resource of Cre Driver Lines for Genetic Targeting of GABAergic Neurons in Cerebral Cortex. Neuron, 71(6), 995–1013.

Tejeda, H. A. et al. Pathway- and Cell-Specific Kappa-Opioid Receptor Modulation of Excitation-Inhibition Balance Differentially Gates D1 and D2 Accumbens Neuron Activity. Neuron 93, 147–163 (2017).

Tervo, D.G. et al. (2016). A Designer AAV Variant Permits Efficient Retrograde Access to Projection Neurons. Neuron, 92(2), 372-382.

Thoeni, S., Loureiro, M., O'Connor, E. C. & Lüscher, C. Depression of Accumbal to Lateral Hypothalamic Synapses Gates Overeating. Neuron 107, 158-172.e4 (2020).

Tian, J., Huang, R., Cohen, J.Y., Osakada, F., Kobak, D., Machens, C.K., Callaway, E.M., Uchida, N., Watabe-Uchida, M. (2016). Distributed and mixed information in monosynaptic inputs to dopamine neurons. Neuron, 91(6), 1374-1389.

Toll, L., Bruchas, M.R., Calo', G., Cox, B.M., & Zaveri, N.T. (2016). Nociceptin/Orphanin FQ Receptor Structure, Signaling, Ligands, Functions, and Interactions with Opioid Systems. Pharmacological Reviews, 68(2), 419-457.

Vachez, Y. M. et al. Ventral arkypallidal neurons inhibit accumbal firing to promote reward consumption. Nat Neurosci 24, 379–390 (2021).

van Zessen, R., Phillips, J.L., Budygin, E.A. & Stuber, G.D. (2012). Activation of VTA GABA neurons disrupts reward consumption. Neuron, 73, 1184-1194.

Volkow, N.D., Wang, G.-J., & Baler, R.D. (2011). Reward, dopamine and the control of food intake: implications for obesity. Trends in Cognitive Sciences, 15(1), 37-46.

Wallace, D. L. et al. The influence of DeltaFosB in the nucleus accumbens on natural reward-related behavior. J Neurosci 28, 10272–10277 (2008).

Watabe-Uchida, M., Zhu, L., Ogawa, S.K., Vamanrao, A. & Uchida, N. (2012). Whole-brain mapping of direct inputs to midbrain dopamine neurons. Neuron, 74, 858-873.

Witkin, J.M., Statnick, M.A., Rorick-Kehn, L.M., Pintar, J.E., Ansonoff, M., Chen, Y., Tucker, R.C., Ciccocioppo, R. (2014). The biology of nociceptin/orphanin FQ (N/OFQ) related to obesity, stress, anxiety, mood, and drug dependence. Pharmacol Ther 141, 283-299.

Zaveri, N.T. (2011). The Nociceptin/Orphanin FQ Receptor (NOP) as a Target for Drug Abuse Medications. Current Topics in Medicinal Chemistry, 11(9), 1151-1156.

Zhao, R.J., Woo, R.S., Jeong, M.S., Shin, B.S., Kim, D.G., Kim, K.W. (2003). Orphanin FQ/nociceptin blocks methamphetamine place preference in rats. Neuroreport, 14(18), 2383-2385.

Zheng, F., Grandy, D.K., Johnson, S.W. (2002). Actions of orphanin FQ/nociceptin on rat ventral tegmental area neurons in vitro. Br J Pharmacol. 136(7), 1065-1071.