Dopaminergic Circuits for Compulsion: The Dorsomedial Striatum in Punishment-Resistant Reward-Seeking

BY

JILLIAN LEIGH SEILER
B.S., University of Michigan, 2013
M.A., University of Illinois at Chicago, 2017

THESIS

Submitted as partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Psychology
in the Graduate College of the
University of Illinois at Chicago, 2021

Defense Committee:

Dr. David Wirtshafter, Chair and Academic Advisor
Dr. Michael Ragozzino
Dr. Jamie Roitman
Dr. Amy Lasek, Psychiatry
Dr. Talia Lerner, Northwestern University, Research Advisor

ACKNOWLEDGEMENTS

I feel extraordinarily fortunate to have so many people who helped me immensely to achieve my goal of completing my doctorate. These acknowledgments cannot come close to explaining how much your support has meant to me throughout this period and how grateful I am to each of you.

First, thank you to my committee members for helping me to hone this project into a dissertation and for your time and support. Thank you to my academic advisor, Dr. Dave Wirtshafter, for always lending an ear and guiding me through some of the bumpier parts of my journey. Your true passion for science has inspired me and kept me grounded. To all of the BNS faculty, thank you for your guidance over the past many years—you have all had a hand in helping me grow.

Most of all, thank you to my research advisor, Dr. Talia Lerner, who took a chance on me when she took me in as a graduate student in her new lab. At that time, I had no idea how fortunate I was to get Dr. Lerner as a mentor. Her mentorship encouraged me to believe in my own abilities, as well as develop new ones, and pushed me to become not only a better scientist and scholar, but a better person.

Thank you to my Lerner labmates—for making the lab a safe space where I could truly be myself and for providing endless support for things both big and small. To all my former labmates, thank you for everything you taught me, about science and so much more. I have been very blessed to always look forward to walking into the lab each day because of the people that filled it. To my cohort—Eliza, Jeff, and Sami—there is no one else I could have gotten through this with. Thanks for reminding me when to go to class.

ACKNOWLEDGEMENTS (continued)

Thank you to my incredible friends. I am endlessly grateful for all that you do for me every day. Thank you for listening to me complain about rodents, giving me a shoulder to cry on, bringing me emergency ice cream, and just getting it. You have given me much needed balance and helped me remember to celebrate even the smallest wins.

To my family—none of this would have been possible without all of you. You fostered my curiosity, taught me to shoot for the moon and that I could do anything I set my mind to, helped me make those dreams a reality, and supported me every step of the way. Thanks for always listening, even when you had no idea what I was talking about.

Finally, to my husband Gabe. Thank you for believing in me even when I did not believe in myself. Thank you for all the sacrifices you have made to help me achieve my dreams. And, of course, for being the best practice audience.

JLS

CONTRIBUTIONS OF AUTHORS

Chapters II and III have been submitted as a preprint and is currently available as “Dopamine Signaling in the Dorsomedial Striatum Promotes Compulsive Behavior” at BioRxiv. I (JLS) served as the primary author and Dr. Caitlin V Cosme, Venus N Sherathiya, Joseph M Bianco, Abigael Bridgemohan, and Dr. Talia N Lerner were co-authors. Contributions were as follows: Conceptualization, Methodology, Software, Validation, and Project Administration, J.L.S., C.V.C., and T.N.L.; Investigation, J.L.S., C.V.C., and A.S.B.; Data Curation and Visualization, J.L.S., J.M.B., A.S.B., and T.N.L.; Formal Analysis, J.L.S., V.N.S. and J.M.B..; Writing - Original Draft, J.L.S.; Resources, Writing - Review & Editing, Supervision, and Funding Acquisition, T.N.L.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE PAGE

LIST OF ABBREVIATIONS

ΔF/F	change in fluorescence over baseline
AAV	adeno-associated virus
ANOVA	analysis of variance
AP	anterior-posterior
AUC	area under the curve
ASD	autism spectrum disorder
BLA	basal lateral amygdala
D1/D2/D3	dopamine receptor expressing neuron subtype 1/2/3
DAT	dopamine transporter
DLS	dorsolateral striatum
DMS	dorsomedial striatum
DSM	Diagnostic and Statistical Manual of Mental Disorders
DPR	delayed punishment resistant
DV	dorsal-ventral
FR	fixed ratio
GABA	γ-aminobutyric acid
IP	intraperitoneal
LED	light-emitting diode
LH	lateral hypothalamus
LiCl	lithium chloride
LTP	long-term potentiation
MAD	mean absolute deviation
ML	medial-lateral

LIST OF ABBREVIATIONS (continued)

MSN	medium spiny neuron
NAc	nucleus accumbens
NP	nosepoke
OCD	obsessive compulsive disorder
OFC	orbito-frontal cortex
PBS	phosphate buffered saline
PE	port entry
PET	positron emission tomography
PFA	paraformaldehyde
PFC	prefrontal cortex
PR	punishment resistant
PS	punishment sensitive
PSTH	peri-stimulus time histogram
RI	random interval
RR	random ratio
SD	standard deviation
SEM	standard error of the mean
SNc	substantia nigra pars compacta
SPECT	single-photon emission computed tomography
SQ	subcutaneous
SUD	substance use disorder
TH	tyrosine hydroxylase

LIST OF ABBREVIATIONS (continued)

TTL	transistor-transistor logic
VTA	ventral tegmental area

SUMMARY

The dorsal striatum is involved in the control of habit formation, with a hypothesized shift from dorsomedial striatum (DMS) to dorsolateral striatum (DLS) as animals move from goal-directed behavior to habits. Though we often think of “bad” habits, habitual responding is a useful strategy that allows us to free ourselves from the cognitive load of worrying about which motor sequence we need to employ to walk to the kitchen for a snack. In psychological terms, a habit is defined as behavior controlled by stimulus-response associations, which are insensitive to degradations of action-outcome contingencies or to devaluations of the outcome itself. Compulsivity, while related to and often linked to habitual behavior, is defined as behavior which is resistant to punishment. Compulsive behavior is especially relevant to understanding diseases like substance use disorder (SUD) and obsessive-compulsive disorder (OCD), where, for example, drug seeking persists despite negative effects on physical health. Indeed, six of the eleven criteria for SUD relate to persistent drug seeking despite negative outcomes (American Psychiatric Association, 2013). The reason for resistance to punishment in compulsive behavior is unclear and may be due to inability to learn new action-outcome contingencies or because drive for reward is so strong that it is “worth” risking negative outcomes. While a few studies have shown potential links between compulsivity and the dorsal striatum, the exact involvement of this area is far from well-understood (Giuliano et al., 2019; Lipton et al., 2019; Lüscher et al., 2020).

The studies presented here aim to elucidate the role of nigrostriatal dopamine in punishment-resistant reward-seeking. We first used fluorescent calcium indicators to measure activity of dopamine terminals simultaneously in the dorsomedial (DMS) and

dorsolateral (DLS) striatum as animals were trained on an operant task known to produce habit-like responding for sucrose pellets (Derusso et al., 2010; Wiltgen et al., 2012). Animals were probed for punishment-resistant reward-seeking at the beginning and end of training. We found large individual variability in the behavior of animals and that a subset of animals naturally became punishment-resistant. After analysis of the neural recordings, we found that we could predict this punishment-resistance based on the activity of the dopamine axons in DMS, but not DLS.

We next wanted to see if we could cause animals to become more or less resistant to punishment by manipulating the dopamine terminal activity in both DMS and DLS using optogenetics. We found that, consistent with our neural recordings, exciting DMS dopamine terminals did, in fact, cause animals to become more punishment-resistant faster, whereas exciting DLS dopamine terminals had no effect. Conversely, inhibiting DMS dopamine terminals slowed learning of the task and led to fewer punishment-resistant animals. Together, these results demonstrate a clear role of DMS in compulsive-like behavior and lay the groundwork for future studies examining how DMS is engaged in diseases involving compulsion.

I. INTRODUCTION

A. Background

The impact of substance use disorder (SUD), or drug addiction, in the US and globally is massive. From a public health perspective, overdose deaths have been steadily on the rise for the past decade and account for more deaths annually than car accidents, suicide, or gun-related deaths (Centers for Disease Control, 2019). The National Institute on Drug Abuse estimates the average annual cost of drug abuse to be more than $740 billion (National Institute on Drug Abuse, 2017). And of course, these statistics say nothing of the personal impact that SUD has on those afflicted or their family members. While the immediate mechanisms of drugs of abuse have been understood for quite some time, much is left to be understood about the circuitry involved in compulsive drug seeking, such as is seen in SUD (American Psychiatric Association, 2013). At the most basic level, drugs, such as cocaine and amphetamine, act by altering dopamine levels in the brain. Cocaine, for example, exerts its primary effects by blocking the reuptake of dopamine through its transporter DAT, perhaps most significantly, in the nucleus accumbens (NAc; Ritz et al. 1988). As the mechanism of action for more and more drugs of abuse proved to increase levels of dopamine in the brain, researchers believed that they had found the mechanism of addiction—and thus began “the dopamine theory of addiction.” Armed with this knowledge, pharmaceutical companies invested significant financial resources into efforts to treat people suffering from drug addiction with pharmacological blockade of dopamine. Yet, these treatments had little success in human subjects. Conversely, patients with Parkinson’s disease whom are taking dopamine receptor agonists report side effects such as compulsive

eating and gambling addiction, consistent with the dopamine theory of addiction (Marques et al., 2018; Napier et al., 2020). Interestingly, naltrexone, a medication that has been, at least marginally, effective in treating some forms of addiction, like alcoholism, does not directly impact dopamine. The failures and successes of treatment for addiction thus far have led to a chasm between basic science research and human treatment which has puzzled scientists. Much research has emerged as an attempt to bridge this disconnect, the vast majority of which has remained focused on the mesolimbic dopamine pathway between the ventral tegmental area (VTA) and NAc or ventral striatum. Yet less is known about the involvement of the neighboring nigrostriatal dopamine pathway, from the substantia nigra pars compacta (SNc) to the dorsal striatum, in substance use, despite evidence implicating its participation in later stages of substance use disorder. Indeed, compulsive behavior and addiction research have largely been disconnected from one another, despite the clear overlap in behavioral phenotype. This work aims to provide context for the involvement of nigrostriatal dopamine in compulsive reward seeking, a hallmark of substance use disorder.

1. Ventral striatum and addiction

Though the first known record of the effects of drugs comes from Hippocrates in the 4th Century BCE, their precise mechanism in the brain was first described in the 1970's and early 1980's. Olds and Milner had long since discovered that animals would press a lever in order to receive stimulation to brain regions and, subsequently this area was found to contain dopamine neurons (Crow, 1972; Olds & Milner, 1954). They had also seen that stimulant drugs increased animals willingness to work for this self-

stimulation (Stein, 1964). Researchers theorized that these “pleasure” centers (as they were incorrectly deemed at the time), must also be involved in the hedonic effects of drugs. The first concrete evidence for this dopamine theory of addiction came from studies showing that giving neuroleptics, known dopamine receptor antagonists, blocked the rewarding effects of stimulant drugs (de Wit & Wise, 1977; R A Wise & Bozarth, 1987). Shortly thereafter, with the invention of microdialysis, the exact location of this effect was discovered to be the ventral striatum—specifically the nucleus accumbens (NAc, Di Chiara & Imperato, 1988).

While a pure theory of positive reinforcement via dopamine pathways made sense of the initial reinforcement of drugs, it could not fully explain the mechanism of drug addiction. Evidence showed that as drug seeking became compulsive, subjects reported progressively less pleasure (ie liking), despite increased cravings and desire for the drug (ie wanting; Berridge & Robinson, 1998). Thus, Berridge and Robinson put forth the incentive-sensitization theory of drug addiction, which posits that through activation of the dopamine system, incentive-salience is attributed to drugs, meaning a user becomes highly motivated toward drugs, as one would be motivated to obtain natural rewards, like food (Robinson & Berridge, 1993). The sensitization component of this theory suggests that with repeated drug use, neural plasticity occurs and causes sensitization to drugs. In combination, this keeps users craving and seeking the drug in a compulsive fashion, despite loss of initial hedonic properties.

Ungless et al, provided evidence to support the idea that drugs of abuse alter mesolimbic circuitry by demonstrating that even a single dose of cocaine caused long-term potentiation (LTP) in the ventral tegmental area (VTA) that lasted several days

(Ungless et al., 2001). Indeed, there are many sites of plasticity within, and connecting to, the mesolimbic pathway that ensue following drug use—such as connections from prefrontal cortex (PFC) and to basolateral amygdala (BLA; Robinson and Kolb 2004; Pascoli et al. 2014; Terrier, Lüscher, and Pascoli 2016). What though, do these changes at the synaptic level actually mean in regards to the control of behavior with prolonged drug use and addiction?

Several seminal studies from Nora Volkow in the 1990s added to the understanding of neural mechanisms of addiction. Using positron emission tomography (PET) and single-photon emission computed tomography (SPECT) technology in human subjects and C11- raclopride, a radiolabeled selective dopamine receptor expressing neuron subtype 2/3 (D2/D3) antagonist, researchers were able to detect changes in dopamine activity corresponding to specific time-locked events (Volkow et al., 1994). Initially, much of this work confirmed previous theories of the role of dopamine in the ventral striatum in addiction, showing that the magnitude of increase in dopamine in NAc was associated with subjective ratings of the euphoric “high” participants felt when taking amphetamine (Laruelle & Abi-Dargham, 1999). However, although a correlation was observed for amphetamine there was no evidence for an increase in dopamine as a direct result of other drugs with known abuse potential such as ketamine and alcohol (Aalto et al., 2002; Yoder et al., 2007). Moreover, Volkow and colleagues showed that drug users had lower striatal dopamine receptor availability than controls and that individuals with lower D2 receptor availability reported higher subjective pleasure ratings from drugs. Studies also revealed a “blunting” effect in drug users, where there was a decrease in dopamine release in response to drugs as

compared to non-users (Martinez et al., 2007, 2011; Volkow et al., 1997). Both of these factors run contrary to the idea that drugs of abuse act by increasing dopamine. Important for this work, Volkow and colleagues found that activity in the dorsal striatum, but *not* the ventral striatum, was correlated with subjective reports of craving in cocaine addicted subjects (Volkow et al., 2006). This key piece of evidence suggests that, while perhaps the initial mechanism of action of drug of abuse relies on mesolimbic circuitry, later compulsive drug-craving in substance use disorder (SUD) may rely on nigrostriatal circuits.

2. Dorsal striatum and habit

Though it may be comparatively less studied than the ventral striatum in the field of addiction research, the dorsal striatum has been well-studied in its own right. Primarily, studies have linked habitual behavior and motor skill learning to the dorsal striatum, or the caudate and putamen in humans. Initial evidence came from a series of lesion and inactivation studies which demonstrated that lesioning or blocking activity in the dorsomedial striatum (DMS) led to a decrease in goal-directed behavior (Gremel & Costa, 2013; Yin, Knowlton, et al., 2005; Yin, Ostlund, et al., 2005) whereas lesioning or blocking activity in the dorsolateral striatum (DLS) prevented the formation of habits (Yin et al., 2004, 2006). Similarly, lesions of the dopamine neurons in the substantia nigra pars compacta (SNc), which project to dorsolateral striatum, also blocked habit formation (Faure et al., 2005). These researchers use the same principles that we know lead to habits in humans, repetition and uncertainty, in an effort to push their animals toward habitual responding for a desired outcome, typically a sugar pellet. A host of

papers have shown that, to this end, certain schedules of reinforcement, namely random interval ("RI"), are better than others (random ratio or "RR" for example) at reliably inducing habits (Derusso et al., 2010; Gremel & Costa, 2013; Wiltgen et al., 2012; Yin, Ostlund, et al., 2005). The initial group of studies demonstrating the roles of DMS and DLS in habit all used outcome devaluation to probe for habits—either by sating animals with the specific reinforcer being used in the task or by pairing that reinforcer with LiCl to cause a conditioned taste aversion. In both scenarios, the outcome becomes less valuable to the animal. If the animal is acting habitually, they should be insensitive to this devaluation. In the study by Yin, Ostlund et al., they also probe animals with action-outcome degradation, where the contingency by which animals have learned to get the reinforcer is broken or altered in some way, such that responding habitually will not result in the learned outcome (Yin, Ostlund, et al., 2005). Therefore, it will take longer for habitual animals to learn the new action-outcome contingency or generally inhibit their behavior. Based on these lesion and inactivation studies, it seems that as animals become more habitual, they transition from DMS- to DLS-mediated behavioral control.

Beyond these foundational studies, work has also been done to more closely examine the control that DMS and DLS have over both goal-directed and habitual behavior. In tasks where animals worked for drugs, such as cocaine or alcohol, blockade or inactivation of DMS dopamine signaling caused a decrease in drug seeking behavior early on in training, but had no effect once animals had gone through extended training on the task (Corbit et al., 2012; J. E. Murray, Belin, et al., 2012). Blocking dopamine signaling in the DLS, on the other hand, caused habitual animals to become

sensitive to outcome devaluation once again and decreased overall seeking for drugs late in training, but not early on (Corbit et al., 2012; Giuliano et al., 2019; Hodebourg et al., 2018; J. E. Murray, Belin, et al., 2012; Pacchioni et al., 2011; Vanderschuren et al., 2005; Zapata et al., 2010). The effects seen following inactivation of DMS or DLS were also dependent on the task structure—where inactivating DMS decreased performance on fixed ratio (FR) tasks but not progressive ratio, and vice versa (J. E. Murray, Belin, et al., 2012; Spoelder et al., 2017; Veeneman et al., 2012). Task structure has also been demonstrated to differentially involve DMS or DLS. DMS lesions decrease cognitive flexibility and reversal learning on tasks that require the animal to keep track of outcomes, such as a probabilistic maze task (Castañé et al., 2010; Grospe et al., 2018; Palencia & Ragozzino, 2004). Additionally, DMS but not DLS showed task-related responses to unpredicted and predicted reward stimuli after training (Brown et al., 2011). In contrast, lesions of DLS have no effect on performance of this type of task but do impair performance on egocentric and instrumental tasks (Braun et al., 2015; Faure et al., 2005; Seip-Cammack et al., 2017). By and large, these studies have supported the initial findings by Yin and colleagues in defining the roles of DMS and DLS.

Though the results of inactivating or lesioning DMS and DLS produce fairly consistent results in regards to behavior, data from *in vivo* electrophysiological recordings of dorsal striatum are less clear. Some studies have shown initial engagement and later disengagement of DMS neurons over the course of operant training and the converse pattern with DLS activity emerging later in training (Thorn et al., 2010; Yin et al., 2009). Other work, however, has shown that DMS activity continues throughout training and remains even when the animals had transitioned to habitual

responding on the task (Gremel & Costa, 2013; Vandaele Y et al., 2019). These studies also did not find a consensus as to when DMS or DLS was engaged in response to the operant task itself. Thus, more work is necessary to parse out the real-time activity of subregions of the dorsal striatum across training.

3. Dorsal striatum and compulsion

While this transition from DMS to DLS has been well documented with the aforementioned lesion and inactivation studies, several studies have brought up issues with this model that require further clarification. In a paper by Singer et al., they demonstrated through an elaborate puzzle task, which prevented the formation of stimulus-response associations, that animals could develop addiction-like phenotypes for cocaine—as measured by insensitivity to shock-pairing, escalation of behavior, and cue-evoked reinstatement—without the development of “habit” (Singer et al., 2017). Importantly, while these animals were not habitual, they were resistant to adverse consequences, and were therefore compulsive. Further, the study proves that this compulsive drug seeking was not DLS-dependent, as the behavior persisted following pharmacological blockade of DLS. This study provides an indication that perhaps compulsivity and stimulus-response habit associations are not always formed in tandem, can be studied independently and may rely on different circuitry.

The findings linking cue-induced craving to dorsal striatum, as well as work demonstrating reorganization of inputs to SNc following cocaine exposure, and others, have hinted at a major role for nigrostriatal dopamine in addiction drug addiction and

compulsive drug seeking that is yet to be fully uncovered (Beaudoin et al., 2018; Volkow et al., 2006).

4. A neural mechanism for the progression to compulsion

How exactly does activity in ventral striatum associated with the initial rewarding effects of drugs of abuse, transition to dorsal striatum activity in compulsive abuse of drugs? Suzanne Haber and colleagues identified a potential anatomical pathway linking the midbrain and the striatum in primates (Haber et al., 2000). Based on this tracing work and a similar study in rodents, the so-called “ascending spiral” model has been posited as a functional means of connecting the mesolimbic dopamine pathway to the nigrostriatal dopamine pathway (Ikemoto, 2007). As shown in Figure 1, activity in VTA dopamine neurons project to NAc, where GABAergic medium-spiny neurons (MSN) project back to slightly more lateral parts of VTA onto GABA interneurons, which in turn disinhibit lateral VTA dopamine neurons. This pattern continues, eventually connecting NAc shell to lateral SNc, and so forth. While some of these connections, namely the reciprocal connections between DMS and SNc and DLS and SNc, have been demonstrated *in vitro*, much of this circuit has yet to be mapped functionally (Lerner et al., 2015). Still, it provides a likely mechanism by which neural circuits of the rewarding effects of drugs may engage more dorsolateral circuits for habit and compulsion. A study by Belin and Everitt provided a key piece of evidence in support of the involvement of this serial connection in drug seeking by showing that disconnections between the ventral and dorsal striatum did decrease drug seeking in animals with extensive training (Belin & Everitt, 2008). Further, Willuhn et al. showed that as an

animal undergoes weeks of cocaine self-administration, dopamine activity decreases in the NAc and increases in the DLS (Willuhn et al., 2012). Moreover, they demonstrated that blocking NAc activity early in training prevented DLS activity from ever occurring. Alternatively, compulsive behavior alone might not require earlier activation of ventral striatum and activity in the dorsal striatum may come online independently or through top-down activation of cortical inputs, like orbito-frontal cortex (OFC). Evidence from studies in which animals optogenetically self-stimulated VTA dopamine neurons have shown that those that become compulsive in their responding (again defined as behavior that persisted in the face of aversive consequences) had enhanced activity in OFC projections to DMS as compared to animals that did not become compulsive (Pascoli et al., 2018). This increased strength of the OFC to DMS connection was also observed in animals that developed compulsive drug seeking for amphetamine and cocaine (Hu et al., 2019a; Wall et al., 2019). Overall, though several neural mechanisms have been proposed or implicated in the transition to compulsion, a close circuit level analysis is lacking.

5. Natural rewards versus drugs

Do drugs of abuse uniquely hijack neural circuits to create compulsive behavior, or can compulsions develop “naturally” for rewards such as sugar? Animals working for natural rewards may engage dorsal striatal circuitry differently than those working for drugs. Indeed, several studies found effects of blocking dorsal striatum on drug seeking but not sucrose seeking (Vanderschuren et al., 2004; Yager et al., 2018). We do not know if the effects seen in studies where animals are working for drugs are due to

immediate effects of these drugs on the dopamine neurons, a cumulative effect of prolonged drug daily drug use, or a synergistic combination of both (Norman & Tsibulsky, 2006; R. A. Wise et al., 1995). Nor do we know, if either of the latter options is true, how long any alterations to this circuit might persist. Further, there is some evidence to suggest that the specific neurons which encode choice for drug are not the same as those that encode choice for natural rewards, suggesting that these pathways may be separable on a cellular level (Carelli et al., 2000; Carelli & Wondolowski, 2003). To date, drugs and natural rewards have been used nearly interchangeably in studying the behavioral correlates of the dorsal striatum. A careful analysis and comparison of these two may clear up outstanding contradictions in this literature and tease apart different patterns of neural activity in the nigrostriatal pathway.

6. Individual variability in habit and compulsion

Several studies have found distinct groups of individuals that looked more compulsive or habitual than others (DePoy et al., 2016; Giuliano et al., 2019; McKim et al., 2016; Siciliano et al., 2019; Yager et al., 2015). This is consistent with the idea that only a relatively small proportion of drug users will develop substance use disorder (around 10-20% for cocaine; Anthony, Warner, & Kessler, 1994) or of those who gamble will become problem-gamblers. By systematically analyzing these individual differences, specific traits can be correlated with neural differences or other behavioral outcomes. For example, in the study by Giuliano et al., animals that had the greatest reduction in behavior following DLS inactivation were the same animals that showed the most compulsive behavior later on in training (Giuliano et al., 2019). Depoy et al. found

nearly opposite effects of sex on the relationship between escalation of drug-taking and later inflexible behavior, where males that escalated drug intake in adolescence showed more habit-like behavior in adulthood, but the opposite was true for females (DePoy et al., 2016). Had sex not been considered in this study, the effect would have washed out and this important finding could have been missed.

In order to capture the variability in a population of animals, some researchers have used a three-criteria model of addiction (Deroche-Gamonet et al., 2004). Animals are considered to display addiction-like behavior when they continue to work for drug despite it being signaled to the animal as unavailable, they have extreme motivation for the drug as measured in a progressive ratio task, and are resistant to punishment (i.e., compulsive). This model aims to increase the external validity of rodent work in capturing SUD according to the Diagnostic and Statistical Manual of Mental Disorders (DSM) definition, but may actually serve to erase important individual differences between animals that meet some, but not all, of these criteria.

7. Adolescence is a period of vulnerability

Adolescence is an age period marked by increased risk-taking and decreased inhibition. Ethologically, it makes sense for an animal to explore outside its comfort zone at the age of sexual maturity to meet potential mates. Yet for humans, this comes with risky decision making in other domains (Kelley et al., 2004; Spear, 2000). Teenagers, ages 16-19, are three times more likely than drivers 20 and over to be involved in a fatal car crash (Centers for Disease Control, 2017). Adolescents have the highest rate of

experimental drug use (Compton et al., 2007). Though not all of those that experiment with drugs develop SUD, the average age of first use in those who do develop SUD is between 15-19 for illicit drugs, alcohol, and tobacco (Giovino, 1999; Wagner & Anthony, 2002).

Decision making and risk strategies are largely controlled by the prefrontal cortex (PFC) and orbitofrontal cortex (OFC; Harlow, 1868; E. A. Murray, O'Doherty, & Schoenbaum, 2007). These are also some of the last regions of the brain to develop and mature. Thus, during adolescence, there is an imbalance of excitation and inhibition in both of these areas and a lack of control over the structures they normally regulate. This lack of top-down control from frontal brain regions, could therefore promote not only risky and exploratory, but also compulsive behavior, as is seen in SUD. Exposure to drugs in adolescence has been shown to significantly impact risky decision making and behavioral flexibility, even later in life, largely via these prefrontal cortices (Amodeo et al., 2017; Jacobs-Brichford et al., 2019; McMurray et al., 2016). The connection between OFC and the dorsal striatum has been shown to be involved specifically in control over compulsive drug-seeking in multiple studies (Ersche et al., 2011; Gremel et al., 2016; Gremel & Costa, 2013; Hu et al., 2019a; Pascoli et al., 2018; Siciliano et al., 2019). Studies have shown that both cocaine and amphetamine exposure can promote later habitual behavior, but it is not clear if the timing of drug exposure is critical in this effect (DePoy et al., 2016; Nelson & Killcross, 2006). The combination of the changing brain of adolescents combined with the manifestations of risky behaviors during this time, make adolescence a unique period, in which drugs of abuse may have a synergistic interaction on the developing brain into adulthood.

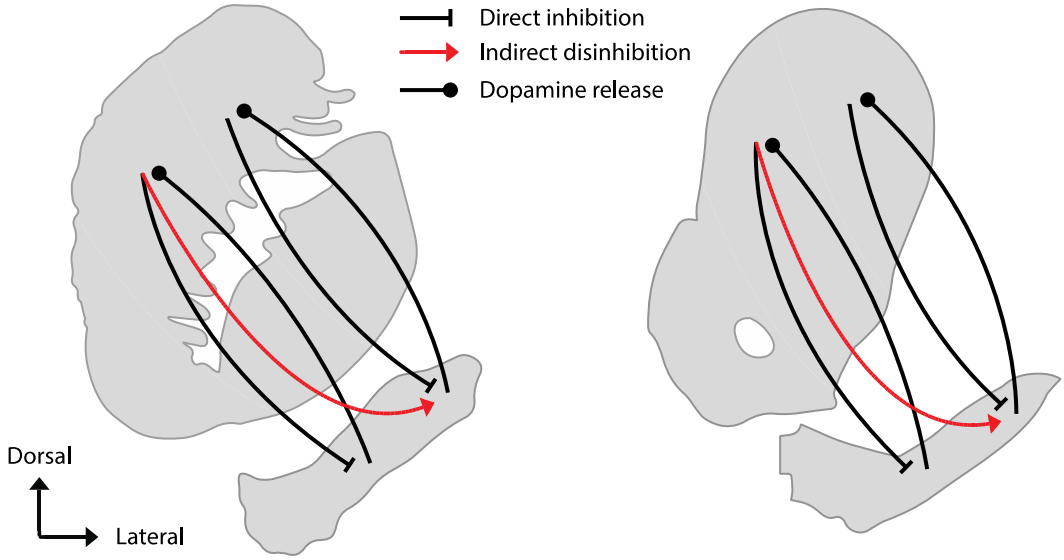

Figure 1. Ascending spiral model. A. Schematic of proposed ascending spiral model (based on (Haber, Fudge, and McFarland 2000)) in a primate brain. B. Schematic of proposed ascending spiral model (based on (Haber, Fudge, and McFarland 2000)) in a rodent brain.

II. CHAPTER TWO: A role for dopaminergic signaling in the dorsal striatum during the development of compulsive behavior

A. Rationale

Compulsion is at the heart of many psychological disorders like addiction, obsessive-compulsive disorder (OCD), and autism spectrum disorder (ASD), where people face an inability to stop behaviors such as drug use, binge-eating, hand-washing, or self-stimulation in the face of a wide variety of consequences. Identifying the exact neural circuits that underlie compulsion may provide new treatment opportunities for the destructive behaviors often associated with disorders of habit.

While previous literature has investigated the neural mechanisms of these diseases in the context of habit formation, the precise relationship between compulsion and habit remains unclear. Perhaps some of the conflict that remains unresolved in the field of habit may be attributable to imprecise definitions of compulsion and habit.

Significant work has been conducted to identify the brain regions involved in different aspects of learning and habit formation, revealing a major role for the nigrostriatal pathway. Throughout the basal ganglia, GABAergic connections to dopamine neurons lead to a series of disinhibtions and shift control from one region to the next over the course of habit formation. To date, the engagement of DMS and DLS specifically through this spiral, have not been demonstrated in-vivo.

Recent advances in technology have allowed us to examine the activity of neural populations more easily in awake and behaving animals (C. K. Kim et al., 2016). This study utilized dual site in-vivo calcium imaging to examine activity in DMS and DLS-

projecting dopamine terminals at multiple time points to reveal the exact dynamics and relationship between these regions as animals transition their behavioral strategies. Further, while it appears that compulsion and habit can be dissociated behaviorally, this technique also allowed us to ask whether the circuits that underlie habit and compulsion are the same or different. What patterns of activity underlie compulsive responding? Can they be isolated from those that underlie habitual responding? The study aims to record real-time activity in these regions across training on a random interval schedule of reinforcement that is known to promote habit (Derusso *et al.*, 2010; Wiltgen *et al.*, 2012) and conduct behavioral probes to measure both compulsivity, measured as punishment-resistant reward seeking, and habit, measured as omission-resistant reward seeking. This study design allowed us to correlate the behavior of individual animals with their neural activity and thereby account for variation in strategy and learning in a population of animals. Together, these results should help to elucidate a pattern of activity in the dorsal striatum that is associated with compulsive behavior.

B. Materials and Methods

1. Subjects

Male and female *WT* (C57BL/6J) and (DAT)::IRES-Cre knockin mice (JAX006660) were obtained from The Jackson Laboratory and crossed in house. Only heterozygote transgenic mice, obtained by backcrossing to C57BL/6J wildtypes, were used for experiments. Littermates of the same sex were randomly assigned to experimental groups (fiber photometry-14 males, 22 females; DMS excitatory optogenetics- 20 males, 19 females; DMS inhibitory optogenetics- 13 males, 13 females; DLS excitatory optogenetics- 18 males, 18 females). Adult mice at least 10

weeks of age were used in all experiments. Mice were group housed under a conventional 12h light cycle (dark from 7:00pm to 7:00am) with *ad libitum* access to food and water prior to operant training. All experiments were approved by the Northwestern University Institutional Animal Care and Use Committee.

2. Methods Details

Operant Behavior

Mice were food restricted to 85% of ad libitum body weight for the duration of operant training. Mice were given one day of habituation to operant chambers (Med Associates) and tethering with patch cords (Doric Lenses) for one hour. They were then trained to retrieve food rewards (45 mg purified pellet, Bio-Serv) from a magazine port. For this magazine training, pellets were delivered to the port on a random interval (RI60) schedule non-contingently for one hour. Next, operant training began, with all training sessions lasting one hour or until 50 rewards had been earned. Mice were trained to associate nosepoking with reward on a fixed ratio (FR1) schedule where both nosepokes delivered a reward. They had to retrieve the reward (as measured by making a port entry following a rewarded nosepoke) before they could earn the next reward. After a mouse showed a preference for one nosepoke (>25 rewards on that side; average of 3.06 days), they were trained on FR1 on their preferred side only, with nosepokes on the other side having no consequence, until they received >30 rewards for a minimum of two consecutive days (average of 5.87 days). Mice that did not reached this criterion after 14 days of FR1 training (mean+2 SD), were removed from the study. Mice passing the FR1 criterion were then moved to either a random interval

(n=36) or random ratio (n=7) schedule of reinforcement. Mice on the random interval schedule were trained on RI30 until they earned >30 rewards in one hour (average of 2.33 days), and then trained on RI60. Mice on a random ratio schedule of reinforcement were trained on RR10 until they earned >30 rewards in one hour (average of 2.71 days), and then trained on RR20 (Fig. 2A). For random interval and random ratio schedules, a normal distribution centered around the number indicated in the name of the schedule was used to create the schedule. The range for RI30 was from 15-45s, RI60 from 30-90s, RR10 from 6-14 nosepokes, and RR20 from 14-28 nosepokes.

Shock Probe

Mice were subjected to a footshock probe early and late in training (Fig. 2B) to evaluate their levels of punishment-resistance reward-seeking. These probes were performed under an FR1 schedule of reinforcement where a mild footshock (0.2mA, 1s) was paired with a subset of rewarded nosepokes on a RR3 schedule, so that, on average, every third rewarded nosepoke was paired with a footshock. During shock probes, the session ended after 60 minutes or a mouse was inactive (no nosepokes on the rewarded side) for >10 minutes. There was no maximum number of rewards.

Omission Probe

A subset of mice (n=20) were returned to RI60/RR20 training after the late footshock probe until their nosepoke rates returned to pre-shock levels. They then received a single omission probe session where they had to withhold nosepoking for 20 seconds in

order to receive a single reward pellet. A nosepoke reset the 20 second timer. Each session ended after a mouse received 50 rewards or 60 minutes had elapsed.

Fear Conditioning

A trace fear conditioning paradigm, adapted from Lugo, Smith, and Holley (2014) was used in a naïve cohort of wild-type mice (n=13) to verify that our shock intensity (0.2mA) is aversive to the mice (Lugo et al., 2014). Mice were randomly assigned to cued or non-cued groups. On the first day, mice received 12 tone only or tone-shock pairings (2900 Hz tone) in a standard operant chamber (Med Associates). The next day, mice were placed in a different context (using white walls, white plastic flooring, and vanilla scent) and 12 tones were presented. All sessions were recorded using Med Associates Video Monitor software.

Stereotaxic Surgery

Viral infusions and optic fiber implant surgeries took place under isoflurane anesthesia (Henry Schein). Mice were anesthetized in an isoflurane induction chamber at 3-4% isoflurane, and then injected with buprenorphine SR (Zoopharm, 0.5 mg/kg s.q.) and carpofen (Zoetis, 5 mg/kg s.q.) prior to the start of surgery. Mice were placed on a stereotaxic frame (Stoetling) and hair was removed from the scalp using Nair. The skin was cleaned with alcohol and a povidone-iodine solution prior to incision. The scalp was opened using a sterile scalpel and holes were drilled in the skull at the appropriate stereotaxic coordinates. Viruses were infused at 100 nl/min through a blunt 33-gauge

injection needle using a syringe pump (World Precision Instruments). The needle was left in place for 5 min following the end of the injection, then slowly retracted to avoid leakage up the injection tract. Implants were secured to the skull with Metabond (Parkell) and Flow-it ALC blue light-curing dental epoxy (Pentron). After surgery, mice were allowed to recover until ambulatory on a heated pad, then returned to their homecage with moistened chow or DietGel available. Mice then recovered for three weeks before behavioral experiments began.

Fiber Photometry

Mice for fiber photometry experiments received infusions of 1µl of AAV5-CAG-FLEX-jGCaMP7b-WPRE (1.02e13 vg/mL, Addgene, lot 18-429) into lateral SNc (AP -3.1, ML 1.3, DV -4.2) in one hemisphere and medial SNc (AP -3.1, ML 0.8, DV -4.7) in the other. Hemispheres were counterbalanced between mice. Fiber optic implants (Doric Lenses; 400 µm, 0.48 NA) were placed above DMS (AP 0.8, ML 1.5, DV -2.8) and DLS (AP -0.1, ML 2.8, DV -3.5). The DMS implant was placed in the hemisphere receiving a medial SNc viral injection, while the DLS implant was placed in the hemisphere receiving a lateral SNc viral injection. Calcium signals from dopamine terminals in DMS and DLS were recorded during RI30, on the first and last days of RI60/RR20 training as well as on both footshock probes for each mouse. All recordings were done using a fiber photometry rig with optical components from Doric lenses controlled by a real-time processor from Tucker Davis Technologies (TDT; RZ5P). TDT Synapse software was used for data acquisition. 465nm and 405nm LEDs were modulated at 211 Hz and 330 Hz, respectively, for DMS probes. 465nm and 405nm LEDs were modulated at 450 Hz

and 270 Hz, respectively for DLS probes. LED currents were adjusted in order to return a voltage between 150-200mV for each signal, were offset by 5 mA, were demodulated using a 4 Hz lowpass frequency filter. Behavioral timestamps, e.g. for nosepokes and port entries, were fed into the real-time processor as TTL signals from the operant chambers (MED Associates) for alignment with the neural data.

Transcardial Perfusions.

Mice received lethal i.p. injections of Euthasol (Virbac, 1mg/kg) a combination of sodium pentobarbital (390 mg/ml) and sodium phenytoin (50 mg/ml), to induce a smooth and rapid onset of unconsciousness and death. Once unresponsive to a firm toe pinch, an incision was made up the middle of the body cavity. An injection needle was inserted into the left ventricle of the heart, the right atrium was punctured and solution (PBS followed by 4% PFA) was infused as the mouse was exsanguinated. The mouse was then decapitated and its brain was removed and fixed overnight at 4°C in 4% PFA.

Histology

After perfusion and fixation, brains were transferred to a solution of 30% sucrose in PBS, where they were stored for at least two overnights at 4°C before sectioning. Tissue was sectioned on a freezing microtome (Leica) at 30 µm, stored in cryoprotectant (30% sucrose, 30% ethylene glycol, 1% polyvinyl pyrrolidone in PB) at 4°C until immunostaining. Tyrosine hydroxlase (TH) staining was performed on free floating sections, which were blocked with 3% normal goat serum in PBS-T for 1 hour at room temperature, then stained with 1:500 primary antibody (Aves Labs, Cat No. TYH) in blocking solution at 4°C overnight. Secondary staining was performed using 1:500

goat anti-chicken Alexa Fluor 647 secondary antibody (Life Technologies, Cat. No. A-21449). Anti-GFP staining was performed on free floating sections to amplify signals from GCaMP7b. This staining was performed by blocking in 3% normal goat serum in PBS-T for 1 hour at room temperature, then using 1:500 primary antibody conjugated directly to Alexa Fluor 488 (Life Technologies, Cat. No. A-21311) in blocking solution at 4°C overnight. Tissue was mounted on slides in PBS and coverslips were secured with Fluoromont-G (Southern Biotech). Slides were imaged using a fluorescent microscope (Keyence BZ-X800) with 5x and 40x air immersion objectives. Probe placements were determined by comparing to the Mouse Brain Atlas (Franklin & Paxinos, 2008). GCaMP neurons expressing YFP were counted and colocalized with TH+ neurons using ImageJ software.

3. Quantification and Statistical Analysis

Behavioral Analysis

Cue-evoked freezing during fear conditioning was scored manually by two blind observers from a recording of the fear conditioning test session using EthoVision software (Noldus). Scores from the two observers were averaged. Freezing was measured throughout the session as a mouse remaining still for more than two seconds. For all other studies, behavioral data was collected automatically by MED-PC software (Med Associates). To sort mice into PR, DPR, and PS groups, we calculated the percent change in shocks received from the early to late shock probe for each mouse. Mice in the top quartile of changers (who increased the number of shocks received by greater than 85%) were classified as delayed punishment resistant (DPR;

n=9). The remaining mice were sorted by a median split, with mice receiving more than 13 shocks on the first probe classified as punishment resistant (PR; n=9) and those earning fewer as punishment sensitive (PS; n=18, total n=36, Fig. 4A). The subset of mice that received the omission probe were also sorted by a median split of omission completion time (time to 50 rewards), with mice taking more than 29 minutes classified as long omission (n=10) and those taking less time as short omission (n=10, Fig. 5H). Plots in Fig. 4O were generated by plotting a segmental linear regression with lines for the average slope of nosepokes/minute across FR1, RI30, and RI60 training to reveal escalation of nosepoke behavior. The shaded area shows the 95% confidence band surrounding each slope. This analysis was done using GraphPad (Prism) software. Inter-reward intervals were calculated as the time from a rewarded nosepoke to the subsequent rewarded nosepoke (Fig. 4L) on each day of RI60 training. A frequency distribution was created and plotted using GraphPad (Prism) software. Plots in Fig 3B-C were generated by binning the number of nosepokes per five minutes during probe sessions and dividing by nosepokes in the same five-minute bin on the most recent day of RI60/RR20 training.

Fiber Photometry Analysis

All analysis was done using custom MATLAB (Mathworks) and Python code. Raw data from 465nm and 405nm channels were passed through a zero-phase digital filter (filtfilt function in Matlab) and a least-squares linear fit (parameters derived with polyfit function) was applied to the 405nm control signal to align it to the 465nm signal. $\Delta F/F$ was calculated with the following formula: (465nm signal - fitted 405nm signal) / (fitted

405nm signal). To facilitate comparisons across animals, z-scores were calculated by subtracting the mean ΔF/F calculated across the entire session and dividing by the standard deviation. Peri-stimulus time histograms (PSTHs) were created using the TTL timestamps corresponding to behavioral events. Maximum and minimum peak values and locations from PSTHs in main figures were generated using max and min functions in MATLAB for the 1.5 seconds following behavioral event (ie nosepoke, port entry). AUC was calculated using trap function in MATLAB. We used a customized logic for peak detection in Figures 7 and 9 adapted from Holly et al (2019) and Muir et al (2018) (Holly et al., 2019; Muir et al., 2018). Events having amplitudes greater than the summation of a median of 30 seconds moving window and two times median absolute deviation (MADs), were filtered out and the median of the resultant trace was calculated. Peaks having local maxima greater than three times MADs of the resultant trace above the median were considered as events.

Statistical Methods

Statistical analysis was done using Prism 9 software (GraphPad). One and two-way ANOVAs, or mixed effects analyses were performed with Tukey's multiple comparisons and Bonferroni post-hoc analyses when statistically significant main effects or interactions were found. A Kolmogorov-Smirnov test was used to compare the distributions of inter-reward intervals. One RI60 mouse was excluded from the fiber photometry study due to improper fiber placement.

C. Results

1. A random interval, but not random ratio, schedule of reinforcement promotes punishment-resistant reward-seeking

We first determined whether training paradigms used to elicit habitual responding also elicit punishment-resistant reward-seeking. Previous literature has demonstrated that a random interval (RI60) schedule of reinforcement, but not a random ratio (RR20) schedule, promotes habit (Derusso et al., 2010; Gremel & Costa, 2013; Wiltgen et al., 2012; Yin, Ostlund, et al., 2005). We therefore compared training on RI60 and RR20 reinforcement schedules to assess whether these schedules have differential effects on punishment-resistant reward-seeking. After initial magazine training and pre-training on a fixed ratio (FR1) schedule, mice were transitioned to either RI30 or RR10 schedules, and then finally to RI60 or RR20 schedules (Fig. 2A). We performed probes for punishment-resistant reward-seeking at an early and late time point: the first after 1-2 days of RI60/RR20 training and the second after 13-14 days of RI60/RR20 training. During the shock probe sessions, nosepokes were accompanied by a ⅓ risk of mild shock (0.2mA, 1s; Fig. 2B). The shock intensity was chosen based on previous studies of punishment-resistant reward-seeking (Harada et al., 2019). We verified that this shock intensity is aversive to the mice in a fear conditioning paradigm in which 12 tone-shock pairings were delivered. The next day, mice that had received tone-shock pairings showed increased freezing to the tone compared with mice that had only been exposed to the tone on the previous day (Fig. 3A; unpaired t test, $p<0.01$). To test whether punishment-resistant reward-seeking developed in tandem with another test of habit-like behavioral inflexibility, a subset of mice were also tested at the end of training on an omission probe (Fig. 2A; Derusso et al., 2010; Rossi & Yin, 2012; Yu et al., 2009).

In the omission probe, mice were required to withhold nosepokes to receive rewards, reversing the previously learned contingency (Fig. 2C).

We observed a significant main effect of schedule (RI60 vs RR20) on the number of shocks mice were willing to receive on the shock probes (two-way ANOVA, $F_{1,45}=6.31$, $p<0.05$) as well as an interaction of schedule and training time ($F_{1,45}=4.54$, $p<0.05$). Furthermore, after extended training on the RI60, but not the RR20, schedule, mice increased the number of shocks they were willing to receive (Fig. 2D; Bonferroni, $p<0.001$). During the second shock probe, both RI60- and RR20-trained mice initially continued to nosepoke at the same rates relative to their training baseline, but the RR20-trained mice reduced their nosepoke rates to near zero by the end of the session (Fig. 3B; two-way ANOVA, main effect of training time, $F_{4.426,181.5}=4.99$, $p<0.001$, Bonferroni, $p<0.05$ for bins after 50 minutes). The RR20-trained mice were also more willing than RI60-trained mice to explore alternative actions during the second shock probe session, as indicated by a higher fraction of nosepokes made at the inactive port (Fig. 3D-E; unpaired t-test, $p<0.05$).

On the omission probe, RI60-trained mice took longer to complete the session than mice trained on RR20 (Fig. 2E; unpaired t-test, $p<0.05$). RR20-trained mice almost immediately stopped nosepoking during the omission probe session, whereas RI60-trained mice continued to nosepoke for much longer periods of time, albeit with large variability between individuals (Fig. 3C; two-way ANOVA, interaction of schedule and time, $F_{11,275}=2.22$, $p<0.05$). RR20-trained mice therefore maintain a higher level of flexibility in their behavior when presented with positive punishments and when presented with a reversed contingency as compared to RI60-trained mice.

It is not clear what differences between RI60 and RR20 training might elicit differences in the development of punishment-resistant or omission-resistant reward-seeking. As previously reported (Gremel & Costa, 2013; Wiltgen et al., 2012), RI60 and RR20 training schedules provoked approximately equivalent rates of nosepoking (Fig. 3G). However, we also observed that RI60-trained mice made fewer nosepokes per reward (Fig. 3H; mixed-effects analysis, $F_{1,42}$=21.70, p<0.0001) and earned significantly more rewards per training session, on average, than RR20-trained mice (Fig. 3I; unpaired t-test, p<0.0001). Therefore, these different task structures have different effort demands and incur different reward histories, which could influence learning trajectories.

In both the shock and omission probes, RI60-trained mice showed significant individual variability (omission F-test to compare variances, $F_{19,5}$=305.90, p<0.0001; shock 1 F test to compare variances $F_{37,8}$=5.88, p<0.01; shock 2 F test to compare variances $F_{37,7}$= 12.08, p<0.001). The variability in punishment-resistance was not due to variation in body weight (Fig. 3J). We wondered whether the same individuals who withstood a high number of shocks also took longer to learn the omission contingency. Looking at data from individual animals tested on both probes, we found that these two measures were not significantly correlated (Fig. 2F; r=0.42, ns). Thus, although the RI60 schedule promotes inflexible behavior in the form of both punishment-resistant and omission-resistant reward-seeking, these two phenomena do not necessarily occur in the same individuals and their development may rely on different brain circuits.

2. Three behavioral phenotypes emerge with extended RI60 training

The large variation in behavior induced by RI60 training caused us to wonder whether individual mice were taking different strategies to “solve” the RI60 task, which could lead to differences in the development of punishment-resistant reward-seeking. To analyze whether differences in performance in the shock probes were associated with different behavior in the RI60 task, we divided the RI60-trained mice into three analysis groups based on shock probe performance: “punishment resistant” (PR) mice that tolerated a high level of shocks in both the first and second probes (25% of mice), “delayed punishment resistant” (DPR) mice that increased the number of shocks they would tolerate from the first to the second probe (25% of mice), and “punishment sensitive” (PS) mice that would not tolerate many shocks on either the first or second probes (50% of mice; Fig. 4A; see methods for sorting criteria). As expected based on this sorting, a mixed effects analysis showed a significant interaction of phenotype and training time ($F_{3, 40}$ = 24.18, $p<0.0001$), with a Bonferroni test revealing that only DPR mice showed a significant difference in shocks received between the early and late shock probes (Fig. 4B; $p<0.0001$).

When mice were divided into these groups, an analysis of behavior in RI60 sessions showed interesting differences (Figs. 4, 5). Both PR and DPR mice had higher rates of nosepoking than PS mice (Fig. 4C; mixed effects analysis, main effect of training time $F_{5.236,156.3}=9.79$, $p<0.0001$, main effect of phenotype $F_{2,33}=18.59$, $p<0.0001$, interaction $F_{26,388}=3.28$, $p<0.0001$). PR mice also took significantly longer to complete the omission probe in comparison to PS mice (Fig. 4D; one-way ANOVA, $F_{2,18}=4.36$, $p<0.05$, Tukey’s multiple comparison, PR=51.50 ± 17s vs PS=27.92 ± 13.65s, $p<0.05$). Meanwhile, PS mice were more likely to explore the inactive nosepoke

at the end of training and during the second shock probe session (Fig. 4E; mixed-effects analysis, main effect of phenotype, $F_{2,33}$=3.79, Tukey's multiple comparisons test, RI60 days 13 and 14 for PS vs DPR, $p<0.05$; Fig. 4F; one-way ANOVA for shock 2, $F_{2,18}$=5.36, Tukey's multiple comparison, for PS vs DPR, $p<0.05$, second shock for PS vs PR, $p<0.05$; Fig. 4G; no significant effect during omission).

At least some of the variation between the PR, DPR and PS mice was attributable to sex differences. PR mice were more likely to be male, whereas PS mice were more likely to be female. The DPR group was evenly split between the sexes (Fig. 5A). In general, as a group, male mice tolerated more shocks on both shock probes than female mice (Fig. 5B) and had higher nosepoke rates during RI60 (Fig. 5C). However, variance in punishment-resistant reward-seeking was not fully explained by sex. A three-way ANOVA of phenotype x sex x time revealed that more of the variance in shocks received was accounted for by phenotype than by sex (15.09 vs 0.03%, respectively). Nevertheless, given these sex differences, we were careful to include a balance of both male and female mice going forward in all our experiments.

Port entry rates and rates of rewards earned per minute did not differ among the PR, DPR, and PS groups (Fig. 4H-I). As a result, PS mice are more "efficient," making fewer nosepokes per reward than PR and DPR mice (Fig. 4J; mixed-effects analysis, main effect of training time $F_{4.65,87.86}$=2.42, $p<0.05$, main effect of phenotype $F_{2,21}$=22.77, $p<0.0001$, interaction $F_{24,227}$=1.76, $p<0.05$). To achieve such efficiency, a mouse needs to understand when a nosepoke is likely to lead to reward. In the RI60 task, a mouse is unlikely to receive a reward if it has recently received one (on average, it will have to wait 60 seconds before its efforts become fruitful again). PS mice waited an average of

38 ± 15 seconds to resume nosepoking after retrieving a reward, significantly longer than the other groups (Fig. 4K; Tukey's multiple comparison, PR=21 ± 5, $p<0.01$; DPR=22 ± 4, $p<0.01$). We also looked at whether mice made "extra" nosepokes after a rewarded nosepoke, before going to the port to collect their reward. PS mice made significantly fewer extra nosepokes before reward retrieval than PR mice (Fig. 4L; one-way ANOVA, $F_{2,33}=4.24$, $p<0.05$, Tukey's multiple comparison, PS=1.28 ± 0.85, PR= 3.23 ± 1.65, $p<0.05$, DPR= 2.1 ± 2.63), which helped maximize the efficiency of their behavior.

Why would PR and DPR mice expend more effort than necessary to earn rewards? On an RI60 reinforcement schedule, the average interval between available rewards is 60 seconds, but there is a normal distribution around this average. To ensure a nosepoke is likely to be rewarded, a mouse should wait to nosepoke, but to "catch" all instances when reward becomes available in the shortest possible amount of time, mice must nosepoke constantly. In other words, if mice are willing to expend the effort, high rates of nosepoking yield rewards at shorter intervals. Indeed, when we looked at the distribution of inter-reward intervals for PR, DPR and PS mice, we found that the inter-reward intervals for PS mice skew longer (Fig. 4M). The average inter-reward interval for PR is 71.2 ± 32.3 seconds and for DPR mice is 69.58 ± 35.18 seconds, while the average for PS mice is 79.77 ± 26.54 seconds (K-S Test, PR vs PS, $p<0.0001$; DPR vs PS, $p<0.0001$). PR and DPR mice thus are better able to track the probability of reward availability imposed in the task, whereas PS mice have longer lags before they earn available rewards. Although we did not observe differences in the total number of rewards earned per minute on any particular day of RI60 training (Fig. 4I), when the

number of rewards earned per session was averaged over all days, we found that PS received slightly fewer rewards (Fig. 4N; one-way ANOVA, $F_{2,33}$=4.45, $p < 0.05$, Tukey's multiple comparison, PS= 39.91 ± 7.39, PR= 45.33 ± 4.33, $p<0.05$). This difference illustrates the advantage of the high-nosepoking approach: maximize rewards at any cost.

We observed that PR and DPR mice take similar reward-seeking strategies in the RI60 task, so to examine differences between them, we looked at earlier training data from FR1 and RI30 sessions. We observed that PR mice escalate their nosepoking more rapidly than DPR mice, as soon as they enter RI30 training (the first RI schedule they encounter; Fig. 4O). DPR mice then escalate their nosepoking throughout RI60 training, concurrent with the development of punishment-resistance. Thus, the timing of nosepoke escalation is a key behavioral predictor of mice who will show punishment-resistant reward-seeking when probed. Importantly, this escalation emerges in PR mice prior to any experience of punishment, suggesting that some mice may have a predisposition towards developing punishment-resistance that is related to their initial reward-seeking strategy.

3. Dopamine axon signals in the DMS predict punishment-resistant reward-seeking

To better understand the neural circuits mediating differences in the development of punishment-resistant reward-seeking, we recorded the activity of dopamine axons in the dorsal striatum during RI/RR training. Dopaminergic projections to the DMS and

DLS are distinct, meaning that dopamine-mediated reinforcement learning can be separately effectuated in these two areas (Ikemoto, 2007; Lerner et al., 2015). We reasoned that by examining the activity of dopaminergic projections to the DMS and DLS during RI/RR training, we could assess whether different dopamine signals were contributing to different aspects of behavior. To record the activity of dopamine axons in the DMS and DLS in freely moving mice, we injected an adeno-associated virus (AAV) expressing cre-dependent GCaMP7b (AAV5-CAG-FLEX-jGCaMP7b-WPRE) into the substantia nigra pars compacta (SNc) of DAT-IRES-cre mice. We then implanted fiber optic probes above the DMS and DLS to record the activity of dopaminergic axons (Fig. 6A, 8A). Dopaminergic axon activity was recorded simultaneously in the DMS and DLS in all mice. Thus, we could be certain that any differences we observed in activity in the two regions were not due to differences in behavior between groups of mice.

To analyze these data, we first examined DMS dopamine axon activity occurring during rewarded and unrewarded nosepokes. We histologically verified that GCaMP was expressed in dopaminergic (TH+) neurons in the medial SNc and found that on average 82.86% of neurons expressing GCaMP also expressed TH (Fig. 6B). We confirmed that probe locations in the DMS were correct (Fig. 6C, 7A). Histology also verified robust GCaMP expression in dopaminergic axons in the DMS (Fig. 6C, inset). We compared DMS dopamine axon activity in RI60-trained mice (PR, DPR, PS; Fig. 6D) and in RR20-trained mice (RR20; Fig 7D) across training. Peaks in DMS dopamine axon activity at the time of a rewarded nosepoke were much clearer in PR and DPR mice than in PS or RR mice, an observation that was not true simply due to poor signal in PS or RR20 mice, as all mice included in the analysis had similar frequencies and

amplitudes of GCaMP events across the entire training session (Fig. 7B-C). A main effect of training time on the frequency of all GCaMP events was observed, but was the same across all groups (Fig. 7B). Clear peaks in response to rewarded nosepokes were observed in PR mice even in the RI30 stage of training, whereas peaks in response to rewarded nosepokes emerged more slowly over the course of extended RI60 training in DPR mice (Fig. 6D). We also noted that mice varied in whether their DMS dopamine axons showed positive or negative responses to unrewarded nosepokes. In RI30 recordings, unrewarded nosepokes tended to give upward deflections in all groups, however, negative deflections appeared over the course of RI60 training, particularly in the PR and DPR groups. The combination of a positive peak in DMS dopamine axon activity for rewarded nosepokes and a negative peak for unrewarded nosepokes creates a notable difference in dopamine axon activity in response to the same motor action depending on the outcome, which is information that could be used to influence future behavior. We calculated a rewarded-unrewarded peak score for each mouse and compared whether the scores were altered by training stage or by behavioral phenotype (PR, DPR, or PS). There was a significant effect of training stage (Fig. 6E; mixed-effects analysis, $F_{1.61,29.8}$=13.83, $p<0.001$), a significant effect of phenotype ($F_{2, 33}$= 8.16, $p<0.01$) and a significant interaction between the two ($F_{4,37}$= 3.29, $p<0.05$).

From these data, it appeared that the peaks in DMS dopamine axon activity tracked with the development of punishment-resistant behavior. We therefore wanted to assess whether an individual's rewarded-unrewarded peak score (regardless of its classification as PR, DPR, or PS) could predict whether it would tolerate shocks in the shock probe sessions. Indeed, the shocks received were significantly correlated with the

DMS rewarded-unrewarded peak score on a mouse-by-mouse basis (Fig. 6F; irrespective of our group classifications; $r=0.53$, $p<0.0001$). In contrast, performance on the omission probe was not correlated with this score (Fig. 6G; $r=0.1$, ns).

We also examined DMS dopamine axon signals surrounding the time of rewarded and unrewarded port entries and noticed ramping activity preceding the rewarded port entries (Fig. 7E-F). Such ramping towards reward has been described primarily in VTA-NAc dopamine circuits (Guru et al., 2020; Hamid et al., 2015; Howe et al., 2013; Kim et al., 2020; Mohebi et al., 2019), but has also been described in DMS dopamine axons in another recent pre-printed study (Hamid et al., 2019). Here, we additionally note that ramping in DMS dopamine axons is more prominent in PR and DPR mice than in PS or RR20 mice (Fig. 7E-F). To encapsulate ramping activity quantitatively, we measured the area under the curve (AUC) of our fiber photometry signal from -5 to 0s relative to the rewarded port entry for each mouse. The AUC was variable, but notably the DPR group of mice showed a significant increase from early RI60 to late RI60 training (Fig. 7F; mixed-effects analysis, interaction of time and phenotype, $F_{3,19}= 4.77$, $p<0.05$, Bonferroni, $p<0.05$).

4. Dopamine signals in the DLS do not predict punishment-resistant or omission-resistant reward-seeking

In recordings from the same mice, we next examined DLS dopamine axon activity occurring during nosepokes to see if it differed from activity in the DMS (Fig. 8A). We verified the expression of virus in SNc (Fig. 8B; 89.56% of GCaMP neurons also expressed TH) and the probe placements in DLS (Fig. 8C, 9A). We also noted that the DLS signals in all groups of mice had similar frequencies and amplitudes of GCaMP

events (Fig. 9B-C). DLS dopamine axon signals at the time of a rewarded nosepoke differed from the DMS dopamine axon signals in that they had both an immediate component and a prolonged, delayed component (Fig. 8D, 9D). DLS dopamine axon signals following an unrewarded nosepoke were variable. During RI30, there tended to be positive deflections in the signal for unrewarded nosepokes, however, later in training negative deflections emerged. To test whether these DLS dopamine axon signals in response to nosepokes bore any relationship to the punishment-resistant phenotype of the mice, we calculated a rewarded-unrewarded peak score as we had for the DMS dopamine axon signals. A mixed-effects analysis revealed a significant effect of training stage (mixed-effects analysis, $F_{2,35}$= 3.53, $p<0.05$), indicating that the signals do change over the course of training, but no significant effect of group. For completeness, we also looked at whether these signals correlated on an individual basis with shocks received in the shock probe or omission completion time and found no correlations (Fig. 8F-G; r=-0.01, ns and r=-0.21, ns, respectively).

We also assessed whether the prolonged elevation of DLS dopamine axon activity after a rewarded nosepoke, which was often sustained for 10 seconds or more, was related to the development of punishment- or omission-resistance. We quantified the prolonged activity as the area under the curve (AUC) from 2-10 seconds after the nosepoke. AUC did not correlate with the performance of individuals on the shock or omission probes (Fig. 9E-F) nor did it differ statistically across groups or training time (Fig. 9G).

We were surprised that we could not observe a correlation of DLS dopamine axon signals with behavior, despite the fact that they changed over time. To get another

view of these signals, we tried grouping mice by their performance on the omission probe rather than the shock probe. However, we saw no obvious differences in the development of dopamine axon signals in either the DMS or the DLS depending on omission probe performance (Fig. 9H).

Finally, we also looked at DLS dopamine axon signals aligned to the time of port entries (Fig. 9I). Here, we observed very weak ramping prior to a rewarded port entry, with the peak of the signal occurring after port entry. Quantification of the AUC of our fiber photometry signal from -5s to 0s relative to the rewarded port entry did not show evidence that pre-port entry activity tracked with the behavioral phenotype (Fig. 9J). The lack of ramping in DLS dopamine axons in comparison to DMS dopamine axons we observed is notably similar to what was observed in another recent study, even though the imaging methods and behaviors used in that study are distinct from ours (Hamid et al., 2019). The functional significance of this difference in ramping activity between the DMS and DLS dopamine axons will be an important area for future investigation.

D. Discussion

Compulsive behavior is a defining feature of disorders such as substance use disorder (SUD) and obsessive-compulsive disorder (OCD). There is some evidence suggesting that corticostriatal circuits control the expression of established compulsions, but little is known about the mechanisms regulating the development of compulsions (Lüscher & Janak, 2021). We hypothesized that dopamine – a key neuromodulator regulating corticostriatal synaptic plasticity – could play a role in sculpting the emergence of compulsive behavior, defined as punishment-resistant reward-seeking. However, it was unclear whether the dopamine signals most relevant punishment-

resistant reward-seeking would lie in the DMS, which is linked to action-outcome learning, or the DLS, which is linked to habit formation. We addressed this question by using dual-site fiber photometry to record the activity of dopamine axons in the dorsomedial striatum (DMS) and dorsolateral striatum (DLS) during a task (RI60) that promotes punishment-resistant reward-seeking. We identified task-relevant increases in DMS dopamine axon activity as a key feature of neural circuit activity predicting which individual animals will develop punishment-resistance. Specifically, DMS dopamine axon activity that effectively discriminated between rewarded and unrewarded nosepokes was correlated with the development of punishment-resistant reward-seeking.

We were surprised to see that DMS dopamine signaling promotes punishment-resistant reward-seeking, since typically DMS activity is thought to be linked to goal-directed behavior and in some cases has even been shown to decrease with extended habit training (J. E. Murray, Everitt, et al., 2012; Vandaele Y & Janak PH, 2021; Willuhn et al., 2012; Yin, Ostlund, et al., 2005). Nevertheless, this observation fits with certain aspects of the previous literature. First, optogenetic stimulation of direct-pathway DMS neurons is reinforcing (Kravitz et al., 2012). The reinforcing effect persists even on a day of extinction training, indicating that the activation of DMS circuits is linked to inflexible behavior. Second, stimulation of SNc dopamine neurons, which project to DMS and DLS, supports reinforcement (Corbett & Wise, 1980; Ilango et al., 2014; Keiflin et al., 2017; Reynolds et al., 2001; Rossi et al., 2013; Saunders et al., 2018; Roy A. Wise, 1981). However, unlike ventral tegmental area (VTA) dopamine neurons, which support reinforcement by adding incentive salience to a reward-predicting cue

(Robinson & Berridge, 1993), SNc dopamine neurons support reinforcement via the invigoration of movement (Crow, 1972; Saunders et al., 2018). Thus, DMS dopamine signaling is perfectly situated to translate histories of rewarded actions into the compulsive performance of those actions. Third, changes in DMS circuit function have been implicated in stereotyped, perhaps compulsive, grooming behaviors (Cromwell & Berridge, 1996; Kalueff et al., 2016). Sapap3-/- mice, a model of OCD that displays overgrooming, are more likely to form habits after training on RI60 (Burguiere et al., 2013; Hadjas et al., 2019), although it is unknown how Sapap3-/- mice would perform on a test for compulsion as in our study design. The answer is one worthy of future investigation as it could shed light on the neural circuit mechanisms of OCD-like compulsions.

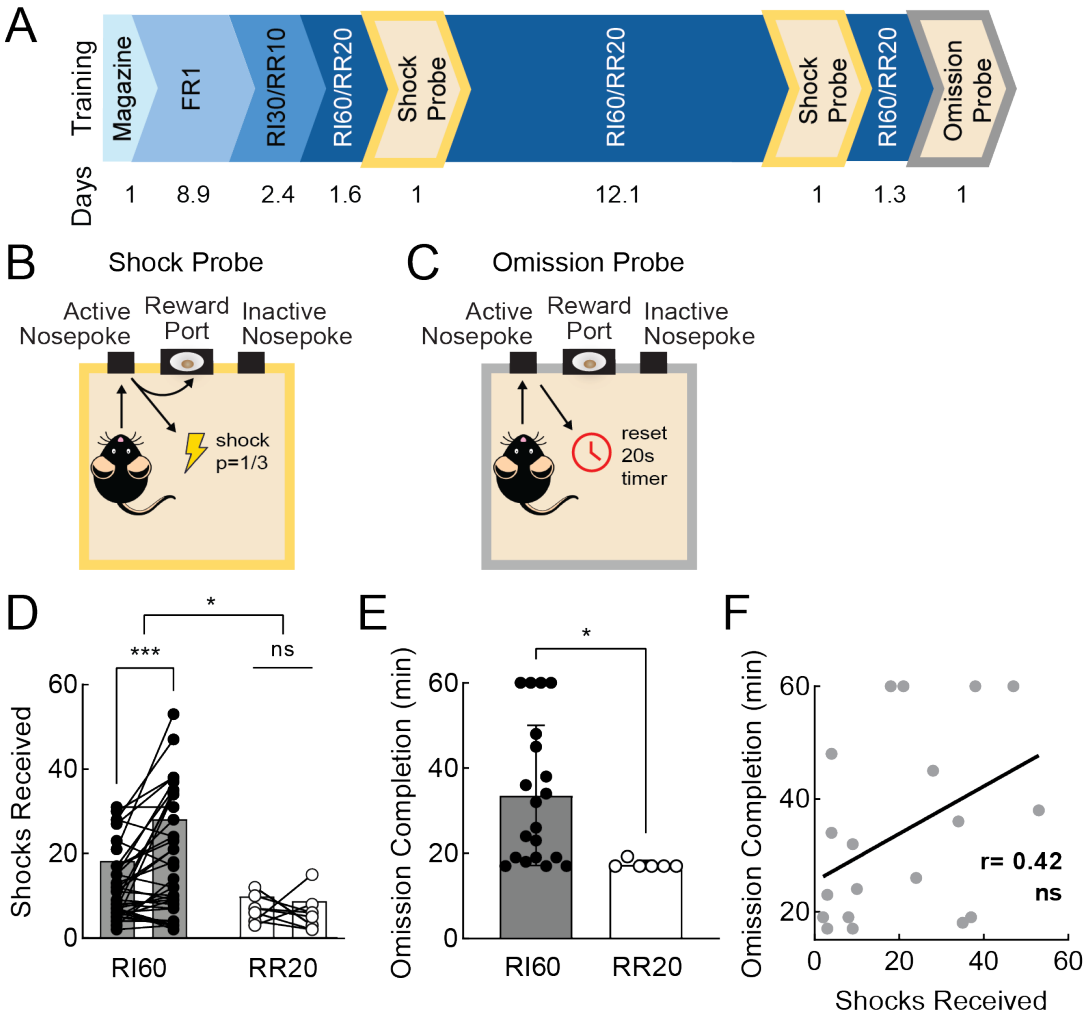

Figure 2. A random interval, but not random ratio, schedule of reinforcement promotes compulsive reward-seeking.

A. Timeline of operant training and probes. Animals progressed from a fixed ratio to a random ratio or random interval schedule of reinforcement. A shock probe occurred both early and late in RI60/RR20 training and an omission probe occurred at the end of training. Average number of days in each stage of training is given below. **B.** Schematic of shock probe: active nosepokes had a ⅓ probability of incurring a shock consequence. **C.** Schematic of omission probe: rewards were delivered when mice refrained from nosepoking at the previously active port for 20s. **D.** Shocks received on early and late shock probes for RI60-trained (black; n=36) and RR20-trained (white;

n=8) mice. Bars represent mean; points represent individuals. *p<0.05, ***p<0.001. **E.** Average time to complete the omission probe (earn 50 rewards; max 60 minutes) for RI60-trained (black; n=20) and RR20-trained (white; n=7) mice. Error bars represent SD. *p<0.05 **F.** Correlation between shocks received on the late shock probe and omission completion time for all RI60-trained mice tested in both probes (r=0.42, ns).

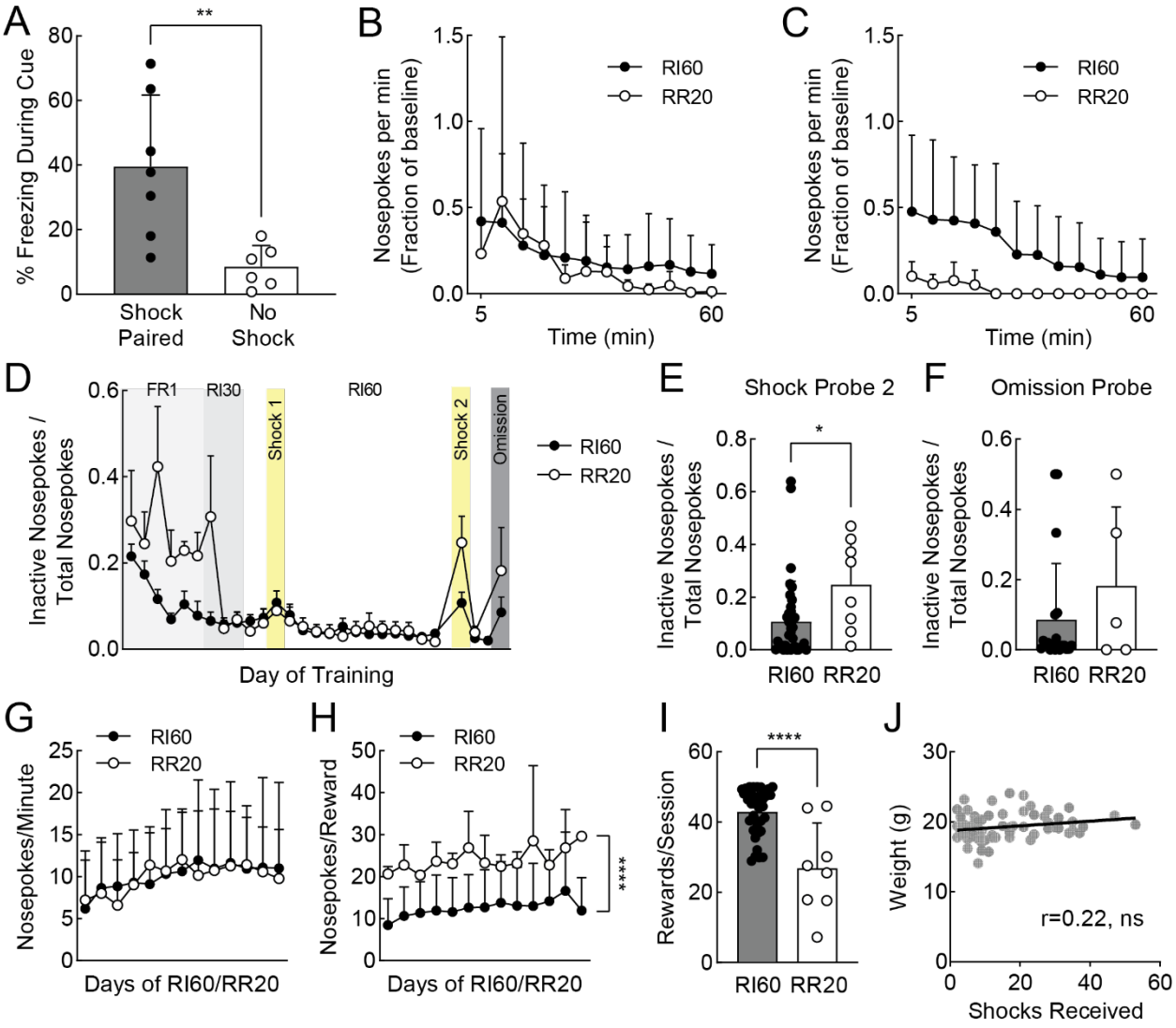

Figure 3.

A. Percent time spent freezing during a tone cue that had been paired with a 1s 0.2mA shock (black) or no shock (white). Error bars represent SD **p<0.01 **B.** Average nosepokes per minute on the second shock probe represented as a fraction of nosepokes made per minute during the most recent RI60 (black) or RR20 (white) session. Points represent 5 minute bins. Error bars represent SD. **C.** Average nosepokes per min on the omission probe represented as a fraction of nosepokes made per minute during the most recent RI60/RR20 session. Points represent 5 minute bins.

Error bars represent SD. **D.** Average nosepokes on the inactive port as a fraction of the total nosepokes across days of training and probe sessions. Error bars represent SD. **E.** Average nosepokes on the inactive port as a fraction of the total nosepokes during the second shock probe session. Error bars represent SD *p<0.05. **F.** Average nosepokes on the inactive port as a fraction of the total nosepokes during the omission probe session. Error bars represent SD. **G.** Average nosepokes per minute across days of RI60/RR20 training. Error bars represent SD. **H.** Average nosepokes per reward earned across days of RI60/RR20 training. Error bars represent SD. **I.** Average rewards earned per session of RI60/RR20 training. Error bars represent SD ****p<0.0001. **J.** Correlation between weight of animals (grams) and number of shocks received during shock sessions for RI60-trained animals (r=0.22, ns).

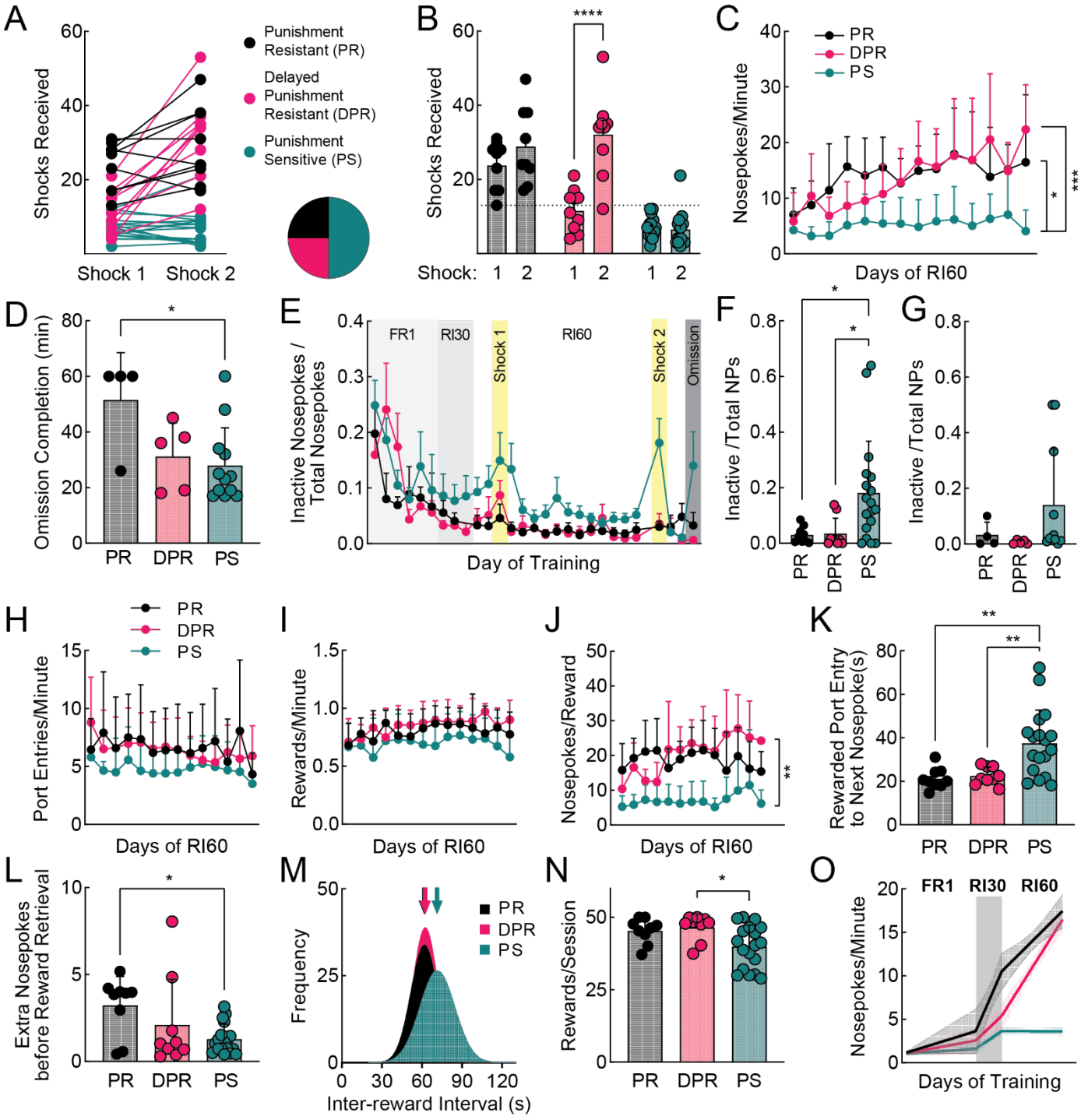

Figure 4. Three punishment-related phenotypes emerge with extended RI60 training.

A. Shocks received by mice trained under RI60 only (same data as shown in Fig. 2D).

Mice were classified as punishment resistant (PR; black), delayed punishment resistant

(DPR; pink), or punishment sensitive (PS; teal) based on the number of shocks they received during the early and late shock probes (see Methods). **B.** Average shocks received on early and late shock probes for each phenotype. Error bars represent SD ****p<0.0001. **C.** Average nosepokes per minute across days of RI60 training. Error bars represent SD, PR vs PS: *p<0.01, DPR vs PS: ***p<0.001. **D.** Average time to complete the omission probe (earn 50 rewards; max 60 minutes). Error bars represent SD *p<0.05. **E.** Average nosepokes on the inactive port as a fraction of the total nosepokes across days of training and probe sessions. Error bars represent SD. **F.** Average nosepokes on the inactive port as a fraction of the total nosepokes during the second shock probe session. Error bars represent SD *p<0.05. **G.** Average nosepokes on the inactive port as a fraction of the total nosepokes during the omission probe session. Error bars represent SD. **H.** Average number of port entries per minute across days of RI60 training. Error bars represent SD. **I.** Average number of rewards earned per minute across days of RI60 training. Error bars represent SD. **J.** Average nosepokes made per reward across days of RI60 training. Error bars represent SD **p<0.01. **K.** Average time from a rewarded port entry to the next nosepoke. Error bars represent SD **p<0.01. **L.** Average unrewarded nosepokes made following a rewarded nosepoke, prior to a rewarded port entry. Error bars represent SD *p<0.05. **M.** Distribution of inter-reward interval times in seconds for each group. Arrows represent mean. **N.** Average rewards earned per RI60 session. Error bars represent SD *p<0.05. **O.** Segmental linear regression showing the slope of nosepokes made per minute in FR1, RI30, and RI60 schedules. Shaded region represents 95% confidence bands.

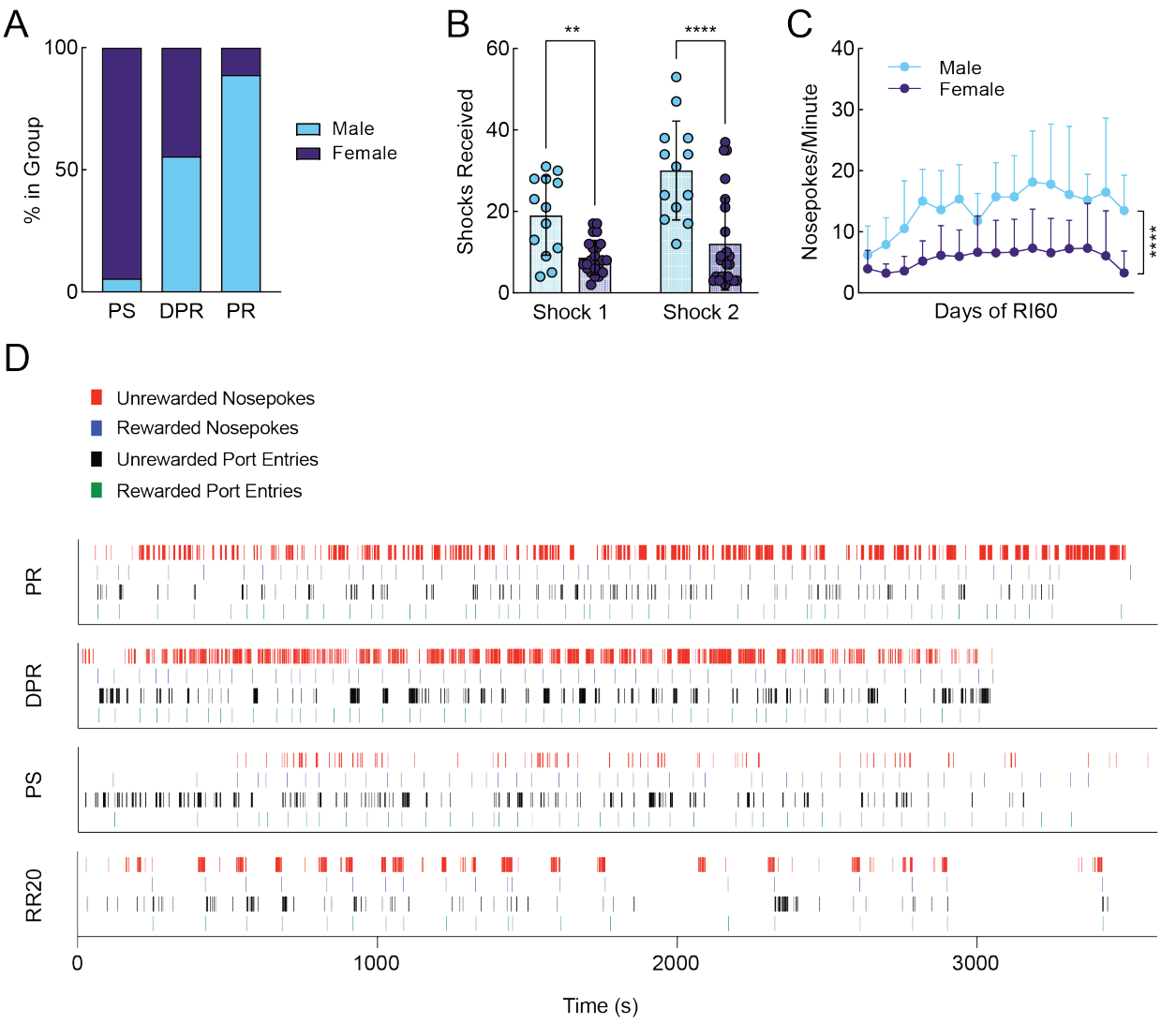

Figure 5.

A. Percentage of male (blue) and female (purple) mice in each behavioral phenotype. **B.** Average shocks received on early and late shock probes for each sex. Error bars represent SD **C.** Average nosepokes per minute across days of RI60 training. Error bars represent SD ****p<0.0001. **D.** Examples of behavioral timestamps recorded over the course of a full RI60 or RR20 session for a representative mouse from each group.

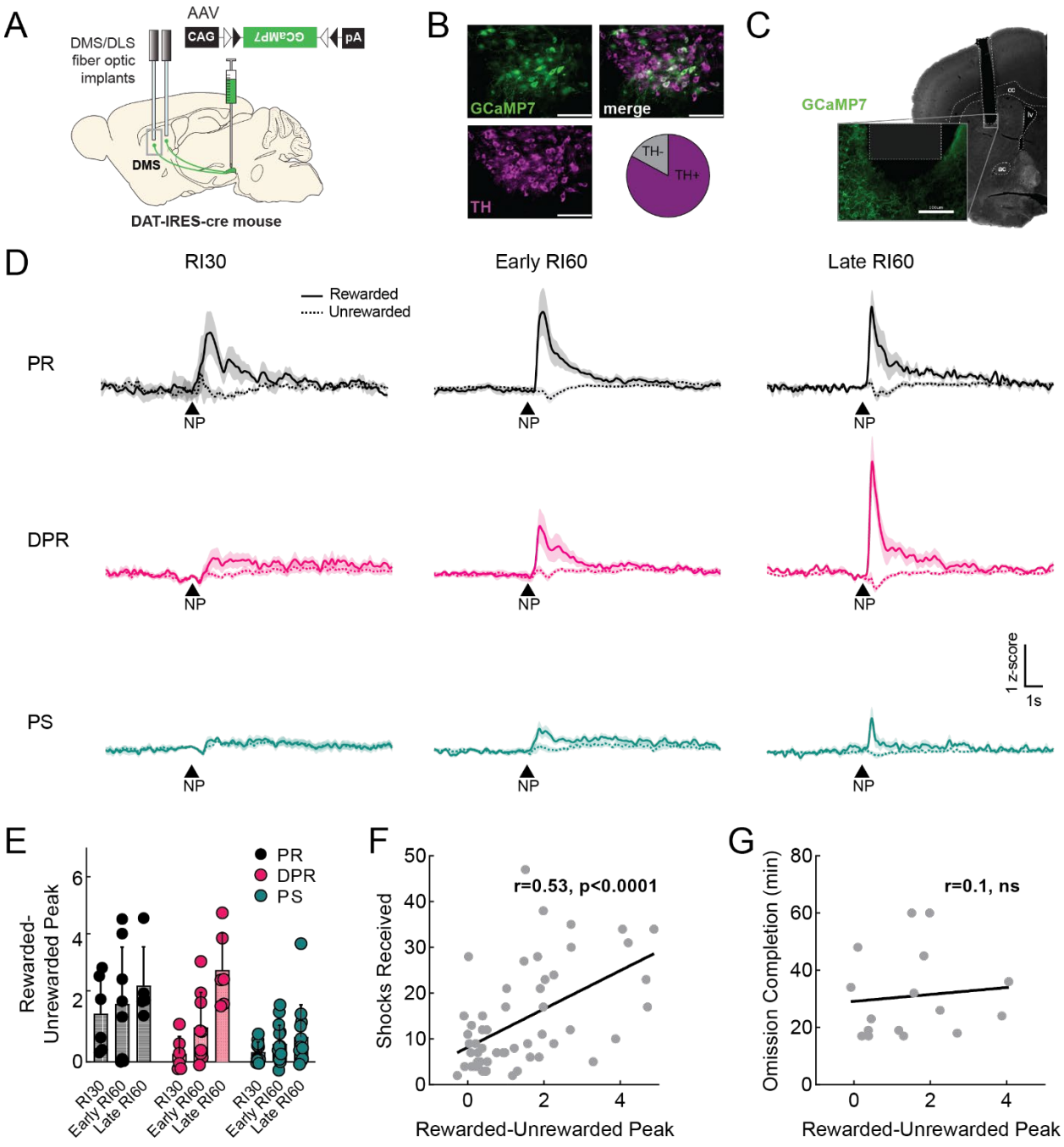

Figure 6. Dopamine axon signals in the DMS predict punishment-resistant reward-seeking.

A. Viral injection and probe placement strategy for recording from dopamine terminals in the DMS and DLS. DMS is highlighted here. **B.** High magnification (40x) images of SNc showing GCaMP7b expression in green, TH positive cells in magenta, and the merged image. Scale bars are 100μm. Quantification of GCaMP7b-expressing cells that are

TH+ is shown; n=572 cells. **C.** Representative image showing probe placement in DMS. Area of magnification shows GCaMP7b expression in dopaminergic axons near the probe site in green. Scale bar is 100μm. cc=corpus callosum, lv=lateral ventricle, ac=anterior commissure. **D.** Peri-stimulus time histograms (PSTHs) showing the average signal from DMS dopamine terminals at the time of rewarded (solid) and unrewarded (dashed) nosepokes (NP) for each phenotype during RI30 training, early in training (RI60 day one or two), and late in training (RI60 day eleven or twelve). Shaded region represents SEM. Punishment resistant (PR; black), delayed punishment resistant (DPR; pink), or punishment sensitive (PS; teal). **E.** Quantification of the average "difference score" for DMS dopamine terminal signals observed in response to rewarded and unrewarded nosepokes (calculated as the peak at the time of rewarded nosepokes minus the peak at the time of unrewarded nosepokes). Error bars represent SD. **F.** Correlation of shocks received in shock probe sessions and "difference score" in DMS dopamine terminals ($r=0.53$, $p<0.0001$). **G.** Correlation of time to complete omission and "difference score" in DMS dopamine terminals ($r=0.1$, ns).

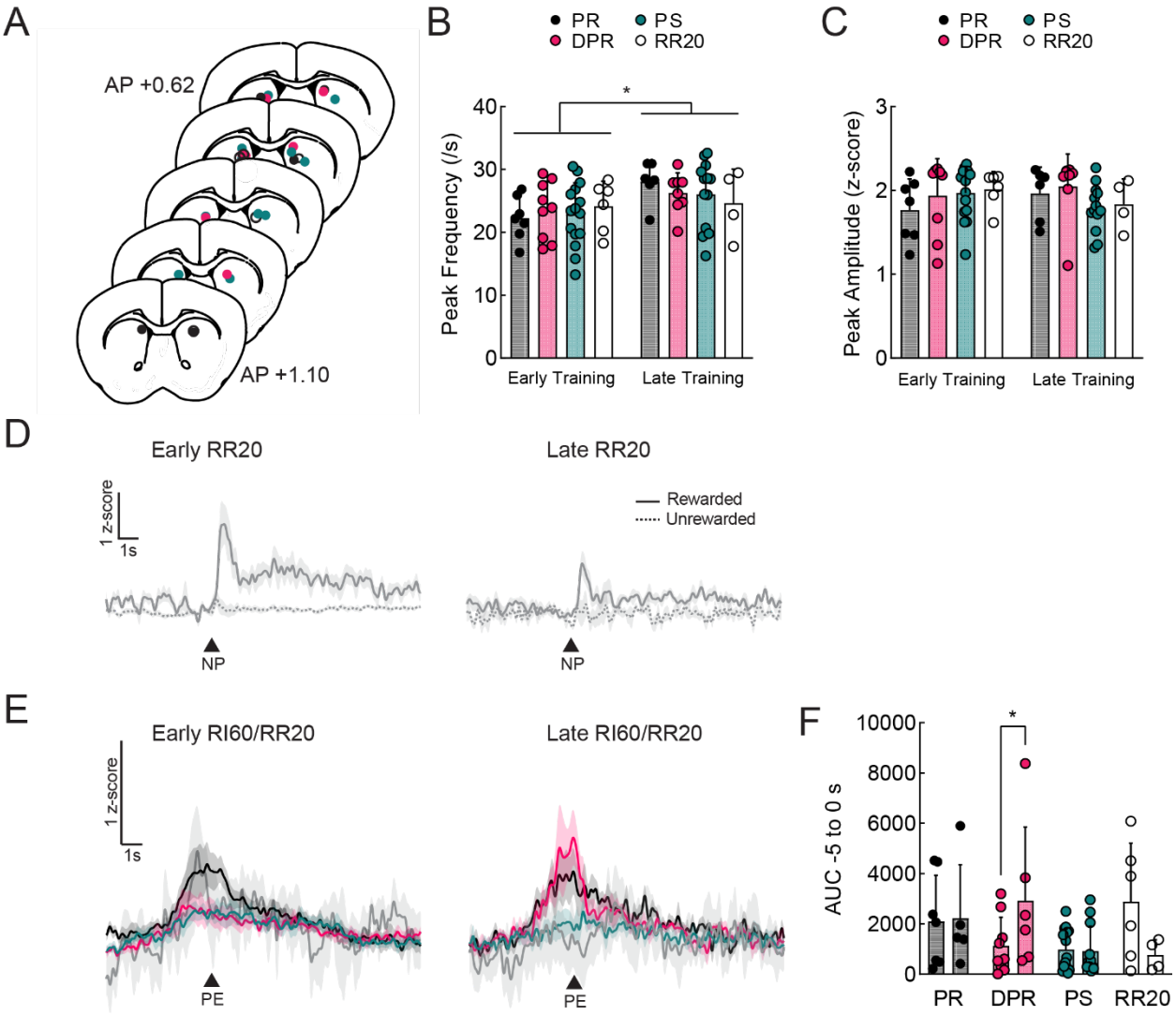

Figure 7.

A. Probe placements in DMS for all mice included in Figure 3. **B.** Average frequency of peaks detected in the DMS dopamine terminal signal for all groups. Error bars represent SD *p<0.05. Punishment resistant (PR; black), delayed punishment resistant (DPR; pink), or punishment sensitive (PS; teal), RR20-trained (RR20; white). **C.** Average amplitude of detected peaks in the DMS dopamine terminal signal for all groups. Error bars represent SD. **D.** Peri-stimulus time histograms (PSTHs) showing the average signal from DMS dopamine terminals at the time of rewarded (solid) and unrewarded (dashed) nosepokes (NP) for RR20-trained mice early in training (RR20 day one or two)

and late in training (RR20 day eleven or twelve). Shaded region represents SEM. **E.** Peri-stimulus time histograms (PSTHs) showing the average signal from DMS dopamine terminals at the time of rewarded port entry (PE) for each phenotype early in training (RI60 day one or two) and late in training (RI60 day eleven or twelve). Shaded region represents SEM. **F.** Quantification of average area under the curve (AUC) of the DMS dopamine terminal signal from -5 to 0 seconds, relative to a rewarded port entry. Error bars represent SD *$p<0.05$.

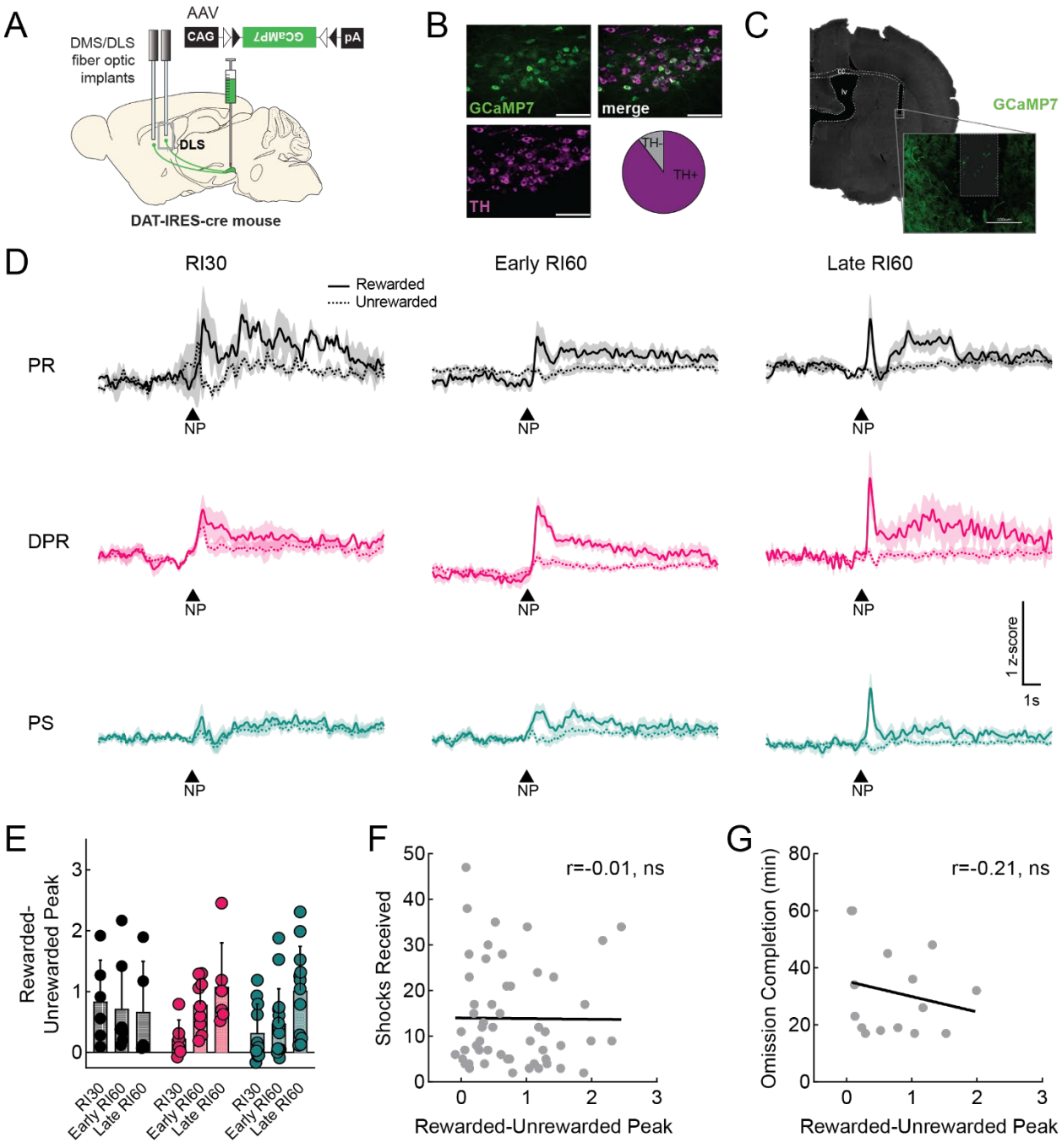

Figure 8. Dopamine signals in the DLS do not predict punishment-resistant or omission-resistant reward-seeking.

A. Viral injection and probe placement strategy for recording from dopamine terminals in the DMS and DLS. DLS is highlighted here. **B.** High magnification (40x) images of SNc showing GCaMP7b expression in green, TH positive cells in magenta, and the merged

image. Scale bars are 100μm. Quantification of GCaMP7b-expressing cells that are TH+ is shown; n=193 cells. **C.** Representative image showing probe placement in DLS. Area of magnification shows GCaMP7b expression in dopaminergic axons near the probe site in green. Scale bar is 100μm. cc=corpus callosum, lv=lateral ventricle. **D.** Peri-stimulus time histograms (PSTHs) showing the average signal from DLS dopamine terminals at the time of rewarded (solid) and unrewarded (dashed) nosepokes (NP) for each phenotype during RI30 training, early in training (RI60 day one or two), and late in training (RI60 day eleven or twelve). Shaded region represents SEM. Punishment resistant (PR; black), delayed punishment resistant (DPR; pink), or punishment sensitive (PS; teal). **E.** Quantification of the average "difference score" for DLS dopamine terminal signals observed in response to rewarded and unrewarded nosepokes (calculated as the peak at the time of rewarded nosepokes minus the peak at the time of unrewarded nosepokes). Error bars represent SD. **F.** Correlation of shocks received in shock probe sessions and "difference score" in DLS dopamine terminals (r=-0.01, ns). **G.** Correlation of time to complete omission and "difference score" in DLS dopamine terminals (r=-0.21, ns).

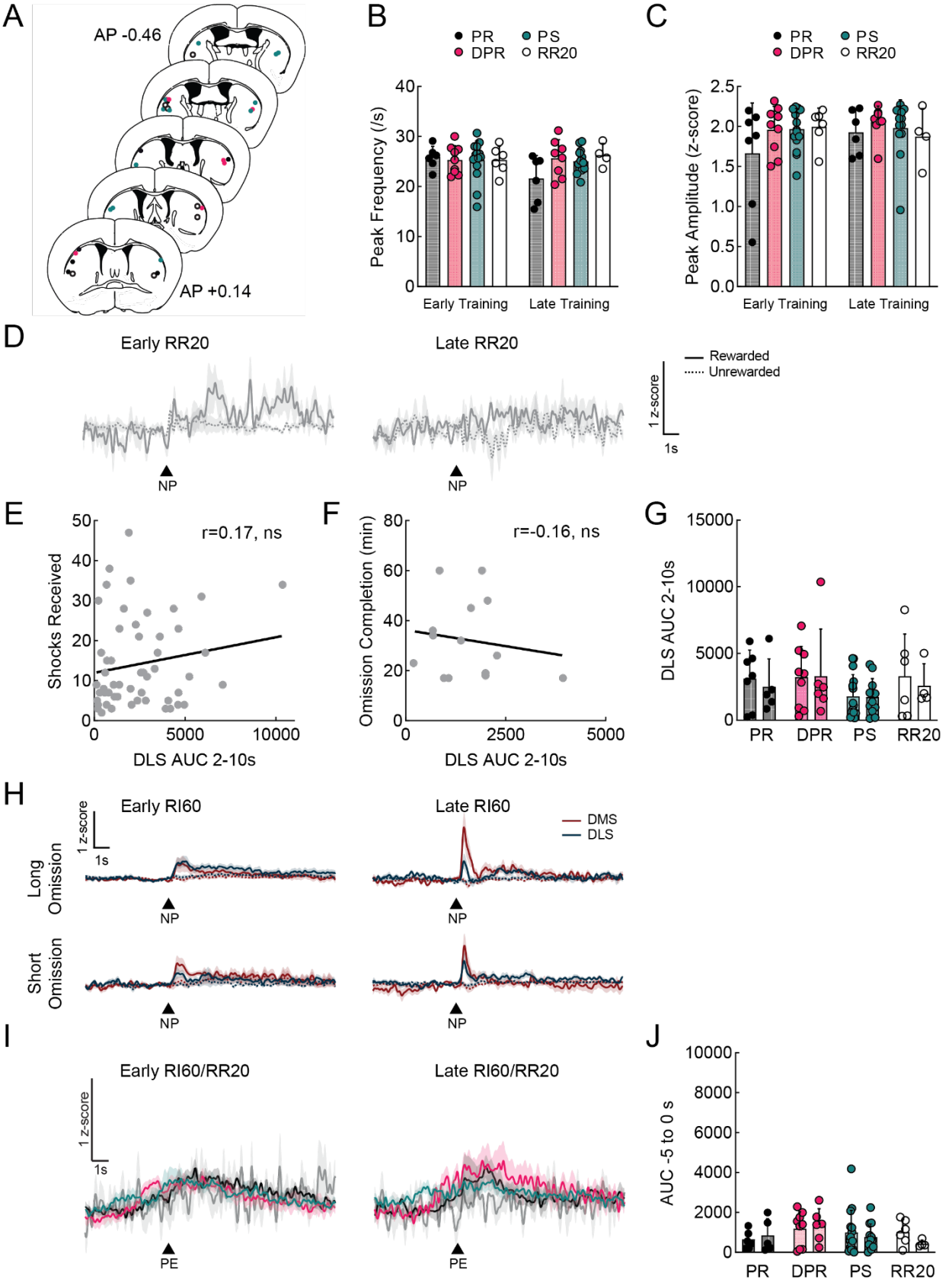
A
AP -0.46
AP +0.14
B
PR
PS
DPR
RR20
Peak Frequency (/s)
Early Training
Late Training
C
Peak Amplitude (z-score)
D
Early RR20
Late RR20
Rewarded
Unrewarded
1 z-score
1s
NP
E
r=0.17, ns
Shocks Received
DLS AUC 2-10s
F
r=-0.16, ns
Omission Completion (min)
G
DLS AUC 2-10s
H
Early RI60
Late RI60
DMS
DLS
Long Omission
Short Omission
I
Early RI60/RR20
Late RI60/RR20
PE
J
AUC -5 to 0 s

Figure 9.

A. Probe placements in DLS for all mice included in Figure 4. **B.** Average frequency of peaks detected in the DLS dopamine terminal signal for all groups. Error bars represent SD. Punishment resistant (PR; black), delayed punishment resistant (DPR; pink), or punishment sensitive (PS; teal), RR20-trained (RR20; white). **C.** Average amplitude of detected peaks in the DLS dopamine terminal signal for all groups. Error bars represent SD. **D.** Peri-stimulus time histograms (PSTHs) showing the average signal from DLS dopamine terminals at the time of rewarded (solid) and unrewarded (dashed) nosepokes (NP) for RR20-trained mice early in training (RR20 day one or two) and late in training (RR20 day eleven or twelve). Shaded region represents SEM. **E.** Correlation of shocks received and average area under the curve of DLS dopamine terminal signal from 2 to 10s, relative to the rewarded nosepoke ($r=0.17$, ns). **F.** Correlation of omission completion time and average area under the curve of DLS dopamine terminal signal from 2 to 10s, relative to the rewarded nosepoke ($r=-0.16$, ns). **G.** Quantification of average area under the curve (AUC) of the DLS dopamine terminal signal from 2 to 10 seconds, relative to the rewarded nosepoke. Error bars represent SD. **H.** Peri-stimulus time histograms (PSTHs) showing the average signal from DMS (red) and DLS (blue) dopamine terminals at the time of rewarded (solid) and unrewarded (dashed) nosepokes (NP) for animals sorted by omission completion time early in training (RI60 day one or two) and late in training (RI60 day eleven or twelve). Shaded region represents SEM. **I.** Peri-stimulus time histograms (PSTHs) showing the average signal from DLS dopamine terminals at the time of rewarded port entry (PE) for each phenotype early in training (RI60 day one or two) and late in training (RI60 day eleven

or twelve). Shaded region represents SEM. **F.** Quantification of average area under the curve (AUC) of the DLS dopamine terminal signal from -5 to 0 seconds relative to a rewarded port entry. Error bars represent SD.

III. CHAPTER THREE: Examining a causal role of dorsal striatal dopamine activity in compulsive behavior

A. Rationale

The results of Chapter II showed a correlation between dopamine terminal activity in DMS and punishment-resistant reward-seeking. In order to further investigate this finding, we wanted to test if there was, in fact, a causal role of dopamine activity in DMS on punishment-resistant reward-seeking. To accomplish this, we used both excited and inhibited dopamine terminals in DMS using optogenetics while animals underwent the same RI60 training. The strongest DMS dopamine terminal signals correlated with punishment-resistant reward-seeking were seen at the time of a rewarded nosepoke. Therefore, we wanted to artificially create this signal in DMS (or enhance, in the case of those animals that would naturally develop large DMS peaks) from early on in training in our excitatory DMS stimulation experiment or prevent it, in the case of our inhibitory DMS stimulation experiment. We also included a “scrambled” stimulation group, defined more carefully in methods, to see if any effect we may see of excitation or inhibition was unique to the specific temporal pattern we observed in Chapter II, or if increasing/decreasing DMS terminal activity in general was sufficient to demonstrate a causal relationship with punishment-resistant reward-seeking.

Although we did not find that DLS terminal activity was predictive of either punishment-resistant reward-seeking or omission resistant reward seeking, we also wanted to uncover what, if any, causal role DLS dopamine terminal stimulation might

have on our measures of compulsive-like and habit-like behavior. As discussed in the literature review, DLS is classically thought to be associated with habitual, inflexible behavior, and thus our lack of findings in Chapter II were somewhat surprising (Faure et al., 2005; Yin et al., 2004, 2006). Again, to test if a causal manipulation might elucidate some role for DLS dopamine not observed in Chapter II, we excited DLS dopamine terminals using optogenetics while animals underwent the same RI60 training.

B. Materials and Methods

Male and female *WT* (C57BL/6J) and (DAT)::IRES-Cre knockin mice (JAX006660) were obtained from The Jackson Laboratory and crossed in house. Only heterozygote transgenic mice, obtained by backcrossing to C57BL/6J wildtypes, were used for experiments. Littermates of the same sex were randomly assigned to experimental groups (DMS excitatory optogenetics- 20 males, 19 females; DMS inhibitory optogenetics- 13 males, 13 females; DLS excitatory optogenetics- 18 males, 18 females). Adult mice at least 10 weeks of age were used in all experiments. Mice were group housed under a conventional 12h light cycle (dark from 7:00pm to 7:00am) with *ad libitum* access to food and water prior to operant training. All experiments were approved by the Northwestern University Institutional Animal Care and Use Committee.

1. Methods Details

Operant Behavior

Mice were food restricted to 85% of ad libitum body weight for the duration of operant training. Mice were given one day of habituation to operant chambers (Med Associates)

and tethering with patch cords (Doric Lenses) for one hour. They were then trained to retrieve food rewards (45 mg purified pellet, Bio-Serv) from a magazine port. For this magazine training, pellets were delivered to the port on a random interval (RI60) schedule non-contingently for one hour. Next, operant training began, with all training sessions lasting one hour or until 50 rewards had been earned. Mice were trained to associate nosepoking with reward on a fixed ratio (FR1) schedule where both nosepokes delivered a reward. They had to retrieve the reward (as measured by making a port entry following a rewarded nosepoke) before they could earn the next reward. After a mouse showed a preference for one nosepoke (>25 rewards on that side; average of 3.06 days), they were trained on FR1 on their preferred side only, with nosepokes on the other side having no consequence, until they received >30 rewards for a minimum of two consecutive days (average of 5.87 days). Mice that did not reached this criterion after 14 days of FR1 training (mean+2 SD), were removed from the study. Mice passing the FR1 criterion were then moved to a random interval schedule of reinforcement. Mice on the random interval schedule were trained on RI30 until they earned >30 rewards in one hour (average of 2.33 days), and then trained on RI60. For random interval schedules, a normal distribution centered around the number indicated in the name of the schedule was used to create the schedule. The range for RI30 was from 15-45s, RI60 from 30-90s.

Shock Probe

Mice were subjected to a footshock probe early and late in training (Fig. 2B) to evaluate their levels of punishment-resistance reward-seeking. These probes were performed

under an FR1 schedule of reinforcement where a mild footshock (0.2mA, 1s) was paired with a subset of rewarded nosepokes on a RR3 schedule, so that, on average, every third rewarded nosepoke was paired with a footshock. During shock probes, the session ended after 60 minutes or a mouse was inactive (no nosepokes on the rewarded side) for >10 minutes. There was no maximum number of rewards. Mice were attached to patch cords, but no optogenetic stimulation was given during shock probe sessions.

Omission Probe

Mice were returned to RI60 training after the late footshock probe until their nosepoke rates returned to pre-shock levels. They then received a single omission probe session where they had to withhold nosepoking for 20 seconds in order to receive a single reward pellet. A nosepoke reset the 20 second timer. Each session ended after a mouse received 50 rewards or 60 minutes had elapsed. Mice were attached to patch cords, but no optogenetic stimulation was given during omission probe sessions.

Stereotaxic Surgery

Viral infusions and optic fiber implant surgeries took place under isoflurane anesthesia (Henry Schein). Mice were anesthetized in an isoflurane induction chamber at 3-4% isoflurane, and then injected with buprenorphine SR (Zoopharm, 0.5 mg/kg s.q.) and carpofen (Zoetis, 5 mg/kg s.q.) prior to the start of surgery. Mice were placed on a stereotaxic frame (Stoetling) and hair was removed from the scalp using Nair. The skin was cleaned with alcohol and a povidone-iodine solution prior to incision. The scalp was

opened using a sterile scalpel and holes were drilled in the skull at the appropriate stereotaxic coordinates. Viruses were infused at 100 nl/min through a blunt 33-gauge injection needle using a syringe pump (World Precision Instruments). The needle was left in place for 5 min following the end of the injection, then slowly retracted to avoid leakage up the injection tract. Implants were secured to the skull with Metabond (Parkell) and Flow-it ALC blue light-curing dental epoxy (Pentron). After surgery, mice were allowed to recover until ambulatory on a heated pad, then returned to their homecage with moistened chow or DietGel available. Mice then recovered for three weeks before behavioral experiments began.

Excitatory Optogenetic Stimulation

Mice for DMS (Fig. 10) and DLS (Fig. 14) excitatory optogenetics experiments received 1 µl of AAV5-EF1α-DIO-hChR2(H134R)-EYFP (3.3e13 GC/mL, Addgene, lot v17652) or the control fluorophore-only virus AAV5-EF1α-DIO-EYFP (3.5e12 virus molecules/mL, UNC Vector Core, lot AV4310K) in medial (AP -3.1, ML 0.8, DV -4.7) or lateral SNc (AP -3.1, ML 1.3, DV -4.2) and a single fiber optic implant (Prizmatix; 250µm core, 0.66 NA) over ipsilateral DMS (AP 0.8, ML 1.5, DV -2.8) or DLS (AP -0.1, ML 2.8, DV -3.5). Hemispheres were counterbalanced between mice. During operant training (beginning with FR1), each rewarded nosepoke was paired with a train of blue light (460nm, 1s, 20 Hz, 15 mW) generated by an LED light source and pulse generator (Prizmatix). A subset of mice (“ChR2 Scrambled”) received the same train of light but paired with random nosepokes on a separate RI60 schedule.

Inhibitory Optogenetic Stimulation

Mice for DMS inhibitory optogenetics experiments (Fig. 12) received 1 µl per side of AAV5-EF1α-DIO-eNpHR3.0-EYFP (1.1e13 GC/mL, Addgene, lot v32533) or the control fluorophore-only virus AAV5-EF1α-DIO-EYFP (3.5e12 virus molecules/mL, UNC Vector Core, lot AV4310K) in bilateral medial SNc (AP -3.1, ML 0.8, DV -4.7) and bilateral fiber optic implants (Prizmatix; 500µm core, 0.66 NA) in DMS (AP 0.8, ML ±1.5, DV -2.8). During operant training (beginning with FR1), each rewarded nosepoke was paired with a continuous pulse of orange/red light (625nm, 1s, 15 mW) generated by an LED light source and pulse generator (Prizmatix). A subset of mice ("NpHR Scrambled") received the same continuous pulse of light but paired with random nosepokes on a separate RI60 schedule.

Transcardial Perfusions.

Mice received lethal i.p. injections of Euthasol (Virbac, 1mg/kg) a combination of sodium pentobarbital (390 mg/ml) and sodium phenytoin (50 mg/ml), to induce a smooth and rapid onset of unconsciousness and death. Once unresponsive to a firm toe pinch, an incision was made up the middle of the body cavity. An injection needle was inserted into the left ventricle of the heart, the right atrium was punctured and solution (PBS followed by 4% PFA) was infused as the mouse was exsanguinated. The mouse was then decapitated and its brain was removed and fixed overnight at 4°C in 4% PFA.

Histology

After perfusion and fixation, brains were transferred to a solution of 30% sucrose in PBS, where they were stored for at least two overnights at 4°C before sectioning. Tissue was sectioned on a freezing microtome (Leica) at 30 μm, stored in cryoprotectant (30% sucrose, 30% ethylene glycol, 1% polyvinyl pyrrolidone in PB) at 4°C until immunostaining. Tyrosine hydroxlase (TH) staining was performed on free floating sections, which were blocked with 3% normal goat serum in PBS-T for 1 hour at room temperature, then stained with 1:500 primary antibody (Aves Labs, Cat No. TYH) in blocking solution at 4°C overnight. Secondary staining was performed using 1:500 goat anti-chicken Alexa Fluor 647 secondary antibody (Life Technologies, Cat. No. A-21449). This staining was performed by blocking in 3% normal goat serum in PBS-T for 1 hour at room temperature, then using 1:500 primary antibody conjugated directly to Alexa Fluor 488 (Life Technologies, Cat. No. A-21311) in blocking solution at 4°C overnight. Tissue was mounted on slides in PBS and coverslips were secured with Fluoromont-G (Southern Biotech). Slides were imaged using a fluorescent microscope (Keyence BZ-X800) with 5x and 40x air immersion objectives. Probe placements were determined by comparing to the Mouse Brain Atlas (Franklin & Paxinos, 2008).

2. Quantification and Statistical Analysis

Behavioral Analysis

Behavioral data was collected automatically by MED-PC software (Med Associates). To sort mice into PR, DPR, and PS groups, we calculated the percent change in shocks received from the early to late shock probe for each mouse. Mice in the top quartile of changers (who increased the number of shocks received by greater than 85%) were

classified as delayed punishment resistant. The remaining mice were sorted by a median split, with mice receiving more than 13 shocks on the first probe classified as punishment resistant and those earning fewer as punishment sensitive. Plots in Fig. 10H, 12H, and 14H were generated by plotting a segmental linear regression with lines for the average slope of nosepokes/minute across FR1, RI30, and RI60 training to reveal escalation of nosepoke behavior. The shaded area shows the 95% confidence band surrounding each slope. This analysis was done using GraphPad (Prism) software.

Statistical Methods

Statistical analysis was done using Prism 9 software (GraphPad). One and two-way ANOVAs, or mixed effects analyses were performed with Tukey's multiple comparisons and Bonferroni post-hoc analyses when statistically significant main effects or interactions were found. A total of six mice were excluded from the optogenetics studies—five due to improper probe placement and one because of illness. All n values listed above do not include these mice.

C. Results

1. Optogenetic excitation of dopamine terminals in DMS at the time of a rewarded nosepoke accelerates the development of punishment-resistant reward-seeking

Since peaks in DMS dopamine axon activity in response to rewarded nosepokes predicted the development of punishment-resistant reward-seeking, we decided to test if

stimulation of DMS dopamine axons at the time of a rewarded nosepoke could cause punishment-resistant behavior to emerge. We used the excitatory opsin ChR2 to stimulate DMS dopamine axons. An AAV expressing cre-dependent ChR2 (AAV5-EF1α-DIO-hChR2(H134R)-EYFP) was injected into the SNc of DAT-IRES-cre mice to express ChR2 specifically in dopamine neurons. A fiber optic probe was placed above the DMS to allow light stimulation of dopamine terminals only within the DMS (Fig. 10A). Previous work has shown that DMS-projecting dopamine neurons have minimal collateralization outside the DMS (Lerner et al., 2015), so this stimulation should be specific even if back-propagating action potentials are generated. We verified that this strategy led to the expression of ChR2 and EYFP in dopamine (TH+) neurons (Fig. 10B-C) and that fiber optics were appropriately placed in the DMS, matching the coordinates used for fiber photometry recordings (Fig. 11A).

Beginning with FR1 training, ChR2 animals received 1s, 20Hz trains of light stimulation on every rewarded nosepoke (Fig. 10D-E). Importantly, stimulation of DMS dopamine terminals was given during FR1/RI30/RI60 training but not during shock probes when punishment-resistant reward seeking was assessed. Therefore, any effects of the stimulation on probe performance are not due to acute effects of dopamine terminal stimulation, but are caused by differences in learning during the training sessions. In addition to animals expressing ChR2 and receiving stimulation on rewarded nosepokes, there were two control groups: EYFP controls that received a fluorophore-only control virus injected into the SNc (AAV5-EF1α-DIO-EYFP) and the same pattern of light stimulation, and “scrambled” controls that received the ChR2 virus, but had their DMS dopamine terminals stimulated on a random subset of nosepokes

rather than on rewarded nosepokes. The scrambled control was important as these mice received the same amount of dopamine terminal stimulation as the ChR2 group and this stimulation reinforced the same action (a nosepoke). Therefore, the only difference between the ChR2 and ChR2 Scrambled groups was whether or not the dopamine terminal stimulation they received boosted vs degraded the ability of the natural DMS dopamine signal to differentiate between externally rewarded and unrewarded actions.

All three groups (ChR2, EYFP, and ChR2 Scrambled) learned the FR1 task and were advanced to the RI schedules (Fig. 10F). However, mice in the ChR2 Scrambled group took significantly longer to perform to criterion on the FR1, perhaps indicating that scrambled stimulation caused an initial learning impairment (Fig. 10G; unpaired t-test, $p<0.05$). After acquiring FR1, however, all mice performed the RI60 task and received approximately the same number of rewards per minute and per session (Fig. 11B-C). ChR2 mice escalated their nosepoking much faster than the two groups of control mice during early training (FR1 and RI30), then leveled or even dropped their nosepoking rates over the course of extended RI60 training (Fig 10H). When tested on the first shock probe after minimal RI60 training, ChR2 mice were significantly more resistant to punishment than EYFP mice (Fig. 10I; Tukey's multiple comparison, $p<0.05$). This difference faded with extended training as punishment-resistant behavior also began to emerge naturally in the control groups. We categorized mice from this experiment as PR, DPR and PS using our previously defined criteria. We found that ChR2 mice were extremely likely to be categorized as PR, whereas the EYFP and ChR2 Scrambled mice were distributed as expected across groups (Fig. 10J). Notably, under ChR2

stimulation, 100% of male mice were PR mice (Fig. 11D). A large majority of female mice (~71%) were also categorized as PR – despite the fact they were unlikely to be PR in our GCaMP, EYFP, and ChR2 Scrambled groups – indicating that DMS dopamine terminal stimulation on rewarded nosepokes can drive both sexes to quickly develop punishment-resistant reward-seeking (Fig. 11D, cf. Figs. 4A, 5A).

We also looked to see whether DMS dopamine terminal stimulation influenced performance on a final omission test after extended RI60 training. It did not. ChR2, EYFP, and ChR2 Scrambled mice all took the same amount of time to complete the omission probe (Fig. 10K). These data confirm that the correlations we observed between DMS dopamine axon activity and the development of punishment-resistant behavior are causal, and specific to the development of punishment-resistant behavior. The fact that DMS dopamine terminal stimulation did not affect omission-resistance suggests that these two forms of inflexible behavior can be supported by distinct neural circuits.

2. Optogenetic inhibition of dopamine terminals in DMS interferes with action-outcome learning

Promoting DMS dopamine activity in response to rewarded nosepokes accelerated the development of punishment-resistant reward-seeking. We therefore wanted to ask the opposite question: would inhibiting DMS dopamine activity delay its development? For this experiment, we performed bilateral inhibition of DMS dopamine axons using the inhibitory opsin eNpHR3.0 (Gradinaru et al., 2010). An AAV expressing

cre-dependent NpHR (AAV5-EF1α-DIO-eNpHR3.0-EYFP) or a fluorophore-only control virus (AAV5-EF1α-DIO-EYFP) was injected into the SNc of DAT-IRES-cre mice to express NpHR (or EYFP) specifically in dopamine neurons. Fiber optic probes were placed above the DMS to allow light delivery (Fig. 12A). We verified that this strategy led to the expression of NpHR and EYFP in dopamine (TH+) neurons (Fig. 12B-C) and that fiber optics were appropriately placed in the DMS (Fig. 13A).

Mice were divided into three groups, as for the stimulation experiment. NpHR mice received a 1s continuous pulse of light on every rewarded nosepoke (Fig. 12D-E). EYFP mice received the same light stimulation, but lacked NpHR. NpHR Scrambled mice received a 1s continuous pulse of light on a random subset of nosepokes. We chose to begin the light delivery on FR1 to parallel the design of the excitatory optogenetics experiment (Fig. 10), however, DMS dopamine terminal inhibition resulted in a learning deficit. In other experiments, 100% of mice had quickly reached criterion for FR1 performance, showing that they were able to learn the association of their action (nosepoke) with an outcome (delivery of a pellet). However, only 75% of NpHR mice and 78% of NpHR Scrambled mice reached criterion (Fig. 12F). The other 25% and 22% of mice, respectively, were dropped from the study after more than 14 days (mean+2SD) of unsuccessful FR1 training. Of the mice that did pass our FR1 criterion, the NpHR mice required more days of training than EYFP control mice to reach FR1 criterion (Fig. 12G; unpaired t-test, $p<0.05$). Both of the groups receiving inhibition of their DMS dopamine terminals also displayed a reduced escalation in nosepoke rates over training (Fig. 12H). On a day-by-day basis, all groups received rewards at approximately the same rate (Fig. 13B), although when pooled over days of RI60,

NpHR mice received slightly fewer rewards than EYFP mice (Fig. 13C; $p<0.05$). Overall, we conclude that optogenetic inhibition of dopamine terminals in DMS at the time of the nosepoking action interferes with action-outcome learning in a manner that is not specific to the time of the rewarded nosepoke.

For the mice that passed FR1 and were able to complete the experiment, we tested punishment-resistance on the shock probe both early and late in training. We did not observe significant differences between groups in the total number of shocks received (Fig. 12I). However, when we sorted the mice into PR, DPR and PS groups, we noted that the NpHR group had an increased incidence of PS mice, while the NpHR Scrambled group had an increased incidence of PR mice, compared to EYFP controls (Fig. 12J). The effects of NpHR inhibition on PR/DPR/PS phenotype are driven by stark effects in male mice (Fig. S6D). We also tested these mice on the omission probe at the end of training, but did not observe significant differences in omission time (Fig. 12K). These data suggest that after initial learning delays due to DMS dopamine terminal inhibition are overcome, inhibition specifically on rewarded nosepokes may also delay the development of punishment-resistance, particularly in male mice.

3. Optogenetic excitation of dopamine terminals in DLS at the time of a rewarded nosepoke does not influence instrumental learning or behavioral flexibility

Peaks in DMS dopamine axon activity in response to rewarded nosepokes predicted the development of punishment-resistant reward-seeking (Fig. 6F), but peaks

in DLS dopamine axon activity did not (Fig. 8F). Therefore, we hypothesized that the stimulation of DLS dopamine terminals following rewarded nosepokes would not affect the development of punishment-resistant reward-seeking. Similarly, since peaks in DLS dopamine axon activity did not correlate with omission completion time, we hypothesized that DLS dopamine terminal stimulation would not affect omission completion time. Nevertheless, we decided to test the effects of DLS dopamine terminal stimulation as a counterpoint to the effects of DMS dopamine terminal stimulation (Fig. 10, 11). In other words, can stimulating *any* dopamine signal boost the development of punishment resistance, or is this effect specific to DMS?

We performed the same experiment as in Figure 10, but targeting DLS instead of DMS. An AAV expressing cre-dependent ChR2 (AAV5-EF1α-DIO-hChR2(H134R)-EYFP) or a fluorophore-only control virus (AAV5-EF1α-DIO-EYFP) was injected into the SNc of DAT-IRES-cre mice to express ChR2 or EYFP specifically in dopamine neurons. A fiber optic probe was placed above the DLS to allow light stimulation of dopamine terminals (Fig. 14A). Fiber placements and appropriate ChR2 and EYFP expression were verified by histology (Fig. 14B-C, 15A). Light stimulation (1s, 20Hz) was delivered during training sessions, beginning with FR1 (Fig. 14D-E). All mice in this experiment quickly reached FR1 criterion and advanced to RI training (Fig. 14F-G). However, in contrast to the DMS dopamine stimulation experiment, all groups of mice (ChR2, EYFP, and ChR2 Scrambled) behaved similarly. All groups of mice escalated their nosepoking at the same rates (Fig. 14H). All received similar numbers of shocks in both of the shock probes (Fig. 14I) and had similar distributions of PR, DPR, and PS mice (Fig. 14J). No significant differences were observed in omission completion time at the end of the

experiment (Fig. 14K). We conclude that DLS dopamine terminal stimulation immediately following a rewarded nosepoke does not influence the development of punishment-resistant or omission-resistant reward-seeking behavior. However, the behavioral consequence of the more prolonged/delayed activity of DLS dopamine axons we observed following rewarded nosepokes (Fig. 8D) remains to be explored in future experiments.

D. Discussion

We hypothesized that it is the discrimination of rewarded vs unrewarded actions by DMS dopamine – not the general dopamine-mediated reinforcement of the nosepoking action – that drives compulsion. To test this hypothesis, we optogenetically stimulated dopamine terminals in the DMS during learning. By creating peaks of DMS dopamine axon activity on rewarded nosepokes, we accelerated the development of punishment-resistant reward-seeking without influencing another form of inflexible responding (omission) more closely related to habit formation. Stimulating DMS dopamine terminals on random nosepokes did not affect the behavior, indicating that the timing of DMS dopamine stimulation with respect to external outcomes was important. Stimulating DLS dopamine terminals on rewarded nosepokes did not accelerate the development of punishment-resistant reward-seeking. These results support a model in which a properly timed DMS dopamine signal is specifically linked to the emergence of punishment-resistant reward-seeking. It also suggests that the normal association of compulsive and habitual responding after extended RI60 training is due to a common upstream driver of DMS and DLS function and not to a promotion of

habitual responding through DMS circuits. While we note that this finding is in conflict with the commonly cited “Ascending Spiral” hypothesis (Haber et al., 2000; Yin & Knowlton, 2006), we also recognize that it was not a direct test. Further investigations of the Ascending Spiral Hypothesis as it relates to the association of DMS and DLS function are necessary and will help to functionally connect the initial mesolimbic signaling associated with goal-directed reward-seeking to later punishment-resistant or outcome-insensitive reward-seeking.

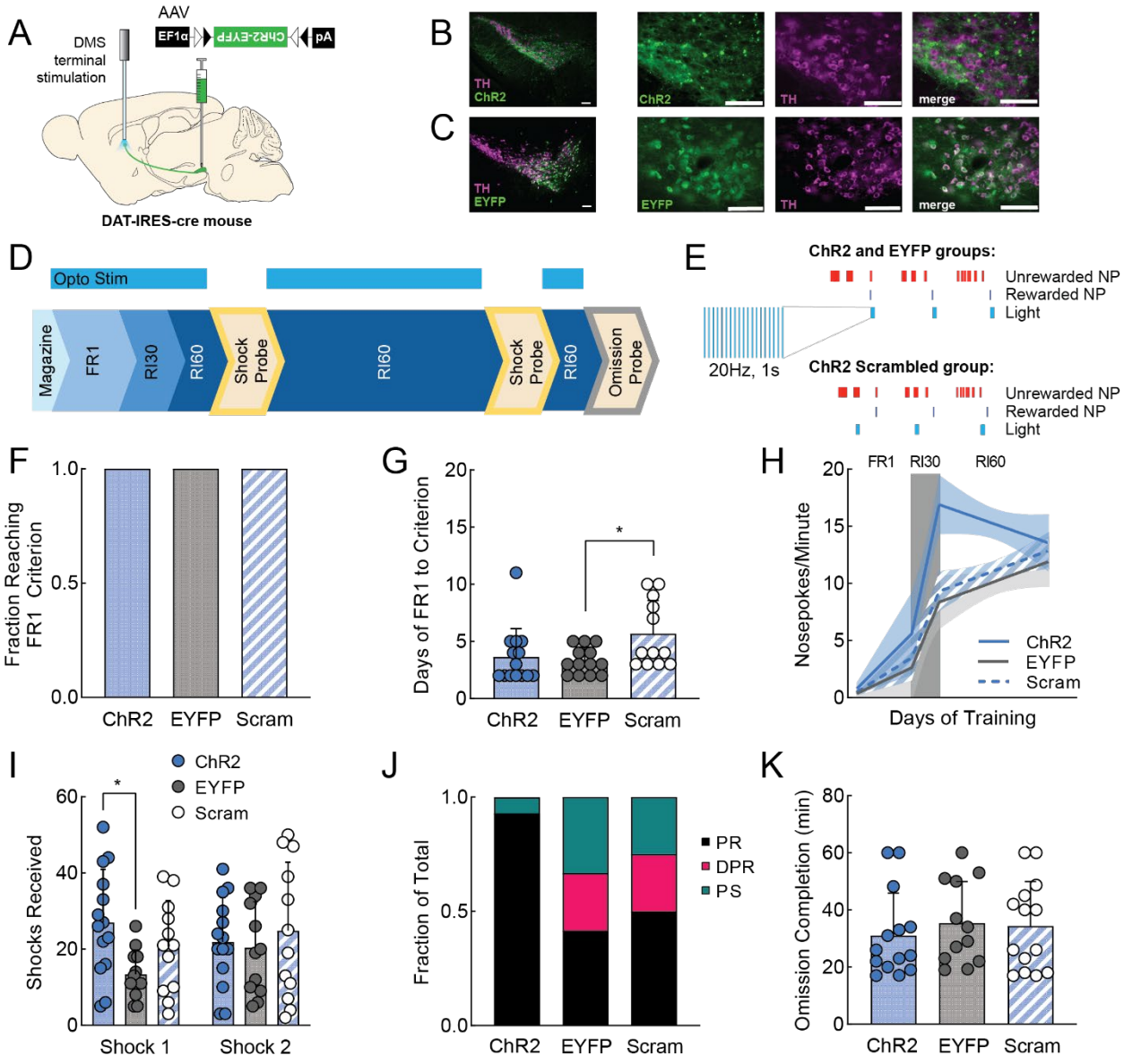

Figure 10. Optogenetic excitation of dopamine terminals in DMS at the time of a rewarded nosepoke accelerates the development of punishment-resistant reward-seeking.

A. Viral injection and probe placement strategy for stimulation of dopamine terminals in the DMS. **B.** Low (10x) and high (40x) magnification images of SNc showing ChR2-EYFP expression in green, TH positive cells in magenta, and the merged image. Scale bars are 100μm. **C.** Low (10x) and high (40x) magnification images of SNc showing EYFP expression in green, TH positive cells in magenta, and the merged image. Scale

bars are 100μm**. D.** Training timeline showing sessions during which optogenetic stimulation was delivered (FR1, RI30, RI60). **E.** Schematic of stimulation parameters. A 1s, 20Hz burst of stimulation was paired with rewarded nosepokes for ChR2 and EYFP groups and the same stimulation was paired with a random subset of nosepokes for ChR2 Scrambled animals. **F.** Fraction of mice in each group that reached the criterion to move on from FR1 training (see methods). ChR2 (blue; n=14), EYFP (gray; n=13), ChR2 scrambled (blue stripe; n=12). **G.** Average days required for animals in each group to reach criterion to move on from FR1. Error bars represent SD *p<0.05. **H.** Segmental linear regression showing the slope of nosepokes made per minute in FR1, RI30, and RI60 schedules. Shaded region represents 95% confidence bands. **I.** Average shocks received on early and late shock probes for each group. Error bars represent SD *p<0.05. **J.** Fraction of each behavioral phenotype (punishment resistant=black, delayed punishment resistant=pink, punishment sensitive=teal) in each group. **K.** Average time to complete the omission probe (earn 50 rewards; max 60 minutes) for each group. Errors bars represent SD.

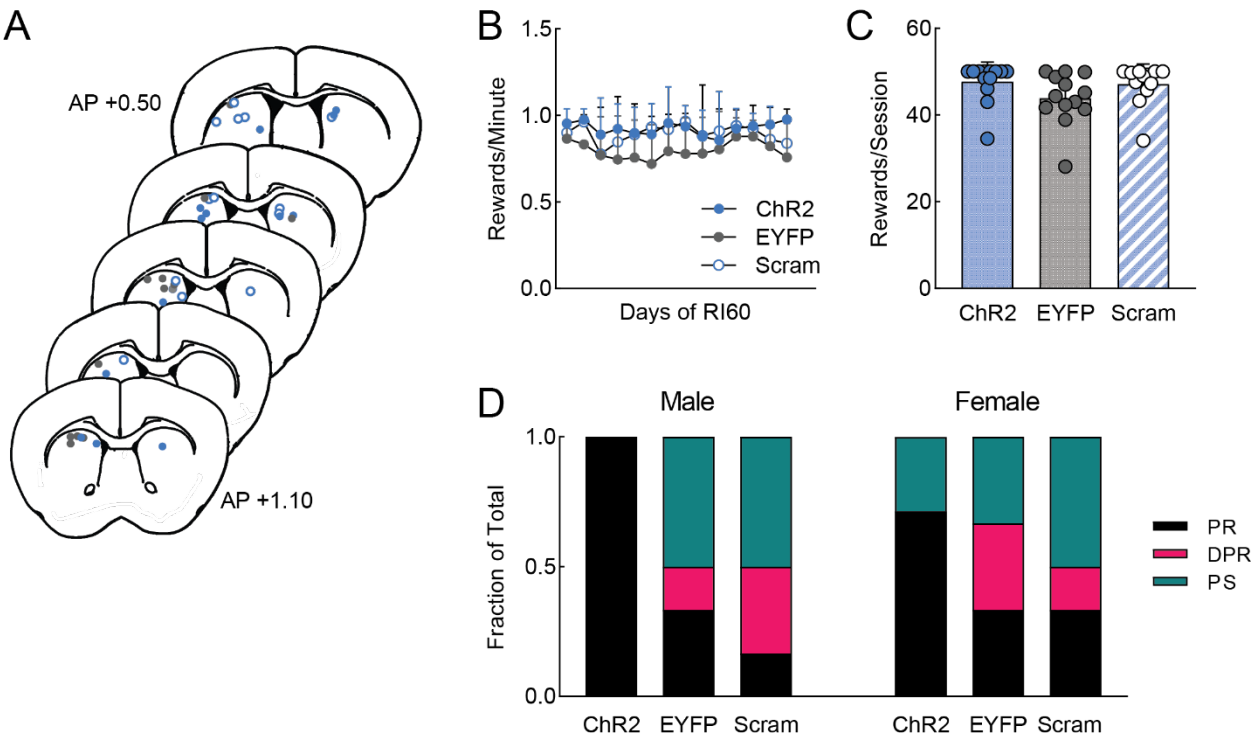

Figure 11.

A. Probe placements in DMS for all mice included in Figure 5. ChR2 (blue), EYFP (gray), ChR2 scrambled (blue outline). **B.** Average number of rewards earned per minute across days of RI60 training. Error bars represent SD. **C.** Average rewards earned per day of RI60 training. Error bars represent SD. **D.** Fraction of each behavioral phenotype (punishment resistant=black, delayed punishment resistant=pink, punishment sensitive=teal) in each manipulation divided by sex.

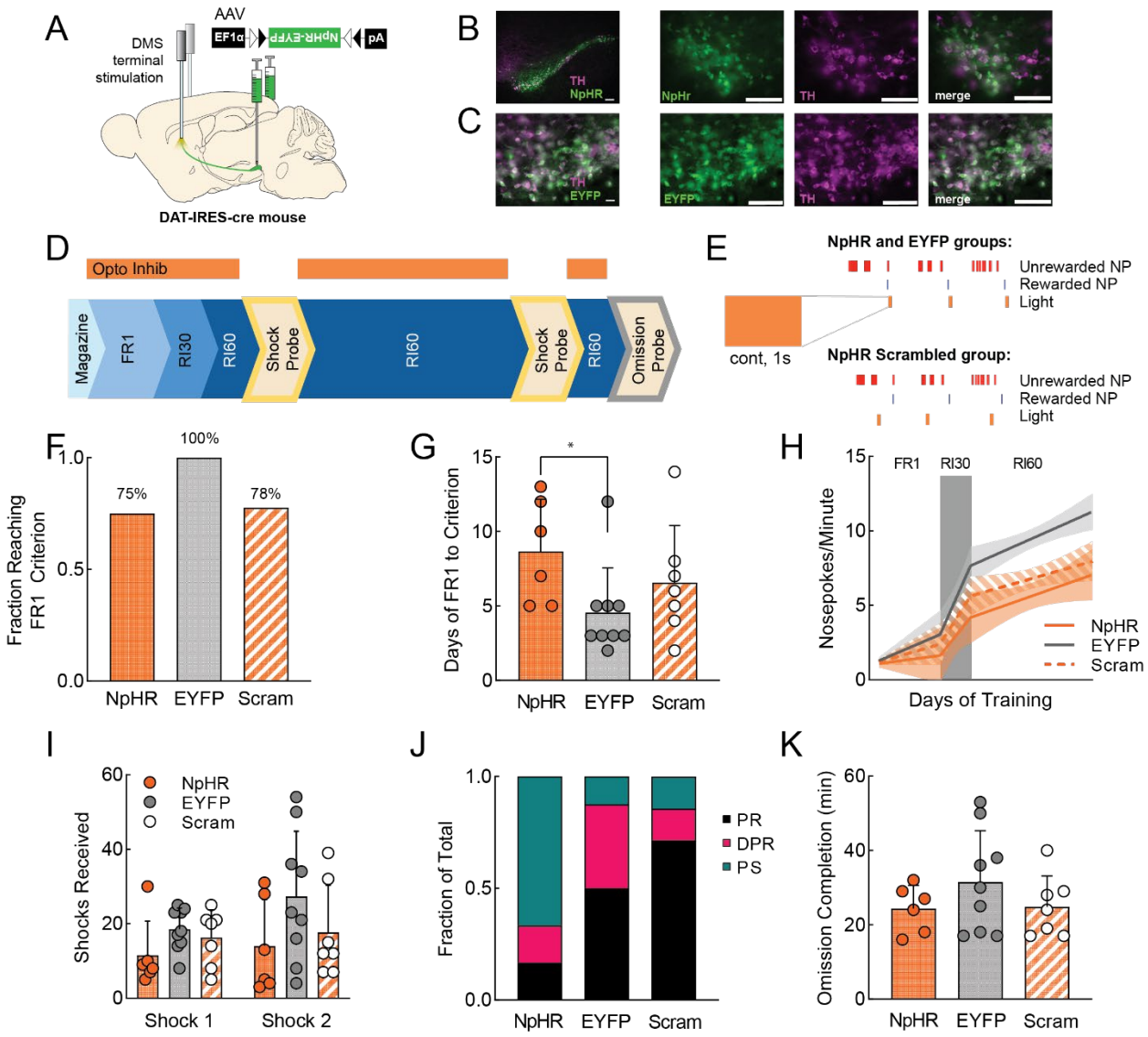

Figure 12. Optogenetic inhibition of dopamine terminals in DMS interferes with action-outcome learning.

A. Viral injection and probe placement strategy for inhibition of dopamine terminals in bilateral DMS. **B.** Low (10x) and high magnification (40x) images of SNc showing NpHR-EYFP expression in green, TH positive cells in magenta, and the merged image. Scale bars are 100µm**. C.** Low (10x) and high magnification (40x) images of SNc showing EYFP expression in green, TH positive cells in magenta, and the merged image. Scale bars are 100µm. **D.** Training timeline showing sessions during which

optogenetic stimulation was delivered (FR1, RI30, RI60). **E.** Schematic of stimulation parameters. A 1s continuous light delivery was paired with rewarded nosepokes for NpHR and EYFP groups and the same light was paired with a random subset of nosepokes for NpHR Scrambled animals. **F.** Fraction of mice in each group that reached the criterion to move on from FR1 training (see methods). NpHR (orange; n=6 of 8), EYFP (gray; n=9), NpHR scrambled (orange stripe; n=7 of 9). **G.** Average days required for animals in each group to reach criterion to move on from FR1 (for those that did). Error bars represent SD *$p<0.05$. **H.** Segmental linear regression showing the slope of nosepokes made per minute in FR1, RI30, and RI60 schedules. Shaded region represents 95% confidence bands. **I.** Average shocks received on early and late shock probes for each group. Error bars represent SD. **J.** Fraction of each behavioral phenotype (punishment resistant=black, delayed punishment resistant=pink, punishment sensitive=teal) in each group. **K.** Average time to complete the omission probe (earn 50 rewards; max 60 minutes) for each group. Errors bars represent SD.

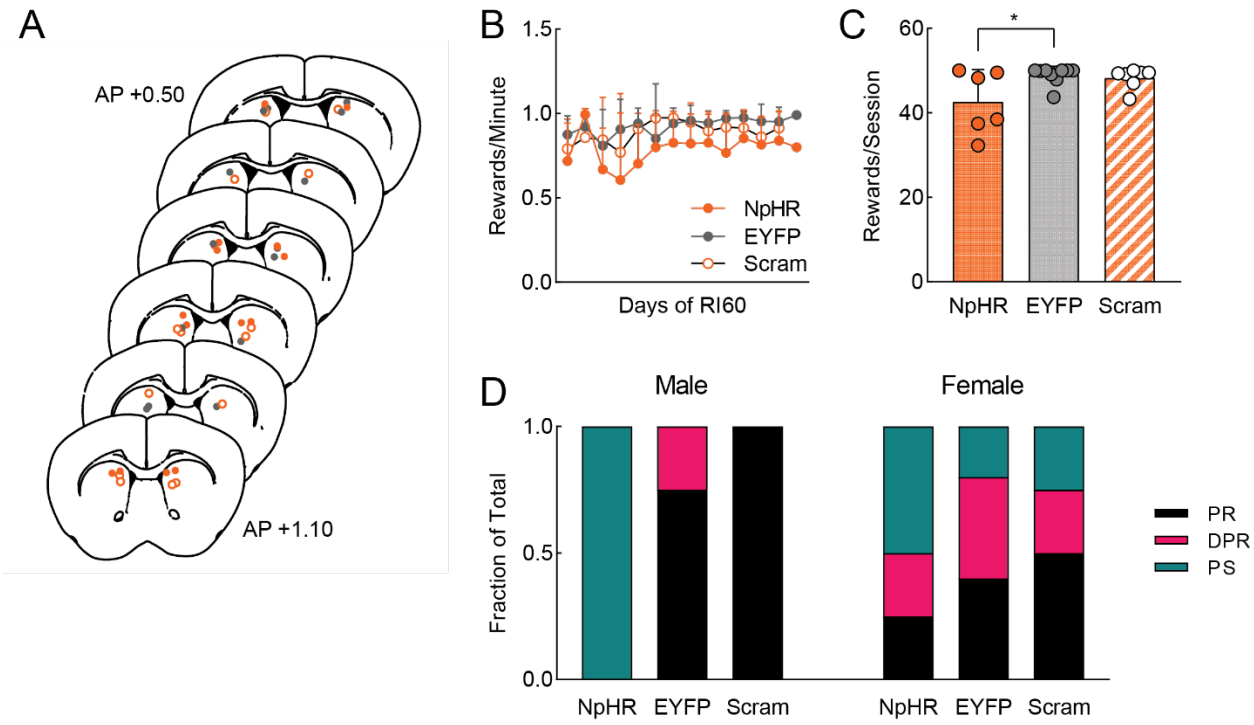

Figure 13.

A. Probe placements in DMS for all mice included in Figure 6. ChR2 (orange), EYFP (gray), ChR2 scrambled (orange outline). **B.** Average number of rewards earned per minute across days of RI60 training. Error bars represent SD. **C.** Average rewards earned per day of RI60 training. Error bars represent SD *p<0.05. **D.** Fraction of each behavioral phenotype (punishment resistant=black, delayed punishment resistant=pink, punishment sensitive=teal) in each manipulation divided by sex.

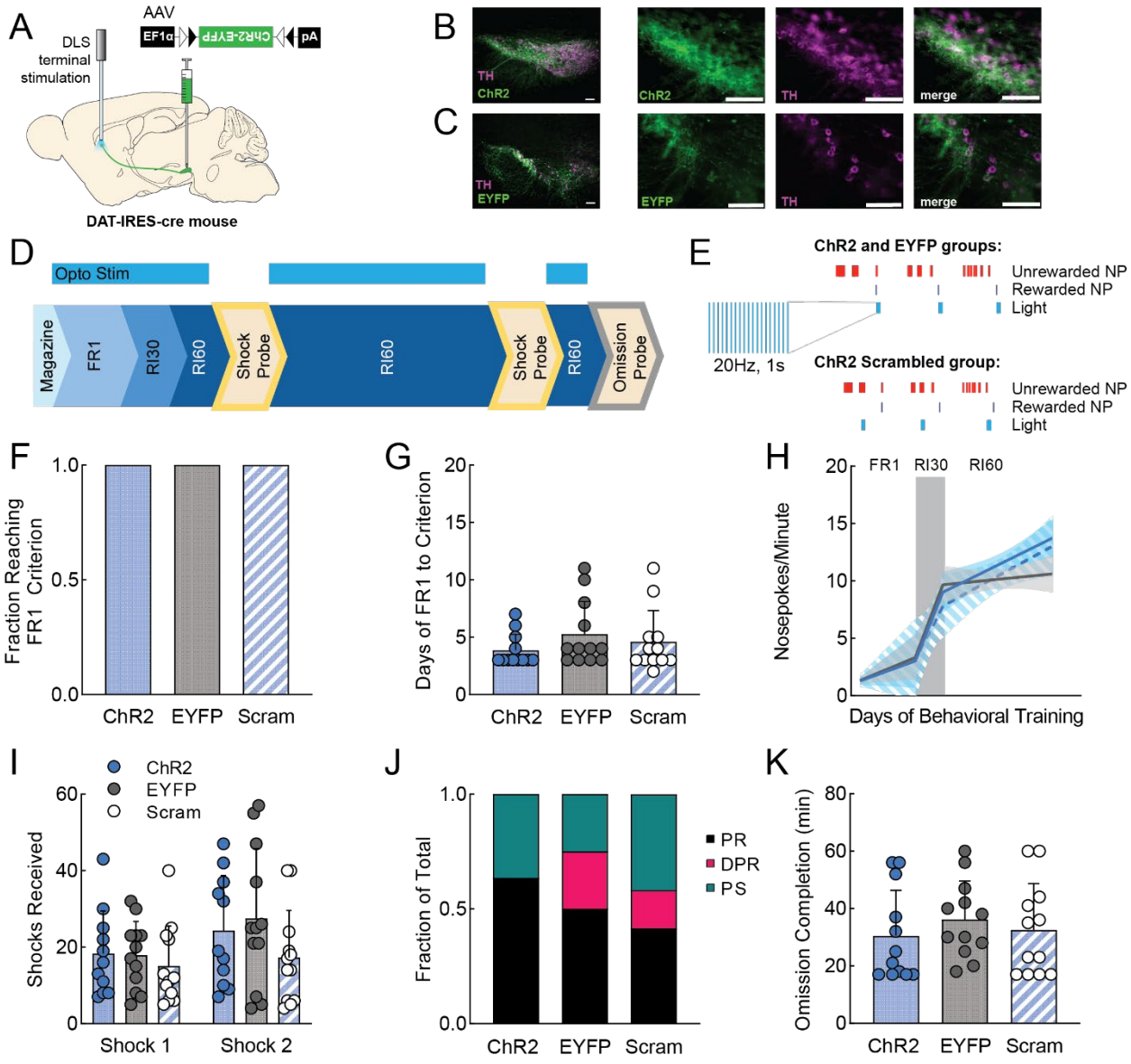

Figure 14. Optogenetic excitation of dopamine terminals in DLS at the time of a rewarded nosepoke does not influence instrumental learning or behavioral flexibility.

A. Viral injection and probe placement strategy for stimulation of dopamine terminals in the DLS. **B.** Low (10x) and high (40x) magnification images of SNc showing ChR2-EYFP expression in green, TH positive cells in magenta, and the merged image. Scale bars are 100μm. **C.** Low (10x) and high (40x) magnification images of SNc showing EYFP expression in green, TH positive cells in magenta, and the merged image. Scale

bars are 100μm. **D.** Training timeline showing sessions during which optogenetic stimulation was delivered (FR1, RI30, RI60). **E.** Schematic of stimulation parameters. A 1s, 20Hz burst of stimulation was paired with rewarded nosepokes for ChR2 and EYFP groups and the same stimulation was paired with a random subset of nosepokes for ChR2 Scrambled animals. **F.** Fraction of mice in each group that reached the criterion to move on from FR1 training (see methods). ChR2 (blue; n=11), EYFP (gray; n=12), ChR2 scrambled (blue stripe; n=12). **G.** Average days required for animals in each group to reach criterion to move on from FR1. Error bars represent SD. **H.** Segmental linear regression showing the slope of nosepokes made per minute in FR1, RI30, and RI60 schedules. Shaded region represents 95% confidence bands**. I.** Average shocks received on early and late shock probes for each group. Error bars represent SD. **J.** Fraction of each behavioral phenotype (punishment resistant=black, delayed punishment resistant=pink, punishment sensitive=teal) in each group. **K.** Average time to complete the omission probe (earn 50 rewards; max 60 minutes) for each group.

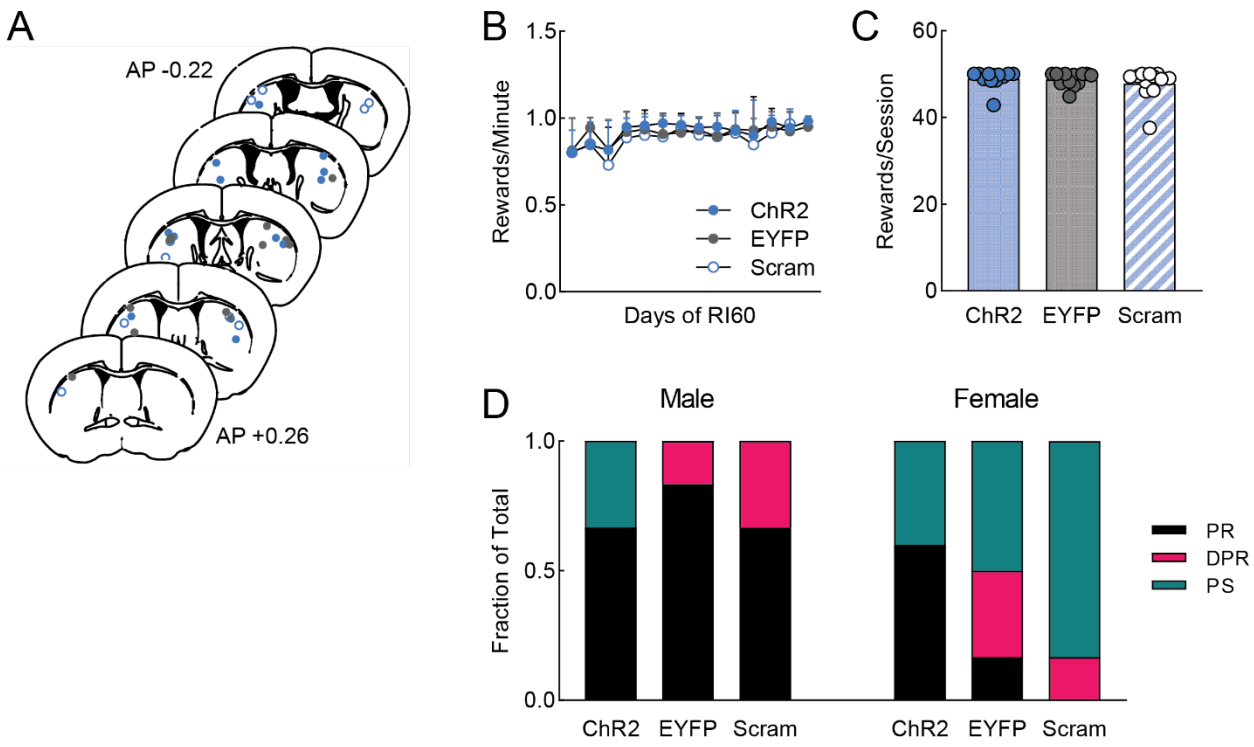

Figure 15.

A. Probe placements in DLS for all mice included in Figure 7. ChR2 (blue), EYFP (gray), ChR2 scrambled (blue outline). **B.** Average number of rewards earned per minute across days of RI60 training. Error bars represent SD. **C.** Average rewards earned per day of RI60 training. Error bars represent SD. **D.** Fraction of each behavioral phenotype (punishment resistant=black, delayed punishment resistant=pink, punishment sensitive=teal) in each manipulation divided by sex.

IV. CONCLUSIONS

The finding that DMS dopamine signaling promotes the development of punishment-resistant reward-seeking fits well with observations indicating that OFC inputs to the DMS are important for this behavior. Punishment-resistance in a paradigm in which animals self-stimulate their VTA dopamine neurons depends on enhanced excitability of OFC and potentiation of the OFC-DMS pathway (Pascoli et al., 2015, 2018). Increased OFC-DMS activity has also been associated with punishment-resistant methamphetamine-seeking (Hu et al., 2019b). Further, the OFC-DMS pathway is strengthened by repeated non-contingent injections of cocaine (Bariselli et al., 2020), and OFC neurons represent cocaine preference in cocaine-preferring rats (Guillem & Ahmed, 2018), indicating that the OFC-DMS circuit is one that could be co-opted by addictive drugs to provoke compulsive use. Our results do not directly test inputs from OFC. However, together with these previous findings, they suggest that an over-strengthening of the OFC-DMS pathway, perhaps by dopamine-dependent synaptic plasticity mechanisms, could promote punishment-resistance on a goal-directed rather than a habitual basis (Gremel et al., 2016; Gremel & Costa, 2013; Hogarth, 2020; Lüscher et al., 2020). Future work should also examine the specific vulnerability of the adolescent brain, in which cortical control is underdeveloped, and therefore might more easily lead to habit and compulsion.

Although some previous studies have observed a progression from habit to compulsion after extended training (Everitt & Robbins, 2005, 2016; Giuliano et al., 2019) we suggest that these observations could be due to a common upstream driver of DMS and DLS function, rather than a direct and necessary link between habit formation and

the development of compulsive behavior. In our experiments, extended RI60 training led to habit formation (as it has been previously documented to do; Derusso et al., 2010; Yin et al., 2006) in addition to compulsion. Thus, these two endpoints could easily be confused. However, by analyzing individual differences in behavior, we determined that the development of habits and compulsions do not inevitably develop together, consistent with the findings of Singer et al. and others (2017; Olmstead et al., 2001). Nevertheless, our results do not rule out the possibility that there are both DMS- and DLS-dependent routes to developing punishment-resistance, which could be invoked under different circumstances (e.g. different training schedules, or different modalities of reward).

Unanswered Questions

Further studies are needed on several topics. First, we need to more explicitly examine the relationship between DLS dopamine signals and habit. Previous approaches using lesions and pharmacology are suggestive of a relationship, but do not provide temporal specificity. Here, we observed novel temporal dynamics in the DLS dopamine signal, which differentiate it from the signal in the DMS. The DLS dopamine signal has an immediate peak following a rewarded nosepoke as well as prolonged activity above baseline. We did not find strong correlations of this signal with individual behavior, but alternative tests for habit, such as outcome devaluation tests, may reveal a stronger association in future studies. DLS dopamine may also be involved in other tasks. For example, one recent study linked high levels of extracellular dopamine in the DLS with high impulsivity in a delay-discounting task (Moreno et al., 2021). Another study linked a molecularly-defined population of dopamine neurons that primarily

projects to the DLS (Aldh1a1+ dopamine neurons) to motor learning on the accelerating rotarod (Wu et al., 2019). These examples highlight how much more parameter space there is still to be explored in terms of the relationship between DLS dopamine and behavior.

Second, future work should examine temporal patterns. As a control for our optogenetics experiment, we included a group of mice that received DMS dopamine terminal stimulation on random nosepokes. Creating peaks in DMS dopamine on random nosepokes did not have the same effect as creating these peaks on rewarded nosepokes. The random stimulation data indicate that the temporal pattern of dopamine activity in the DMS matters, but it remains to be determined *why* the pattern is important. For example, future studies could examine whether the pattern of cortical inputs to the DMS is different during rewarded vs unrewarded nosepokes. If different cortical inputs to the DMS are active during rewarded vs unrewarded nosepokes, dopamine release at these distinct times would reinforce the strength of different corticostriatal synapses.

Third, further studies could amplify our understanding of the DMS mechanism we observed, and clarify how it relates to the finding that projections from the lateral hypothalamus (LH) to the VTA could bidirectionally control compulsive sucrose seeking as in a previous study (Nieh et al., 2015). We do not know how this LH-VTA circuit might interact with DMS dopamine signaling; however, there are several possibilities. Most simply, some VTA dopamine neurons project to the DMS (Beier et al., 2015; M. W. Howe & Dombeck, 2016). Additionally, DMS-projecting dopamine neurons receive

inputs from LH, VTA and NAc, any of which could easily interconnect the circuitry (Lerner et al., 2015).

Larger Implications

To understand behavior, we need to grapple with individual differences. Not all mice (or people) who try drugs become compulsive users. Large individual variability in compulsivity has been observed in animals working for drugs such as cocaine and alcohol (Giuliano et al., 2019; Siciliano et al., 2019). Our findings underline the necessity of this approach; we saw large individual differences as mice worked for natural rewards. We identified one reason for this variability: the different strategies used by individual animals to deal with uncertainty in the RI60 schedule. In our study, some mice maximized their rate of reward retrieval by nosepoking at high rates, while other mice conserved effort. This finding also sheds light on the poorly understood question of how punishment-resistant reward-seeking emerges over time. In addition to differences in neural signals, we also observed different behavioral strategies prior to the experience of punishment, which allowed us to predict which animals will go on to become punishment-resistant. This pattern suggests there is a predisposition to towards developing punishment-resistance present in individuals before they confront punishments, rather than a stochastic process occurring during the experience of punishment, to generate punishment-resistance.

One source of individual variability in our studies was sex: male mice were more likely to be punishment-resistant than females. Nevertheless, sex could not fully explain individual variability. Furthermore, the correlation between DMS dopamine axon

signaling and the development of punishment-resistance was not sex-dependent, and DMS dopamine terminal stimulation could induce both sexes to transition to punishment-resistance. We therefore suspect that sex differences influencing differences in the likelihood of males and females to develop punishment resistance occur upstream of dopamine neurons.

Understanding compulsive behavior in humans and rodents also requires understanding whether or not there is variability associated with the nature of the rewarding substance. In this study, we examined the behavior of mice during learning to pursue sucrose rewards. Compulsive sucrose-seeking is relatively understudied; most previous studies have examined compulsive drug-seeking. It is important to understand behavior across natural and manufactured rewards of various kinds. By studying compulsive seeking of a natural reward, we can better understand the evolutionary context under which this behavior developed, perhaps working to promote what might be called “grit” in the face of life’s inevitable challenges. Furthermore, understanding how compulsive drug-seeking and compulsive sucrose-seeking relate to each other can elucidate how concepts from SUD should be applied to our understanding other disorders like eating disorders, gambling disorders, or OCD (among others). While the circuit mechanism promoting compulsive sucrose-seeking we have identified is exciting, we need to ascertain whether it would similarly drive the development of compulsive drug-seeking. Conversely, previously identified circuits for compulsive drug-seeking might or might not impact compulsive sucrose-seeking. In fact, one behavioral study directly examined whether the development of compulsive drug-seeking and compulsive sucrose-seeking were correlated and found that they were NOT: animals that became

compulsive for one type of reward did not necessarily become compulsive for the other type (Datta et al., 2018). Thus, it is imperative to examine in more detail whether all rewards do or do not activate the same circuits for compulsivity.

In summary, we have identified DMS dopamine signaling as a key part of the circuitry that drives the emergence of compulsive behavior in the context of natural reward-seeking. The data presented here set the stage for interesting new studies in a variety of areas. Examining how the mechanisms we have identified contribute to the etiology of disorders such as SUD is of particular importance for translational impact.

V. CITED LITERATURE

Aalto, S., Hirvonen, J., Kajander, J., Scheinin, H., Någren, K., Vilkman, H., Gustafsson, L., Syvälahti, E., & Hietala, J. (2002). Ketamine does not decrease striatal dopamine D2 receptor binding in man. *Psychopharmacology*, *164*(4), 401–406. https://doi.org/10.1007/s00213-002-1236-6

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Association. https://doi.org/10.1176/appi.books.9780890425596

Amodeo, L. R., Jacobs-Brichford, E., McMurray, M. S., & Roitman, J. D. (2017). Acute and long-term effects of adolescent methylphenidate on decision-making and dopamine receptor mRNA expression in the orbitofrontal cortex. *Behavioural Brain Research*, *324*, 100–108. https://doi.org/10.1016/j.bbr.2017.02.019

Anthony, J., Warner, L., & Kessler, R. (1994). Comparative epidemiology of dependence on tobacco, alcohol, controlled substances, and inhalants: Basic findings from the National Comorbidity Survey | Health Care Policy | Harvard Medical School. *Experimental and Clinical Psychopharmacology*. https://www.hcp.med.harvard.edu/publications/comparative-epidemiology-dependence-tobacco-alcohol-controlled-substances-and-inhalants

Bariselli, S., Miyazaki, N. L., Creed, M. C., & Kravitz, A. V. (2020). Orbitofrontal-striatal potentiation underlies cocaine-induced hyperactivity. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-17763-8

Beaudoin, G. M. J., Gomez, J. A., Perkins, J., Bland, J. L., Petko, A. K., & Paladini, C. A. (2018). Cocaine Selectively Reorganizes Excitatory Inputs to Substantia Nigra Pars Compacta Dopamine Neurons. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *38*(5), 1151–1159. https://doi.org/10.1523/JNEUROSCI.1975-17.2017

Beier, K. T. T., Steinberg, E. E. E., Deloach, K. E. E., Xie, S., Miyamichi, K., Schwarz, L., Gao, X. J. J., Kremer, E. J. J., Malenka, R. C. C., & Luo, L. (2015). Circuit Architecture of VTA Dopamine Neurons Revealed by Systematic Input-Output Mapping. *Cell*, *162*(3), 622–634. https://doi.org/10.1016/J.CELL.2015.07.015

Belin, D., & Everitt, B. J. (2008). Cocaine Seeking Habits Depend upon Dopamine-Dependent Serial Connectivity Linking the Ventral with the Dorsal Striatum. *Neuron*, *57*(3), 432–441. https://doi.org/10.1016/J.NEURON.2007.12.019

Berridge, K. C., & Robinson, T. E. (1998). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Research. Brain Research Reviews*, *28*(3), 309–369. http://www.ncbi.nlm.nih.gov/pubmed/9858756

Braun, A. A., Amos-Kroohs, R. M., Gutierrez, A., Lundgren, K. H., Seroogy, K. B., Skelton, M. R., Vorhees, C. V., & Williams, M. T. (2015). Dopamine depletion in either the dorsomedial or dorsolateral striatum impairs egocentric Cincinnati water maze performance while sparing allocentric Morris water maze learning.

Neurobiology of Learning and Memory, *118*. https://doi.org/10.1016/j.nlm.2014.10.009

Brown, H. D., Mccutcheon, J. E., Cone, J. J., Ragozzino, M. E., & Roitman, M. F. (2011). Primary food reward and reward-predictive stimuli evoke different patterns of phasic dopamine signaling throughout the striatum. *European Journal of Neuroscience*, *34*(12), 1997–2006. https://doi.org/10.1111/j.1460-9568.2011.07914.x

Burguiere, E., Monteiro, P., Feng, G., & Graybiel, A. M. (2013). Optogenetic Stimulation of Lateral Orbitofronto-Striatal Pathway Suppresses Compulsive Behaviors. *Science*, *340*(6137), 1243–1246. https://doi.org/10.1126/science.1232380

Carelli, R. M., Ijames, S. G., & Crumling, A. J. (2000). Evidence that separate neural circuits the nucleus accumbens encode cocaine versus “natural” (water and food) reward. *Journal of Neuroscience*, *20*(11), 4255–4266. https://doi.org/10.1523/JNEUROSCI.20-11-04255.2000

Carelli, R. M., & Wondolowski, J. (2003). Selective Encoding of Cocaine versus Natural Rewards by Nucleus Accumbens Neurons Is Not Related to Chronic Drug Exposure. *Journal of Neuroscience*, *23*(35), 11214–11223. https://doi.org/10.1523/jneurosci.23-35-11214.2003

Castañé, A., Theobald, D. E. H., & Robbins, T. W. (2010). Selective lesions of the dorsomedial striatum impair serial spatial reversal learning in rats. *Behavioural Brain Research*, *210*(1), 74–83. https://doi.org/10.1016/J.BBR.2010.02.017

Centers for Disease Control. (n.d.). *NVSS - Mortality Data*. Retrieved June 15, 2021, from https://www.cdc.gov/nchs/nvss/deaths.htm?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fnchs%2Fdeaths.htm

Centers for Disease Control. (2017). *Older Adult Drivers | Motor Vehicle Safety | CDC Injury Center*. https://www.cdc.gov/motorvehiclesafety/teen_drivers/index.html

Chiara, G. Di, & Imperato, A. (1988). Drugs abused by humans preferentially increase synaptic dopamine concentrations in the mesolimbic system of freely moving rats. *Proceedings of the National Academy of Sciences*, *85*(14), 5274–5278. https://doi.org/10.1073/PNAS.85.14.5274

Compton, W. M., Thomas, Y. F., Stinson, F. S., & Grant, B. F. (2007). Prevalence, correlates, disability, and comorbidity of DSM-IV drug abuse and dependence in the United States: results from the national epidemiologic survey on alcohol and related conditions. *Archives of General Psychiatry*, *64*(5), 566–576. https://doi.org/10.1001/archpsyc.64.5.566

Corbett, D., & Wise, R. A. (1980). Intracranial self-stimulation in relation to the ascending dopaminergic systems of the midbrain: A moveable electrode mapping study. *Brain Research*, *185*(1), 1–15. https://doi.org/10.1016/0006-8993(80)90666-6

Corbit, L. H., Nie, H., & Janak, P. H. (2012). Habitual Alcohol Seeking: Time Course and the Contribution of Subregions of the Dorsal Striatum. *Biological Psychiatry*, *72*(5), 389–395. https://doi.org/10.1016/j.biopsych.2012.02.024

Cromwell, H. C., & Berridge, K. C. (1996). Implementation of action sequences by a neostriatal site: a lesion mapping study of grooming syntax. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *16*(10), 3444–3458. http://www.ncbi.nlm.nih.gov/pubmed/8627378

Crow, T. J. (1972). A map of the rat mesencephalon for electrical self-stimulation. *Brain Research*, *36*(2), 265–273. https://doi.org/10.1016/0006-8993(72)90734-2

Datta, U., Martini, M., & Sun, W. (2018). Different functional domains measured by cocaine self-administration under the progressive-ratio and punishment schedules in male Wistar rats. *Psychopharmacology*, *235*(3), 897–907. https://doi.org/10.1007/s00213-017-4808-1

de Wit, H., & Wise, R. A. (1977). Blockade of cocaine reinforcement in rats with the dopamine receptor blocker pimozide, but not with the noradrenergic blockers phentolamine or phenoxybenzamine. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *31*(4), 195–203. https://doi.org/10.1037/h0081662

DePoy, L. M., Allen, A. G., & Gourley, S. L. (2016). Adolescent cocaine self-administration induces habit behavior in adulthood: sex differences and structural consequences. *Translational Psychiatry*, *6*(8), e875–e875. https://doi.org/10.1038/tp.2016.150

Deroche-Gamonet, V., Belin, D., & Piazza, P. V. (2004). Evidence for addiction-like behavior in the rat. *Science (New York, N.Y.)*, *305*(5686), 1014–1017. https://doi.org/10.1126/science.1099020

Derusso, A. L., Fan, D., Gupta, J., Shelest, O., Costa, R. M., & Yin, H. H. (2010). Instrumental uncertainty as a determinant of behavior under interval schedules of reinforcement. *Frontiers in Integrative Neuroscience*, *4*. https://doi.org/10.3389/fnint.2010.00017

Ersche, K. D., Barnes, A., Jones, P. S., Morein-Zamir, S., Robbins, T. W., & Bullmore, E. T. (2011). Abnormal structure of frontostriatal brain systems is associated with aspects of impulsivity and compulsivity in cocaine dependence. *Brain*, *134*(7), 2013–2024. https://doi.org/10.1093/brain/awr138

Everitt, B. J., & Robbins, T. W. (2005). Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nature Neuroscience*, *8*(11), 1481–1489. https://doi.org/10.1038/nn1579

Everitt, B. J., & Robbins, T. W. (2016). Drug Addiction: Updating Actions to Habits to Compulsions Ten Years On. *Annual Review of Psychology*, *67*(1), 23–50. https://doi.org/10.1146/annurev-psych-122414-033457

Faure, A., Haberland, U., Condé, F., & El Massioui, N. (2005). Lesion to the Nigrostriatal Dopamine System Disrupts Stimulus-Response Habit Formation.

Journal of Neuroscience, *25*(11), 2771–2780. https://doi.org/10.1523/JNEUROSCI.3894-04.2005

Franklin, K. B. J., & Paxinos, G. (2008). *The mouse brain in stereotaxic coordinates*. Boston.

Giovino, G. (1999). Epidemiology of tobacco use among US adolescents. *Nicotine & Tobacco Research*, *1*(0), S31–S40. https://doi.org/10.1080/14622299050011571

Giuliano, C., Belin, D., & Everitt, B. J. (2019). Compulsive Alcohol Seeking Results from a Failure to Disengage Dorsolateral Striatal Control over Behavior. *The Journal of Neuroscience*, *39*(9), 2615–2618. https://doi.org/10.1523/JNEUROSCI.2615-18.2018

Gradinaru, V., Zhang, F., Ramakrishnan, C., Mattis, J., Prakash, R., Diester, I., Goshen, I., Thompson, K. R., & Deisseroth, K. (2010). Molecular and Cellular Approaches for Diversifying and Extending Optogenetics. *Cell*, *141*(1), 154–165. https://doi.org/10.1016/j.cell.2010.02.037

Gremel, C. M., Chancey, J. H., Atwood, B. K., Luo, G., Neve, R., Ramakrishnan, C., Deisseroth, K., Lovinger, D. M., & Costa, R. M. (2016). Endocannabinoid Modulation of Orbitostriatal Circuits Gates Habit Formation. *Neuron*, *90*(6), 1312–1324. https://doi.org/10.1016/j.neuron.2016.04.043

Gremel, C. M., & Costa, R. M. (2013). Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. *Nature Communications*, *4*(1), 2264. https://doi.org/10.1038/ncomms3264

Grospe, G. M., Baker, P. M., & Ragozzino, M. E. (2018). Cognitive Flexibility Deficits Following 6-OHDA Lesions of the Rat Dorsomedial Striatum. *Neuroscience*, *374*, 80–90. https://doi.org/10.1016/j.neuroscience.2018.01.032

Guillem, K., & Ahmed, S. H. (2018). Preference for Cocaine is Represented in the Orbitofrontal Cortex by an Increased Proportion of Cocaine Use-Coding Neurons. *Cerebral Cortex*, *28*(3), 819–832. https://doi.org/10.1093/cercor/bhw398

Guru, A., Seo, C., Post, R., Kullakanda, D., Schaffer, J., & Warden, M. (2020). Ramping activity in midbrain dopamine neurons signifies the use of a cognitive map. *BioRxiv*. https://doi.org/10.1101/2020.05.21.108886

Haber, S. N., Fudge, J. L., & McFarland, N. R. (2000). Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *20*(6), 2369–2382. http://www.ncbi.nlm.nih.gov/pubmed/10704511

Hadjas, L. C., Lüscher, C., & Simmler, L. D. (2019). Aberrant habit formation in the Sapap3-knockout mouse model of obsessive-compulsive disorder. *Scientific Reports*, *9*(1), 12061. https://doi.org/10.1038/s41598-019-48637-9

Hamid, A. A., Frank, M. J., & Moore, C. I. (2019). Dopamine waves as a mechanism for spatiotemporal credit assignment. *BioRxiv*, 729640. https://doi.org/10.1101/729640

Hamid, A. A., Pettibone, J. R., Mabrouk, O. S., Hetrick, V. L., Schmidt, R., Vander Weele, C. M., Kennedy, R. T., Aragona, B. J., & Berke, J. D. (2015). Mesolimbic dopamine signals the value of work. *Nature Neuroscience*. https://doi.org/10.1038/nn.4173

Harada, M., Hiver, A., Pascoli, V., & Lüscher, C. (2019). Cortico-striatal synaptic plasticity underlying compulsive reward seeking. *BioRxiv*, 789495. https://doi.org/10.1101/789495

Harlow, J. M. (1868). Passage of an iron rod through the head. In *Journal of Neuropsychiatry and Clinical Neurosciences* (Vol. 11, Issue 2, pp. 281–283). https://doi.org/10.1176/jnp.11.2.281

Hodebourg, R., Murray, J. E., Fouyssac, M., Puaud, M., Everitt, B. J., & Belin, D. (2018). Heroin seeking becomes dependent on dorsal striatal dopaminergic mechanisms and can be decreased by N-acetylcysteine. *European Journal of Neuroscience*. https://doi.org/10.1111/ejn.13894

Hogarth, L. (2020). Addiction is driven by excessive goal-directed drug choice under negative affect: translational critique of habit and compulsion theory. *Neuropsychopharmacology*, 1–18. https://doi.org/10.1038/s41386-020-0600-8

Holly, E. N., Davatolhagh, M. F., Choi, K., Alabi, O. O., Vargas Cifuentes, L., & Fuccillo, M. V. (2019). Striatal Low-Threshold Spiking Interneurons Regulate Goal-Directed Learning. *Neuron*, *103*(1), 92-101.e6. https://doi.org/10.1016/J.NEURON.2019.04.016

Howe, M. W., & Dombeck, D. A. (2016). Rapid signalling in distinct dopaminergic axons during locomotion and reward. *Nature*. https://doi.org/10.1038/nature18942

Howe, Mark W., Tierney, P. L., Sandberg, S. G., Phillips, P. E. M., & Graybiel, A. M. (2013). Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. *Nature*, *500*(7464), 575–579. https://doi.org/10.1038/nature12475

Hu, Y., Salmeron, B. J., Krasnova, I. N., Gu, H., Lu, H., Bonci, A., Cadet, J. L., Stein, E. A., & Yang, Y. (2019a). Compulsive drug use is associated with imbalance of orbitofrontal- and prelimbic-striatal circuits in punishment-resistant individuals. *Proceedings of the National Academy of Sciences*, *116*(18), 9066–9071. https://doi.org/10.1073/PNAS.1819978116

Hu, Y., Salmeron, B. J., Krasnova, I. N., Gu, H., Lu, H., Bonci, A., Cadet, J. L., Stein, E. A., & Yang, Y. (2019b). Compulsive drug use is associated with imbalance of orbitofrontal- And prelimbic-striatal circuits in punishment-resistant individuals. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(18), 9066–9071. https://doi.org/10.1073/pnas.1819978116

Ikemoto, S. (2007). Dopamine reward circuitry: Two projection systems from the ventral midbrain to the nucleus accumbens–olfactory tubercle complex. *Brain Research Reviews*, *56*(1), 27–78. https://doi.org/10.1016/j.brainresrev.2007.05.004

Ilango, A., Kesner, A. J., Keller, K. L., Stuber, G. D., Bonci, A., & Ikemoto, S. (2014).

Similar Roles of Substantia Nigra and Ventral Tegmental Dopamine Neurons in Reward and Aversion. *The Journal of Neuroscience*, *34*(3), 817–822. https://doi.org/10.1523/JNEUROSCI.1703-13.2014

Jacobs-Brichford, E., Manson, K. F., & Roitman, J. D. (2019). Effects of chronic cannabinoid exposure during adolescence on reward preference and mPFC activation in adulthood. *Physiology and Behavior*. https://doi.org/10.1016/j.physbeh.2018.12.006

Kalueff, A. V., Stewart, A. M., Song, C., Berridge, K. C., Graybiel, A. M., & Fentress, J. C. (2016). Neurobiology of rodent self-grooming and its value for translational neuroscience. *Nature Reviews Neuroscience*, *17*(1), 45–59. https://doi.org/10.1038/nrn.2015.8

Keiflin, R., Pribut, H. J., Shah, N. B., & Janak, P. H. (2017). Phasic Activation of Ventral Tegmental, but not Substantia Nigra, Dopamine Neurons Promotes Model-Based Pavlovian Reward Learning. *BioRxiv*, 232678. https://doi.org/10.1101/232678

Kelley, A. E., Schochet, T., & Landry, C. F. (2004). Risk taking and novelty seeking in adolescence. *Annals of the New York Academy of Sciences*, *1021*, 27–32. https://doi.org/10.1196/annals.1308.003

Kim, C. K., Yang, S. J., Pichamoorthy, N., Young, N. P., Kauvar, I., Jennings, J. H., Lerner, T. N., Berndt, A., Lee, S. Y., Ramakrishnan, C., Davidson, T. J., Inoue, M., Bito, H., & Deisseroth, K. (2016). Simultaneous fast measurement of circuit dynamics at multiple sites across the mammalian brain. *Nature Methods*. https://doi.org/10.1038/nmeth.3770

Kim, H. G. R., Malik, A. N., Mikhael, J. G., Bech, P., Tsutsui-Kimura, I., Sun, F., Zhang, Y., Li, Y., Watabe-Uchida, M., Gershman, S. J., & Uchida, N. (2020). A Unified Framework for Dopamine Signals across Timescales. *Cell*, *183*(6), 1600-1616.e25. https://doi.org/10.1016/j.cell.2020.11.013

Kravitz, A. V, Tye, L. D., & Kreitzer, A. C. (2012). Distinct roles for direct and indirect pathway striatal neurons in reinforcement. *Nature Neuroscience*, *15*(6), 816–818. https://doi.org/10.1038/nn.3100

Laruelle, M., & Abi-Dargham, A. (1999). Dopamine as the wind of the psychotic fire: New evidence from brain imaging studies. *Journal of Psychopharmacology*, *13*(4), 358–371. https://doi.org/10.1177/026988119901300405

Lerner, T. N., Shilyansky, C., Davidson, T. J., Evans, K. E., Beier, K. T., Zalocusky, K. A., Crow, A. K., Malenka, R. C., Luo, L., Tomer, R., & Deisseroth, K. (2015). Intact-Brain Analyses Reveal Distinct Information Carried by SNc Dopamine Subcircuits. *Cell*, *162*(3), 635–647. https://doi.org/10.1016/j.cell.2015.07.014

Lipton, D. M., Gonzales, B. J., & Citri, A. (2019). Dorsal Striatal Circuits for Habits, Compulsions and Addictions. *Frontiers in Systems Neuroscience*, *13*, 28. https://doi.org/10.3389/fnsys.2019.00028

Lugo, J. N., Smith, G. D., & Holley, A. J. (2014). Trace fear conditioning in mice. *Journal*

of Visualized Experiments : JoVE, *85*. https://doi.org/10.3791/51180

Lüscher, C., & Janak, P. H. (2021). Consolidating the Circuit Model for Addiction. *Annual Review of Neuroscience*, *44*(1), annurev-neuro-092920-123905. https://doi.org/10.1146/annurev-neuro-092920-123905

Lüscher, C., Robbins, T. W., & Everitt, B. J. (2020). The transition to compulsion in addiction. *Nature Reviews Neuroscience*, 1–17. https://doi.org/10.1038/s41583-020-0289-z

Marques, A., Durif, F., & Fernagut, P. O. (2018). Impulse control disorders in Parkinson's disease. In *Journal of Neural Transmission*. https://doi.org/10.1007/s00702-018-1870-8

Martinez, D., Carpenter, K. M., Liu, F., Slifstein, M., Broft, A., Friedman, A. C., Kumar, D., Van Heertum, R., Kleber, H. D., & Nunes, E. (2011). Imaging dopamine transmission in cocaine dependence: link between neurochemistry and response to treatment. *The American Journal of Psychiatry*, *168*(6), 634–641. https://doi.org/10.1176/appi.ajp.2010.10050748

Martinez, D., Narendran, R., Foltin, R. W., Slifstein, M., Hwang, D.-R., Broft, A., Huang, Y., Cooper, T. B., Fischman, M. W., Kleber, H. D., & Laruelle, M. (2007). Amphetamine-induced dopamine release: markedly blunted in cocaine dependence and predictive of the choice to self-administer cocaine. *The American Journal of Psychiatry*, *164*(4), 622–629. https://doi.org/10.1176/ajp.2007.164.4.622

McKim, T. H., Bauer, D. J., & Boettiger, C. A. (2016). Addiction History Associates with the Propensity to Form Habits. *Journal of Cognitive Neuroscience*, *28*(7), 1024–1038. https://doi.org/10.1162/jocn_a_00953

McMurray, M. S., Amodeo, L. R., & Roitman, J. D. (2016). Consequences of Adolescent Ethanol Consumption on Risk Preference and Orbitofrontal Cortex Encoding of Reward. *Neuropsychopharmacology*. https://doi.org/10.1038/npp.2015.288

Mohebi, A., Pettibone, J. R., Hamid, A. A., Wong, J.-M. T., Vinson, L. T., Patriarchi, T., Tian, L., Kennedy, R. T., & Berke, J. D. (2019). Dissociable dopamine dynamics for learning and motivation. *Nature*, 1. https://doi.org/10.1038/s41586-019-1235-y

Moreno, M., Azocar, V., Vergés, A., & Fuentealba, J. A. (2021). High impulsive choice is accompanied by an increase in dopamine release in rat dorsolateral striatum. *Behavioural Brain Research*, *405*. https://doi.org/10.1016/j.bbr.2021.113199

Muir, J., Lorsch, Z. S., Ramakrishnan, C., Deisseroth, K., Nestler, E. J., Calipari, E. S., & Bagot, R. C. (2018). In Vivo Fiber Photometry Reveals Signature of Future Stress Susceptibility in Nucleus Accumbens. *Neuropsychopharmacology*, *43*(2), 255–263. https://doi.org/10.1038/npp.2017.122

Murray, E. A., O'Doherty, J. P., & Schoenbaum, G. (2007). What we know and do not know about the functions of the orbitofrontal cortex after 20 years of cross-species studies. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *27*(31), 8166–8169. https://doi.org/10.1523/JNEUROSCI.1556-

07.2007

Murray, J. E., Belin, D., & Everitt, B. J. (2012). Double dissociation of the dorsomedial and dorsolateral striatal control over the acquisition and performance of cocaine seeking. *Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology*, *37*(11), 2456–2466. https://doi.org/10.1038/npp.2012.104

Murray, J. E., Everitt, B. J., & Belin, D. (2012). N-Acetylcysteine reduces early- and late-stage cocaine seeking without affecting cocaine taking in rats. *Addiction Biology*, *17*(2), 437–440. https://doi.org/10.1111/j.1369-1600.2011.00330.x

Napier, T. C., Kirby, A., & Persons, A. L. (2020). The role of dopamine pharmacotherapy and addiction-like behaviors in Parkinson's disease. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 109942. https://doi.org/10.1016/j.pnpbp.2020.109942

National Institute on Drug Abuse. (2017). *Costs of Substance Abuse | NIDA Archives*. https://archives.drugabuse.gov/trends-statistics/costs-substance-abuse#supplemental-references-for-economic-costs

Nelson, A., & Killcross, S. (2006). Amphetamine Exposure Enhances Habit Formation. *Journal of Neuroscience*, *26*(14), 3805–3812. https://doi.org/10.1523/JNEUROSCI.4305-05.2006

Nieh, E. H., Matthews, G. A., Allsop, S. A., Presbrey, K. N., Leppla, C. A., Wichmann, R., Neve, R., Wildes, C. P., & Tye, K. M. (2015). Decoding Neural Circuits that Control Compulsive Sucrose Seeking. *Cell*, *160*(3), 528–541. https://doi.org/10.1016/j.cell.2015.01.003

Norman, A. B., & Tsibulsky, V. L. (2006). The compulsion zone: A pharmacological theory of acquired cocaine self-administration. *Brain Research*, *1116*(1), 143–152. https://doi.org/10.1016/j.brainres.2006.07.092

Olds, J., & Milner, P. M. (1954). Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *Journal of Comparative and Physiological Psychology*, *47*(6), 419–427. http://www.ncbi.nlm.nih.gov/pubmed/13233369

Olmstead, M. C., Lafond, M. V., Everitt, B. J., & Dickinson, A. (2001). Cocaine seeking by rats is a goal-directed action. *Behavioral Neuroscience*, *115*(2), 394–402. https://doi.org/10.1037/0735-7044.115.2.394

Pacchioni, A. M., Gabriele, A., & See, R. E. (2011). Dorsal striatum mediation of cocaine-seeking after withdrawal from short or long daily access cocaine self-administration in rats. *Behavioural Brain Research*, *218*(2), 296–300. https://doi.org/10.1016/j.bbr.2010.12.014

Palencia, C. A., & Ragozzino, M. E. (2004). The influence of NMDA receptors in the dorsomedial striatum on response reversal learning. *Neurobiology of Learning and Memory*, *82*(2), 81–89. https://doi.org/10.1016/j.nlm.2004.04.004

Pascoli, V., Hiver, A., Van Zessen, R., Loureiro, M., Achargui, R., Harada, M., Flakowski, J., & Lüscher, C. (2018). Stochastic synaptic plasticity underlying compulsion in a model of addiction. *Nature*, *564*(7736), 366–371. https://doi.org/10.1038/s41586-018-0789-4

Pascoli, V., Terrier, J., Espallergues, J., Valjent, E., O'connor, E. C., & Lüscher, C. (2014). Contrasting forms of cocaine-evoked plasticity control components of relapse. *Nature*. https://doi.org/10.1038/nature13257

Pascoli, V., Terrier, J., Hiver, A., & Lüscher, C. (2015). Sufficiency of Mesolimbic Dopamine Neuron Stimulation for the Progression to Addiction. *Neuron*, *88*(5), 1054–1066. https://doi.org/10.1016/j.neuron.2015.10.017

Reynolds, J. N. J., Hyland, B. I., & Wickens, J. R. (2001). A cellular mechanism of reward-related learning. *Nature*, *413*(6851), 67–70. https://doi.org/10.1038/35092560

Ritz, M. C., Lamb, R. J., Goldberg, S. R., & Kuhar, M. J. (1988). Cocaine self-administration appears to be mediated by dopamine uptake inhibition. *Progress in Neuropsychopharmacology and Biological Psychiatry*, *12*(2–3), 233–239. https://doi.org/10.1016/0278-5846(88)90040-1

Robinson, T. E., & Berridge, K. C. (1993). The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain Research. Brain Research Reviews*, *18*(3), 247–291. https://doi.org/10.1016/0165-0173(93)90013-P

Robinson, T. E., & Kolb, B. (2004). Structural plasticity associated with exposure to drugs of abuse. *Neuropharmacology*. https://doi.org/10.1016/j.neuropharm.2004.06.025

Rossi, M. A., Sukharnikova, T., Hayrapetyan, V. Y., Yang, L., & Yin, H. H. (2013). Operant Self-Stimulation of Dopamine Neurons in the Substantia Nigra. *PLoS ONE*, *8*(6), e65799. https://doi.org/10.1371/journal.pone.0065799

Rossi, M. A., & Yin, H. H. (2012). Methods for Studying Habitual Behavior in Mice. In *Current Protocols in Neuroscience: Vol. Chapter 8* (p. Unit 8.29). John Wiley & Sons, Inc. https://doi.org/10.1002/0471142301.ns0829s60

Saunders, B. T., Richard, J. M., Margolis, E. B., & Janak, P. H. (2018). Dopamine neurons create Pavlovian conditioned stimuli with circuit-defined motivational properties. *Nature Neuroscience*, *21*(8), 1072–1083. https://doi.org/10.1038/s41593-018-0191-4

Seip-Cammack, K. M., Young, J. J., Young, M. E., & Shapiro, M. L. (2017). Partial lesion of the nigrostriatal dopamine pathway in rats impairs egocentric learning but not spatial learning or behavioral flexibility. *Behavioral Neuroscience*, *131*(2). https://doi.org/10.1037/bne0000189

Siciliano, C. A., Noamany, H., Chang, C.-J., Brown, A. R., Chen, X., Leible, D., Lee, J. J., Wang, J., Vernon, A. N., Vander Weele, C. M., Kimchi, E. Y., Heiman, M., & Tye, K. M. (2019). A cortical-brainstem circuit predicts and governs compulsive

alcohol drinking. *Science (New York, N.Y.)*, *366*(6468), 1008–1012. https://doi.org/10.1126/science.aay1186

Singer, B. F., Fadanelli, M., Kawa, A. B., & Robinson, T. E. (2017). Are cocaine-seeking "habits" necessary for the development of addiction-like behavior in rats? *The Journal of Neuroscience*, *38*(1), 2458–17. https://doi.org/10.1523/JNEUROSCI.2458-17.2017

Spear, L. P. (2000). The adolescent brain and age-related behavioral manifestations. *Neuroscience and Biobehavioral Reviews*, *24*(4), 417–463. https://doi.org/10.1016/S0149-7634(00)00014-2

Spoelder, M., Hesseling, P., Styles, M., Baars, A. M., Lozeman-van 't Klooster, J. G., Lesscher, H. M. B., & Vanderschuren, L. J. M. J. (2017). Dopaminergic neurotransmission in ventral and dorsal striatum differentially modulates alcohol reinforcement. *European Journal of Neuroscience*, *45*(1), 147–158. https://doi.org/10.1111/ejn.13358

Stein, L. (1964). SELF-STIMULATION OF THE BRAIN AND THE CENTRAL STIMULANT ACTION OF. *Federation Proceedings*, *23*, 836–850. http://www.ncbi.nlm.nih.gov/pubmed/14195468

Terrier, J., Lüscher, C., & Pascoli, V. (2016). Cell-Type Specific Insertion of GluA2-Lacking AMPARs with Cocaine Exposure Leading to Sensitization, Cue-Induced Seeking, and Incubation of Craving. *Neuropsychopharmacology*. https://doi.org/10.1038/npp.2015.345

Thorn, C. A., Atallah, H., Howe, M., & Graybiel, A. M. (2010). Differential Dynamics of Activity Changes in Dorsolateral and Dorsomedial Striatal Loops during Learning. *Neuron*. https://doi.org/10.1016/j.neuron.2010.04.036

Ungless, M. A., Whistler, J. L., Malenka, R. C., & Bonci, A. (2001). Single cocaine exposure in vivo induces long-term potentiation in dopamine neurons. *Nature*, *411*(6837), 583–587. https://doi.org/10.1038/35079077

Vandaele Y, & Janak PH. (2021). Unveiling the neural correlates of habit in the dorsal striatum. *BioRxiv*, 2021.04.03.438314. https://doi.org/10.1101/2021.04.03.438314

Vandaele Y, Mahajan, N. R., Ottenheimer, D. J., Richard, J. M., Mysore, S. P., & Janak, P. H. (2019). Distinct recruitment of dorsomedial and dorsolateral striatum erodes with extended training. *ELife*. https://doi.org/10.7554/eLife.49536

Vanderschuren, L. J. M. J., Di Ciano, P., & Everitt, B. J. (2005). *Involvement ofthe Dorsal Striatum in Cue-Controlled Cocaine Seeking*. *25*(38). https://doi.org/10.1523/JNEUROSCI.0925-05.2005

Vanderschuren, L. J. M. J., Everitt, B. J., Dalley, J. W., Robbins, T. W., & Everitt, B. J. (2004). *Drug Seeking Becomes Compulsive After Prolonged Cocaine Self-Administration*. *305*(5686). https://doi.org/10.1126/science.1098975

Veeneman, M. M. J., Broekhoven, M. H., Damsteegt, R., & Vanderschuren, L. J. M. J. (2012). Distinct Contributions of Dopamine in the Dorsolateral Striatum and

Nucleus Accumbens Shell to the Reinforcing Properties of Cocaine. *Neuropsychopharmacology*, *37*(2), 487–498. https://doi.org/10.1038/npp.2011.209

Volkow, N. D., Wang, G. -J, Fowler, J. S., Logan, J., Schlyer, D., Hitzemann, R., Lieberman, J., Angrist, B., Pappas, N., MacGregor, R., Burr, G., Cooper, T., & Wolf, A. P. (1994). Imaging endogenous dopamine competition with [11C]raclopride in the human brain. *Synapse*, *16*(4), 255–262. https://doi.org/10.1002/syn.890160402

Volkow, N. D., Wang, G. J., Fowler, J. S., Logan, J., Gatley, S. J., Hitzemann, R., Chen, A. D., Dewey, S. L., & Pappas, N. (1997). Decreased striatal dopaminergic responsiveness in detoxified cocaine-dependent subjects. *Nature*, *386*(6627), 830–833. https://doi.org/10.1038/386830a0

Volkow, N. D., Wang, G. J. G.-J., Telang, F., Fowler, J. S., Logan, J., Childress, A.-R. A. R., Jayne, M., Ma, Y., & Wong, C. (2006). Cocaine Cues and Dopamine in Dorsal Striatum: Mechanism of Craving in Cocaine Addiction. *Journal of Neuroscience*, *26*(24), 6583–6588. https://doi.org/10.1523/JNEUROSCI.1544-06.2006

Wagner, F. A., & Anthony, J. C. (2002). From first drug use to drug dependence: Developmental periods of risk for dependence upon marijuana, cocaine, and alcohol. *Neuropsychopharmacology*, *26*(4), 479–488. https://doi.org/10.1016/S0893-133X(01)00367-0

Wall, N. R., Neumann, P. A., Beier, K. T., Mokhtari, A. K., Luo, L., & Malenka, R. C. (2019). Complementary Genetic Targeting and Monosynaptic Input Mapping Reveal Recruitment and Refinement of Distributed Corticostriatal Ensembles by Cocaine. *Neuron*, *104*(5), 916-930.e5. https://doi.org/10.1016/J.NEURON.2019.10.032

Willuhn, I., Burgeno, L. M., Everitt, B. J., & Phillips, P. E. M. (2012). Hierarchical recruitment of phasic dopamine signaling in the striatum during the progression of cocaine use. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(50), 20703–20708. https://doi.org/10.1073/pnas.1213460109

Wiltgen, B. J., Sinclair, C., Lane, C., Barrows, F., Molina, M., & Chabanon-Hicks, C. (2012). The Effect of Ratio and Interval Training on Pavlovian-Instrumental Transfer in Mice. *PLoS ONE*, *7*(10), e48227. https://doi.org/10.1371/journal.pone.0048227

Wise, R. A., Leeb, K., Pocock, D., Newton, P., Burnette, B., & Justice, J. B. (1995). Fluctuations in nucleus accumbens dopamine concentration during intravenous cocaine self-administration in rats. *Psychopharmacology*, *120*(1), 10–20. https://doi.org/10.1007/BF02246140

Wise, R A, & Bozarth, M. A. (1987). A psychomotor stimulant theory of addiction. *Psychological Review*, *94*(4), 469–492. http://www.ncbi.nlm.nih.gov/pubmed/3317472

Wise, Roy A. (1981). Intracranial self-stimulation: mapping against the lateral boundaries of the dopaminergic cells of the substantia nigra. *Brain Research*, *213*(1), 190–194. https://doi.org/10.1016/0006-8993(81)91260-9

Wu, J., Kung, J., Dong, J., Chang, L., Xie, C., Habib, A., Hawes, S., Yang, N., Chen, V., Liu, Z., Evans, R., Liang, B., Sun, L., Ding, J., Yu, J., Saez-Atienzar, S., Tang, B., Khaliq, Z., Lin, D. T., … Cai, H. (2019). Distinct Connectivity and Functionality of Aldehyde Dehydrogenase 1a1-Positive Nigrostriatal Dopaminergic Neurons in Motor Learning. *Cell Reports*, *28*(5), 1167-1181.e7. https://doi.org/10.1016/j.celrep.2019.06.095

Yager, L. M., Garcia, A. F., Donckels, E. A., & Ferguson, S. M. (2018). Chemogenetic inhibition of direct pathway striatal neurons normalizes pathological, cue-induced reinstatement of drug-seeking in rats. *Addiction Biology*. https://doi.org/10.1111/adb.12594

Yager, L. M., Pitchers, K. K., Flagel, S. B., & Robinson, T. E. (2015). Individual Variation in the Motivational and Neurobiological Effects of an Opioid Cue. *Neuropsychopharmacology*. https://doi.org/10.1038/npp.2014.314

Yin, H. H., & Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nature Reviews Neuroscience*, *7*(6), 464–476. https://doi.org/10.1038/nrn1919

Yin, H. H., Knowlton, B. J., & Balleine, B. W. (2004). Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *European Journal of Neuroscience*, *19*(1), 181–189. https://doi.org/10.1111/j.1460-9568.2004.03095.x

Yin, H. H., Knowlton, B. J., & Balleine, B. W. (2005). Blockade of NMDA receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning. *European Journal of Neuroscience*, *22*(2), 505–512. https://doi.org/10.1111/j.1460-9568.2005.04219.x

Yin, H. H., Knowlton, B. J., & Balleine, B. W. (2006). Inactivation of dorsolateral striatum enhances sensitivity to changes in the action-outcome contingency in instrumental conditioning. *Behavioural Brain Research*, *166*(2), 189–196. https://doi.org/10.1016/j.bbr.2005.07.012

Yin, H. H., Mulcare, S. P., Hilário, M. R. F., Clouse, E., Holloway, T., Davis, M. I., Hansson, A. C., Lovinger, D. M., & Costa, R. M. (2009). Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. *Nature Neuroscience*. https://doi.org/10.1038/nn.2261

Yin, H. H., Ostlund, S. B., Knowlton, B. J., & Balleine, B. W. (2005). The role of the dorsomedial striatum in instrumental conditioning. *European Journal of Neuroscience*, *22*(2), 513–523. https://doi.org/10.1111/j.1460-9568.2005.04218.x

Yoder, K. K., Constantinescu, C. C., Kareken, D. A., Normandin, M. D., Cheng, T. E., O'Connor, S. J., & Morris, E. D. (2007). Heterogeneous effects of alcohol on dopamine release in the striatum: A PET study. *Alcoholism: Clinical and Experimental Research*, *31*(6), 965–973. https://doi.org/10.1111/j.1530-0277.2007.00390.x

Yu, C., Gupta, J., Chen, J.-F., & Yin, H. H. (2009). Genetic Deletion of A2A Adenosine Receptors in the Striatum Selectively Impairs Habit Formation. *Journal of*

Neuroscience, *29*(48), 15100–15103. https://doi.org/10.1523/JNEUROSCI.4215-09.2009

Zapata, A., Minney, V. L., & Shippenberg, T. S. (2010). Shift from Goal-Directed to Habitual Cocaine Seeking after Prolonged Experience in Rats. *Journal of Neuroscience*, *30*(46), 15457–15463. https://doi.org/10.1523/JNEUROSCI.4072-10.2010

VI. VITA

Jillian L Seiler
847.525.4442
jillianseiler@gmail.com

Dept of Physiology, Ward 5-120
Northwestern University, Feinberg School of Medicine
303 E. Chicago Ave, Chicago, IL 60611

Education

PhD Candidate, University of Illinois at Chicago. Chicago, IL — 2014-July 2021 (expected)
Behavioral Neuroscience, Psychology Department.
Minor: Psychopharmacoepidemiology
Research Advisor: Dr. Talia Lerner, Academic Advisor: Dr. Dave Wirtshafter

M.A., Thesis: Effects of physiological need states on affect and reward sensitivity — 2017
Advisor: Dr. Mitch Roitman

B.S. University of Michigan. Ann Arbor, MI — 2009-2013
Biopsychology, Cognition and Neuroscience
Senior Thesis: *In vivo* identification of dopamine transmission during reward omission
Advisor: Dr. Brandon Aragona

Experience

Visiting Pre-Doctoral Fellow, Northwestern University, Feinberg School of Medicine — 2018-present
Pending Dissertation Title: Neural circuit drivers of compulsive behavior
Advisor: Dr. Talia Lerner

Teaching Assistant, University of Illinois at Chicago — 2014-present
14 semesters. Classes Taught: Laboratory in Behavioral Neuroscience,
Behavioral Neuroscience, Learning and Conditioning,
Laboratory in Cognitive Neuroscience

Graduate Research Assistant, University of Illinois at Chicago — 2014-2018
Advisor: Dr. Mitch Roitman

Laboratory Manager and Technician, University of Michigan — 2012-2014
Advisor: Drs. Brandon Aragona & Terry Robinson

Laboratory Technician, University of Illinois at Chicago — 2012
Advisor: Dr. Mitch Roitman

Undergraduate Research Assistant, University of Michigan — 2010-2013
Advisor: Dr. Brandon Aragona

Publications

Seiler, JL, Cosme, CV, Sherathiya VN, Bianco, JM, Bridgemohan, AS, Lerner, TN (2020). Dopamine signaling in the dorsomedial striatum promotes compulsive behavior. *Preprint available at BioRxiv

Lerner, TN, Holloway AL, **Seiler, JL** (2021). Dopamine, Updated: Reward Prediction Error and Beyond. *Current Opinion in Neurobiology, 67*, 123-130. doi:10.1016/j.conb.2020.10.012

Singer, BF, Guptaroy, B, Austin, CJ, Wohl, I, Lovic, V, **Seiler, JL,** Vaughan, RA, Gnegy, ME, Robinson, TE, Aragona, BJ. (2016). Individual variation in incentive salience attribution and accumbens dopamine transporter expression and function. European Journal of Neuroscience, 43: 662–670. doi: 10.1111/ejn.13134

Porter-Stranksy, KA, **Seiler, JL,** Day, JJ, and Aragona, BJ (2013). Development of behavioral preferences for the optimal choice following unexpected reward omission is mediated by a reduction of D2-like receptor tone in the nucleus accumbens. European Journal of Neuroscience, 38: 2572-2588. doj: 10.1111/ejn.12253

Invited Talks

Seiler, JL & Lerner, TL. Laboratory in Integrative Neuroscience Symposium, 2021, online. "Dopamine Signaling in the Dorsomedial Striatum Promotes Compulsive Behavior."

Seiler, JL & Lerner, TL. Women in Learning: Women to Watch Series, 2021, online. "Dopamine Signaling in the Dorsomedial Striatum Promotes Compulsive Behavior."

Seiler, JL, Cross Program Conference, University of Illinois at Chicago, 2017, Chicago, IL. "Effects of hunger and its mimetics on reward sensitivity."

Seiler, JL, Laboratory in Integrative Neuroscience Symposium, University of Illinois at Chicago, 2017, Chicago, IL. "Effects of hunger and its mimetics on reward sensitivity."

Seiler, JL, Department of Psychology Honors Research Symposium, 2013, University of Michigan, Ann Arbor, MI. "*In-vivo* identification of dopamine transmission during reward omission."

Poster Abstracts

Seiler, JL, Cosme, CV, Lerner, TL. Catecholamines Gordon Research Conference, 2019, Newry, ME. "Habit formation in the medial nigrostriatal pathway."

Seiler, JL, Conway, SM, Roitman, MF, Society for Neuroscience Annual Meeting, 2016, San Diego, CA. "Effects of pharmacologically induced physiological need states on reward processing through intracranial self-stimulation."

Singer, BF, Guptaroy, B, Austin, CJ, Wohl, I, Lovic, V, **Seiler, JL,** Vaughan, RA, Gnegy, ME, Robinson, TE, Aragona, BJ, Society for Neuroscience Annual Meeting, 2015. Chicago, IL. "Individual variation in incentive salience attribution and accumbens dopamine transporter expression and function."

Singer, BF, Austin, CJ, Wohl, I, Guptaroy, B, Lovic, V, **Seiler, JL,** Bryan, MA, Gnegy, ME, Aragona, BJ, Robinson, TE, Society for Neuroscience Annual Meeting, 2014, Washington, DC. "Differential dopamine signaling in sign-tracker and goal-tracker rats following amphetamine."

Porter-Stransky, KA, **Seiler, JL,** Mabrouk, OS, Kennedy, RT, and Aragona, BJ, Society for Neuroscience Annual Meeting, 2013, San Diego, CA. "Phasic and tonic dopamine transmission dynamics in the nucleus accumbens during unexpected reward omission."

Porter-Stransky, KA, **Seiler, JL,** Klinger, MC, and Aragona, BJ, Society for Neuroscience Annual Meeting, 2012, New Orleans, LA. "Disrupting D2, but not D1, tone within the nucleus accumbens core impacts the rapid establishment of behavioral preferences resulting from unexpected reward omission."

Experimental Skills

Recording neural signals using fiber photometry
Optical manipulation of neuronal circuits (optogenetics)
Detection of *in vivo* dopamine release using fast-scan cyclic voltammetry (FSCV)
High efficiency detection of extracellular molecules using *in vivo* microdialysis
Stereotaxic surgeries; craniotomy, viral infusion, and implantation
Behavioral pharmacology
Intracranial self-stimulation
Immunohistochemistry and fluorescent microscopy
Operant and classical conditioning design
Vaginal lavaging and identification of estrous stage in female rats to monitor cycle

Other Skills

Planning and management of experiments
Critical thinking and synthesis of complex data sets
Dissemination of results to science community and broader public
Training and mentorship of undergraduate research assistants and post-bac students (see below)
Software proficiency: SPSS, GraphPad, Microsoft Office, Tarheel CV, MedPC, Arduino, MATLAB, Noldus Ethovision, ImageJ, Adobe Illustrator
Languages: English, conversational Spanish

Mentorship

Abigael Bridgemohan, undergraduate Honors thesis student	2019-2021
Currently: University College of London, Masters in Prenatal Genetics and Fetal Medicine	
Jack Lavey, undergraduate research assistant	2019-2021
Joseph Bianco, post-bac research assistant	2018-2021
Blake Berkowitz, summer research assistant	2018
Sophia Khan, undergraduate research assistant	2016
Cate Soane, undergraduate research assistant	2015-2017
Currently: Data analyst at United Airlines	
Melanie Schweir, undergraduate research assistant	2013-2014
Currently: Western Michigan University, OBGYN resident	
Daniel Saffei, undergraduate research assistant	2013-2014

Awards

Laboratory in Integrative Neuroscience (LIN) Presentation Award	2021
Professional Development & Technical Award	2021
LAS Travel Award	2014, 2016, 2019
University Honors	2010

Membership/Service

Faculty Search Committee Member	2020-present
Women in Learning	2020-present
Graduate Women in Science (GWIS)	2019-present
Expanding Your Horizons (EYH), Board Member	2016-present
Society for Neuroscience	2012-present
Females Excelling More in Math, Engineering, and Sciences	2012-2013
Order of Omega	2011-2013
Delta Phi Epsilon Sorority, Executive Board	2009-2013

References

Dr. Talia Lerner (graduate research advisor): talia.lerner@northwestern.edu | 312-503-5596
Dr. Dave Wirtshafter (graduate academic advisor): davew@uic.edu | 312-413-2631
Dr. Jamie Roitman (committee member): jroitman@uic.edu | 312-355-1458
Dr. Michael Ragozzino (committee member & department chair): mrago@uic.edu | 312-996-2836