Dopamine D2 Receptors Modulate the Cholinergic Pause and Flexible Learning

Kelly Marie Martyniuk

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2022

Abstract

Dopamine D2 Receptors Modulate the Cholinergic Pause and Flexible Learning

Kelly Marie Martyniuk

Animals respond to changes in the environment and internal states to modify their behavior. The basal ganglia, including the striatum contribute to action selection by integrating sensory, motor and reward information. Therefore, dysregulation of striatal function is common in many neuropsychiatric disorders, including Parkinson's disease, Huntington disease, schizophrenia, and addiction. Here, using fiber photometry, pharmacology, and behavioral approaches in transgenic mice, I explored the cellular and circuit mechanisms underlying key striatal functions. In **Chapter 1**, I begin by presenting the existing literature on the anatomy and physiology of the striatum. Next, I review the important functions of the striatum. Within this general review, I highlight the specific roles that striatal (DA) and acetylcholine (ACh) play in striatal circuitry and function. In **Chapter 2**, I demonstrate the naturally evoked ACh dip has a DA component and a non-DA component. Specifically, I show that DA via cholinergic DA D2 receptors (D2Rs) modulate the length of the ACh dip and rebound ACh levels following the dip. In addition, I show that DA coordinates the activity between DA and ACh during behavior. Finally, I present data that supports a role for ACh in motivated behavior. In **Chapter 3**, I show that cholinergic D2Rs are not necessary for reward learning but do facilitate reversal learning in a probabilistic choice task. In addition, I show that changes in DA and ACh levels contribute to reversal learning in a probabilistic choice task. Finally, in **Chapter 4**, I discuss the general conclusions and study implications, as well as future directions.

Table of Contents

List of Figures

Acknowledgments

This work would not have been possible without the help and support of many people. During my PhD, I had the opportunity to work with so many incredible scientists.

To my thesis mentor, Christoph Kellendonk, thank you for welcoming me into your lab and providing a supportive and inclusive environment to do this work. Your advice and guidance have been indispensable throughout my PhD journey. You taught me how to think critically, persevere through failure and rejection, and appreciate the small victories. For this, I am a better scientist. Thank you.

To my thesis committee (Christine Ann Denny, David Sulzer, Jonathan Javitch and Peter Balsam), thank you for all your guidance and keeping me on track throughout my PhD. Specifically, to Christine, thank you for your unconditional support both professionally and personally. You provided the motherload of advice from when I first interviewed at Columbia and throughout the last 5 years. Thank you.

To the current members (Pedro Olivetti, Arturo Torres-Herraez, Joshua Taliaferro, Joseph Vallarin, Julia Greenwald, and Daniel Lowes), and past members (Marie Labouesse, Laura Benoît, Troels Rahbek-Clemmensen, Eduardo Gallo, and Sarah Canetta), I have learned so much from all of you and thoroughly enjoyed working alongside each of you. To Marie and Troels, thank you both for welcoming into the lab and helping me get my initial project up and running. To Pedro, thank you for being my unofficial psychiatrist and supporting me throughout the years both scientifically and personally. To Arturo, thank for all your analytical help, you taught me how to code and eventually enjoy it, kind of. Thank you all.

To the NB&B program directors past and current (Carol, Ai, Darcy, Wes, Aniruddha, and Ken) and the administrators (Rozanna, Alla, and Sarah), thank you for welcoming me into the

incredible Columbia community. To Wes, your advice and support helped me through difficult times, and I am thankful for your kindness and confidence. The cone of silence will never be broken. Thank you.

To my undergraduate mentor, Brenda Bloodgood, I decided to pursue a PhD in neuroscience because of you. Thank you for your inspiration and guidance over the years. To past members of the Bloodgood lab (Andrea Hartzell and Laura Sancho), thank you for instilling in me a love and passion for science.

To my friends and colleagues that I met through Columbia (Luke, Avery, Josie, and Bri), This journey would not be possible without your support and a lot less fun. I cherish you deeply.

To my family (Dad, Deanna, Grandma May, Grandpa Neil, Aunt Anna, Uncle Tom, Brooke, Brian), thank you for cheering me on from afar. To my parents-not-parents, Tom and Wendy Ashe, thank you for your unconditional love and support throughout this journey. To my friends in New York (Macayla, Erin, Joey, and Colleen), thank you for making my time in New York enjoyable and memorable. To my best friends since high school (Jackie, Paris, Brittany, and Collin), I love you weirdos. Thank you. To my best friend and partner, Matt, thank you for being my biggest fan and supporting me through the final stages of my PhD. To my pets, Whiskey and Bella, thank you for providing me with comfort and joy throughout this process. Thank you.

To all my colleagues, friends, family, mentors, and mentees near and far, I am eternally grateful for the support that you provided throughout the years. Thank you.

Dedication

I dedicate this thesis work to the unsung heroes: the mice. Without them none of this work would be possible. Thank you, mice.

Chapter 1: Introduction

The striatum is the primary input of the basal ganglia and is intimately involved in action selection, motor learning, motivation, reinforcement learning, behavioral flexibility, and reward processing (Aoki et al., 2015; Aosaki et al., 1994a; Bradfield et al., 2013; Cachope et al., 2012; Carvalho Poyraz et al., 2016; Gallo et al., 2018; Gerfen & Surmeier, 2011; Joshua et al., 2008; Kellendonk et al., 2006; Schultz et al., 1997). Given this, dysregulation of striatal function is common in several neurological disorders including, Parkinson's disease, Huntington disease, obsessive compulsive disorder and addiction (Albin et al., 1989; DeLong, 1990; Graybiel, 2000, 2008; Hyman et al., 2006). Therefore, examining striatal physiology is crucial for understanding normal striatal function and in disease. Here, I will describe the anatomy, physiology, and primary functions of the striatum.

1.1 Anatomy and physiology of the striatum

1.1.1 Regions of the striatum

As the primary input into the basal ganglia, the striatum receives excitatory glutamatergic input from various cortical and thalamic regions and dopaminergic inputs from the midbrain (Figure 1.1A) (Bolam et al., 2000; Kincaid et al., 1998; Smith et al., 1994). In primates, the dorsal striatum is separated by the internal capsule into two distinct nuclei, the caudate and the putamen. Like primates, rodents have two distinct dorsal striatal regions, the dorsal medial and dorsal lateral, which are anatomically and functionally different from each other (Joel & Weiner, 1994; Parent & Hazrati, 1995; Yin & Knowlton, 2006). A third region, the ventral striatum (nucleus accumbens) has distinct properties and inputs from the dorsal striatum (Figure 1.1B) (Nicola, 2007).

1.1.2 Striatal neurons

Classically, striatal neurons have been categorized into two main groups: spiny projections neurons (SPNs) and aspiny interneurons. SPNs constitute ~95% of all striatal neurons, are GABAergic and project out of the striatum (DiFiglia et al., 1976; Kemp & Powell, 1971; Kita & Kitai, 1988; Liles, 1974; Precht & Yoshida, 1971; Preston et al., 1980; Yoshida & Precht, 1971). The remaining 5% of striatal neurons consist of large aspiny cholinergic interneurons (CINs) and various GABAergic interneurons (Kawaguchi, 1993; Kawaguchi et al., 1995). Here, I will further discuss the properties of SPNs and CINs.

Striatal projection neurons (SPNs)

SPNs are the principle striatal neurons. While SPNs have a similar morphology, they can be segregated into two categories classified by their axonal projection targets and expression of neuronal markers. Direct pathway SPNs (dSPNs) project monosynaptically to the output nuclei of the basal ganglia. Indirect pathway SPNs (iSPNs) send their axons to the external segment of the globus pallidus (GPe), which is disynatpically connected to the output nuclei of the basal ganglia (Figure 1.1A) (Smith et al., 1998). In addition to their divergent projection pathways, SPNs also express different neurochemical markers. dSPNs contain dopamine (DA) D1 receptors, while iSPNs express DA D2 receptors (Gerfen et al., 1990; Surmeier et al., 1996). DA receptors are G-protein coupled receptors, but they can be divided into two main categories depending on the activation of downstream targets (Kebabian & Calne, 1979). D1/5 DA receptors are $G_{\alpha s}$- or $G_{\alpha(olf)}$-coupled and are positively linked to adenylyl cyclase, while D2 receptors are $G_{\alpha i}$-coupled and are negatively linked to adenylyl cyclase (Corvol et al., 2001; Hervé et al., 2001; Hervé et al., 1993). Due to this difference in DA receptor expression, DA differentially modulates SPN activity. Local application of DA excites dSPNs and inhibits iSPNs (Gerfen & Surmeier, 2011). In addition, DA

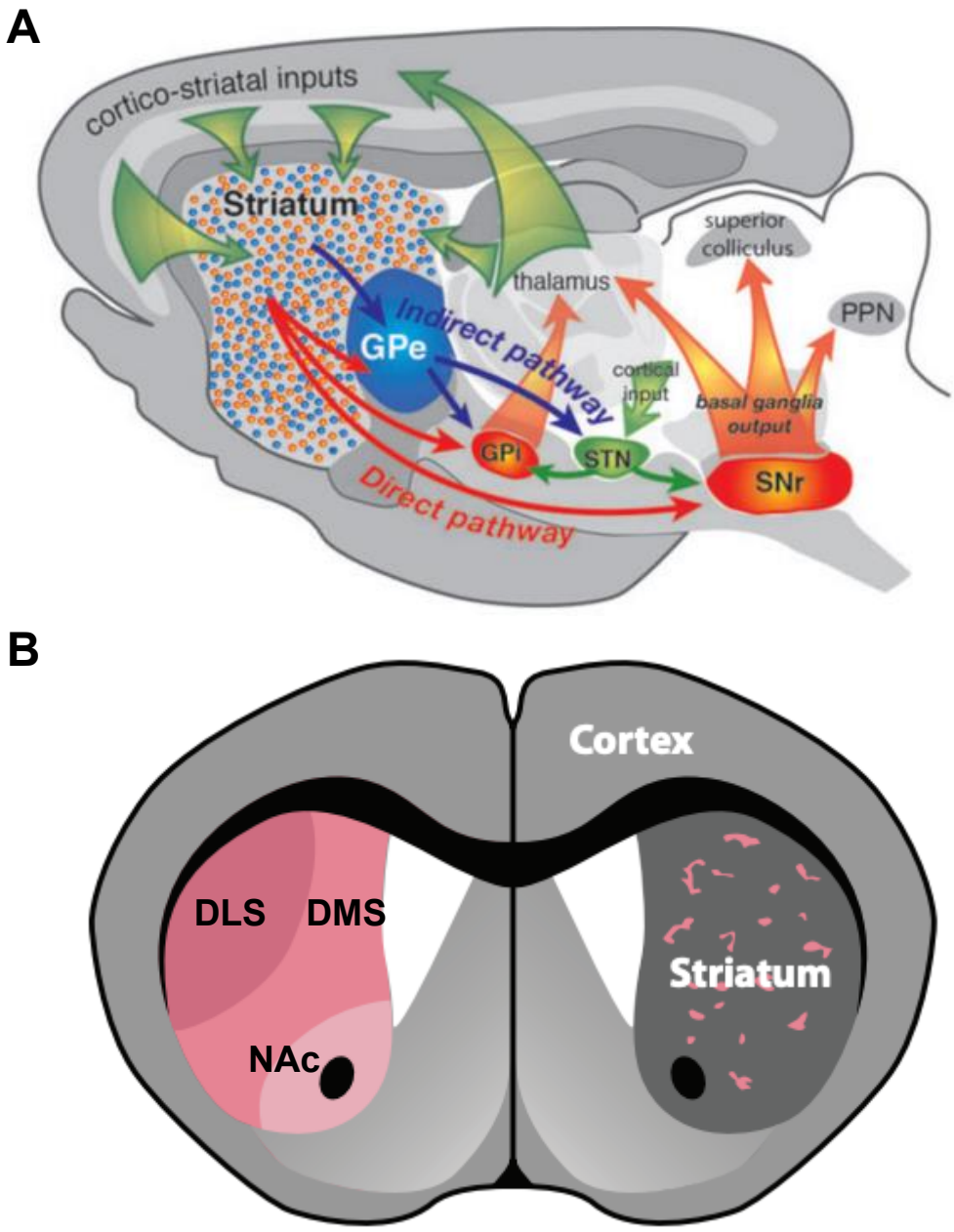

Figure 1.1. Diagram of basal ganglia circuits and regions of the striatum. (A) Diagram of basal ganglia circuits. The striatum receives excitatory inputs from the cortex and thalamus. The striatum also receives dopaminergic input from the midbrain (not shown). Spiny projection neurons (SPNs) differentially project out of the striatum to the output of the basal ganglia. dSPNs (red) that express DA D1 receptors project directly to the SNr, while iSPNs (blue) that express DA D2 receptors, project to the GPe. Adapted from (Gerfen & Surmeier, 2011). **(B)** Diagram of the three striatal regions. The dorsal striatum is divided into two regions, the dorsal medial striatum (DMS) and the dorsal lateral striatum (DLS). The third region is the ventral striatum or nucleus accumbens (NAc). Adapted from (Kreitzer, 2009)

inhibits presynaptic glutamate release via G_i-coupled D2 receptors on corticostriatal projections (Bamford et al., 2004). Finally, SPNs also express different neuropeptides and genes, further segregating their properties and potentially specific function in the striatum.

Cholinergic interneurons (CINs)

CINs represent 1-2% of all striatal neurons, however, their influence is substantial. CINs have large somata (20-50 μm), long-ranging dendritic arborizations and extensive axonal fields (Kawaguchi et al., 1995; Wilson CJ, 1990). CINs are spontaneously active and respond to reward predicting cues and salient outcomes (Apicella et al., 1997; Goldberg & Reynolds, 2011). Their response is typically triphasic, starting with initial burst firing followed by a prolonged pause and a rebound in burst firing. Given their widespread reach and unique firing pattern, CINs are poised to modulate the striatal microcircuitry and overall striatal function.

A key feature of CINs is their pause in tonic firing. Classic studies of "Tonically Active Neurons" (TANs), the presumed counterparts of CINs in non-human primates (NHPs), demonstrate that TANs will pause in firing *in vivo* following a salient cue. When the salient cue is not paired with a reward ~15% of TANs will pause, however, when the salient cue is repeatedly paired with a reward 60-70% of TANs will pause their firing (Figure 1.2A) (Aosaki et al., 1994a). What are the mechanisms underlying the CIN pause? Original studies done in NHPs, demonstrate that the CIN pause is DA dependent. DA lesions and local administration of a D2R antagonist completely abolished the pause (Figure 1.2B) (Aosaki et al., 1994b; Watanabe & Kimura, 1998). This is consistent with more recent studies done in rodents demonstrating that CIN pauses can be induced by local exposure to DA or by stimulation of DA terminals, and are both abolished by D2R antagonists (Augustin et al., 2018; Chuhma et al., 2014; Straub et al., 2014; Wieland et al.,

2014). Furthermore, CIN pauses that are induced by DA or by local stimulation of the striatum are eliminated in a selective CIN D2R knockout mouse (Augustin et al., 2018; Kharkwal et al., 2016).

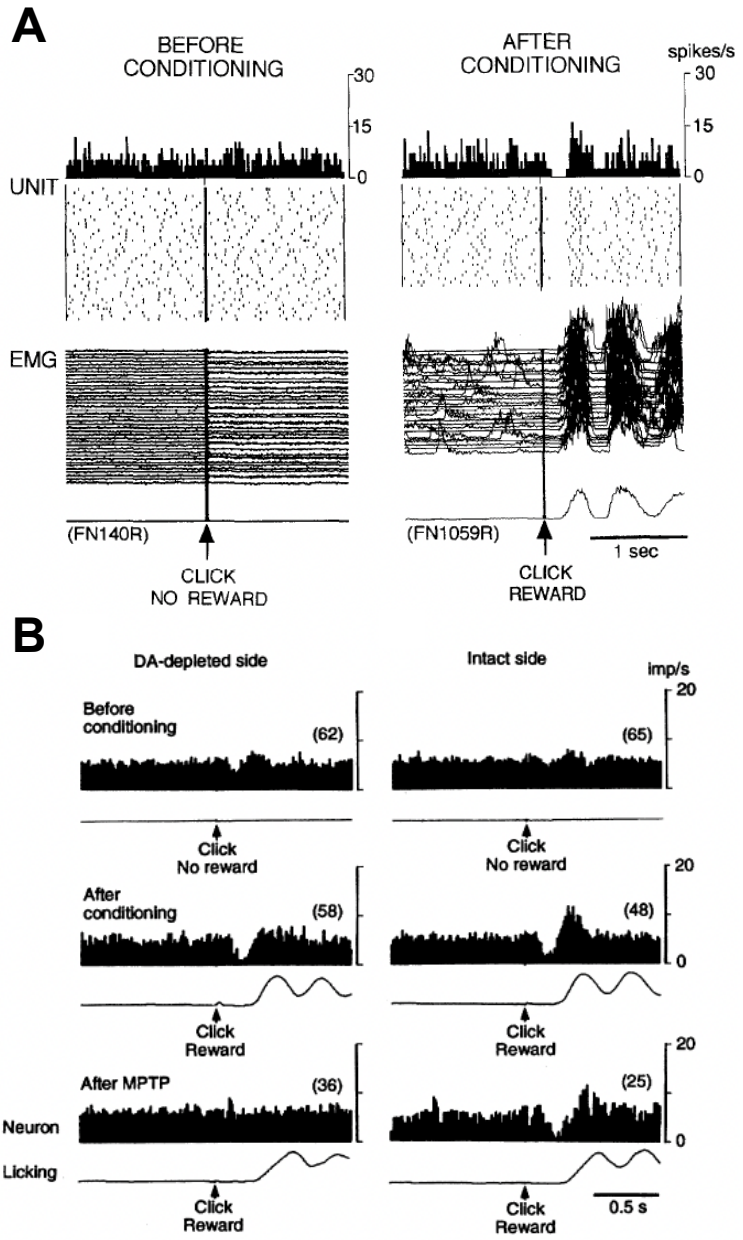

Figure 1.2. Pauses in TAN activity are induced by salient cues and are DA dependent. (A) Single unit recordings of TANs during a behavioral task before (left) and after (right) conditioning. TANs pause their firing to the click after the learned association of the click with reward. EMG traces show increased activity of tongue muscles after conditioning but not before. Adapted from (Aosaki et al., 1994a). **(B)** Population activity of TANs recorded before (top) and after (middle) conditioning, and after MPTP infusion (bottom left) and on the intact side (bottom right). Conditioned pauses in TAN activity are abolished after DA lesions. Adapted from (Aosaki et al., 1994b)

Finally, the CIN pause coincides with changes in DA activity in response to reward-predicting cues (see below) (Schultz, 2007; Schultz et al., 1997). Together, these studies support a role for DA in regulating the CIN pause.

Other studies suggest that the CIN pause is driven by excitatory inputs arising from the thalamus or cortex or by long-range GABAergic projections (Brown et al., 2012; Cover et al., 2019; Ding et al., 2010; Doig et al., 2014; English et al., 2012b; Matsumoto et al., 2000; Zhang et al., 2018). Non-selective stimulation of thalamic inputs to the striatum leads to burst firing in striatal CINs followed by a pause that is blocked by a D2R antagonist (Ding et al., 2010). Consistent with this, selective stimulation of intralaminar nuclei of the thalamus terminals in the striatum causes CIN burst firing followed by a prominent pause. However, the CIN pause was not entirely abolished by D2R antagonism (Cover et al., 2019). Here, the mechanism proposed is that glutamatergic excitation of CINs leads to burst firing, due to their expression of HCN and sodium channels. (Bennett et al., 2000; Jiang & North, 1991). Following firing, CINs go into a pronounced after hyperpolarization state due to activation of calcium channels during the action potential, which triggers a potassium current (I_{Kr}) via activation of inward rectifying potassium (K_v7) channels (Goldberg & Wilson, 2005; Kawaguchi, 1993; Reynolds et al., 2004; Wilson CJ, 1990; Wilson & Goldberg, 2006; Zhang et al., 2018). Together, these studies suggest that the initiation of the CIN pause is driven by excitatory thalamic or cortical inputs.

The cause for this discrepancy could be due to several factors, including the species or techniques used. Notably, the original primate studies showing DA dependence looked at the naturally evoked TAN pauses, while the rodent studies that used stimulation to evoke CIN pauses suggest DA independence. Therefore, further exploration of the mechanisms underlying the naturally evoked CIN pause should provide clarity on this question.

1.1.3 Striatal neuromodulators

The two main neuromodulators in the striatum are dopamine (DA) and acetylcholine (ACh). Both DA and ACh are known for their involvement in key striatal functions including movement and reward processing. Notably, there is growing evidence that DA and ACh regulate the release of each other within the striatum. Here, I will describe each of their key targets and how they regulate striatal circuity and function.

Dopamine (DA)

DA plays an instrumental role in learning, acting as a teaching signal that reflects reward prediction error (Jermey J. Day et al., 2007; Ali Mohebi et al., 2019; Nasser et al., 2017; Schultz et al., 1997; Elizabeth E. Steinberg et al., 2013). This teaching function is encoded by phasic bursting of DA neurons, which induces a rapid but transient increase of extracellular DA. Initially, DA is released in response to an unpredicted reward, but over learning the response shifts away

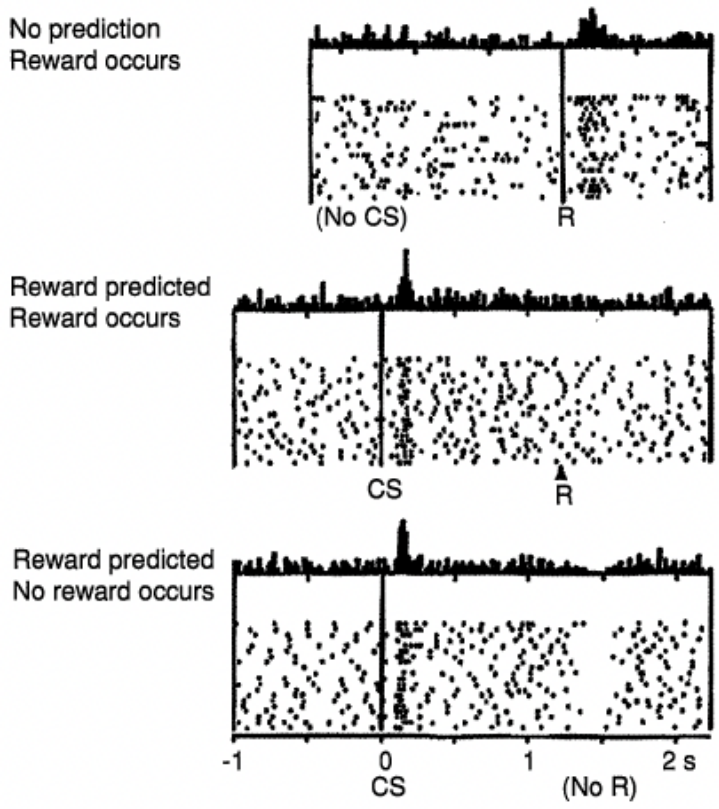

Figure 1.3. DA neurons encode reward prediction errors (RPEs). Before learning (top), an unpredicted reward enhances the activity of DA neurons. After learning (middle), the reward predicting cue elicits an increase in DA firing, while the now expected reward does not. After learning (bottom), unexpected reward omission depresses DA activity at the time when the reward was expected. Adapted from (Schultz et al., 1997).

from the reward outcome and towards the reward predicting cue. In the case of unexpected reward omission, DA neurons are depressed at the exact time the reward was expected (Figure 1.3) (Schultz, 2007; Schultz et al., 1997). On a longer timescale, DA encodes the availability of rewards for specific actions highlighting a role for DA in action selection (Schultz, 2007).

DA neurons originating in the ventral midbrain terminate in the striatum. Dopaminergic inputs from the substania nigra pars compacta (SNc) terminate in the dorsal striatum (Prensa & Parent, 2001), while DA neurons from the ventral tegmental area (VTA) innervate the ventral striatum (Fields et al., 2007). DA modulates striatal circuity and function via DA receptors located on all striatal neurons; however, different neurons express different types of DA receptors. Striatal projection neurons (SPNs) differentially express DA receptors: dSPNs express D1 receptors while iSPNs express D2 receptors, as previously described. GABAergic interneurons primarily express D5 receptors (Centonze et al., 2003; Rivera et al., 2002). Cholinergic interneurons (CINs) express both D2 and D5 receptors (Yan et al., 1997; Yan & Surmeier, 1997), which I will discuss further.

DA inhibits the activity of CINs via DA D2 receptors. Stimulation of DA terminals in the striatum, induces pauses in tonically active CINs that are blocked by D2 antagonists (Chuhma et al., 2014; Straub et al., 2014). DA mediated CIN pauses are also eliminated in a selective D2 receptor knockout mouse (Kharkwal et al., 2016). Furthermore, overexpression of D2 receptors in CINs extends the pause resulting in a deficit in inhibitory learning (Gallo et al., 2021).

Acetylcholine (ACh)

Like DA neurons, CINs modulate their activity in response to reward predicting cues and reward outcomes (see above, Figure 1.2). Tonically active CINs release ACh into the extracellular space. CINs only account for 1-2% of all striatal neurons, however, their extensive axonal fields enable widespread ACh release, which can act locally or have a more expansive modulation via

volume transmission (Bolam et al., 1984; Contant et al., 1996). CINs provide the majority of the ACh within the striatum; however, it is rapidly degraded by extracellular acetylcholinesterase, an enzyme that limits ACh diffusion. ACh regulates striatal neurons via both muscarinic (mAChR) and nicotinic (nAChR) receptors. mAChRs are located on SPNs, CINs, GABAergic interneurons and cortical/thalamic afferents, while nAChRs are primarily expressed on DA terminals and GABAergic interneurons.

The corticostriatal and thalamostriatal circuits that lead to excitation of striatal SPNs is modulated by ACh release (Bolam et al., 2000; Ding et al., 2008; Doig et al., 2010; P. Pakhotin & E. Bracci, 2007; Smith et al., 2004). Tonic release of ACh rapidly inhibits glutamate release from cortical and thalamic inputs via M2-class mAChRs, thereby reducing the excitation onto SPNs (Akaike et al., 1988; Barral et al., 1999; Briggs et al., 1981; Malenka & Kocsis, 1988). At the same time ACh acts directly on SPNs by potentiating their response to glutamate via M1-class mAChR (Ding et al., 2010). Therefore, it is suggested that presynaptic inhibition by ACh serves to tune excitation of these synapses while also priming SPNs to be more responsive to glutamatergic excitation. This mechanism supports a role for ACh in regulating corticostriatal and corticothalamic circuits underlying action selection (for an more in depth review, see (Goldberg et al., 2012). In addition, ACh excites striatal interneurons via nAChRs activation leading to inhibition of SPNs (English et al., 2012a; T. W. Faust et al., 2015; Koós & Tepper, 2002; Witten et al., 2010).

ACh modulates striatal DA release via nAChRs, which are expressed on DA terminals. Stimulation of CINs in both acute brain slices and *in vivo* induces DA release in the striatum (Cachope & Cheer, 2014; Cachope et al., 2012; Cragg, 2006; David Sulzer et al., 2016; Threlfell et al., 2012). However, activation of nAChRs is complex. In acute brain slices, transient loss of

striatal ACh release reduces DA release following electrical stimulation but enhances DA release in response to DA burst firing (Zhou et al., 2001). Inhibition of acetylcholinesterase (AChE), the enzyme that breaks down extracellular ACh, also reduces DA release (Zhou et al., 2001). This result demonstrates excess ACh causes desensitization of nAChRs on DA terminals. Moreover, the effect of nAChR activation has been shown to be dependent on the frequency of DA neuron firing. In response to a single action potential, nAChR activation reduces DA release but enhances DA release during phasic burst firing (Rice & Cragg, 2004; Zhang & Sulzer, 2004; Zhou et al., 2001). This is consistent with *in vivo* studies showing that nicotine at high concentrations will desensitize nAChRs, leading to enhanced DA release when DA neurons are burst firing (Pontieri et al., 1996). Furthermore, synchronized ACh release from CINs can evoke DA release via axo-axonic activation of nAChRs located on DA terminals (Cachope et al., 2012; Threlfell et al., 2012). In addition, ACh enhances DA release in the striatum via M5 mAChRs on DA terminals (Shin et al., 2015). Taken together, these studies demonstrate that ACh directly affects DA release and acts as a frequency-dependent filter on DA release and provides insight on the delicate balance between DA and ACh signaling in the striatum.

Both DA and ACh signals respond to reward outcomes and reward predicting cues and increasing evidence suggests that they co-regulate each other (Cachope & Cheer, 2014; Cachope et al., 2012; Chuhma et al., 2014; Cragg, 2006; Helseth AR, 2021; Kharkwal et al., 2016; Straub et al., 2014; David Sulzer et al., 2016; Threlfell et al., 2012; Yan & Surmeier, 1991). Therefore, further studies that can simultaneously measure task evoked changes in striatal DA and ACh levels during behavior are needed to determine if/how this coregulation drives learning and behavior. This type of experiment is now possible with recent technology advancements in imaging that

allow for DA and ACh levels to be monitored in awake behaving mice (Labouesse & Patriarchi, 2021).

1.2 Functions of the striatum

As the primary input into the basal ganglia, the striatum receives sensory and motor information from the cortex and thalamus as well as modulatory input from dopamine afferents. Using this information the striatum has been shown to be important action selection, motor learning, motivation, reinforcement learning, behavioral flexibility, and reward processing (Aoki et al., 2015; Aosaki et al., 1994a; Bradfield et al., 2013; Cachope et al., 2012; Carvalho Poyraz et al., 2016; Gallo et al., 2018; Gerfen & Surmeier, 2011; Joshua et al., 2008; Kellendonk et al., 2006; Schultz et al., 1997). Here, I will further discuss the striatum's role in motivation/reinforcement learning and behavioral flexibility.

1.2.1 Motivation and reinforcement learning

The striatum regulates motivated behaviors (Carvalho Poyraz et al., 2016; Corbit et al., 2001; Gallo et al., 2018; Lex & Hauber, 2010; Mai et al., 2012; Nowend et al., 2001; Tai et al., 2012; Yin et al., 2005a; Yin et al., 2005b). Motivation can be defined as the process that initiates, guides, and maintains goal-directed behaviors. This process includes a "directional" component, engaging animals to select a behavior that leads to a desirable outcome, which is the basis of reinforcement learning. The "activational" component initiates and sustains the vigor and persistence of actions and depends motivational drive, which incorporates both internal state of the animal and its environmental context (Salamone, 1998; Simpson & Balsam, 2016). Deficits in motivation is common in many psychiatric disorders, including schizophrenia and addiction. Therefore, we can use reinforcement learning as a model to better understand the circuits

underlying motivation in health and in disease. Here, I will describe the key aspects of reinforcement learning.

Reinforcement learning has several different components, allowing the dissection of each component when animals engage in reward-seeking behaviors. Initially, animals must learn that specific actions lead to rewarding outcomes. To do this, animals perform a cost-benefit analysis. The costs can include physical or mental effort, time, or loss of potential beneficial outcomes, while benefits can include the value of reward. Therefore, the value of each cost and each benefit must be computed and encoded. Importantly, the value of reward must be encoded when a goal is achieved and stored for retrieval in the future when achieving that goal becomes necessary again. This is crucial for reinforcement learning where the value of reward is dependent on both prior experience as well as the current state and environmental context. After completion of a cost-benefit analysis, animals learn the necessary behavioral direction and vigor of action needed to repeatedly obtain a rewarding outcome (Simpson & Balsam, 2016).

Dopamine (DA) plays a key role in reinforcement learning. In the 1980's, Wolfram Schultz and colleagues performed single-unit recordings of midbrain DA neurons in primates during an instrumental task involving reach movements for a food reward following auditory and visual cues (Aebischer & Schultz, 1984; Schultz, 1986). These studies report an increase in phasic activity in ~60% of recorded striatal projecting DA neurons that was time-locked to the presentation of the auditory and visual cues. This increase in phasic DA activity also occurred before arm movement and prior to any EMG activity in the arm muscle. Moreover, DA responses were reduced when the visual cue was removed. Less than 1% of recorded DA neurons decreased their activity to the presentation of auditory and visual cues. About 30% of recorded DA neurons increased their phasic activity during reaching movements, however this change in DA activity was slow and less robust

compared to the changes in response to the auditory and visual cues, suggesting that the activity of DA neurons does not report specific movements. In contrast, this is substantial literature showing that changes in DA activity tracks with movement and is important for movement initiation and action selection (for an in depth review, see (Klaus et al., 2019). Later in the 1990's, Schultz and colleagues showed that midbrain DA neurons respond to unpredicted rewards, reward-predicting cues and unexpected reward omission (see above, Figure 1.3) (Schultz et al., 1997). Together, these studies provide evidence that DA is involved in reward prediction error, a model of reinforcement learning, which predicts that phasic DA release into the striatum encodes a discrepancy between the prediction and expectation of a reward outcome (Schultz, 1998).

Like DA neurons, striatal cholinergic interneurons (CINs) are also engaged during reinforcement learning. In the 1980's and 1990's, researchers showed that tonically active neurons (TANs), the presumed counterpart of CINs, respond to conditioned sensory stimuli that signal reward and evoke behavioral reactions in highly trained primates, while not specifically responding to body movements (Apicella et al., 1991; Kimura, 1986, 1992; Kimura et al., 1984). Aosaki and colleagues extracellularly recorded the activity of striatal TANs in primates undergoing classical Pavlovian conditioning, where a click sound was paired with a reward. This study's goals were to determine the activity of TANs during the acquisition phase and to test the maintenance and plasticity of the TANs response over training. Before conditioning, only 17% of recorded TANs responded to the click reward. During training, the number of TANs responding increased and by the end more than 50-70% of recorded TANs responded to the click reward. The responses included a pause in firing ~90 msec after the click and in some cells the pause was proceeded by a brief excitation and in most cells the pause was followed by a rebound excitation (see above, Figure 1.2). These responses were maintained over training and remained even after a

4 week break in training. Moreover, when they trained the primates to receive rewards to a new conditioned stimulus, the recorded TANs were able to switch their response to the new stimulus, indicating the flexibility of TAN responses. These results demonstrate that TANs are actively recruited during the acquisition phase of a reinforcement task as they respond to a reward-paired sensory stimuli that induces a conditioned behavior (Aosaki et al., 1994a). This responsiveness in TANs might modulate the activity of surrounding striatal projection neurons that are engaged in learned sensorimotor behaviors.

Both DA neurons and CINs are engaged in reinforcement learning and changes in their activity are time-locked to each other. However, it remains unknown if they are encoding overlapping or distinct aspects of reinforcement learning. Therefore, further studies that can monitor the activity of DA neurons and CINs at the same time during learning will be crucial for answering this question. Morris and colleagues first developed this approach by conducting single-unit recordings from striatal TANs and midbrain DA neurons in primates while they were subjected to a probabilistic instrumental conditioning task. They found that both DA neurons and TANs responded to reward-related events, including salient cues and reward outcomes. However, only DA neurons and not TANs responded to reward predictability (Morris et al., 2004). This result indicates that while DA neuron and TAN activity is coincident, they encode distinct properties underlying reinforcement learning. This is consistent with DA neurons and TANs encoding differences between reward and aversive events at distinct time points (Joshua et al., 2008). However, it is yet to be determined if they mutually co-regulate each other during reinforced behaviors and whether the change in TAN/CIN activity is necessary for learning; therefore, manipulation of DA and ACh signaling is necessary to answer this question.

1.2.2 Behavioral flexibility

Behavioral flexibility can be defined as the ability to shift response patterns or strategies when there is a change in environmental contingencies. In addition to regions of the frontal cortex, the basal ganglia and specifically the dorsal medial striatum (DMS) has been implicated for its role in behavioral flexibility (Clarke et al., 2008; Ragozzino, 2007; Ragozzino et al., 2002). In addition, several neurochemicals have been implicated for the role in reversal learning, including serotonin, DA, ACh and glutamate (Izquierdo et al., 2017). To study behavioral flexibility in humans, primates and rodents, researchers utilize a set of paradigms collectivity called reversal learning. Reversal learning requires a subject to flexibly shift their behavior when previously learned reward-related contingencies are reversed. Therefore, reversal learning can be used to assess behavioral flexibility when stimulus-outcomes (S-O) or response-outcomes (R-O) change (Izquierdo et al., 2017). Moreover, alterations in striatal dopamine (DA) and acetylcholine (ACh) are common in many neuropsychiatric disorders, including schizophrenia and Parkinson's disease where patients are prone to deficits in cognitive flexibility (Bohnen & Albin, 2009; Howes et al., 2012; Weinstein et al., 2017). Here, I will discuss the studies that implicate the DMS in behavioral flexibility and highlight how DA and ACh regulate reversal learning.

Neuroimaging studies done in humans show the recruitment of both the dorsal (Rogers et al., 2000) and ventral (Cools et al., 2002) regions of striatum during reversal learning, while basal ganglia lesions are linked to impairments in reversal learning. For example, neurotoxic lesions of the dorsal striatum (DMS) impair reversal learning in primates and rats (Braun & Hauber, 2011; Castañé et al., 2010; Clarke et al., 2008; Ragozzino, 2007). The DMS receives input from prefrontal cortex, including the orbitofrontal and prelimbic cortices, which both contribute to behavioral flexibility (Berendse et al., 1992; Dias et al., 1997; Jones & Mishkin, 1972; Nonneman

et al., 1974; Ragozzino, 2007). Several studies have examined how inactivation of the DMS affects acquisition and reversal learning. In one study, experimenters used 2% tetracaine (TET) to inactivate the DMS and tested rat's performance during a response and visual discrimination task. In each task, rats had the choice to make a right or left turn and a visual cue was randomly place in one of the choice arms. In the response task, once a choice was made to turn left or right the rat must continue to turn in the same direction to receive a food reward regardless of the visual cue. In the visual task, rats had to enter the visual cue arm and continue to turn to the visual cue arm regardless of turning in the same direction to receive a food reward. TET inactivation of the DMS did not impair the acquisition of either the response or visual task strategy. However, DMS inactivation did impair a shift between the response and visual strategies. Further analysis of the error pattern revealed that DMS inactivation did not increase preservation errors, but rather caused an increase in regressive errors, indicating an inability to adapt a new strategy (Ragozzino et al., 2002). This finding is consistent with other studies where DMS inactivation and lesions of the DMS cause an impairment in reversal learning (Kolb, 1977; Pisa & Cyr, 1990; Ragozzino & Choi, 2004).

Striatal dopamine (DA) neurons and tonically active neurons (TANs) play a key role in reinforcement learning (see above). Are both DA and ACh also involved in behavioral flexibility? Studies in humans, primates and rodents have demonstrated DA's role in reversal learning. In humans, methylphenidate, a drug commonly used to treat attention deficit disorder (ADD) and attention deficit hyperactivity disorder (ADHD), increases extracellular DA levels via blocking the DA transporter, DAT, and improves reversal learning (Clatworthy et al., 2009). Consistent with this, stimulation of DA neurons in the substania nigra pars compacta (SNc) or ventral tegmental area (VTA) improves performance in spatial reversal learning in rats (Adamantidis et

al., 2011; Rossi et al., 2013), while DA depletion in the striatum causes impairments in reversal but not acquisition learning in marmosets (Clarke et al., 2011). Together, these studies suggest that DA signaling, and the availability of striatal DA optimize reversal learning.

As previously described, DA modulates striatal circuits via different DA receptor subtypes and each subtype has been shown to have different roles in regulating reversal learning. First, DA D2 receptor (D2R) agonists and antagonists impair reversal learning in humans (Mehta et al., 2001), primates (Lee et al., 2007; Smith et al., 1999) and rodents (Boulougouris et al., 2009). This is consistent with studies showing that non-specific D2R ablation in mice also impairs performance in reversal learning in mice (De Steno & Schmauss, 2009; Kruzich & Grandy, 2004; Kruzich et al., 2006). Of note, these D2R ablation studies were not cell-type specific, therefore, studies that selectively delete D2Rs from specific striatal neurons will reveal the role of each D2R expressing cell type in reversal learning. Moreover, low D2R availability correlates with impaired reversal learning in humans (Jocham et al., 2009), primates (Groman et al., 2011) and rodents (Laughlin et al., 2011). These studies indicate that reversal learning relies on the optimal balance of D2R function. In contrast, antagonism of DA D1 receptors (D1Rs) does not impair reversal learning in primates (Lee et al., 2007).

Like DA neurons, changes in TAN/CIN activity are associated with initial learning in a reinforcement task (see above) (Joshua et al., 2008; Morris et al., 2004). Despite this, pharmacological and CIN-specific lesion studies in the dorsomedial striatum (DMS) and ventral striatum (NAc) does not reveal a role for striatal ACh during initial learning, but rather impaired performance when task contingencies change (Aoki et al., 2015; Bradfield et al., 2013; Brown et al., 2010; Okada et al., 2014; Okada et al., 2017; Ragozzino et al., 2009). This is consistent with a study where genetic inactivation of the vesicular acetylcholine transporter (vAChT), which

eliminates ACh storage and release in the striatum, causes an impairment in reversal but not acquisition learning (Favier et al., 2020). Several hypotheses have been proposed to explain why CIN activity is sensitive to changes in task contingencies that guide behavioral flexibility. First, one group proposed that CINs "track beliefs of the current state" (Stalnaker et al., 2016). Another group proposed that CIN activity is important for decreasing the interference between old and new memories, which is important for switching actions when task contingencies change (Bradfield et al., 2013).

CIN activity is multiphasic in response to salient cues and reward outcomes (see above). This multiphasic response might differentially regulate the activity of striatal projection neurons (SPNs) and guide animals to stop current actions and explore alternative actions. Researchers have shown that the initial excitation and rebound excitation in the CIN response inhibits SPN activity, however, the consequences of the CIN pause on SPN activity depends on the duration of the pause. Using optogenetic tools to vary the length of the CIN pause, these studies have demonstrated that length of CIN pauses differentially modulates SPN activity. Short pauses (< 500 msec) have a minimal effect on SPN activity, medium pauses (> 500 msec) moderately inhibit SPN activity, while longer pauses (15 sec) strongly disinhibit SPN activity (English et al., 2012b; Witten et al., 2010; Zucca et al., 2018). Therefore, this discrepancy in CIN pause length on regulating SPN activity could indicate that shorter and longer pauses may have different effects on flexible learning. This hypothesis was supported by computational modeling. In this model, varying the lengths of CIN pauses had no effect on acquisition learning but did effect reversal learning (Franklin & Frank, 2015). While shorter CIN pauses facilitate quicker reversal learning, longer pauses slows down reversal learning (Figure 1.4). This model supports the hypothesis that shorter CIN pauses results in decreased activity-dependent plasticity of corticocortical projections, which

allows for the flexibility to explore alternative actions in future trials. In contrast, longer CIN pauses cause stronger disinhibition of SPNs via increased activity-dependent plasticity of corticocortical projections, supporting the preservation of current actions while inhibiting exploration of alternative actions. Therefore, future studies that can manipulate the length of CIN pauses during behavior flexibility tasks will determine if varying pause lengths does affect reversal learning. Moreover, both striatal DA and ACh have been shown to regulate flexible learning, however, it has yet to be determined if they are co-regulating each other or acting independently

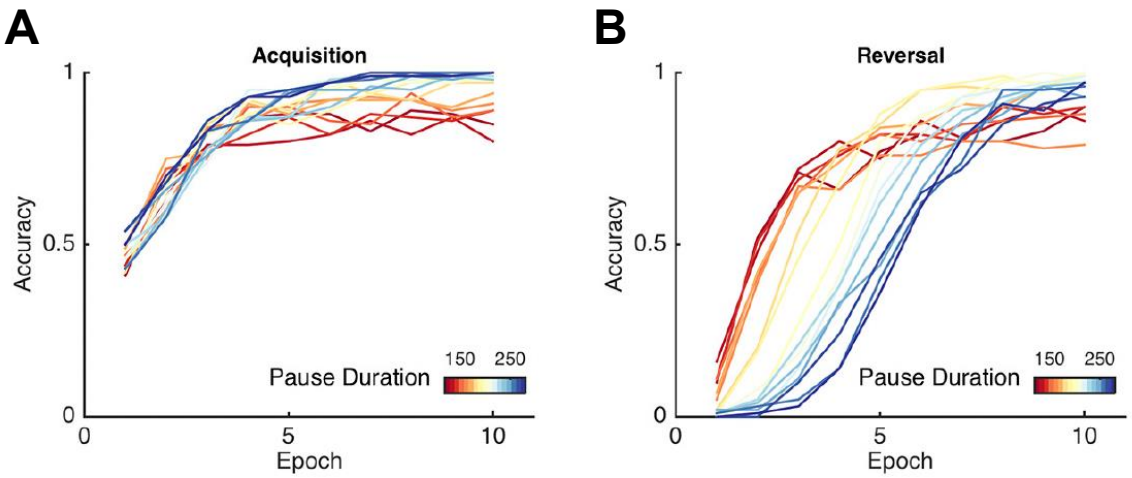

Figure 1.4. Computational modeling of CIN pauses predicts performance during reversal learning. (A) Varying CIN pause duration does not affect acquisition learning. **(B)** Shorter CIN pauses enhances reversal learning, while longer pauses slows down reversal learning. Adapted from (Franklin & Frank, 2015).

of each other to guide behavioral flexibility.

1.3 Conclusions

There is growing evidence that striatal dopamine (DA) and acetylcholine (ACh) mutually co-regulate each other. In this chapter, I have described the extensive research showing that DA and ACh signals overlap during key striatal-dependent functions, including reinforcement and flexible learning. I have also outlined the mechanisms by which DA and ACh regulate they release of each other. However, several questions remain. First, do striatal DA and ACh co-regulate each other to promote learning? Second is the multiphasic response of CINs entirely regulated by DA

via its receptors on CINs? This question remains open and controversial as the literature supports a role for DA as well other striatal inputs, including thalamus and cortex in regulating the CIN pause. Third, what are the behavioral consequences of varying CIN pause durations? Future work will need to explore the mechanisms underlying different pause lengths and how learning is affected. Fourth, how does varying CIN pause lengths affect DA signaling? While the literature has described the effect of CIN pause lengths on corticostriatal projection plasticity, it remains unknown how DA signaling is affected. Answering these questions will be vital to understanding how co-regulation of DA and ACh signaling affects striatal function and learning. Ultimately, this work could enhance our understanding of several neuropsychiatric disorders and provide a framework for better therapeutic strategies.

This thesis describes work done to address some of these key questions. First, a new imaging strategy to simultaneously monitor the activity of striatal DA and ACh during behavior was used to determine how DA regulates the ACh dip in a reinforcement task. I found that the duration, but not the initiation, of the ACh dip is dependent on striatal DA via proper CIN DA receptor (D2R) function. In addition, the rebound in ACh levels are also dependent on DA. These results confirm that there is a DA-dependent and a non-DA-dependent component of the CIN pause. Moreover, I found that CIN D2Rs are necessary for the coordinated activity between DA and ACh. Next, I explored the consequences of a shorter ACh dip on behavior. First, I found that the size of the ACh dip inversely correlates with behavioral responding: the larger the dip the faster the mice pressed a lever in a reinforcement task. This result suggests a role for ACh in motivated behavior. Finally, I found that selective ablation of CIN D2Rs shortens the ACh dip and facilitates reversal learning in a probabilistic choice task. This finding indicates that the length of the ACh dip is important for flexible learning, supporting the computational model by Franklin and Frank

where shorter CIN pauses promote reversal learning compared to longer pauses (see above) (Franklin & Frank, 2015).

Altogether, these findings reveal the importance of optimal balance between striatal DA and ACh signaling underlying key striatal functions. This work also provides hope for the development of new treatment strategies for many neuropsychiatric disorders.

Lastly, as part of this thesis work as outlined in Appendix A, we examined the role of D2Rs in indirect pathway spiny projection neurons (iSPNs) in the ventral striatum (NAc). The goal of this study was to determine the behavioral consequences of upregulation of D2Rs in NAc iSPNs. We found that upregulation of D2Rs in iSPNs did not affect Pavlovian conditioning or extinction and did not impair No-Go learning. Our results support previous findings that D2Rs in NAc iSPNs play a specific role in motivation by balancing cost/benefit computations but do not necessarily affect associative reward learning (Gallo et al., 2018; Martyniuk et al., 2021).

Chapter 2: Dopamine D2Rs Coordinate Cue-Evoked Changes in Striatal Acetylcholine Levels

2.1 Abstract

In the striatum, acetylcholine (ACh) neuron activity is modulated co-incident with dopamine (DA) release in response to unpredicted rewards and reward predicting cues and both neuromodulators are thought to regulate each other. While this co-regulation has been studied using stimulation studies, the existence of this mutual regulation *in vivo* during natural behavior is still largely unexplored. One long-standing controversy has been whether striatal DA is responsible for the induction of the cholinergic pause or whether D2R modulate a pause that is induced by other mechanisms. Here, we used genetically encoded sensors in combination with pharmacological and genetic inactivation of D2Rs from cholinergic interneurons (CINs) to simultaneously measure ACh and DA levels after CIN D2R inactivation. We found that CIN D2Rs are not necessary for the induction of cue induced dips in ACh levels but regulate dip lengths and rebound ACh levels. Importantly, D2R inactivation strongly decreased the temporal correlation between DA and Ach signals not only at cue presentation but also during the intertrial interval. This points to a general mechanism by which D2Rs coordinate both signals. At the behavioral level D2R antagonism increased the latency to lever press, which was not observed in CIN-selective D2R knock out mice. This latency correlated with the cue evoked dip length supporting a role of the ACh dip and it's regulation by D2Rs in motivated behavior. Overall, our data indicate that striatal DA coordinates phasic ACh and DA signals via CIN D2Rs, which is important for the regulation of motivated behavior.

* This chapter is under review at eLife. It has been modified from the pre-print version of this publication, which can be found at: **Martyniuk KM**, Torres-Herraez A, Rubinstein M, Labouesse MA, Kellendonk C. Dopamine D2Rs Coordinate Cue-Evoked Changes in Striatal Acetylcholine Levels. bioRxiv. 2021.12.08.471871; dio: http://dio.org/10.1101/2021.12.08.471871.

K.M.M., M.A.L., C.K. designed the experiments. K.M.M performed the experiments and analyzed the data. A.T. and M.A.L. assisted in the analysis of the data. K.M.M., A.T., M.A.L., and C.K. interpreted the results and wrote the manuscript. M.R. assisted in writing the manuscript.

2.2 Introduction

Dopamine (DA) plays a key role in learning, serving as a teaching signal that reflects reward prediction error (Jermey J. Day et al., 2007; Ali Mohebi et al., 2019; Nasser et al., 2017; Schultz et al., 1997; Elizabeth E. Steinberg et al., 2013). This teaching function is encoded in the phasic bursting of DA neurons, which induces a rapid but transient increase of extracellular DA. DA is initially released in response to an unpredicted reward, but with learning the response shifts away from the reward outcome towards reward predicting cues (Schultz, 2007; Schultz et al., 1997).

Like DA neurons, cholinergic interneurons (CINs) in rodents and their presumed counterparts, "tonically active neurons" (TANs), in primates modulate their activity in response to reward predicting cues and salient outcomes. CINs represent about 1-2% of the neurons in the striatum and regulate mental processes including reinforcement learning, action selection, associative learning, and cognitive flexibility (Aoki et al., 2015; Bradfield et al., 2013; Joshua et al., 2008; Matamales et al., 2016; Maurice et al., 2015; Morris et al., 2004; Okada et al., 2014). Pharmacogenetic inhibition of CINs in the NAc also increases the influence of appetitive cues on instrumental actions pointing to a role of striatal ACh in motivation (Collins et al., 2019). CINs are tonically active and show a multiphasic response to salient and conditioned stimuli that can

include a short excitation followed by a prominent pause and rebound excitation (Aosaki et al., 1994b; Aosaki T, 1994b; Apicella, 2007; Apicella et al., 2009; Apicella et al., 2011). This multiphasic response in CIN firing coincides with phasic activation of midbrain DA neurons that terminate in the striatum (Joshua et al., 2008; Morris et al., 2004; Schultz, 2007; Schultz et al., 1997). Furthermore, there is increasing evidence that DA and ACh regulate each other within the striatum (Cachope & Cheer, 2014; Cachope et al., 2012; Chuhma et al., 2014; Cragg, 2006; Helseth AR, 2021; Kharkwal et al., 2016; Straub et al., 2014; David Sulzer et al., 2016; Threlfell et al., 2012; Yan & Surmeier, 1991).

Here, we will focus on the DA-regulation of the multiphasic ACh response. One long-standing discussion in this regard has been whether the cholinergic pause is dependent on DA via DA D2R receptor (D2R) mediated inhibition of CINs. Early evidence that the CIN pause is DA dependent originate from studies in non-human primates (NHPs). *In vivo* electrophysiological recordings from TANs have revealed a pronounced pause in firing to a reward-predicting stimulus. This pause was entirely abolished by 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) lesions of DA neurons and local administration of a D2R antagonist (Aosaki et al., 1994b; Watanabe & Kimura, 1998). Consistent with this, more recent slice physiology studies in rodents have shown that pauses in CIN activity can be induced by local application of DA or DA terminal stimulation, which both are eliminated by pharmacological blockade of D2Rs (Augustin et al., 2018; Chuhma et al., 2014; Straub et al., 2014; Wieland et al., 2014). Additionally, optogenetic stimulation of NAc DA terminals results in a pause in CIN firing and this pause is prolonged when D2Rs are selectively overexpressed in CINs (Gallo et al., 2021). Lastly, pauses generated by DA or local stimulation of the striatum are eliminated in a selective CIN D2 knockout mouse (Augustin et al., 2018; Kharkwal et al., 2016). Taken together, the slice physiology experiments provide

evidence that the CIN pause can be induced by DA activation in a CIN D2R dependent manner while the NHP studies show the necessity for DA and D2Rs for the generation of the pause.

However, more recent evidence suggests that the CIN pause is not induced by DA but by cortical, thalamic, or long-range GABAergic inputs (Brown et al., 2012; Cover et al., 2019; Ding et al., 2010; Doig et al., 2014; English et al., 2012b; Matsumoto et al., 2000; Zhang et al., 2018). Consistent with this, stimulation of cortical and thalamic inputs to the striatum in slices or *in vivo* induces a triphasic cholinergic pause. One model suggests that the cholinergic pause is generated by intrinsic properties of CINs. When CINs come out of the early glutamatergic excitation, voltage-gated potassium channels (Kv7.2/7.3) open and induce an after-hyperpolarization that induces the pause. In this model DA plays a role in augmenting the intrinsically induced pause (Zhang et al., 2018). Consistent with this, thalamo-striatal stimulation induced a pause that was shortened but not fully abolished by a D2R antagonist (Cover et al., 2019). However, in earlier influential slice physiology experiments, the pause induced by thalamic stimulation was fully blocked by D2R antagonism suggesting that activation of DA release from intrastriatal DA terminals was responsible for pause generation (Ding et al., 2010).

One limitation of the mechanistic studies in rodents has been that they relied on stimulation experiments rather than on DA evoked by natural stimuli. While the early NHP studies suggested necessity for DA in inducing the pause during behavior, these studies lacked the cellular specificity for excluding the possibility that the effects of pharmacological D2R blockage were due to inhibiting D2Rs on CINs vs other neuronal populations.

Here, we used genetically encoded biosensors (Labouesse & Patriarchi, 2021) to simultaneously monitor DA and ACh in the dorsal striatum during behavior in mice with pharmacological blockade and/or selective ablation of D2Rs from CINs. Using this approach, we

addressed the question of whether the natural stimulus induced pause is fully dependent on DA or not. We first determined whether changes in DA and ACh levels occur simultaneously to reward-predicting stimuli in mice as has been shown in NHPs via electrophysiological recordings of DA and TAN neurons (Morris et al., 2004). *In vivo* imaging of ACh and DA levels revealed cue-induced decreases in striatal ACh and increases in DA levels, confirming the ability to measure concomitant ACh dips and DA peaks with functional imaging. Using a Pavlovian learning task, we confirmed that both signals co-occur and develop in parallel during the training of the task. Using a simpler reinforcement task that enables better quantification of the neuromodulator signals, we quantified cue-induced changes in DA and ACh after manipulating D2R function. We found that selective ablation of D2Rs from CIN or blocking D2Rs in control mice with the selective D2R antagonist eticlopride did not abolish the stimulus-induced decrease in ACh levels. Rather it shortened the length of dip and enhanced ACh rebound levels in a dose dependent manner. This indicates that DA is necessary for controlling the overall shape of the ACh dip. During simultaneous recordings experiments, the relationship between DA and ACh was strongest in response to reward predicting cues but still present during the intertrial interval supporting a general mechanism by which DA coordinates ACh levels. At the behavioral level, D2R antagonism increased the latency to lever press to a reward-paired lever, but this relationship was abolished when we inactivated CIN D2Rs. Moreover, cue evoked changes in ACh levels correlated with the latency to press, altogether supporting a role of the ACh dip in motivated behavior.

2.3 Results

The goal of this study was to determine if behaviorally induced decreases in ACh levels are entirely dependent on CIN D2Rs. To address this question, we used genetically encoded

sensors in combination with pharmacological and genetic inactivation of D2Rs from CINs to simultaneously measure ACh and DA levels after CIN D2R inactivation.

2.3.1 GACh3.0 allows for measuring fast decreases in task evoked ACh levels

First, we validated our experimental approach to simultaneously image changes in ACh and DA levels within the same mouse (Figure 2.1A) during an instrumental task, continuous reinforcement (CRF) (Figure 2.1B). We aligned our photometry signals to the lever extension, which with training becomes a reward predicting cue. After 3 days of training, we observed an increase in DA (red) and a decrease in ACh (blue) at lever extension presentation (Figure 2.1C). To confirm that the fluorescent indicator, GACh3.0, is measuring changes in ACh levels (and not movement artefacts or electrical noise), we measured the GACh3.0 signal in the presence of 15 mg/kg scopolamine, a M3R muscarinic antagonist, which targets the GACh3.0 parent receptor (M3R). We found that scopolamine blocked the detection of ACh levels in response to lever extension when compared to saline injections (Figure 2.1D).

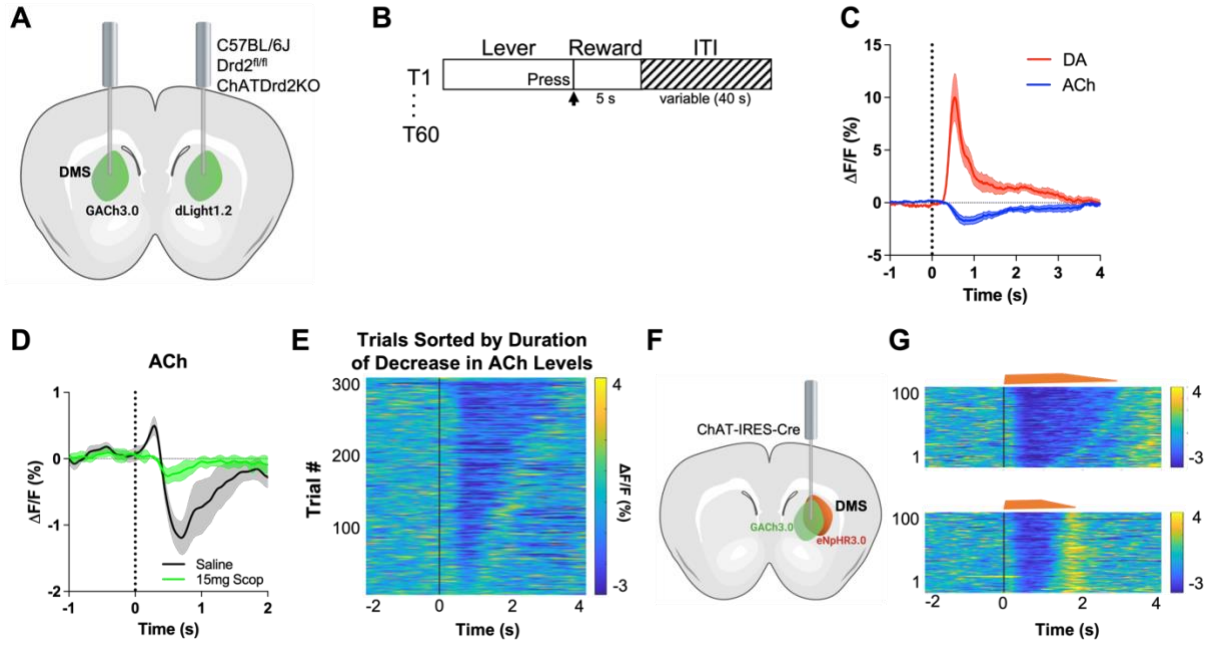

Figure 2.1. GACh3.0 reliably measures fast decreases in ACh during an instrumental task. (A) Schematic of the surgery setup. All mice were injected with both biosensor viruses (GACh3.0 and dLight1.2) in separate hemispheres of the DMS and counterbalanced across mice. Fiber photometry lenses were bilaterally implanted at the site of viral injection to simultaneously monitor ACh and DA in the same mouse. **(B)** CRF task design. Mice were trained to press a lever to retrieve a milk reward for 60 trials/day with a variable ITI (40 s). (**C)** Changes in fluorescence (ΔF/F (%)) aligned to lever extension (timepoint= 0 s). DA levels (red) increased and ACh levels (blue) decreased, N= 5 mice in trained mice. **(D)** 15 mg/kg of scopolamine (green), a mAChR antagonist, blunts the initial ACh peak and dip compared to saline (black) confirming that the GACh3.0 sensor is reporting changes in ACh levels. N= 4 mice. **(E)** Heatmap of ACh responses aligned to lever extension (Time= 0 s) and sorted by duration of ACh dip for 300 individual trials (60 trials in 5 mice). **(F)** Schematic of the surgery setup. ChAT-ires-Cre mice were co-injected with GACh3.0 and Cre-dependent halorhodopsin into the DMS and a fiber photometry lens was implanted at the site of viral injection. **(G)** Approximation of trials with short dips (bottom) and long dips (top) using the short and long optogenetic inhibition protocol (100 trials, 20 trials/ 5 mice).

To confirm that the GACh3.0 sensor has the kinetics to measure a rapid decrease in ACh levels, we expressed the inhibitory opsin eNpHR3.0 in ChAT-IRES-Cre mice to selectively inhibit CINs. Lever extension induced decreases in ACh levels within 250 ms showed varying lengths trial by trial (Figure 2.1E). Light activation of eNpHR3.0 induced a dip with even shorter latency (latency to dip onset 206.4 [186.8-226.1] ms, n=5 mice), which was followed by a rebound in ACh levels (Figure 2.1G). The optogenetic experiment was performed in the home cage. This is consistent with CINs displaying rebound activity after injecting hyperpolarizing currents in brain slices (Wilson, 2005) and optogenetic inhibition *in vivo* (English et al., 2012b). Combined, these data show that GACh3.0 can measure fast decreases in ACh levels. It also indicates that ACh release is tightly controlled by CIN neuron activity, while degradation is rapidly regulated by acetylcholinesterase (AChE).

2.3.2 Simultaneous changes in DA and ACh signals in response to a reward predicting stimulus

To determine whether changes in DA and ACh levels in response to reward predicting stimuli are co-incident, we measured the release of DA and ACh during a classic Pavlovian reward learning task (Figure 2.2B). Using fiber photometry and genetically encoded fluorescent indicators, we simultaneously imaged DA and ACh in separate hemispheres within the same animal. On Day 1 of training, we observed an increase in DA (red) and a decrease in ACh (blue)

during unexpected reward following the offset of the CS+. Over training, we saw these changes in both DA and ACh shift to the onset of the CS+ tone, while decreasing to the now expected reward. We did not observe these changes during CS- trials. We then related the changes in DA and ACh to changes in anticipatory nose poking during the CS+ as a measure of learning. We found that both DA and ACh signals correlated well with anticipatory head poking in one animal (Figure

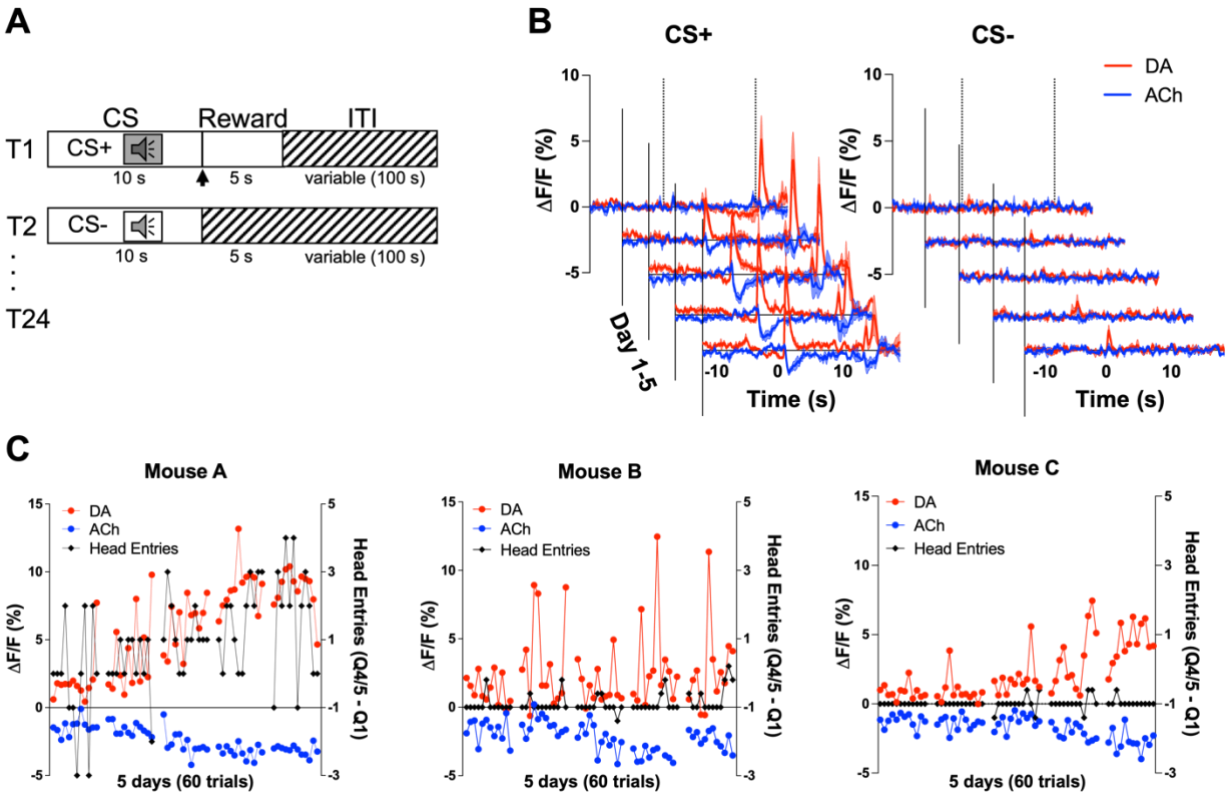

Figure 2.2. Co-development of DA and ACh signals to a reward predicting cue. (A) Pavlovian task design. Mice were trained on 24 (12 CS+, 12 CS-) trials/day for 5 days. Each trial starts with a 10 s tone (CS+ or CS-). At the end of the CS+ a dipper comes up presenting a milk food reward for 5 s. There is an intertrial interval (ITI) variable in length (100 s). **(B)** Changes in fluorescence (ΔF/F (%)) over 5 days of training for DA (red) and ACh (blue) aligned to CS+ (left) and CS- (right) onset. Signals were averaged over 12 CS+ and 12 CS- trials/day, N= 3 mice. **(C)** Maximum change in DA peak (blue) and ACh dip (red) after CS+ onset over 5 days of training (60 trials) for Mouse A (left). Anticipatory responding (black) is calculated as the difference in nose poking during the CS+ quintile with the maximum responses (Q4 or 5) and the first quintile. Correlations between DA and ACh maxima and behavioral responding: r= 0.4, $p<0.002$ and r= -0.41, $p<0.002$, respectively. We did not observe the same correlation between DA/ACh and anticipatory responses in Mouse B (middle) or Mouse C (right).

2.2C). However, other mice did not show any anticipatory responding as this task is non-contingent and head poking is not required to obtain the reward during CS+ trials. These findings indicate that DA and ACh signals co-develop with learning in response to a reward predicting stimulus.

2.3.3 D2 receptor blockade dose dependently decreases ACh dip durations and enhances the rebound in ACh levels

To determine if the cue induced ACh dip is dependent on DA activation of D2Rs, we used the CRF task as it allows for more trials per session aiding the quantification of the signal. After

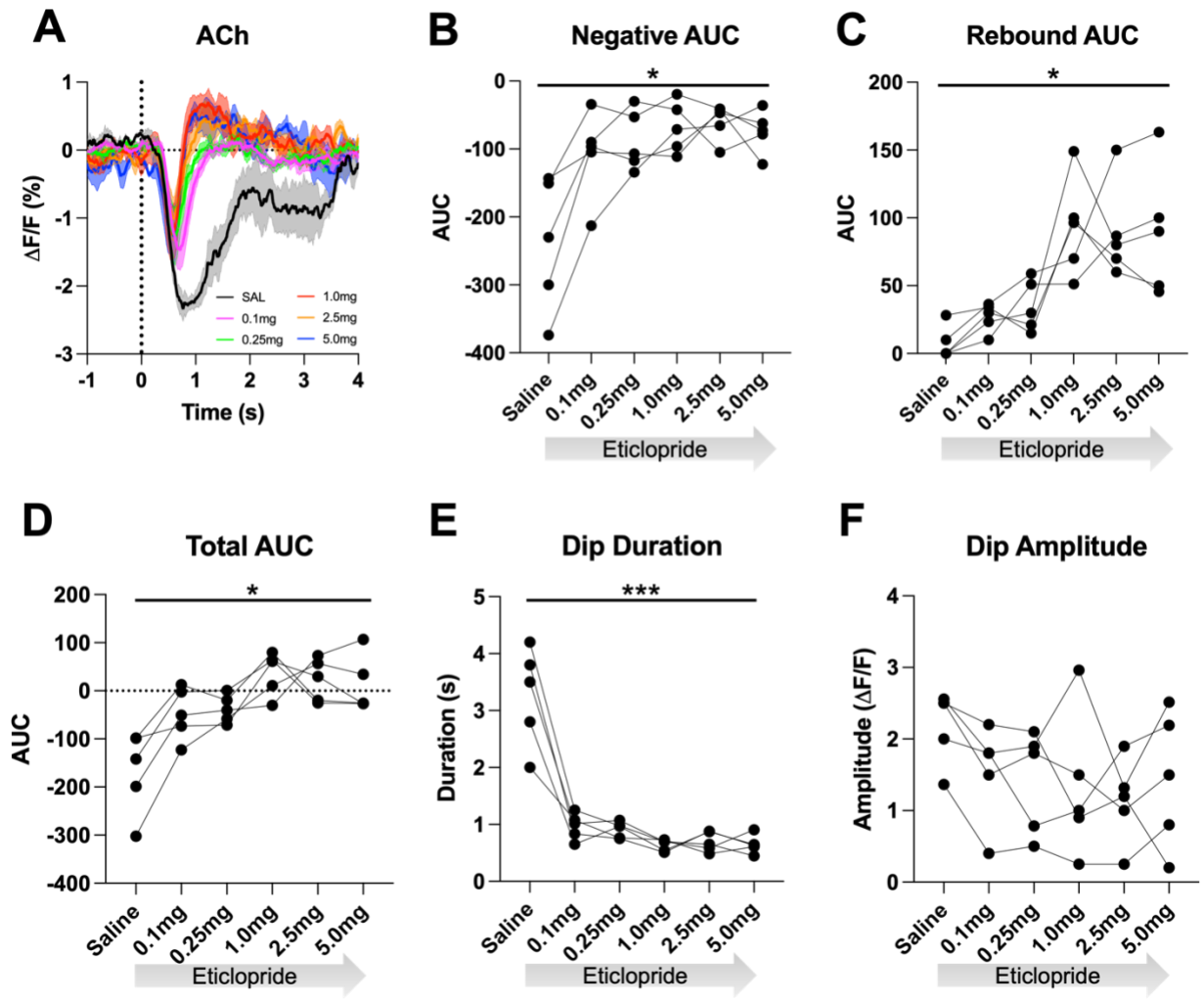

Figure 2.3. D2R antagonism decreases ACh dip duration and enhances rebound. (A) Changes in ACh fluorescence (ΔF/F (%)) aligned to lever extension with saline (black) and increasing doses of eticlopride: 0.1 mg/kg (pink), 0.25 mg/kg (green), 1.0 mg/kg (red), 2.5 mg/kg (orange) and 5.0 mg/kg (blue). **(B)** Negative AUC is reduced by eticlopride in a dose-dependent manner ($F_{(1.694, 6.777)} = 8.756$, $p = 0.0150$). **(C)** The rebound AUC is increased by eticlopride in a dose-dependent manner ($F_{(1.549, 6.197)} = 8.833$, $p = 0.0181$) **(D)** Total AUC is increased by eticlopride in a dose-dependent manner ($F_{(1.612, 6.448)} = 8.724$, $p = 0.0170$). **(E)** Dip duration is decreased by eticlopride in a dose-dependent manner ($F_{(1.392, 5.569)} = 36.37$, $p = 0.0009$). **(F)** The dip amplitude was not affected by eticlopride ($F_{(2.063, 8.251)} = 1.864$, $p = 0.2147$).

systemic delivery of the D2R antagonist eticlopride we found a dose-dependent shortening of the

ACh dip, which uncovered a rebound following the pause (Figure 2.3A). We quantified these changes by calculating the area under the curve (AUC), dip duration and dip amplitude. We found that eticlopride significantly reduced the negative AUC (Figure 2.3B), increased the rebound AUC (Figure 2.3C), increased the total AUC (Figure 2.3D), and decreased the dip duration (Figure 2.3E), while the dip amplitude was not affected by D2R antagonism (Figures 2.3F). This suggests that D2Rs do not participate in the initial induction of the ACh dip but do increase the duration of the

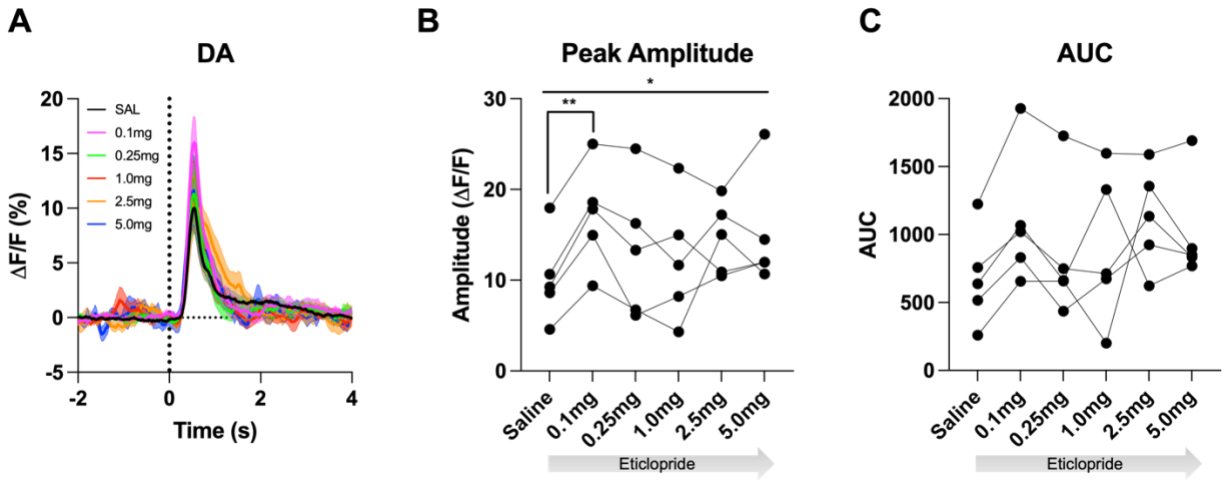

Figure 2.4. D2R antagonism increases phasic DA release. (A) Changes in DA fluorescence (ΔF/F (%)) aligned to lever extension with saline (black) and increasing doses of eticlopride: 0.1 mg/kg (pink), 0.25 mg/kg (green), 1.0 mg/kg (red), 2.5 mg/kg (orange) and 5.0 mg/kg (blue) in C57BL/6J mice. **(B)** Peak amplitude is increased by eticlopride in a dose-dependent manner ($F_{(2.201,\ 8.805)} = 4.268$, $p = 0.0480$). The most prominent increase in peak amplitude is between saline and 0.1 mg/kg ($p = 0.0069$). **(C)** No overall effect of eticlopride on the AUC ($F_{(1.440,\ 5.759)} = 2.347$, $p = 0.1807$).

dip and prevent rebound activity following the ACh dip. Since DA neurons are inhibited by D2 auto-receptors, we also analyzed the effect of D2R antagonism on cue induced DA release and quantified changes in peak amplitude and AUC (Figure 2.4A). We found an overall effect of drug increasing the peak amplitude (Figure 2.4B) with the most prominent increase between saline and 0.1 mg/kg. There was no overall effect of drug on the AUC (Figure 2.4C). These results confirm that blocking D2 auto-receptors on DA neurons increases phasic DA release.

Individual CRF trials revealed varying durations of lever-extension aligned ACh dips that we sorted by lever press latency using a heatmap (Figure 2.5A). Based on this heatmap, we

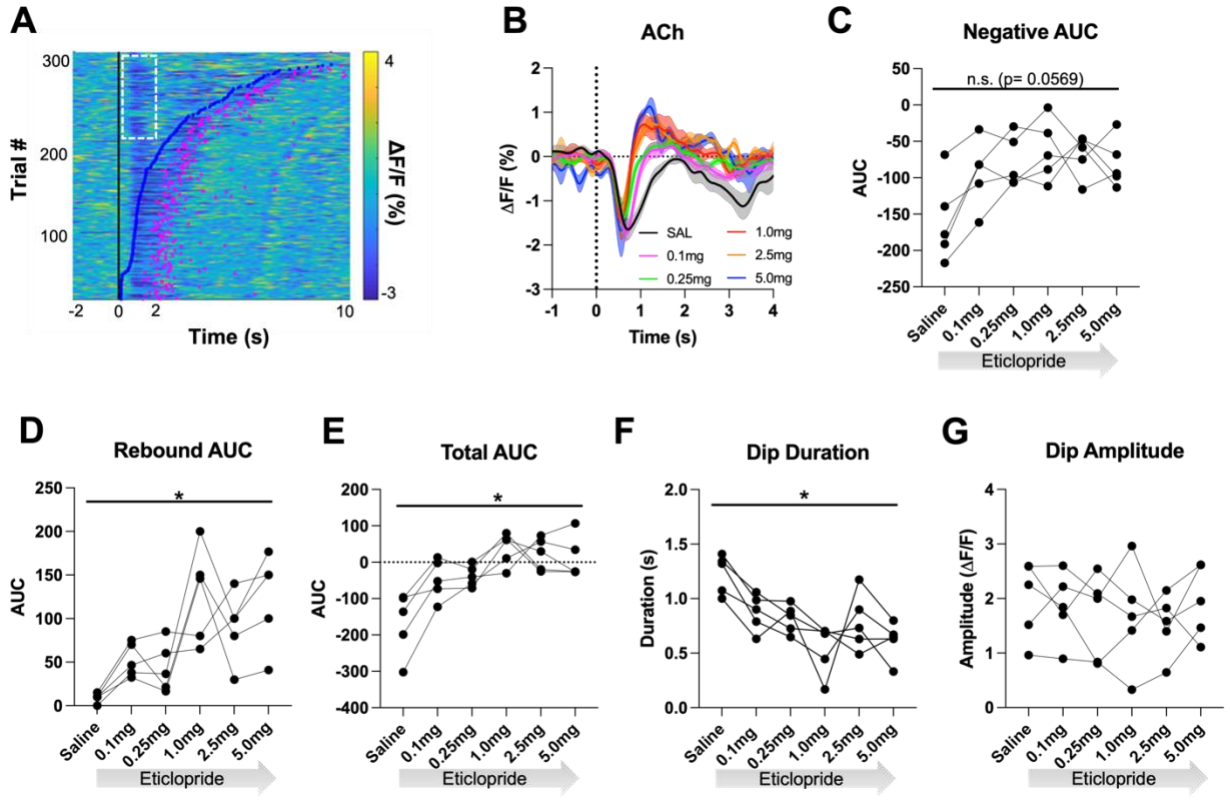

Figure 2.5. D2R antagonism shortens cue evoked ACh dip and enhances rebound. (A) Heatmap of ACh responses aligned to lever extension (Time= 0 s) for 300 individual trials (60 trials in 5 mice) and sorted by response length (bottom). Blue dots show the lever press, and the pink dots show the head entry for each trial. White dashed box represents the cue evoked ACh response to the lever extension where press latencies > 2 s. **(B)** Changes in ACh fluorescence (ΔF/F (%)) aligned to lever extension for only trials with press latencies > 2 s with increasing doses of eticlopride. **(C)** Negative AUC is reduced by eticlopride in a dose-dependent manner ($F_{(2.237,\ 8.950)}$ = 3.911, p = 0.0569). **(D)** Rebound AUC is enhanced by eticlopride in a dose-dependent manner ($F_{(1.667,\ 6.668)}$ = 8.143, p = 0.0184). **(E)** Total AUC was increased by eticlopride in a dose-dependent manner ($F_{(1.597,\ 6.387)}$ = 8.542, p = 0.0182). **(F)** Dip duration was significantly decreased by eticlopride in a dose-dependent manner ($F_{(1.657,\ 6.628)}$ = 6.729, p = 0.0284). **(G)** Eticlopride had no effect on the dip amplitude ($F_{(2.722,\ 10.89)}$ = 0.5379, p = 0.6503).

observed longer dips associated with quick press latencies and two smaller dips with slower press latencies with the second dip co-occurring with the lever press. Thus, for press latencies < 2 s the ACh is a combination of a cue induced and movement associated pause. To separate the cue induced pause from the movement induced pause, we analyzed trials with press latencies > 2 s. We still observed a decrease in the ACh dip duration with increasing doses of eticlopride (Figure

2.5B). Quantification of the negative AUC revealed a non-significant but trending decrease with increasing doses of eticlopride (Figure 2.5C), while the rebound AUC increased (Figure 2.5D), the total AUC increased (Figure 2.5E) and the dip duration (Figures 2.5F) decreased. Eticlopride had no effect on the ACh dip amplitude (Figure 2.5G).

We also examined the effect of D2R antagonism on cue induced DA release for trials with press latencies > 2s (Figure 2.6A). Quantification of DA peak amplitude (Figure 2.6B) and AUC (Figure 2.6C) revealed an overall increase in both measures. Moreover, we found a significant increase between saline and 0.1 mg/kg eticlopride for DA peak amplitude (Figure 2.6B). Taken together, these results demonstrate that the cue induced ACh dip and rebound levels are regulated by cholinergic D2Rs.

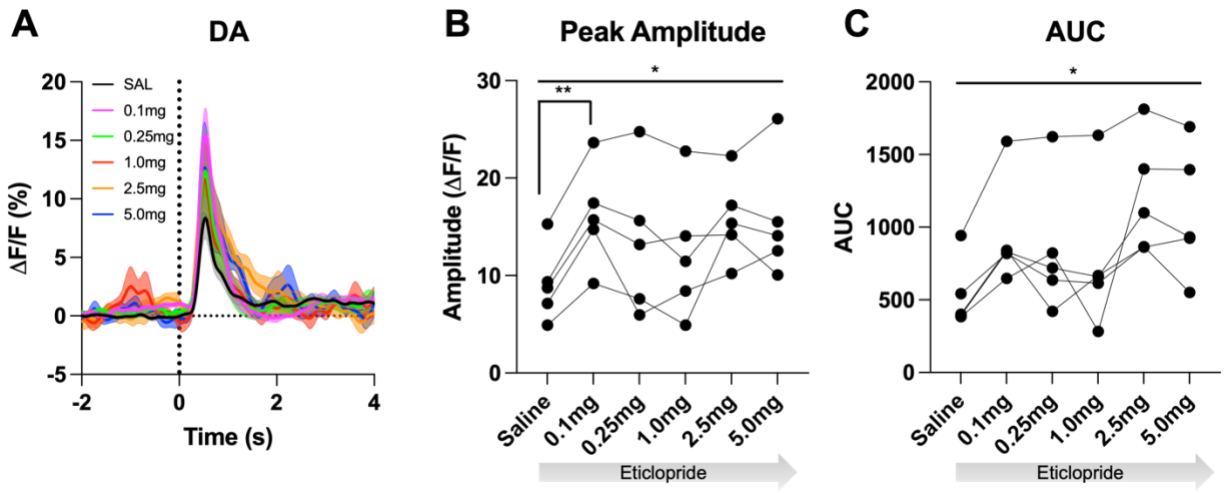

Figure 2.6. D2R antagonism enhances cue evoked DA release for trials with press latencies > 2s. (A) Changes in DA fluorescence (ΔF/F (%)) aligned to lever extension for only trials with press latencies > 2 s with increasing doses of eticlopride in C57BL/6J mice. **(B)** Peak amplitude of DA is increased by eticlopride in a dose-dependent manner ($F_{(2.785, 11.14)} = 5.804$, $p = 0.0133$) with the most prominent increase between saline and 0.1 mg/kg ($p = 0.0099$). **(C)** DA AUC is increased by eticlopride in a dose-dependent manner ($F_{(1.822, 7.288)} = 6.872$, $p = 0.0244$).

2.3.4 D2R blockade decreases negative and enhances positive correlations between DA and ACh

We further determined the relationship between ACh and DA levels within trials using a Pearson's correlation analysis. Using a lag analysis, we temporally shifted the ACh recording behind or in front of the DA recording to identify maximal points of correlation. During CRF trials,

the strongest correlation is a negative correlation (Figure 2.7A, label 1, saline: Pearson's r= -0.475 ± 0.037, N=5) that occurs when ACh lags DA (Lag= -178.92 ms ± 14.38 ms). This negative correlation, which reflects the ACh dip that follows the DA peak, is reduced with eticlopride in a dose-dependent manner (Figure 2.7B). Next, we found a small positive correlation (Figure 2.7A, label 2, saline: Pearson's r= 0.039 ± 0.014) when ACh lags DA (Lag= -1.5 s ± 0.138 s). This positive correlation, which reflects the rebound in ACh, is significantly increased with eticlopride (Figure 2.7C).

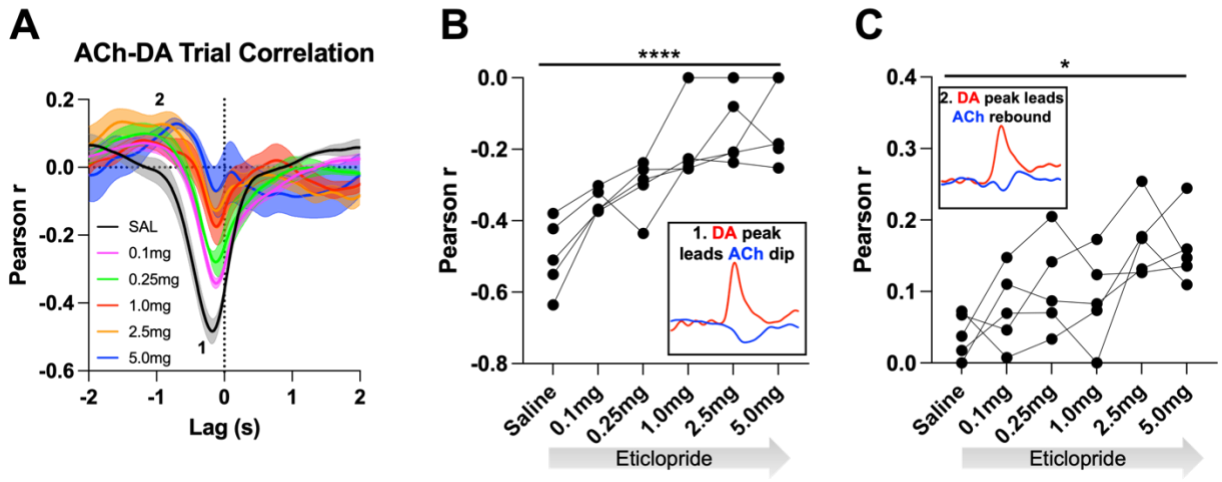

Figure 2.7. Task dependent ACh-DA interactions are altered by D2R antagonism at lever extension. (A) Correlation between ACh and DA during CRF trials with increasing doses of eticlopride in 5 C57BL/6J mice. The ACh signal moved in front of or behind the DA signal to identify points of highest correlation. The first correlation is a negative correlation (1) with ACh lagging DA (Lag= -178.92 ms ± 14.38 ms) and the second correlation is a positive correlation (2) with ACh lagging DA (Lag= -1.5 s ± 0.138 s). **(B)** The negative correlation with the DA peak leading the ACh dip (inset) is significantly reduced dose-dependently by eticlopride ($F_{(3.082, 12.33)} = 18.67$, $p < 0.0001$) **(C)** The positive correlation with the DA peak leading the ACh rebound (inset) is enhanced by eticlopride in a dose-dependent manner ($F_{(2.325, 9.299)} = 4.731$, $p= 0.0346$).

We then analyzed these correlations during the intertrial interval (ITI) to determine whether they are only present during stimulus induced DA/ACh signals or may represent a more general mechanism or coordination (Figure 2.8A). Of note, we looked for any interaction between DA and ACh regardless of event size. Like CRF trials, we observed two correlations during the ITI; DA peak leads ACh dip (Pearson's r= -0.355 ± 0.065 and Lag= -212.34 ms ± 16.91 ms) and DA peak leads ACh peak/rebound (Pearson's r= 0.058 ± 0.021 and Lag= -1.41 s ± 0.19 s). We found that

eticlopride decreases the negative correlation in a dose-dependent manner (Figure 2.8B). Eticlopride also increased the positive correlation, which represents the ACh rebound (Figure 2.8C). These results indicate that DA-ACh correlations are dependent on D2Rs. While they are strong during salient cue presentations the relationship between both signals still exists during the intertrial interval reflecting a general mechanism of co-regulation.

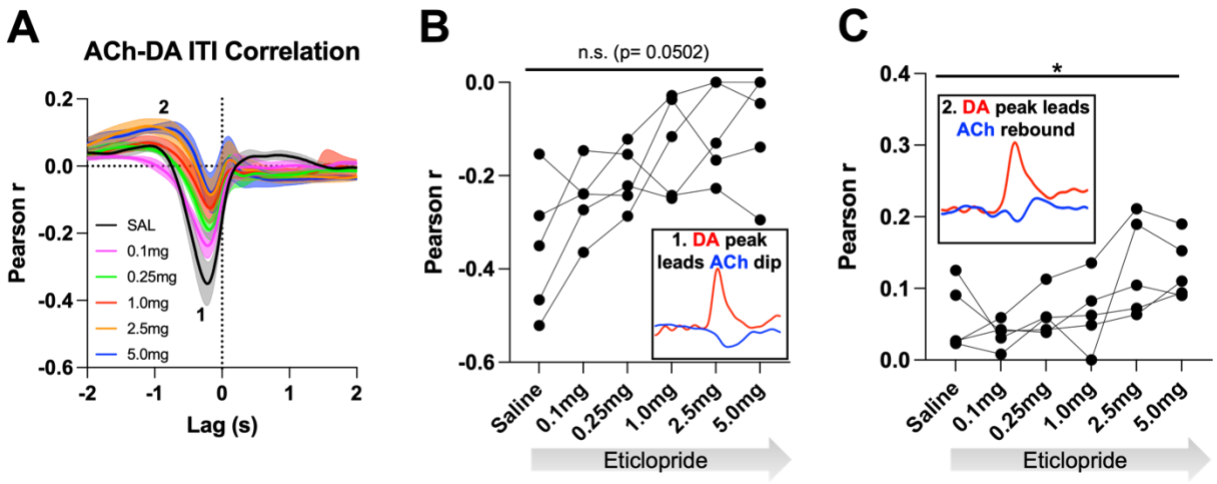

Figure 2.8. General ACh-DA interactions are altered by D2R antagonism during the ITI. (A) Correlation between ACh and DA during the ITI with increasing doses of eticlopride in C57BL/6J mice. We observe the same two correlations during the ITI: a negative correlation (1) with ACh lagging DA (Lag= -212.34 ms ± 16.91 ms) and a positive correlation (2) with ACh lagging DA (Lag= -1.41 s ± 0.19 s) **(B)** The negative correlation with the DA peak leading the ACh dip (inset) is decreased by eticlopride in a dose-dependent manner ($F_{(1.900,\ 7.598)}$ = 4.606, p= 0.0502). **(C)** The positive correlation with the DA peak leading the ACh rebound (inset) is increased dose-dependently by eticlopride ($F_{(2.118,\ 8.474)}$ = 4.873, p= 0.0377).

2.3.5 Genetic inactivation of D2Rs from CINs decreases dip duration

Systemic eticlopride injections block all D2Rs. To determine the specific modulatory role that D2Rs present in CINs play in the cholinergic pause, we used mouse genetics to selectively inactivate D2Rs from CINs (ChATDrd2KO mice). We observed a smaller and shorter ACh dip in ChATDrd2KO mice (Dip amplitude= 0.797 ΔF/F ± 0.131 and Dip duration= 0.931 s ± 0.105 s) compared to control mice (Dip amplitude= 1.50 ΔF/F ± 0.141 and Dip duration= 2.80 s ± 0.682 s) in trials with press latencies > 2 s (Figure 2.9A). Both, the dip amplitude, and dip duration were significantly reduced between the two groups (Figure 2.9B-C). Note that the effects of D2R

deletion differed from the highest dose of eticlopride in that ChATDrd2KO mice showed differences in the dip amplitude while eticlopride did not.

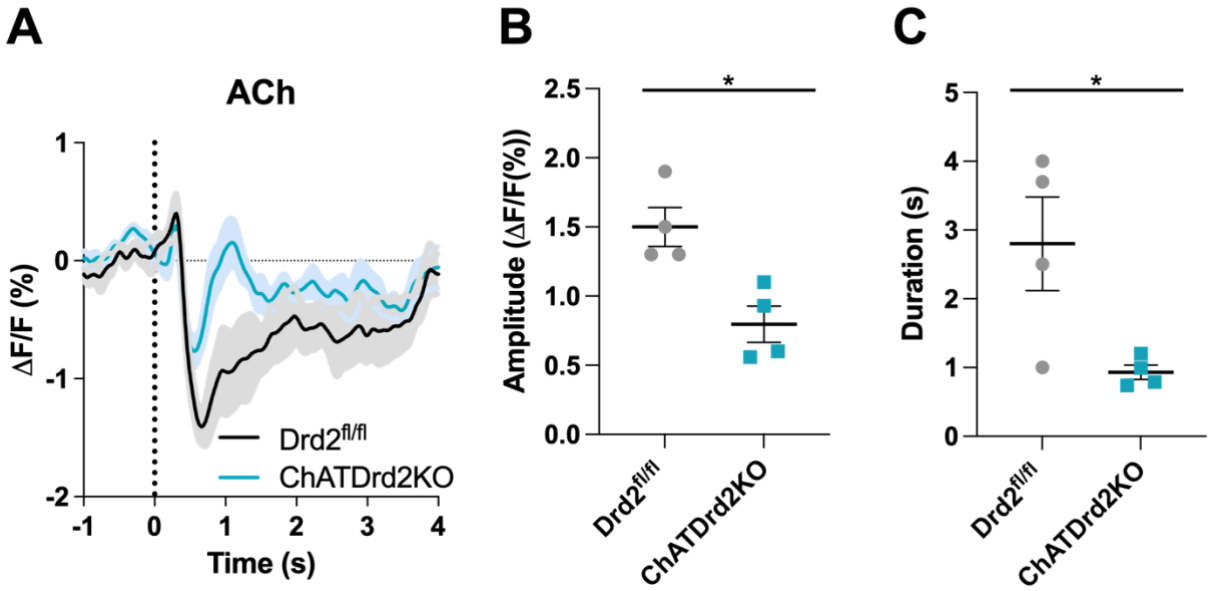

Figure 2.9. Selective D2R ablation from CINs alters the cue evoked ACh dip. (A) Changes in ACh fluorescence (ΔF/F (%)) aligned to lever extension for only trials with press latencies >2 s for Drd2$^{fl/fl}$ control (black) and ChATDrd2KO (blue) mice, N= 4/genotype. **(B)** Dip amplitude is significantly smaller in ChATDrd2KO animals compared to controls (p= 0.0107). **(C)** Dip duration is significantly shorter in ChATDrd2KO mice compared to Drd2$^{fl/fl}$ controls (p= 0.0351).

In contrast to ACh levels, stimulus induced DA release was not altered in ChATDrd2KO mice (Figure 2.10). This result indicates that loss of cholinergic D2Rs does not affect stimulus

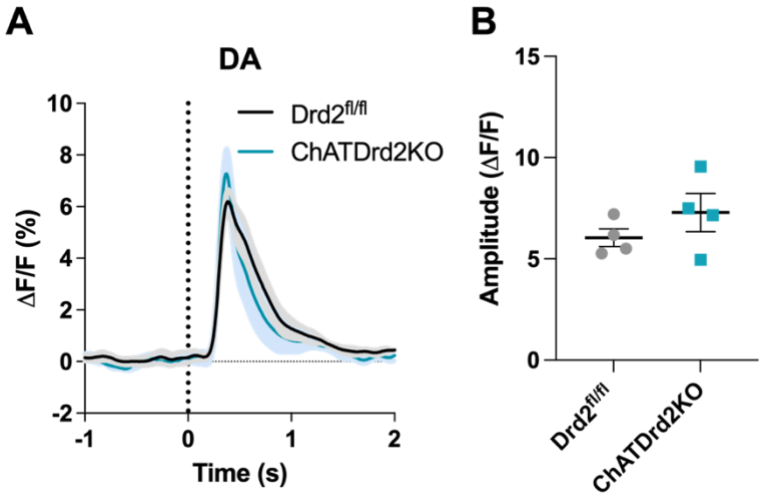

Figure 2.10. Selective D2R ablation from CINs does not alter cue-evoked DA release for trials with press latencies > 2 s. (A) Changes in DA fluorescence (ΔF/F (%)) aligned to lever extension for only trials with press latencies > 2 s for $Drd2^{fl/fl}$ control (black) and ChATDrd2KO (blue) mice, N=4/ genotype. **(B)** The DA peak amplitude is comparable between ChATDrd2KO and $Drd2^{fl/fl}$ control mice (p= 0.2744).

induced DA release and confirm that the effects of DA regulation of the ACh dip are mediated by CIN D2Rs and not an indirect effect by potential changes in DA levels.

2.3.6 DA-mediated changes in ACh levels are dependent on CIN D2Rs

Next, we determined if D2Rs present in CINs are necessary for the effect of D2R antagonism on modulating the ACh dip. Control $Drd2^{fl/fl}$ mice were more sensitive to eticlopride than the C57BL/6J wild-type mice of Figure 4 as they did not complete any trials with the two highest doses, 2.5 mg/kg, and 5.0 mg/kg (Figure 2.11A-F). Quantification of the ACh dip using the 3 lower doses revealed a decrease in the negative AUC (Figure 2.11B), an increase in the rebound AUC (Figure 2.11C), an increase in the total AUC (Figure 2.11D) and a decrease in dip duration (Figure 2.11E) that were comparable to what we measured in the C57BL/6J mice (Figure 2.5). Like the C57BL/6J mice, there was no effect on ACh dip amplitude with eticlopride (Figure 2.11F). In contrast, in ChATDrd2KO mice, we observed no change in the ACh dip with eticlopride (Figure 2.11G) and there was no effect of eticlopride on the negative AUC (Figure 2.11H), the rebound AUC (Figure 2.11I), total AUC (Figure 2.11J), dip duration (Figure 2.11K), or dip amplitude (Figures 2.11L). These results confirm that D2Rs present in CINs are key players in the modulation of the ACh signal elicited by D2R antagonism.

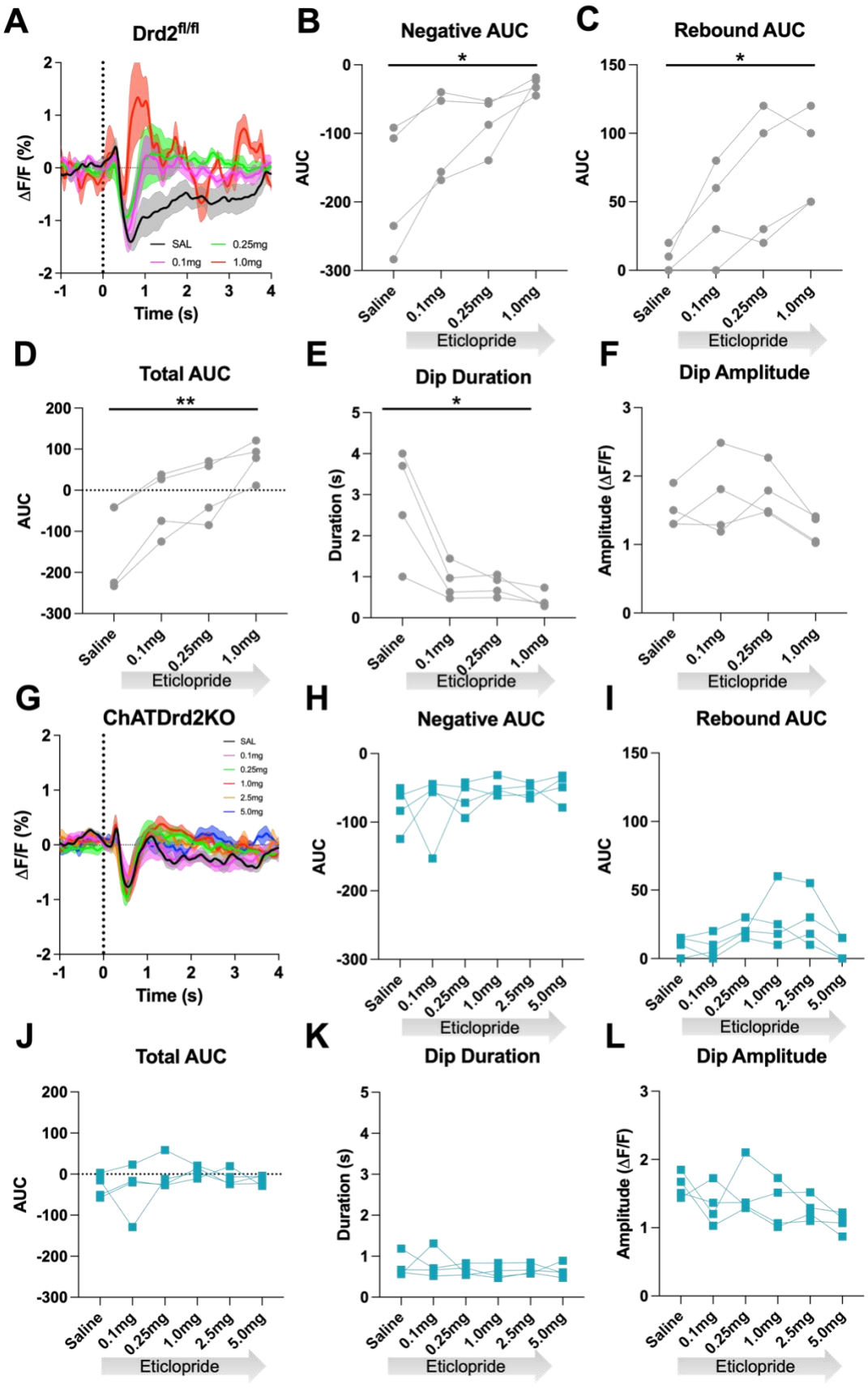
A
Drd2fl/fl
ΔF/F (%)
Time (s)
SAL
0.25mg
0.1mg
1.0mg
B
Negative AUC
AUC
C
Rebound AUC
AUC
Saline
0.1mg
0.25mg
1.0mg
Eticlopride
D
Total AUC
AUC
E
Dip Duration
Duration (s)
F
Dip Amplitude
Amplitude (ΔF/F)
G
ChATDrd2KO
ΔF/F (%)
Time (s)
SAL
0.1mg
0.25mg
1.0mg
2.5mg
5.0mg
H
Negative AUC
AUC
I
Rebound AUC
AUC
J
Total AUC
AUC
K
Dip Duration
Duration (s)
L
Dip Amplitude
Amplitude (ΔF/F)

Figure 2.11. D2R antagonism does not alter the cue evoked ACh dip in ChATDrd2KO mice. (A) Changes in ACh fluorescence (ΔF/F (%)) aligned to lever extension for only trials with press latencies > 2 s for $Drd2^{fl/fl}$ control mice with increasing doses of eticlopride. **(B)** Negative AUC is decreased by eticlopride in a dose-dependent manner ($F_{(1.387, 4.160)} = 8.541$, $p = 0.0381$). **(C)** Rebound AUC is increased by eticlopride in a dose-dependent manner ($F_{(1.642, 4.925)} = 10.21$, $p = 0.0195$). **(D)** Total AUC is increased dose-dependently by eticlopride ($F_{(1.525, 4.676)} = 23.14$, $p = 0.0047$). **(E)** Dip duration is decreased by eticlopride in a dose-dependent manner ($F_{(1.664, 4.992)} = 9.279$, $p = 0.0226$). **(F)** Dip amplitude is not affected by eticlopride ($F_{(1.433, 4.300)} = 6.056$, $p = 0.0606$). **(G)** Changes in ACh fluorescence (ΔF/F (%)) aligned to lever extension for only trials with press latencies > 2 s for ChATDrd2KO mice with increasing doses of eticlopride. **(H)** Negative AUC is not affected by eticlopride ($F_{(1.663, 4.990)} = 0.7919$, $p = 0.4803$). **(I)** Rebound AUC is not affected by eticlopride ($F_{(1.706, 5.119)} = 2.857$, $p = 0.1484$). **(J)** Total AUC is not affected by eticlopride ($F_{(1.844, 5.532)} = 1.079$, $p = 0.3958$). **(K)** Dip duration is not affected by eticlopride ($F_{(1.848, 5.545)} = 0.4380$, $p = 0.6516$). **(L)** Dip amplitude is not affected by eticlopride ($F_{(2.073, 6.219)} = 2.546$, $p = 0.1551$).

2.3.7 DA-mediated changes in DA-ACh correlations are dependent on CIN D2Rs

Next, we assessed the effect of CIN D2Rs in the ACh-DA coregulation, again using Pearson's r correlation analysis and lag analysis. We found that the interaction between DA and ACh was greatly reduced (> 2-fold) in ChATDrd2KO mice compared to $Drd2^{fl/fl}$ controls (Figure 2.12). The negative correlation with ACh lagging DA was significantly smaller in ChATDrd2KO mice during both CRF trials (Figure 2.12A-B) and the ITI (Figure 2.12C-D) compared to $Drd2^{fl/fl}$ controls. We further examined the role of CIN D2Rs in the synchronization of DA and ACh activity using eticlopride to transiently block CIN D2Rs. $Drd2^{fl/fl}$ control mice showed a strong negative correlation during CRF trials (Figure 2.12A: Pearson's r= -0.521 ± 0.038, N=4) with ACh lagging DA (Lag= -167.12ms ±18.82 ms) that was reduced by eticlopride (Figure 2.13B). In addition, a rebound in ACh activity was revealed, which is measured as a positive correlation (Figure 2.13A: Pearson's r= 0.050 ± 0.030, N=4) with ACh lagging DA (Lag= -1.51 s ± 0.138 s) that was increased with increasing doses of eticlopride (Figure 2.13C). During the ITI, the negative correlation between ACh and DA with ACh lagging DA (Figure 2.14A: Lag= -213.81 ms ± 34.14 ms) was attenuated by eticlopride (Figure 2.14B), while the positive correlation with ACh lagging DA (Lag= -1.39 s ± 0.208 s) was not significantly affected by eticlopride (Figure 2.14C). In ChATDrd2KO mice, neither the negative correlation between ACh and DA (Figure 2.15:

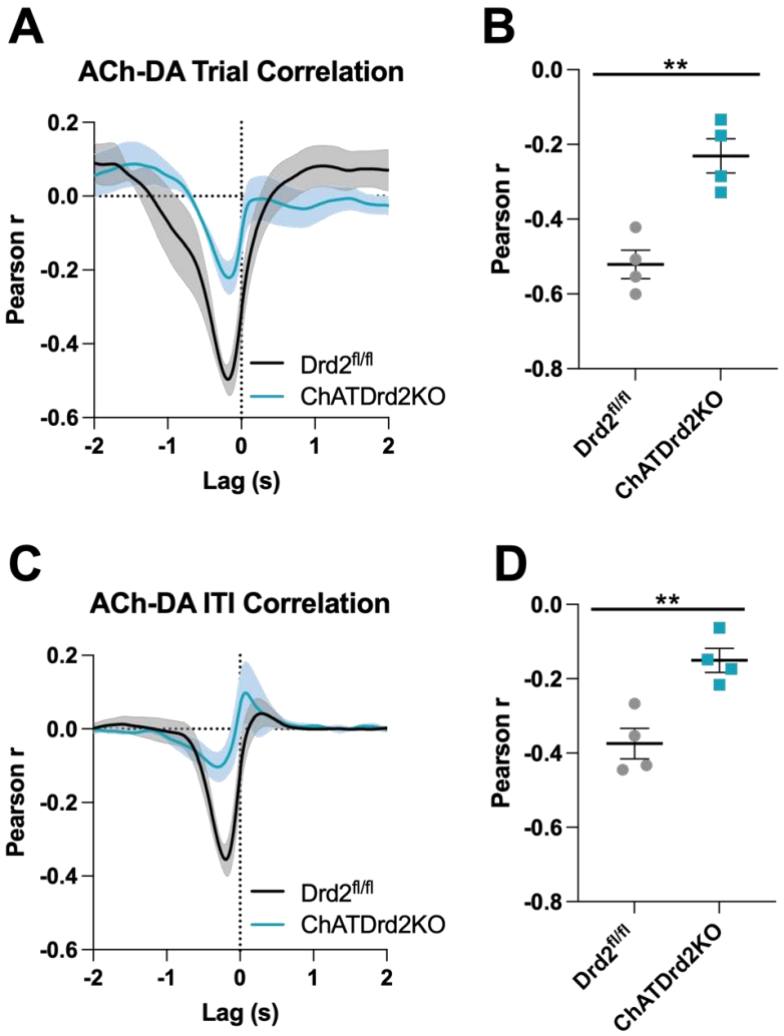

Figure 2.12. ACh-DA interactions are reduced in ChATDrd2KO mice. **(A)** Task evoked correlation between ACh and DA for Drd2$^{fl/fl}$ control (black) and ChATDrd2KO (blue) mice. **(B)** The negative correlation with ACh lagging DA is significantly reduced in ChATDrd2KO mice compared to Drd2$^{fl/fl}$ controls (p= 0.0155). **(C)** Correlation between ACh and DA during the ITI for Drd2$^{fl/fl}$ control (black) and ChATDrd2KO (blue) mice. **(D)** The negative correlation of ACh lagging DA is significantly reduced in ChATDrd2KO mice compared to Drd2$^{fl/fl}$ controls (p= 0.0052).

Pearson's r= -0.231 ± 0.046, N=4) with ACh lagging after DA (Lag= -179.41 ms ± 30.66 ms, N=4) nor the positive correlation with ACh lagging DA (Lag= -1.635 s ± 0.174 s) were affected by eticlopride (Figure 2.15). During the ITI, we observed a smaller negative correlation between ACh and DA (Figure 2.12C: Pearson's r= -0.153 ± 0.034, N=4) with ACh lagging DA (Lag= -282.62 ms ± 93.81 ms) that was significantly smaller in ChATDrd2KO mice compared to control mice

(Figure 2.12D) and was not affected by eticlopride (Figure 2.16). These results indicate that D2Rs in CINs regulate ACh dip and rebound levels.

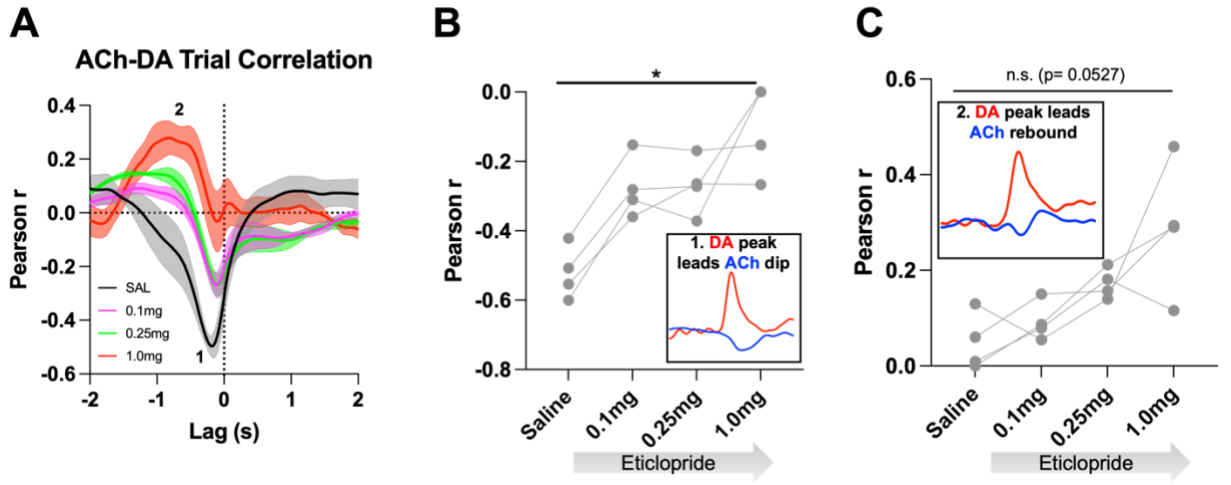

Figure 2.13. D2R antagonism alters cue-evoked ACh-DA interactions at the lever extension. (A) Correlation between ACh and DA during CRF trials with increasing doses of eticlopride in Drd2$^{fl/fl}$ control mice. The ACh signal moved in front of or behind the DA signal to identify points of highest correlation. The first correlation is a negative correlation (1) with ACh lagging DA (Lag= -167.12ms ±18.82 ms) and the second correlation is a positive correlation (2) with ACh lagging DA (Lag= -1.51 s ± 0.138 s) **(B)** The negative correlation with the DA peak leading the ACh dip (inset) is decreased by eticlopride in a dose-dependent manner ($F_{(1.141, 3.424)}$ = 15.48, p = 0.0221). **(C)** The positive correlation with the DA peak leading the ACh rebound (inset) is increased by eticlopride in a dose-dependent manner ($F_{(1.338, 4.014)}$ = 7.088, p = 0.0527).

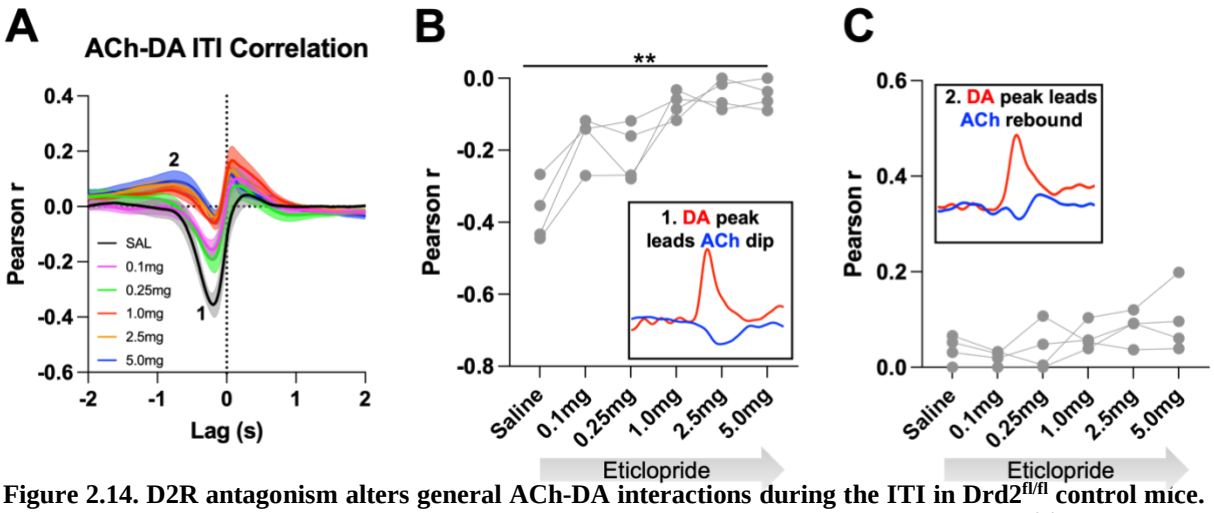

Figure 2.14. D2R antagonism alters general ACh-DA interactions during the ITI in Drd2$^{fl/fl}$ control mice. (A) Correlation between ACh and DA during the ITI with increasing doses of eticlopride in Drd2$^{fl/fl}$ control mice. The ACh signal moved in front of or behind the DA signal to identify points of highest correlation. The first correlation is a negative correlation (1) with ACh lagging DA (Lag= -213.81 ms ± 34.14 ms) and the second correlation is a positive correlation (2) with ACh lagging DA (Lag= -1.39 s ± 0.208 s) **(B)** The negative correlation with the DA peak leading the ACh dip (inset) is dose-dependently reduced by eticlopride ($F_{(2.102, 6.307)}$ = 18.68, p = 0.0021). **(C)** The positive correlation with the DA peak leading the ACh rebound (inset) is not significantly affected by eticlopride ($F_{(1.276, 3.827)}$ = 2.504, p = 0.1966).

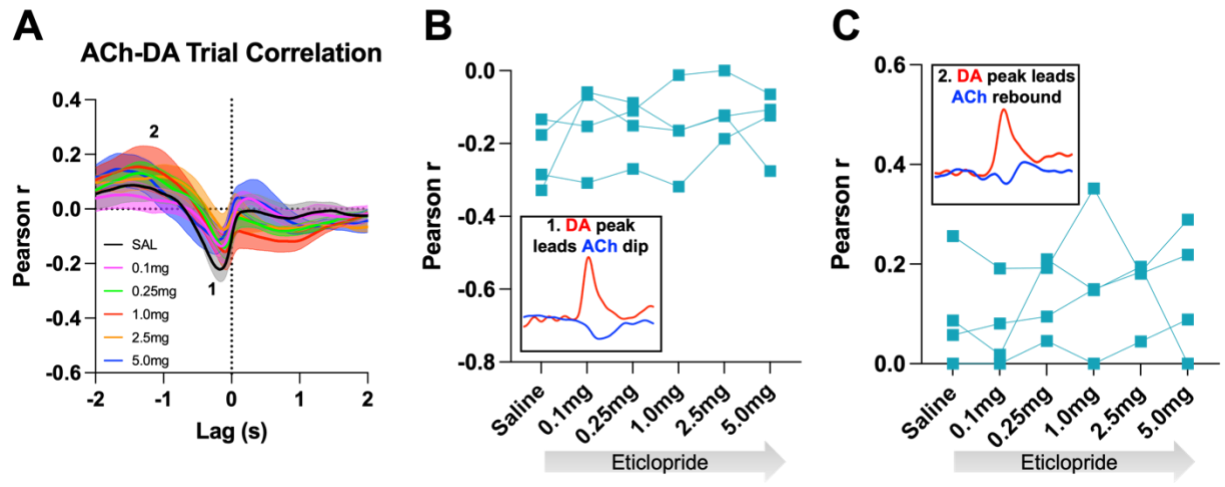

Figure 2.15. D2R antagonism does not alter cue evoked ACh-DA interactions in ChATDrd2KO mice at the lever extension. (A) Correlation between ACh and DA during CRF trials with increasing doses of eticlopride in ChATDrd2KO mice. The ACh signal moved in front of or behind the DA signal to identify points of highest correlation. The first correlation is a negative correlation (1) with ACh lagging DA (Lag= -179.41 ms ± 30.66 ms) and the second correlation is a positive correlation (2) with ACh lagging DA (Lag= -1.635 s ± 0.174 s). **(B)** The negative correlation with the DA peak leading the ACh dip (inset) is not affected by eticlopride ($F_{(1.720, 5.161)}$ = 1.170, p = 0.3682). **(C)** The positive correlation with the DA peak leading the ACh rebound (inset) is not affected by eticlopride ($F_{(2.016, 6.048)}$ = 0.9160, p = 0.4500).

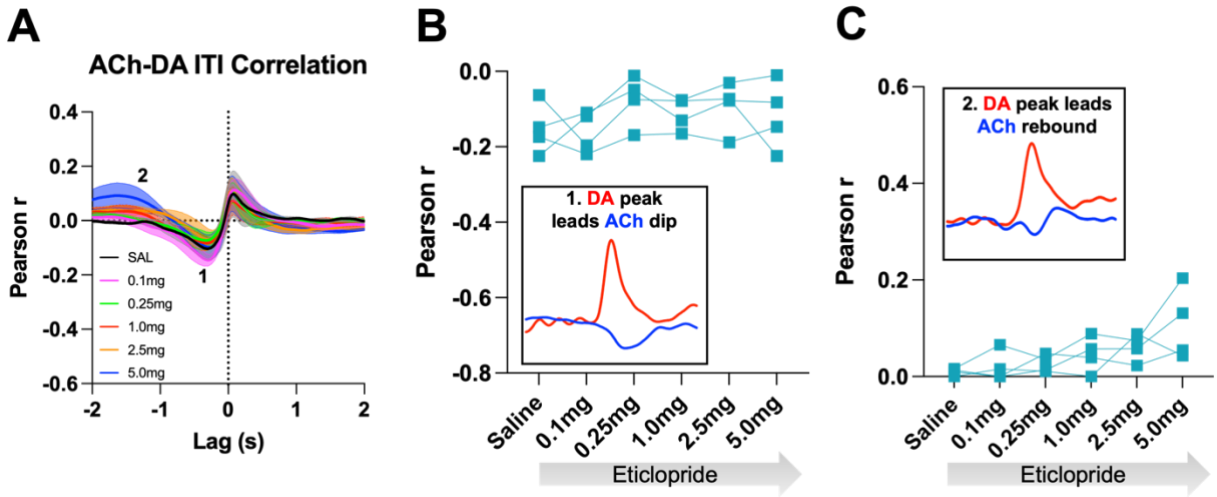

Figure 2.16. D2R antagonism does not alter general ACh-DA interactions in ChATDrd2KO mice during the ITI. (A) Correlation between ACh and DA during CRF trials with increasing doses of eticlopride in ChATDrd2KO mice. The ACh signal moved in front of or behind the DA signal to identify points of highest correlation. The first correlation is a negative correlation (1) with ACh lagging DA (Lag= -282.62 ms ± 93.81 ms) and the second correlation is a positive correlation (2) with ACh lagging DA (Lag= -1.33 s ± 0.189 s) **(B)** The negative correlation with the DA peak leading the ACh dip (inset) is not affected by eticlopride ($F_{(1.508, 4.523)}$ = 1.366, p = 0.3275). **(C)** The positive correlation with the DA peak leading the ACh rebound (inset) is not affected by eticlopride ($F_{(1.747, 5.240)}$ = 4.839, p = 0.0669).

2.3.8 D2R antagonism increases the latency to press in a CIN D2R dependent manner

We then determined if manipulating CIN D2R function affects behavior in the CRF task. Since D2R blockade induces catalepsy (Kharkwal et al., 2016) we wondered whether Drd2 ablation or D2R antagonism alters behavioral responding (latency to press in the task), an indicator of motivated behavior. In C57BL/6J mice, we found that eticlopride significantly increased lever press latency in a dose-dependent manner (Figures 2.17A). Eticlopride had no effect on lever press latency in ChATDrd2KO mice, compared to $Drd2^{fl/fl}$ control mice (Figure 2.17B). Next, we determined if the size of the stimulus induced ACh dip correlates with behavioral responding. To do this, we analyzed the correlation between the AUC and lever press latency for trials with press latencies > 2s to isolate the stimulus induced ACh dip from the lever press associated dip. In C57BL/6J mice, we found a positive correlation between total AUC and press latency (Figure 2.17C). Similarly, in $Drd2^{fl/fl}$ control mice, we found a similar positive correlation between total AUC and press latency (Figure 2.17E). This correlation was mainly driven by the ACh dip as the negative AUC positively correlated with press latency for C57BL/6J (Figure 2.17D) and $Drd2^{fl/fl}$ control (Figure 2.17F) mice. This positive correlation between total AUC and press latency was disrupted in ChATDrd2KO animals (Figure 2.17G), while there remained a weaker positive correlation between the negative AUC and press latency (Figure 2.17H). These data suggest that D2R mediated regulation of cholinergic ACh levels contribute to the regulation of motivated behavior.

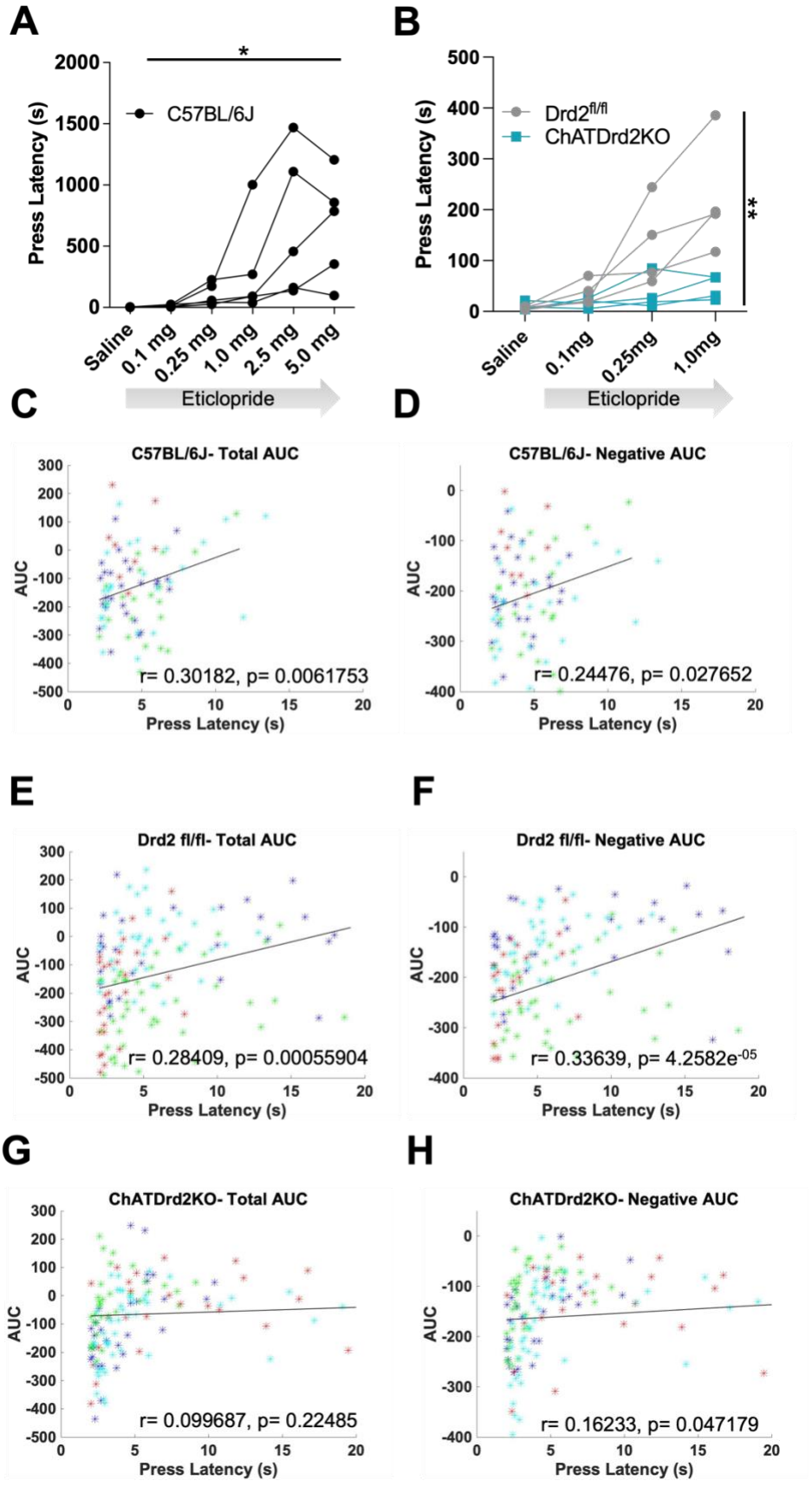
A
B
C
D
E
F
G
H
Press Latency (s)
C57BL/6J
Drd2fl/fl
ChATDrd2KO
Saline
0.1 mg
0.25 mg
1.0 mg
2.5 mg
5.0 mg
0.1mg
0.25mg
1.0mg
Eticlopride
*
**
C57BL/6J- Total AUC
C57BL/6J- Negative AUC
Drd2 fl/fl- Total AUC
Drd2 fl/fl- Negative AUC
ChATDrd2KO- Total AUC
ChATDrd2KO- Negative AUC
AUC
r= 0.30182, p= 0.0061753
r= 0.24476, p= 0.027652
r= 0.28409, p= 0.00055904
r= 0.33639, p= 4.2582e-05
r= 0.099687, p= 0.22485
r= 0.16233, p= 0.047179

Figure 2.17. Behavioral responding correlates with ACh event size but is affected by D2R antagonism and ablation. (A) Lever press latency is increased by eticlopride in a dose-dependent manner in C57BL/6J mice ($F_{(1.383, 5.533)} = 6.369$, $p = 0.0427$). **(B)** D2R antagonism does not increase lever press latency in ChATDrd2KO mice (blue squares) compared to $Drd2^{fl/fl}$ controls (gray circles) ($F_{(3,18)} = 5.664$, $p = 0.0065$, eticlopride x genotype). **(C)** Total AUC positively correlates with lever press latency in C57BL/6J mice ($r= 0.30182$, $p= 0.0061753$). **(D)** Negative AUC positively correlates with lever press latency in C57BL/6J mice ($r= 0.24476$, $p= 0.027652$). **(E)** Total AUC positively correlates with lever press latency in $Drd2^{fl/fl}$ control mice ($r= 0.28409$, $p= 0.00055904$). **(F)** Negative AUC positively correlates with lever press latency in $Drd2^{fl/fl}$ control mice ($r= 0.33639$, $p= 4.2582e^{-05}$). **(G)** Total AUC does not correlate with lever press latency in ChATDrd2KO mice ($r= 0.099687$, $p= 0.22485$). **(H)** Negative AUC positively correlates with lever press latency in ChATDrd2KO mice ($r= 0.16233$, $p= 0.047179$).

2.4 Discussion

Here, we investigated the mechanism by which striatal DA regulates cue induced changes in ACh levels during behavior. Understanding this mechanism is important because both neuromodulators coincidentally signal salient cues or outcomes during learning and motivated behavior and thus DA may regulate behavior via regulating ACh levels (Apicella et al., 1992). Moreover, it addresses the longstanding question of whether the ACh dip is fully dependent on striatal DA.

By simultaneously recording task-evoked DA and ACh levels in mice we made several observations: First, we observed that changes in striatal DA and ACh levels are induced by reward-predicting stimuli and the time locked signals develop in parallel with learning. Second, we found that pharmacological and genetic inactivation of D2Rs does not completely abolish the stimulus-induced dip in ACh, but it does shorten the dip and enhances rebound levels. Third, using correlational analysis, we found a relationship between DA and ACh that was strongest in response to lever extension as a reward predicting cue but still present during the inter-trial interval. This relationship was disrupted by D2R inactivation. Fourth, we found that D2R antagonism increased latency to lever press during behavior, but this was abolished when we inactivated CIN D2Rs.

Lastly, the size of the cue evoked ACh dip but not rebound correlated with lever press latency, even for lever presses that happen long after the cued signal has ended. Altogether, these findings indicate that DA and cholinergic D2Rs are necessary for controlling the shape of the ACh dip and the coordinated activity between DA and ACh during reward driven behaviors. Moreover, the cue induced ACh dip correlates and therefore may drive behavioral responding during motivated behavior.

2.4.1 Cue induced changes in striatal DA and ACh levels are time locked and develop in parallel with learning

The changes in DA and ACh levels that we recorded in the Pavlovian conditioning task are consistent with DA neurons and CINs encoding unexpected rewards and reward-predicting cues (Aosaki T, 1994b; Joshua et al., 2008; Morris et al., 2004; Schultz et al., 1997; Watanabe & Kimura, 1998). Like these previous studies, that assayed neuronal activity, we see a robust increase in DA levels and a decrease in ACh levels to unexpected reward that diminish as the reward becomes expected. These data show that the neurotransmitter levels of both, DA and ACh, follow neuronal activity of their respective neurons with a sub-second kinetics. The fast induction of the ACh dip is particularly striking as it suggests fast degradation or diffusion of ACh.

In addition, we observed similar changes in DA and ACh levels to the conditioned stimulus and not the unconditioned stimulus, which occur in parallel over learning. Our data is consistent with previous studies showing that both DA neurons and CINs respond to salient and conditioned stimuli. Moreover, we found that these changes in DA and ACh levels correlate with behavioral responding. However, this correlation was not observed in all tested mice due to the nature of the Pavlovian task. In the Pavlovian task, animals have the possibility to learn an association between the CS+ and reward. Consequently, some animals may show anticipatory responding during the

CS+ (head poking the reward port). However, because anticipatory behavior is not required to obtain a reward, some of the tested mice did not exhibit anticipatory head poking. Strikingly, the development of a DA and ACh signal over time indicate that these animals are nevertheless learning the stimulus-reward association. In conclusion, these data show that task-evoked changes in ACh and DA levels in mice follow what has been described at the level of neuronal activity level in primates (Joshua et al., 2008; Morris et al., 2004; Schultz et al., 1997)

2.4.2 D2R inactivation in CINs shortens but does not abolish the cue induced ACh dip

Salient and conditioned stimuli are known to induce pauses in CIN and TAN firing in rodents and primates, however, the dependence on DA and CIN D2Rs for pause induction is widely debated (Aosaki et al., 1994b; Morris et al., 2004; Watanabe & Kimura, 1998; Zhang & Cragg, 2017). Thus, to determine the role D2Rs play in modulating the stimulus induced ACh dip, we pharmacologically blocked D2Rs or selectively ablated Drd2 from CINs and measured ACh and DA levels in the CRF task. We found that D2R blockade or ablation shortened the ACh dip, which we quantified by calculating the dip duration. Moreover, in our control mice, D2R blockade also decreased the negative and increased total and rebound AUCs in a dose-dependent manner while the dip amplitude was unaffected. In contrast, D2R blockade had no effect on the stimulus induced ACh dip in ChATDrd2KO mice. This data reveals that the generation of ACh dip is not dependent on CIN D2Rs. Instead, cholinergic D2Rs are important for modulating the duration of the stimulus induced ACh dip. Our data provide clarity on the controversial role that DA plays in the regulation of the ACh dip and suggest that the stimulus induced ACh dip *in vivo* is not entirely DA or D2R-dependent as studies in primates and slice physiology studies have suggested (Aosaki et al., 1994b; Ding et al., 2010; Watanabe & Kimura, 1998). Our data further indicate that slice physiology studies in rodents where optogenetic stimulation of DA terminals or caged DA induced

CIN pauses are abolished by D2R antagonists or CIN-selective ChATDrd2KO mice are not fully capturing the natural pause (Augustin et al., 2018; Chuhma et al., 2014; Kharkwal et al., 2016; Straub et al., 2014; Wieland et al., 2014).

In addition to the effects on dip durations, we found that CIN D2Rs also regulate the level of ACh rebound levels, acting as a mechanism to constrain ACh rebounds after the dip. Currently, it is unknown which role the rebound in ACh plays during behavior. Generally, CINs are thought to inhibit spiny projections neurons (SPNs) via nicotinic activation of local interneurons or via muscarinic M_2/M_4-mediated inhibition of corticostriatal inputs (English et al., 2012b; Thomas W. Faust et al., 2015; Pavel Pakhotin & Enrico Bracci, 2007; Witten et al., 2010). Thus, a larger dip may lead to disinhibition and higher rebound to a stronger inhibition of SPNs. D2R antagonism decreases the first and enhances the second, which may inhibit movement initiation leading to the longer latency in lever pressing. Consistent with this, we observed that the size of the cue induced dip correlates with press latency (the larger the dip the shorter the latency). Surprisingly, this relationship also holds true for lever presses that were performed long after the cue induced ACh signal reverted to normal. This suggests that the cue-evoked dip signals the motivational state of the animal. This finding is consistent with recent inhibition studies in which CIN inhibition in the NAc during Pavlovian to Instrumental Transfer (PIT) enhanced the ability of the pavlovian cue to invigorate behavior (Collins et al., 2019). Note, however, that mice with selective Drd2 ablation do not show a deficit in PIT suggesting a more subtle deficit affecting latencies rather than the level of responding (Gallo et al., 2021).

2.4.3 Cue induced DA levels are unchanged in ChATDrd2KO mice

Slice and *in vivo* stimulation studies have shown that stimulation of ACh release induces DA release via activation of nicotinic receptors on DA terminals (Cachope & Cheer, 2014;

Cachope et al., 2012; David Sulzer et al., 2016; Threlfell et al., 2012). Thus, it is surprising that despite a change in cue induced ACh signal the cue-induced DA signal is largely unchanged. Especially, we hypothesized that the rebound in ACh levels observed in the eticlopride conditions may enhance or prolong the DA peak. Although the early induction of the DA peak is indeed enhanced in the eticlopride condition this is also observed in CIN-selective D2R KO mice, which do not show the eticlopride-associated rebound. We believe that the enhanced initial peak in the eticlopride condition is due to antagonism of D2 auto receptors on DA terminals. It is unclear why there is no additional effect on DA release as a consequence of the ACh rebound but it could be that under the behavioral conditions of the CRF task nicotinic receptors on DA neurons are desensitized.

2.4.4 DA and ACh correlation during task-dependent behaviors

Our approach to simultaneously image both DA and ACh in the same animal allowed us to examine the relationship between these two neuromodulators within trials. In both C57BL/6J and $Drd2^{fl/fl}$ control mice, we identified a strong negative correlation with the cue-induced DA release leading the ACh dip that is attenuated by D2R antagonism. This strong negative correlation between DA and ACh is significantly reduced in ChATDrd2KO mice compared to $Drd2^{fl/fl}$ controls, and ChATDrd2KO mice are unaffected by D2R blockade. We also found a weaker positive correlation with DA leading ACh that is enhanced by D2R antagonism. We believe that this positive correlation reveals the rebound in ACh activity following the dip that is blunted by D2R activation at baseline. This data suggests, as discussed above, that CIN D2Rs not only modulate the ACh dip but also the rebound activity.

Electrical stimulation in brain slices revealed on/off kinetics for dLight1.2 of 9.5 ms and 90 ms, respectively (Labouesse & Patriarchi, 2021). In contrast, the kinetics for the ACh3.0 sensor

after electrical stimulation were slower with on/off kinetics of 90-105 ms and 0.9-3.7 s, respectively (Jing et al., 2020). Here, we measured a decrease in striatal ACh levels in vivo with an onset of 206 ms after optogenetic inhibition of CINs. While this is significantly shorter than the reported on/off kinetics for the ACh3.0 sensor, it is possible that the lag of the cue induced decrease in ACh in relationship to the peak in DA is due to differences in the on/off kinetics of the two sensors and not a natural reflection of when both events occur.

2.4.5 Implications for behavior

What does an altered ACh signal mean for behavior? CIN selective D2R knock out mice learn the Pavlovian task presented in Figure 1 as well as control littermates (data not shown). This suggests that even with a shortened dip mice still can learn cue-reward associations. Similarly, we recently described that enhancing the dip durations by selective overexpression of D2Rs in CINs of the NAc (D2R-OE$_{NAcChAT}$ mice) did not affect Pavlovian learning but was associated with a deficit in Go/No-Go learning (Gallo et al., 2021). Opposite to what we found in this present study using ChATDrd2KO mice, rebound ACh levels are lower in D2R-OE$_{NAcChAT}$ mice. Thus, relatively higher rebound levels in wild-type control mice may suppress SPN activity and responding. However, rebound activity was significantly greater in incorrect No-Go (press) relative to correct No-Go (withhold) trials questioning this hypothesis. Rather we observed that the ACh dip in control mice was different between Go and No-Go trials, in contrast to what we observed in D2R-OE mice. This difference supports the hypothesis that D2R-OE mice did not learn the switch from the Go to the No-Go component of the task due to reduced contrast in ACh dip information between Go and No-Go contingencies. This is consistent with prior studies reporting that striatal ACh is not necessary for initial learning but is important for behavioral

performance when task contingencies change (Aoki et al., 2015; Bradfield et al., 2013; Brown et al., 2010; Favier et al., 2020; Okada et al., 2014; Okada et al., 2017; Ragozzino et al., 2009).

In this present study, we report that D2R antagonism increases lever press latency during CRF in a dose-dependent manner, but this is abolished in ChATDrd2KO mice. Moreover, lever press latency correlated with the AUC of the ACh signal, which was mostly driven by dip size (the larger the dip the shorter the latency) and was observed in trials with latencies longer than the ACh signal durations. As discussed, these data suggest that the cue-evoked dip signals the motivational state of the animal confirming a role of ACh in motivated behavior (Aosaki T, 1994b; Apicella, 2007; Apicella et al., 1991; Collins et al., 2019; Joshua et al., 2008; Kimura et al., 1984; Morris et al., 2004; Nougaret & Ravel, 2015; Ravel et al., 2003; Shimo & Hikosaka, 2001).

However, eticlopride enhanced press latency in controls but not KOs. This result demonstrates that acute inhibition of CIN D2Rs enhances press latency. Surprisingly, we did not measure an increase in press latencies in D2R KO mice, despite the shortened dip duration. This discrepancy may be due to compensation in the regulation of press latencies via other mechanisms that have developed during development or as the consequence of the chronic change in D2R levels. Such compensatory mechanisms may not have the time to develop in response to the acute action of the antagonist. However, the correlation between cue induced dip-duration and press latency could indeed be an epiphenomenon. E.g., D2R antagonism may change tonic ACh levels (Ikarashi et al., 1997) and cue induced changes in ACh, but it is only the first one that regulates press latencies. To address whether the duration of cue induced decrease in ACh levels alter press latencies, we will use optogenetic tools to artificially induce shorter and longer decreases in ACh at cue presentation. If the results are negative, then what we report is likely an epiphenomenon.

In conclusion, our data demonstrate that the stimulus induced ACh dip is multiphasic, encompassing a non-DA component and a DA component. Striatal DA is responsible for confining the temporal boundaries of the ACh dip and preventing rebound excitation via CIN D2R. Notably, we also find a positive correlation between the size of the stimulus induced ACh dip and behavioral responding, which implicates a role for ACh in motivated behaviors. Thus, further dissection of this system will provide a better understanding of the different components of the ACh dip and what each component represents for specific striatal functions and behavior.

2.5 Methods

Animals

Adult male and female C57BL/6J (JAX stock # 000664) mice were bred in house. For control and KO animals: double-transgenic mice were generated by crossing heterozygous ChAT-IRES-Cre mice (Rossi et al., 2011) (JAX stock #031661) to homozygous *Drd2*$^{fl/fl}$ (Drd2$^{fl/fl}$) mice (Bello et al., 2011). Control (Drd2$^{fl/fl}$) and ChATDrd2KO (ChAT-IRES-Cre x Drd2$^{fl/fl}$) mice are littermates, bred in house and back crossed onto C57BL/6J background. Mice were housed 1-4 per cage for most experiments on a 12-hr light/dark cycle, and all experiments were conducted in the light cycle. All experimental procedures were conducted following NIH guidelines and were approved by Institutional Animal Care and Use Committees by Columbia University and the New York State Psychiatric Institute.

Pharmacology

Intraperitoneal injections of saline, eticlopride (Tocris Cat. No. 1847) (0.1, 0.25, 1.0, 2.5 and 5.0 mg/kg) or scopolamine (Tocris Cat. No. 1414) (15 mg/kg) were administered 1hr before behavioral testing. To generate a dose-response curve with eticlopride, saline days alternated with drug days and the drug was administered in order from the lowest to highest dose.

Surgical procedures

Mice (≥ 8 weeks old) were induced with 4% isoflurane and maintained at 1-2% throughout the procedure. Mice were bilaterally injected with 450 nL/hemisphere with either AA5-hSYN-dLight1.2 (Addgene) (Patriarchi et al., 2018) or AA5-hSYN-ACh3.0 (Vigene) (Jing et al., 2020) (also known as GRAB-ACh3.0) into separate hemispheres of the dorsal medial striatum (DMS) using stereotactic Bregma-based coordinates: AP, +1.1mm; ML ± 1.4 mm; DV, -3.1 mm, -3.0 mm and -2.9 mm (150 nL/DV site). The on/off kinetics for the dLight1.2 sensor are 9.5 ms and 90 ms, respectively (Labouesse & Patriarchi, 2021). The on/off kinetics for the ACh3.0 sensor are 90-105 ms and 0.9-3.7 s, respectively (Jing et al., 2020). For the optogenetic inhibition experiment, mice were co-injected unilaterally with AA5-hSYN-ACh3.0 and AAV5-EF1α-DIO- eNpHR3.0 (UNC Vector Core) into the DMS. Following virus injection, 400-µm fiber optic cannulas (Doric, Quebec, Canada) were carefully lowered to a depth of -3.0 mm and fixed in place to the skull with dental cement anchored to machine mini-screws. Groups of mice used for experiments were housed in a counterbalanced fashion that accounted for sex, age, and home cage origin. Cannula-implanted mice began behavioral training 4 weeks after surgery. At the end of experiments, animals were perfused, and brains were processed post-*hoc* to validate virus expression and optic fiber location as in (Gallo et al., 2021).

In vivo fiber photometry and optogenetics

Fiber photometry equipment was set up using two 4-channel LED Driver (Doric) connected to two sets of a 405-nm LED and a 465-nm LED (Doric, cLED_405 and cLED_465). The 405 nm LEDs were passed through 405-410 nm bandpass filters, while the 465 nm LEDs were passed through a 460-490 nm GFP excitation filters using two 6-port Doric minicubes. 405 and 465 LEDs were then coupled to a dichroic mirror to split excitation and emission lights. Two low-

autofluorescence patch cords (400 µm/0.48NA, Doric) arising from the 2 minicubes were attached to the cannulas on the mouse's head and used to collect fluorescence emissions. These signals were filtered through 500-540 nm GFP emission filters via the same minicubes coupled to photodetectors (Doric, gain set to DC Low). Signals were sinusoidally modulated, using Synapse® software and RZ5P Multi I/O Processors (Tucker-Davis Technologies), at 210 Hz and 330 Hz (405 nm and 465 nm, respectively) to allow for low-pass filtering at 3 Hz via a lock-in amplification detector. 405 and 465 nm power at the patch cord were set to 30 µW or below. For acute optogenetic inhibition via eNpHR3.0, amber light (595 nm LED, Doric) was applied through the same optic fiber using a short and long optogenetic protocol: (i) 500 ms square pulses at 1mW + 1s ramp down; (ii) 500 ms + 3s ramp down. The 595 nm light was passed through a 580-680 F2 port (photodetector removed) of the same 6-port minicube (Pisansky et al., 2019). The optogenetic experiment was performed in the home cage.

Photometry data processing

All photometry and behavioral data utilized custom in-house MATLAB analysis scripts. Photometry signals were analyzed as time-locked events aligned to the lever extension (CRF) or tone onset (Pavlovian) of each trial. The 405 nm channel was used to control for potential noise/movement artifacts and the 465 nm channel was used to detect the conformational modulation of either the GACh3.0 sensor by ACh or the dLight1.2 sensor by DA. Both demodulated signals were extracted as a 15 s window surrounding the event, which was denoted as time = 0. Both signals were down sampled by a factor of 10 using a moving window mean. The change in fluorescence, ΔF/F (%), was defined as (F-F0)/F0 x 100, where F represents the fluorescent signal (465 nm) at each time point. F0 was calculated by applying a least-squares linear fit to the 405 nm signal to align with the 465nm signal (Calipari et al., 2016). To normalize signals

across animals and sessions, we calculated a local baseline fluorescence value for each trial using the average of the 5 s period preceding the event and subtracted that from the signal. The daily average GACh3.0 and dLight1.2 traces were calculated using session average traces from individual mice. ACh dip and DA peak amplitudes were calculated as the maximal change of the signal that was at least 1 or 2 STD below or above the local baseline, respectively. ACh dip duration was calculated using the last and the first zero crossings preceding and following the dip. Total AUC was calculated as the area of all three components of the ACh signal (initial peak, dip and rebound). Negative AUC was calculated as the area for only the negative component. Rebound AUC was calculated as the area of the positive component immediately following the dip. The AUC analysis was restricted to a 5 s time window following the task event. Individual CRF trial ($\Delta F/F$ (%)) traces were used for correlation analysis for CRF trials. For the ITI correlation, we examined any interaction between dLight1.2 and GACh3.0 regardless of event size during the variable 40 s of the ITI.

Operant apparatus

Four operant chambers (model Env-307w; Med-Associates, St. Albans, VT) equipped with liquid dippers were used. Each chamber was in a light- and sound-attenuating cabinet equipped with an exhaust fan, which provided 72-dB background white noise in the chamber. The dimensions of the experimental chamber interior were $22 \times 18 \times 13$ cm, with flooring consisting of metal rods placed 0.87 cm apart. A feeder trough was centered on one wall of the chamber. An infrared photocell detector was used to record head entries into the trough. Raising of the dipper inside the trough delivered a drop of evaporated milk reward. A retractable lever was mounted on the same wall as the feeder trough, 5 cm away. A house light located on the wall opposite to trough illuminated the chamber throughout all sessions.

Dipper training

Four weeks after surgery, mice underwent operant training. Mice were weighed daily and food- restricted to 85–90% of baseline weight; water was available *ad libitum*. In the first training session, 20 dipper presentations were separated by a variable inter-trial interval (ITI) and ended after 20 rewards were earned or after 30 min had elapsed, whichever occurred first. Criterion consisted of the mouse making head entries during 20 dipper presentations in one session. In the second training session, criterion was achieved when mice made head entries during 30 of 30 dipper presentations.

Pavlovian conditioning

Mice were trained for 16 consecutive days in a Pavlovian conditioning paradigm, which consisted of 12 conditioned stimulus-positive (CS+) trials and 12 unconditioned stimulus (CS-) trials occurring in a pseudorandom order. Each trial consisted of an 80-dB auditory cue presentation for 10 s of an 8 kHz tone or white noise (counterbalanced between mice) and after cue offset a milk reward was delivered only in CS+ trials, whereas no reward was delivered in CS- trials. There was a 100 s variable intertrial interval, drawn from an exponential distribution of times. Head entries in the food port were recorded throughout the session, and anticipatory head entries during the presentation of the cue were considered the conditioned response. Anticipatory responding was calculated as the difference in nose poking during the CS+ quintile with the maximum response (Q4 or 5) and the first quintile.

Continuous reinforcement schedule (CRF)

For lever press training, lever presses were reinforced on a continuous reinforcement (CRF) schedule. Levers were retracted after each reinforcer and were presented again after a variable ITI (average 40 s). The reward consisted of raising the dipper for 5 s. The session ended

when the mouse earned 60 reinforcements, or one hour elapsed, whichever occurred first. Sessions were repeated daily until mice achieved 60 reinforcements.

Data analysis

Sample sizes were determined by performing statistical power analyses based on effect sizes observed in preliminary data or on similar work in the literature. Statistical analyses were performed using GraphPad Prism 9 (GraphPad), MATLAB (MathWorks). Data are generally expressed as mean ± standard error of the mean (SEM). Paired and unpaired two-tailed Student's t-tests were used to compare 2-group data, as appropriate. Multiple comparisons were evaluated by one- or two-way ANOVA and Bonferroni's post hoc test, when appropriate. In rare cases of values missing in repeated measures samples, the data were analyzed by fitting a mixed effects model, as implemented by Prism 9. Photometry correlation analyses were performed using Pearson's correlation coefficients. A p-value of < 0.05 was considered statistically significant. Behavioral findings were replicated with mice from different litters, ages, or sexes. Investigators were blinded to the genotype of mice during behavioral assays as well as throughout the data analysis. Computer code for data analysis is available on Github (username: kmartyniuk1). Files are titled, "DA-ACh_Dualimaging_CRF" and "DA-ACh_Dualimaging_Pavlovian".

Chapter 3: The Role of Cholinergic Dopamine D2Rs in Behavior and Flexible Learning

3.1 Abstract

In the striatum, cholinergic interneuron (CIN) activity is modulated concomitant with phasic dopamine (DA) neuron activity in response to salient cues and reward outcomes, and both neuromodulators are believed to regulate each other. While this co-regulation has been examined using stimulation studies, the existence of mutual co-regulation during natural behavior is still largely unexplored. Here, we genetically inactivated D2Rs from CINs to determine how loss of DA regulation of CINs affects reward learning as well as flexible learning in mice. We found that loss of CIN D2Rs does not affect acquisition, extinction, or reinstatement in a Pavlovian task. In addition, we found no deficit in Pavlovian Instrumental Transfer (PIT) in mice with selective ablation of CIN D2Rs. However, we did find that loss of CIN D2Rs facilitated reversal learning in a probabilistic choice task, while acquisition was unaffected. We further explored how DA and ACh levels changed in both CIN D2R KO and control mice during acquisition and reversal learning in the probabilistic choice task. We found that changes in DA and ACh level encode the reward outcome during reversal learning.

3.2 Introduction

Dopamine (DA) neurons and cholinergic interneurons (CINs) both modulate their activity to unexpected rewards and reward predicting cues (Aosaki et al., 1994a; Apicella et al., 1997; Goldberg & Reynolds, 2011; Schultz, 2007; Schultz et al., 1997), and growing evidence suggests that they mutually co-regulate the activity of each other (Cachope & Cheer, 2014; Chuhma et al., 2014; Cragg, 2006; Kharkwal et al., 2016; Straub et al., 2014; D. Sulzer et al., 2016; Threlfell et

al., 2012). These studies demonstrate the DA inhibits CIN activity via DA D2 receptors (D2Rs), while CIN enhances phasic DA release via activation of nicotinic acetylcholine receptors (nAChRs). However, this co-regulation was primarily assessed using stimulation studies and the understanding of mutual co-regulation during natural behavioral is largely unknown.

Previous studies have shown that DA neurons and CINs encode distinct components of reward learning (Joshua et al., 2008; Morris et al., 2004). However, it is unclear if manipulating DA's influence on CINs or vice versa will affect reward learning. Here, we focused on the DA-regulation of CINs during reward learning. DA exerts it's influence on CINs via D2Rs, therefore, we genetically inactivated CIN D2Rs and first tested mice in various reward learning tasks. We found that selective loss of CINs D2Rs did not impair acquisition, extinction, or reinstatement in Pavlovian task. These results were not surprising as studies have shown that CINs are not necessary for initial learning, but rather they are they are important when task contingencies change (Aoki et al., 2015; Bradfield et al., 2013; Brown et al., 2010; Okada et al., 2017), which we will test. In addition, we tested if loss of CIN D2Rs affected cue-motivated behavior in a Pavlovian Instrumental Transfer (PIT) task. We found no difference in PIT learning in mice lacking CIN D2Rs compared to controls (Gallo et al., 2021). This finding indicates that loss of CIN D2Rs does not affect cue-induced invigoration of responding for a food reward.

Striatal DA neurons and CINs have been implicated in flexible learning (Adamantidis et al., 2011; Aoki et al., 2015; Bradfield et al., 2013; Brown et al., 2010; Clarke et al., 2011; Favier et al., 2020; Okada et al., 2014; Okada et al., 2017; Ragozzino et al., 2009; Rossi et al., 2013). However, it has yet to be determined if manipulating DA's regulation of CINs and vice versa affects flexible learning. To address this question, we tested mice lacking CIN D2Rs in probabilistic choice task. We found that selective loss of CIN D2Rs facilitates reversal learning

but not acquisition. This result indicates that CIN D2Rs regulate flexible learning when the reward outcome is uncertain. In addition, we monitored both striatal DA and ACh levels during acquisition and reversal learning in the probabilistic choice task. We found that changes in DA and ACh levels respond to the reward outcome during reversal learning.

3.3 Results

The first goal of this study was to determine the behavioral consequences of selectively ablating D2Rs from CINs. To address this question, we tested how CIN D2R inactivation affected reward association tasks, including Pavlovian conditioning, extinction, reinstatement, as well Pavlovian Instrumental Transfer (PIT). The second goal was to determine how loss of CIN D2Rs affect acquisition and reversal learning in a probabilistic choice task. The third goal was to examine both striatal dopamine (DA) and acetylcholine (ACh) levels during flexible learning.

3.3.1 Selective D2R ablation from CINs does not affect Pavlovian conditioning

We tested how selective ablation of D2Rs from CINs (ChATDrd2KO mice) affects associative learning in a classic Pavlovian conditioning task (Figure 3.1A). To determine learning, we measured anticipatory responding (# of head entries into the reward port) during the conditioned stimulus (CS+) and the unconditioned stimulus (CS-). To better visualize anticipatory responding, we plotted head entries during the CS+ and CS- in 2 sec bins. As mice learn the fixed duration of the CS+, their anticipatory response sharply increased over the 10 sec CS+ (Figure 3.1B). We found that both ChATDrd2KO and $Drd2^{fl/fl}$ control mice comparably increased their head entries over the duration of the CS+ and continued this pattern of anticipatory behavior across the 9 days of training (Figure 3.1B). However, we found no difference in anticipatory responding to the CS+ over training between the two groups. ChATDrd2KO and $Drd2^{fl/fl}$ mice did not display

any anticipatory response to the CS- (Figure 3.1C). This result indicates that loss of CIN D2Rs does not affect the acquisition of Pavlovian conditioning.

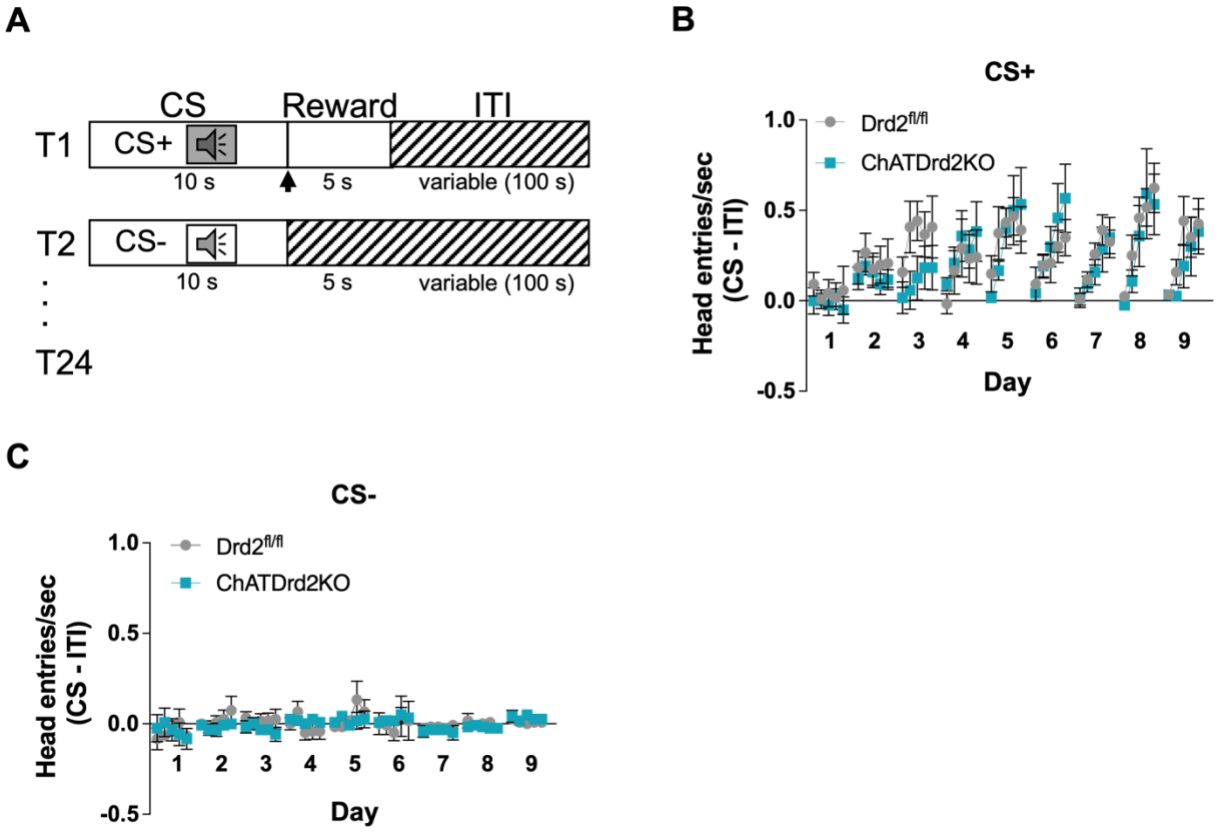

Figure 3.1. Selective ablation of D2Rs from CINs does not affect Pavlovian conditioning. (A) Task design. Mice were trained on 24 (12 CS+, 12 CS-) trials/day for 9 days. Each trial starts with a 10 sec tone (CS+ or CS-). At the end of the CS+, a dipper presents a milk food reward for 5 sec. There is an intertrial interval (ITI) variable in length (100 sec). **(B)** Within each day, the 10 sec CS+ was split into 5 two-second time bins to better visualize the anticipatory response. $Drd2^{fl/fl}$ (gray circles) and ChATDrd2KO (blue squares) mice sharply increased head entries into the reward port over the 10 sec CS+ and continued this anticipatory behavior over the 9 days of training ($F_{(44,352)}$= 5.605, p< 0.0001). There was no effect of genotype on learning ($F_{(44,352)}$= 0.7976, p= 0.8187). **(C)** Both $Drd2^{fl/fl}$ control and ChATDrd2KO mice showed no anticipatory response during the 10 sec CS- ($F_{(44,352)}$= 1.059, p= 0.3771). There was no effect of genotype on learning ($F_{(44,352)}$= 0.6286, p= 0.9693). N= 5/genotype.

3.3.2 Loss of CIN D2Rs does not affect extinction learning or reinstatement of a learned Pavlovian association

Next, we determined if loss of CIN D2Rs affected extinction learning or reinstatement of a learned Pavlovian association. To test the effects of D2R ablation from CINs on extinction learning, we ran our ChatDrd2KO and Drd2$^{fl/fl}$ control mice in a Pavlovian conditioning task (see above, Figure 3.1A), however, this time the CS+ was not followed by a milk reward. We found that both ChATDrd2KO and Drd2$^{fl/fl}$ control mice extinguished their anticipatory response to the

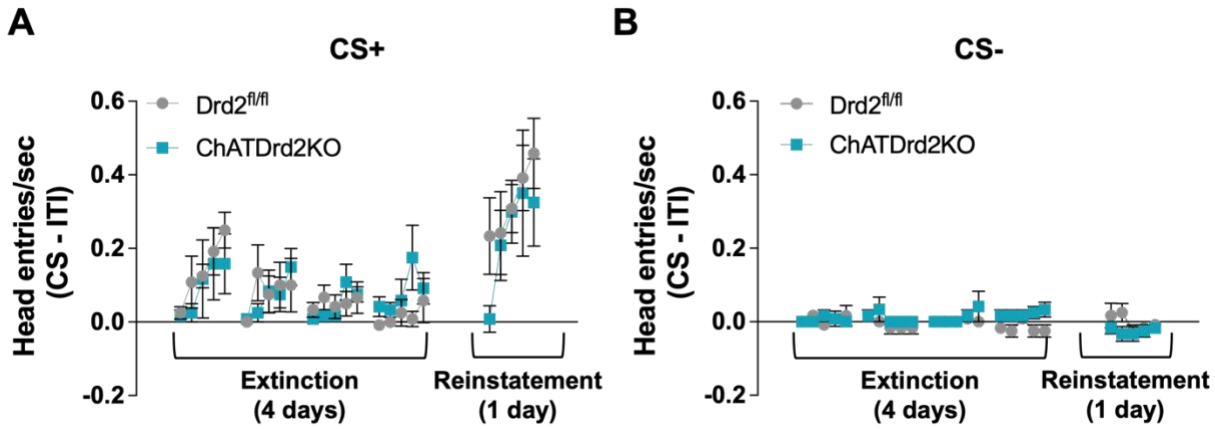

Figure 3.2. Loss of CIN D2Rs does not affect extinction learning or reinstatement of a learned Pavlovian association. (A) Both ChATDrd2KO and Drd2$^{fl/fl}$ control mice comparably reduced anticipatory responding the CS+ over 4 days of Pavlovian extinction (Day effect: $F_{(3.9, 31.20)}$= 3.159, p= .0282. There was no genotype effect on learning (Day x genotype: $F_{(19, 152)}$= 1.065, p= 0.3926). Both groups also similarly increased anticipatory responding to the CS+ during reinstatement of the CS+ (Day effect: $F_{(1.576, 12.61)}$= 4.879, p= 0.0333). There was no effect of genotype on reinstatement learning (Day x genotype: $F_{(4, 32)}$= 0.7939, p= 0.5379). **(B)** Neither ChATDrd2KO nor Drd2$^{fl/fl}$ control mice displayed any anticipatory responding to the CS- over the 4 days of Pavlovian extinction (Day effect: $F_{(3.746, 29.97)}$= 0.7661, p= 0.5481; Day x genotype effect: $F_{(19, 152)}$= 1.175, p= 0.2854) or reinstatement (Day effect: $F_{(2.134, 17.07)}$= 1.793, p= 0.1952; Day x genotype: $F_{(4, 32)}$= 1.574, p= 0.2052). N=5/genotype.

CS+ after 4 days of training (Figure 3.2A). Again, we found no difference in extinction learning between the two groups. In addition, both ChatDrd2KO and Drd2$^{fl/fl}$ control mice did not show any anticipatory response to the CS- (Figure 3.2B). Next, we determined if loss of CIN D2Rs affected reinstatement learning of the CS+ after extinction. We found that ChATDrd2KO and Drd2$^{fl/fl}$ mice comparably reinstated their anticipatory response to the CS+ after 1 day of training (Figure 3.2A). There was no change in anticipatory behavior to the CS- during reinstatement for

ChATDrd2KO or Drd2$^{fl/fl}$ control mice (Figure 3.2B). Together, these results suggest that CIN D2Rs do not regulate extinction learning or reinstatement of a previously learned Pavlovian association.

3.3.3 Loss of CIN D2Rs does not affect Pavlovian Instrumental Transfer

Next, we determined if selective ablation of CIN D2Rs affected Pavlovian Instrumental Transfer (PIT). We found no difference in Pavlovian responding to the CS+ or on instrumental responding on a random-ratio (RR) schedule between ChATDrd2KO and Drd2$^{fl/fl}$ mice (Figure 3.3A-B). During the transfer phase, we found a trending but non-significant increase in lever pressing rate during the CS+ compared to the CS^{Ø} and ITI for both groups, indicating successful PIT (Figure 3.3C). However, the ChATDrd2KO mice displayed similar patterns of responding during the PIT test compared to the Drd2$^{fl/fl}$ control mice. This result suggests that loss of CIN D2Rs does affect cue-induced invigoration of responding for a food reward (Gallo et al., 2021).

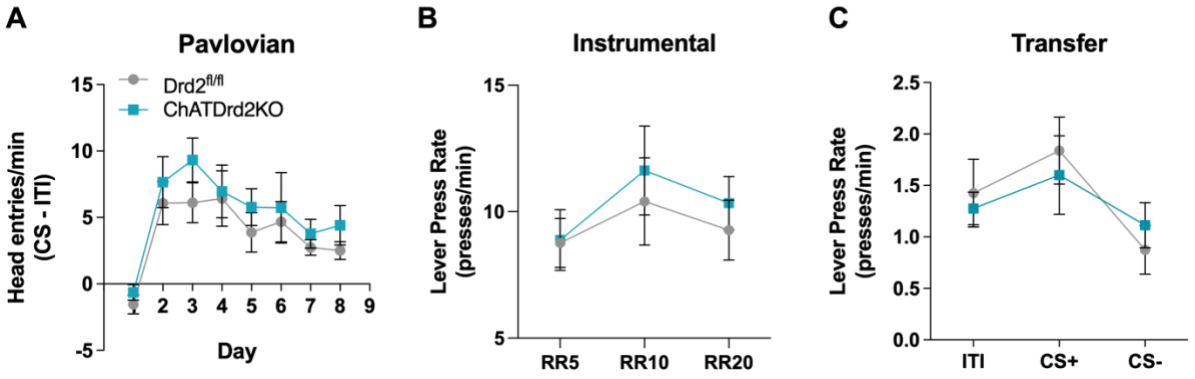

Figure 3.3. Loss of CIN D2Rs does not impact PIT. ChATDrd2KO and Drd2$^{fl/fl}$ mice did not significantly differ in **(A)** Pavlovian responding to the CS+ (genotype effect: $F_{(1,14)}$= 1.018, p= 0.33; genotype x day interaction: $F_{(7, 98)}$= 0.25, p= 0.97), N= 8/genotype, **(B)** Instrumental responding (genotype effect: $F_{(1,12)}$= 0.2834, p= 0.60; genotype x ratio interaction: $F_{(2,24)}$= 0.17, p= 0.84, N= 7/genotype, **(C)** Transfer phase (genotype effect: $F_{(1,12)}$= 0.03, p= 0.86; genotype x CS period interaction: $F_{(2,36)}$= 0.39, p= 0.6778, N= 7/genotype of the assay. Lever pressing did increase during the CS+ compared to the CS^{Ø} and ITI (CS period effect: $F_{(2,36)}$= 3.236, p= 0.0510) for both groups.

3.3.4 Selective loss of CIN D2Rs does not impair acquisition of a probabilistic choice task

To determine how loss of CIN D2Rs affects learning when a reward outcome is uncertain, we tested ChATDrd2KO and Drd2$^{fl/fl}$ control mice in a probabilistic choice task (Figure 3.4). We found no difference in acquisition learning over 5 days of testing in the probabilistic choice task between ChATDrd2KO and Drd2$^{fl/fl}$ control mice (Figure 3.5A). There was no significant difference in the number of sessions required to reach criterion determined for learning (80% "high payoff" lever choice in choice trials over 2 consecutive days) (Figure 3.5B). This result indicates that selective loss of CIN D2Rs does not impact acquisition learning in the probabilistic choice task.

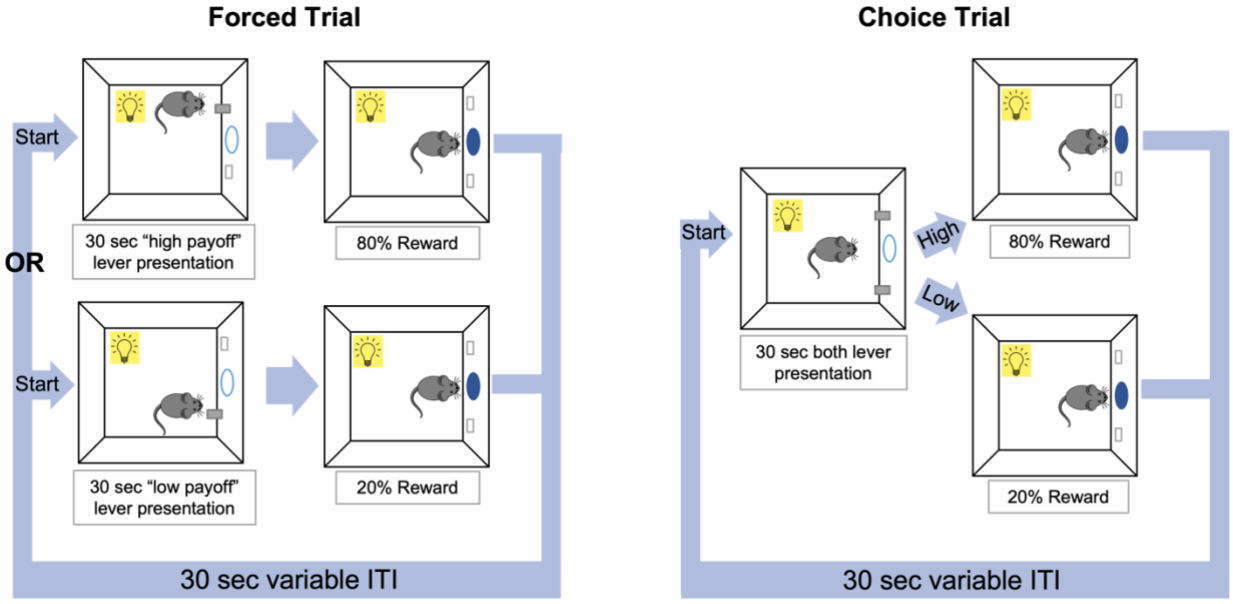

Figure 3.4. Probabilistic choice task design. Each session has 80 trials: 40 forced and 40 choice trials. Trial type was pseudo-randomly ordered with 5 forced and 5 choice trials per bock. In forced trials (left), only 1 lever is presented at a time for 30 sec. The "high payoff" lever grants a reward 80% of the time while the "low payoff" lever only 20% of the time. In choice trials (right), both levers are presented at the same time for 30 sec and the mice can press either the "high payoff" or "low payoff" lever. In both trial types, mice must make a choice to press a lever within 30 sec or the lever(s) are retracted, and a new trial starts. There ITI is variable in length with a mean of 30 sec.

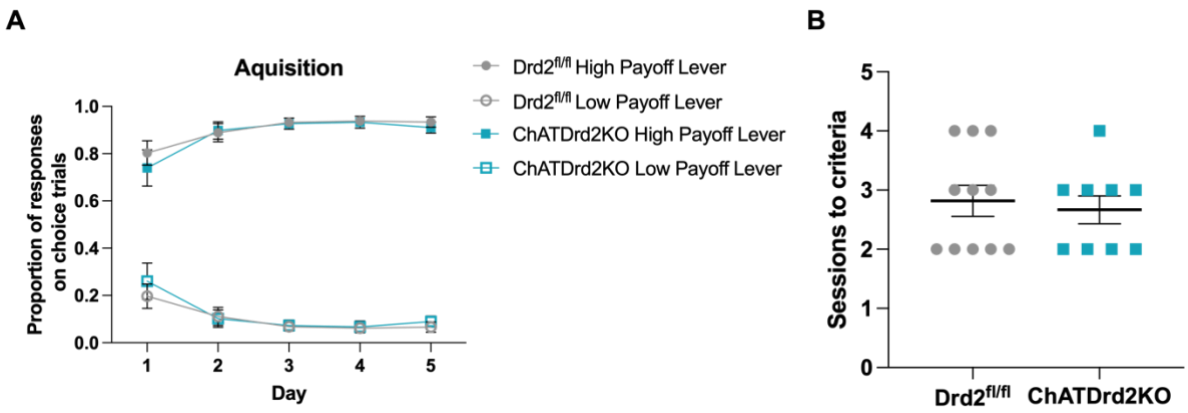

Figure 3.5. Loss of CIN D2Rs does not affect acquisition learning in a probabilistic choice task. (A) ChATDrd2KO and Drd2$^{fl/fl}$ control mice comparably learned the task contingencies (Day effect: $F_{(1.751, 31.52)}$= 10.79, p= 0.0004). There was no difference in learning between the two groups (Day x genotype effect: $F_{(4,72)}$= 0.4631, p= 0.7625, ChATDrd2KO mice (N= 9), Drd2$^{fl/fl}$ mice (N=11). **(B)** No difference in the number of sessions required to reach criterion between the two groups (p= 0.6800, Unpaired t-test).

3.3.5 Selective ablation of D2Rs from CINs facilitates reversal learning in a probabilisitic choice task

Next, we determined if selective loss of CIN D2Rs affected reversal learning in a probabilistic choice task. The task design was the same as acquisition (Figure 3.4), however, the "high payoff" lever and "low payoff" lever were switched. We found that ChATDrd2KO mice were faster at learning the new rule change compared to Drd2$^{fl/fl}$ control mice (Figure 3.6A). In

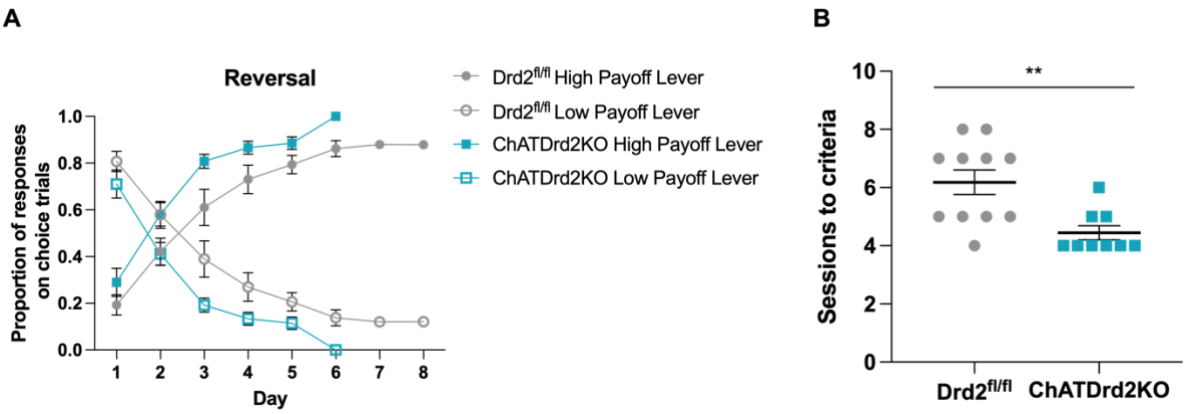

Figure 3.6. Loss of CIN D2Rs facilitates reversal learning in a probabilistic choice task. (A) ChATDrd2 KO mice were faster at reversal learning compared to Drd2$^{fl/fl}$ control mice. **(B)** ChATDrd2 KO mice required significantly few sessions to reach criterion during reversal compared to Drd2$^{fl/fl}$ control mice (p= 0.0035, Unpaired t-test). N=9 for ChATDrd2 KO mice, N=11 for Drd2$^{fl/fl}$ control mice.

addition, ChATDrd2KO mice required significantly fewer sessions to reach criterion compared to Drd2$^{fl/fl}$ control mice (Figure 3.6B). Moreover, we also explored potential differences in win-stay/lose-shift (WSLS) strategies between ChATDrd2KO and control mice during acquisition and reversal in the probabilistic choice task. To do this, we compared advantageous strategies (win-stay, lose-shift) and disadvantageous strategies (win-shift, lose-stay) and generated an adaptive score (win-stay + los-shift) – (win-shift + lose-stay). We found that while both ChATDrd2KO and Drd2$^{fl/fl}$ improve in WSLS strategies during the acquisition phase, there was no difference between the two groups over learning (Figure 3.7A). Interestingly, we found that ChATDrd2KO mice developed advantageous strategies quicker during reversal learning compared to control mice (Figure 3.7B). These results suggest that CIN D2Rs inhibit flexible learning when task contingencies change.

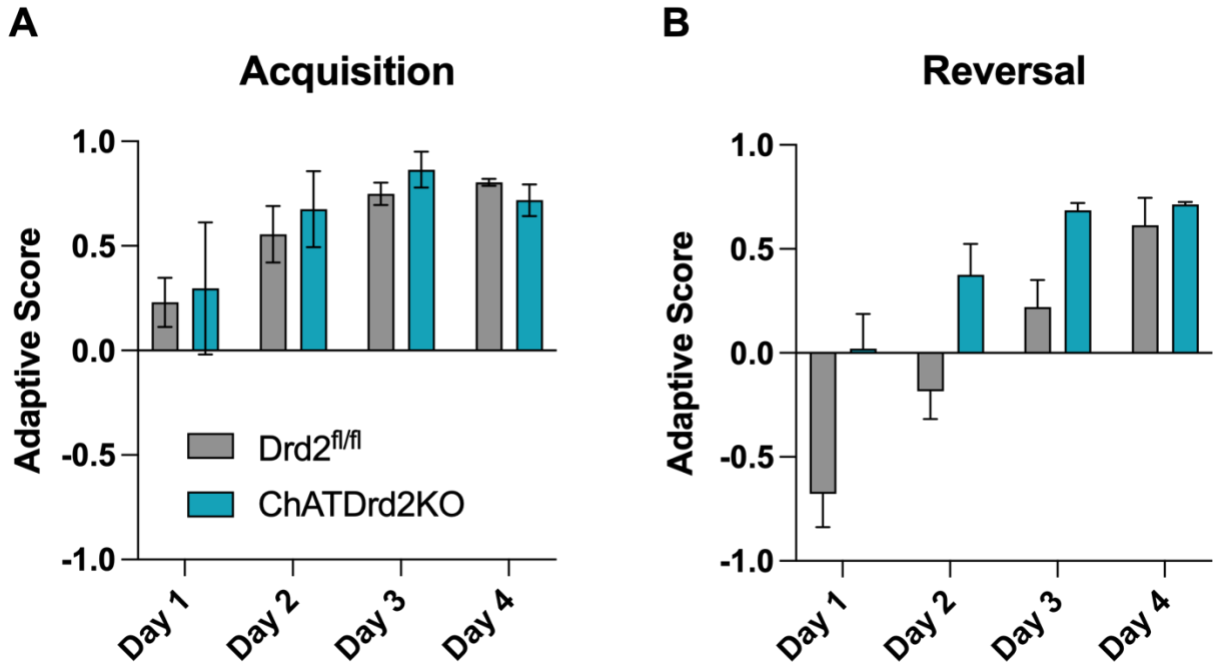

Figure 3.7. ChATDrd2KO mice develop advantageous strategies faster than controls during reversal learning. (A) Both ChATDrd2KO and Drd2$^{fl/fl}$ mice improved WSLS strategies during acquisition (Day effect: $F_{(1.407, 8.445)}$= 11.71, p= 0.0056). There was no effect of genotype on performance (Day x genotype effect: $F_{(3, 18)}$= 0.4456. **(B)** Both ChATDrd2KO and Drd2$^{fl/fl}$ mice improved WSLS strategies during reversal (Day effect: $F_{(1.971, 11.82)}$= 40.93, p< 0.0001. ChATDrd2KO mice improved faster compared to Drd2$^{fl/fl}$ control mice (Day x genotype effect: $F_{(3, 18)}$= 3.544, p= 0.0355).

3.3.6 DA and ACh signals encode reward outcome during acquisition

Next, we explored how DA and ACh levels change during acquisition and reversal learning in the probabilistic choice task. We aligned the ACh and DA signals to the lever press to examine reward outcome (expected reward, unexpected reward, expected reward omission and unexpected reward omission) and compared between ChATDrd2KO and Drd2$^{fl/fl}$ control mice. During the acquisition phase, we observed no differences in ACh levels between ChATDrd2KO and Drd2$^{fl/fl}$ control mice when aligned to the reward outcome (Figure 3.8). However, we observed a decrease

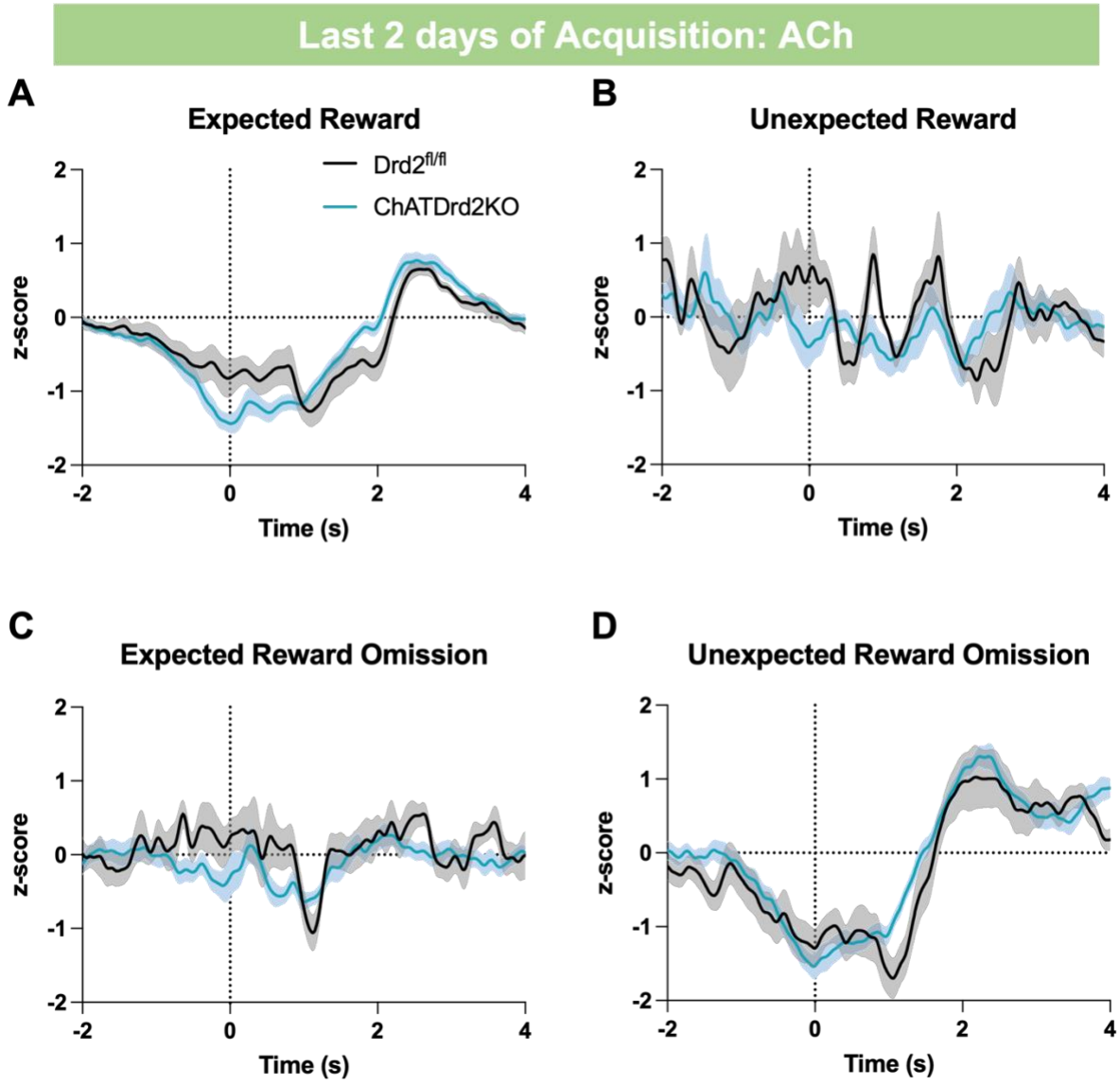

Figure 3.8. ACh levels signal reward expectation and outcome during acquisition. ACh signals were aligned to the lever press (Time= 0 s) for different reward outcomes, **(A)** expected reward, **(B)** unexpected reward, **(C)** expected reward omission, and **(D)** unexpected reward omission. Data is combined from the last 2 days of acquisition. N= 3 for $Drd2^{fl/fl}$ control mice, N=4 for ChATDrd2KO mice.

in ACh levels prior to the lever press (Time= 0 s) in trials where the mice expected to receive a reward, which may signal reward expectation (Figure 3.8A-D). In addition, after the lever press in trials where the mice expected a reward, we observed an increase in ACh levels that may signal the reward outcome (Figure 3.8A-D). Surprisingly, we did not observe any changes in ACh levels in trials where there was an unexpected reward (Figure 3.8B). Moreover, we did not observe any changes in ACh levels in trials where the mice expected to not receive a reward (Figure 3.8C), likely because they have learned that the "low payoff" lever is unlikely to end in a reward. These results indicate the ACh signals reward outcomes during initial learning, and loss of CIN D2Rs does not affect these ACh signals. These results are consistent with our behavioral finding that loss of CIN D2Rs does not affect performance in the acquisition phase of the probabilistic choice task (Figure 3.5).

In addition, we observed no differences in DA levels to reward outcome between ChATDrd2KO and $Drd2^{fl/fl}$ control mice during acquisition (Figure 3.9). However, we see changes in DA levels that are consistent with reward prediction errors. In both groups, DA levels increased before the lever press in trials where a reward was expected (Figure 3.9A-D). In addition, DA levels robustly increased after the lever press (Time= 0 s) in trials with an unexpected reward (Figure 3.9B) and decreased in trials with unexpected reward omission (Figure 3.9D). Notably, increases in DA levels appear to be slightly lower after the lever press (Time= 0 s) in trials with expected reward (Figure 3.9A), compared to trials with unexpected reward (Figure 3.9B). In addition, the decrease in DA levels after lever press in trials with unexpected reward omission (Figure 3.9D) appear to be larger than trials with expected reward omission (Figure 3.9C) These

results are consistent with DA differently responding to expected vs unexpected rewards. Moreover, we observed no difference between DA levels between ChATDrd2KO and Drd2$^{fl/fl}$ control mice, consistent with no difference in performance during the acquisition phase (Figure 3.5).

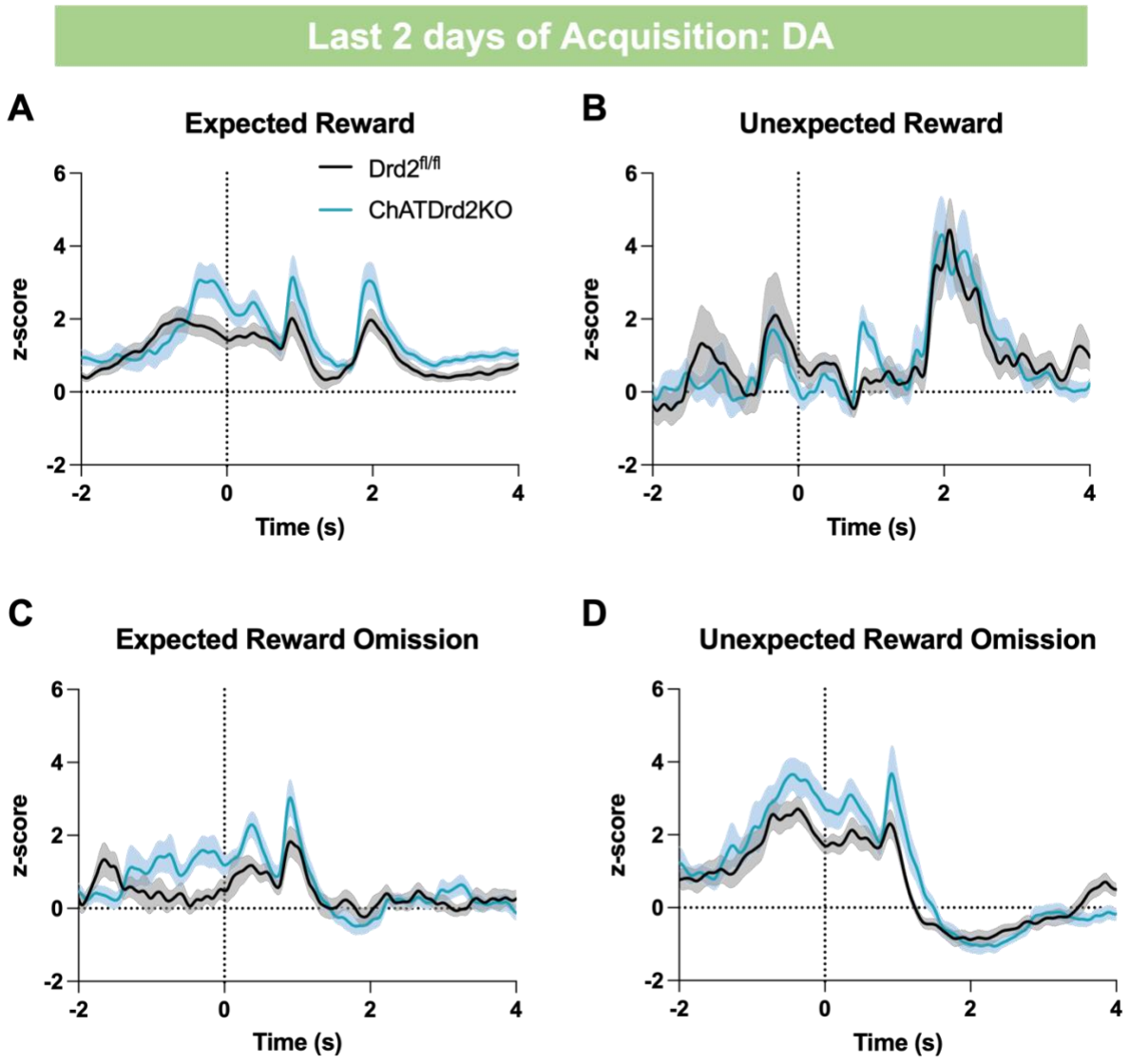

Figure 3.9. DA levels signal reward prediction errors during acquisition. DA signals were aligned to the lever press (Time= 0 s) for different reward outcomes, **(A)** expected reward, **(B)** unexpected reward, **(C)** expected reward omission, and **(D)** unexpected reward omission. Data is combined from the last 2 days of acquisition. N= 3 for Drd2$^{fl/fl}$ control mice, N=4 for ChATDrd2KO mice.

3.3.7 DA and ACh levels singal reward expectation and outcomes in reversal learning

Like acquisition, we observed no difference in ACh levels to reward outcome between ChATDrd2KO and Drd2$^{fl/fl}$ control mice during the first two days of reversal learning (Figure

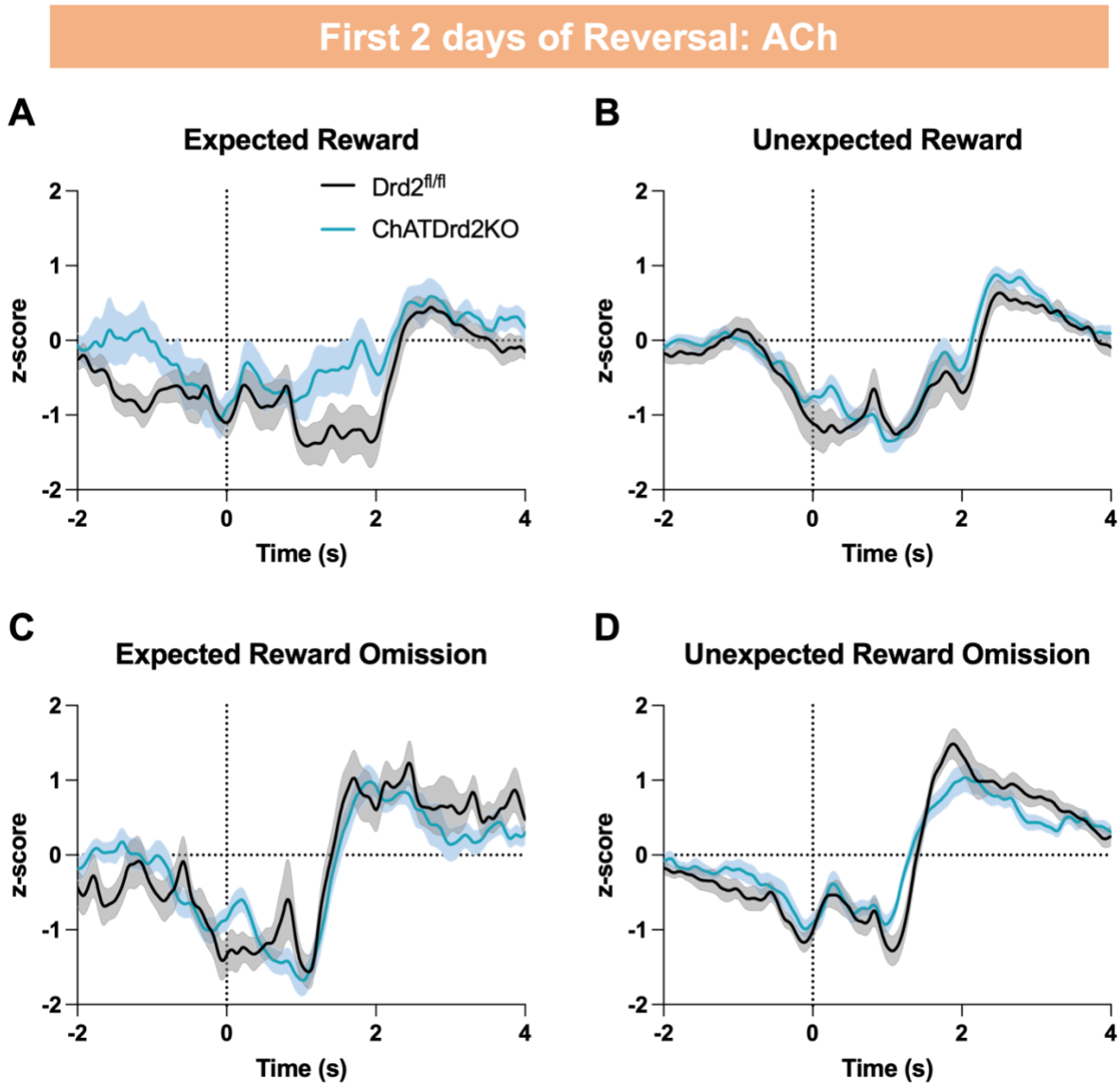

Figure 3.10. ACh levels during the early phase of reversal learning. ACh signals were aligned to the lever press (Time= 0 s) for different reward outcomes, **(A)** expected reward, **(B)** unexpected reward, **(C)** expected reward omission, and **(D)** unexpected reward omission. Data is combined from the first 2 days of reversal. N= 3 for Drd2$^{fl/fl}$ control mice, N=4 for ChATDrd2KO mice.

3.10). In contrast to the acquisition phase, we did observe changes in ACh levels in trials with unexpected reward during early reversal (Figure 3.10B). After the lever press (Time= 0 s) we observed a large decrease followed by an increase in ACh levels in both ChATDrd2KO and

Drd2$^{fl/fl}$ mice (Figure 3.10B). In addition, ACh levels appear to be comparably responding to expected and unexpected reward omission (Figure 3.10C-D), likely due to confusion in the change in task contingencies during early reversal. These results indicate that loss of CIN D2Rs does not affect ACh levels to reward outcome during the early phase of reversal learning. This is consistent with both ChATDrd2KO and Drd2$^{fl/fl}$ mice performing similarly during the first two days of reversal (Figure 3.6A).

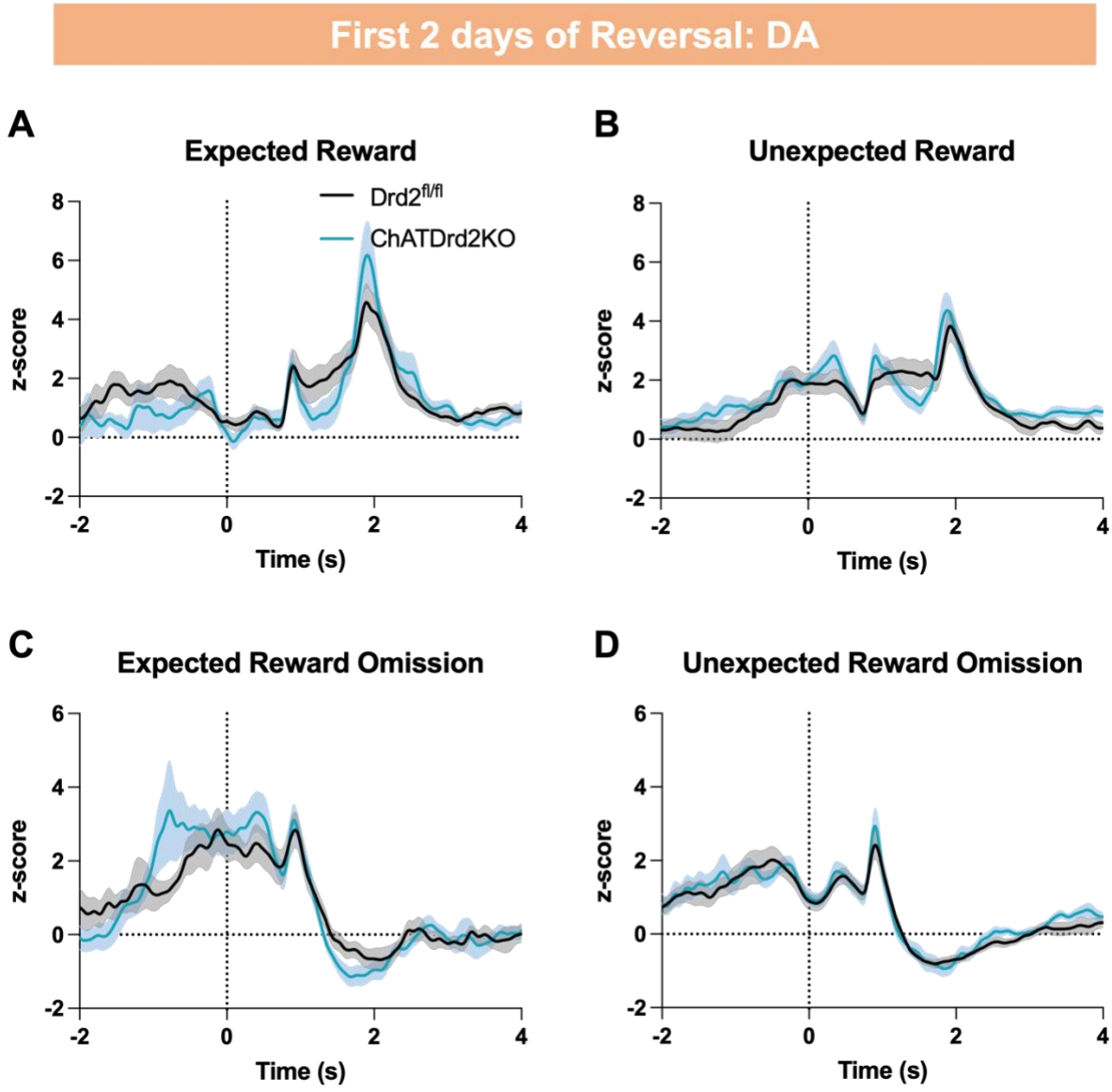

Figure 3.11. DA levels signal reward prediction errors during early reversal learning. DA signals were aligned to the lever press (Time= 0 s) for different reward outcomes, **(A)** expected reward, **(B)** unexpected reward, **(C)** expected reward omission, and **(D)** unexpected reward omission. Data is combined from the first 2 days of reversal. N= 3 for Drd2$^{fl/fl}$ control mice, N=4 for ChATDrd2KO mice.

Like acquisition, DA levels signal reward prediction errors in the early phases of reversal learning (Figure 3.11). We observed increased DA levels in trials with expected and unexpected reward (Figure 3.11A-B) and decreased DA levels in trials with expected and unexpected reward omission (Figure 3.11C-D). Again, we observed no difference in DA levels to reward outcome between ChATDrd2KO and Drd2$^{fl/fl}$ mice. These results support our behavioral data where both groups perform comparably during early reversal learning (Figure 3.6).

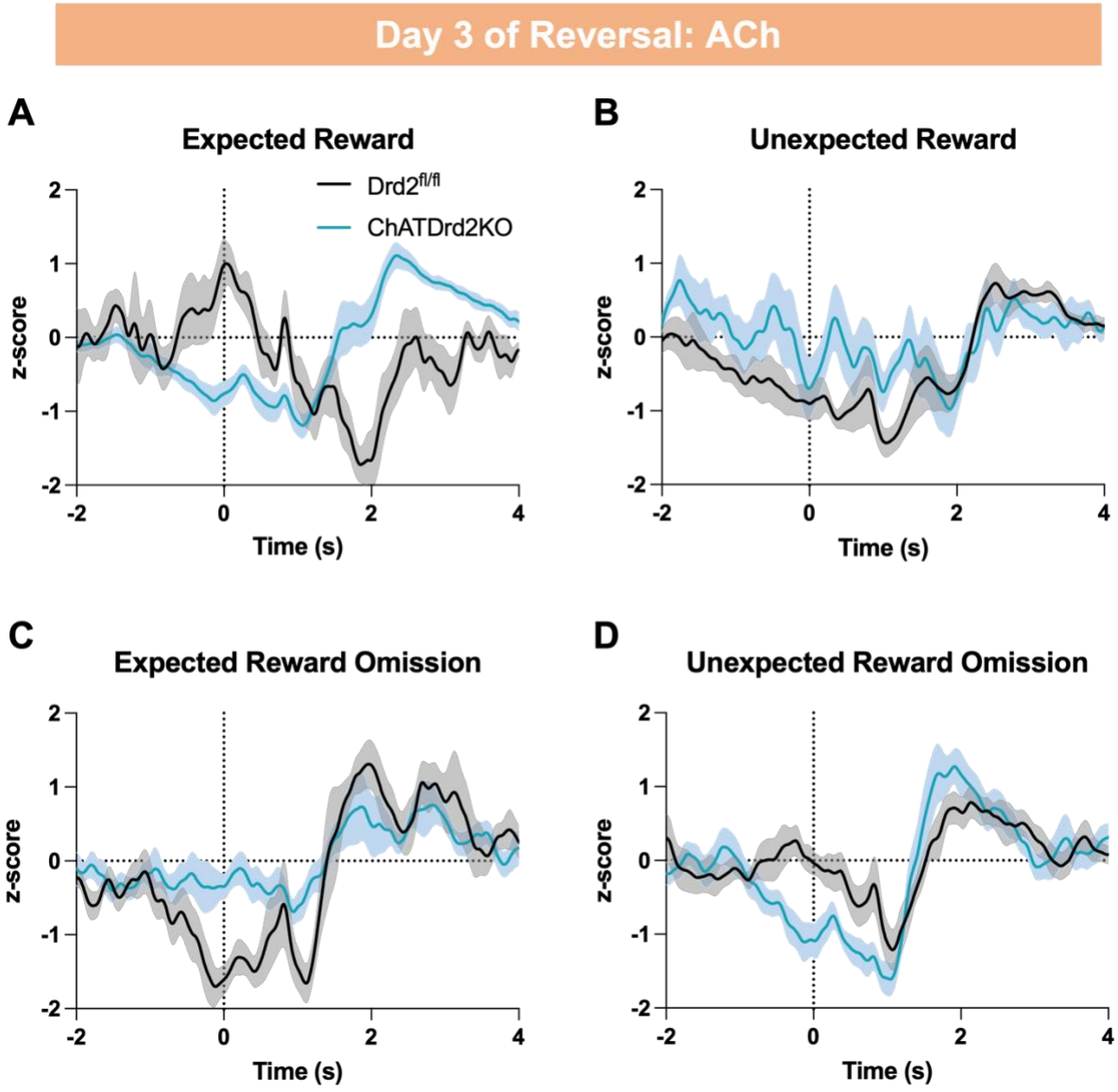

Figure 3.12. ACh levels to reward outcomes differ between ChATDrd2KO and Drd2$^{fl/fl}$ mice on reversal Day 3. ACh signals were aligned to the lever press (Time= 0 s) for different reward outcomes, **(A)** expected reward, **(B)** unexpected reward, **(C)** expected reward omission, and **(D)** unexpected reward omission. N= 3 for Drd2$^{fl/fl}$ control mice, N=4 for ChATDrd2KO mice.

Interestingly, on Day 3 of reversal, we observed differences in ACh levels to the reward outcome between ChATDrd2KO and Drd2$^{fl/fl}$ mice (Figures 3.12). Notably, in ChATDrd2KO mice, we observed larger rebound ACh levels in trials with expected reward (Figure 3.12A) and trials with unexpected reward omission compared to Drd2$^{fl/fl}$ control mice (Figure 3.12D). This result suggests that rebound ACh levels may be important for learning when task contingencies change and supports our behavioral data showing that ChATDrd2KO mice are performing significantly better on Day 3 of reversal compared to Drd2$^{fl/fl}$ controls (Figure 3.6).

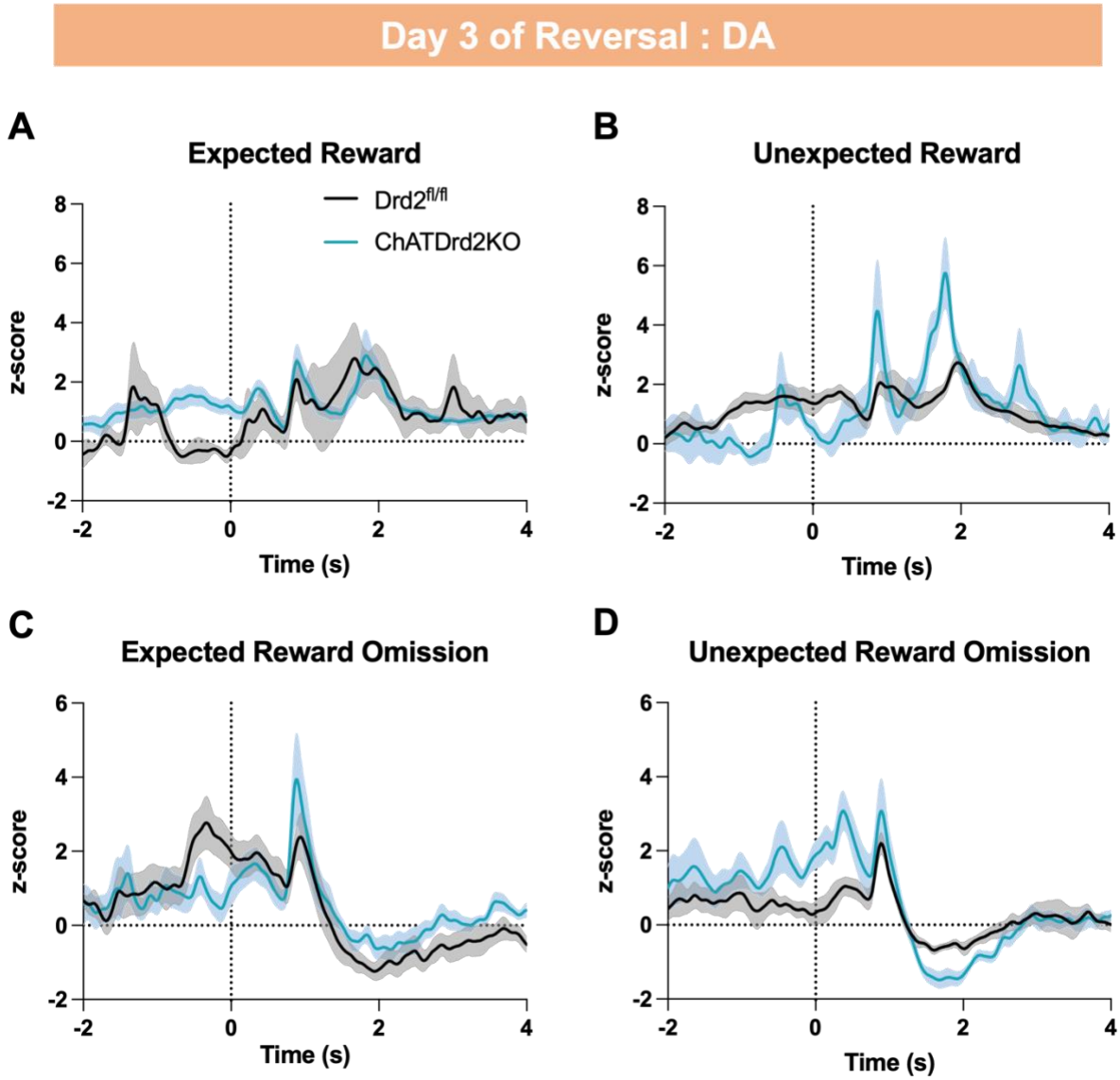

Figure 3.13. DA levels to reward outcome differ between ChATDrd2KO and Drd2$^{fl/fl}$ mice on Day 3 of reversal. DA signals were aligned to the lever press (Time= 0 s) for different reward outcomes, **(A)** expected reward, **(B)** unexpected reward, **(C)** expected reward omission, and **(D)** unexpected reward omission. N= 3 for Drd2$^{fl/fl}$ control mice, N=4 for ChATDrd2KO mice.

In addition, we observed differences in DA levels to reward outcome between ChATDrd2KO and Drd2$^{fl/fl}$ mice on Day 3 of reversal (Figure 3.13). Notably, we observed a larger decrease in DA levels in trials with unexpected reward omission in ChATDrd2KO mice compared to Drd2$^{fl/fl}$ mice (Figure 3.13D), which may aid in faster reversal learning. Again, this result supports our behavioral finding that mice lacking CIN D2Rs are performing better on Day 3 of reversal compared to control (Figure 3.6).

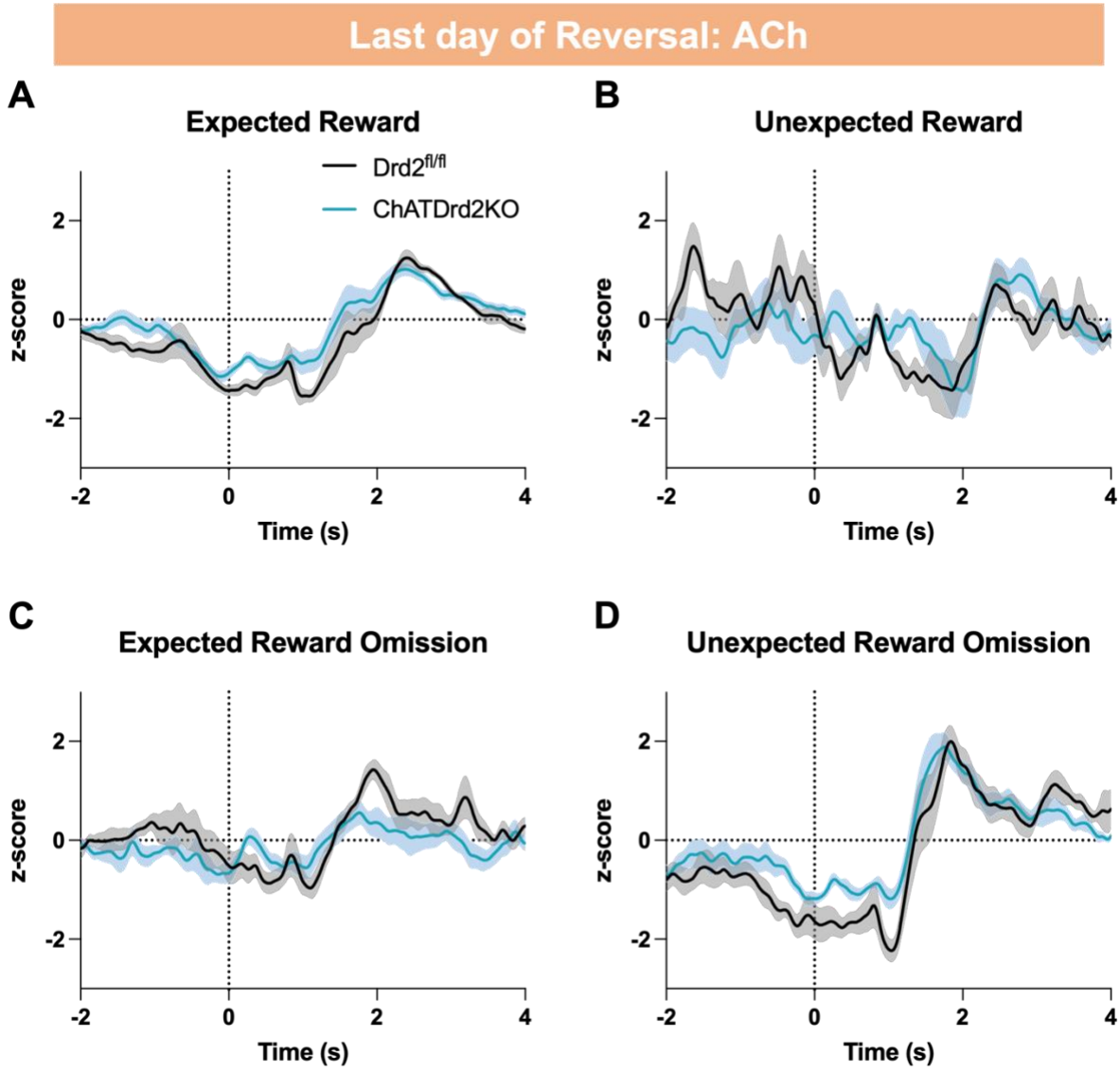

Figure 3.14. ACh levels signal reward outcome on the last day of reversal. ACh signals were aligned to the lever press (Time= 0 s) for different reward outcomes, **(A)** expected reward, **(B)** unexpected reward, **(C)** expected reward omission, and **(D)** unexpected reward omission. N= 3 for Drd2$^{fl/fl}$ control mice, N=4 for ChATDrd2KO mice.

On the last day of reversal learning, ACh levels to reward outcome are comparable between ChATDrd2KO and Drd2$^{fl/fl}$ mice (Figure 3.14). In both groups, we observed an increase in rebound ACh levels in trials with expected reward (Figure 3.14A) and trials with unexpected reward omission (Figure 3.14D). In addition, we observed no differences in DA levels to reward outcome between ChATDrd2KO and Drd2$^{fl/fl}$ mice on the last day of reversal learning (Figure 3.15). These results support our behavioral data showing that at the end of reversal learning both ChATDrd2KO and Drd2$^{fl/fl}$ mice have learned the new rule change (Figure 3.6).

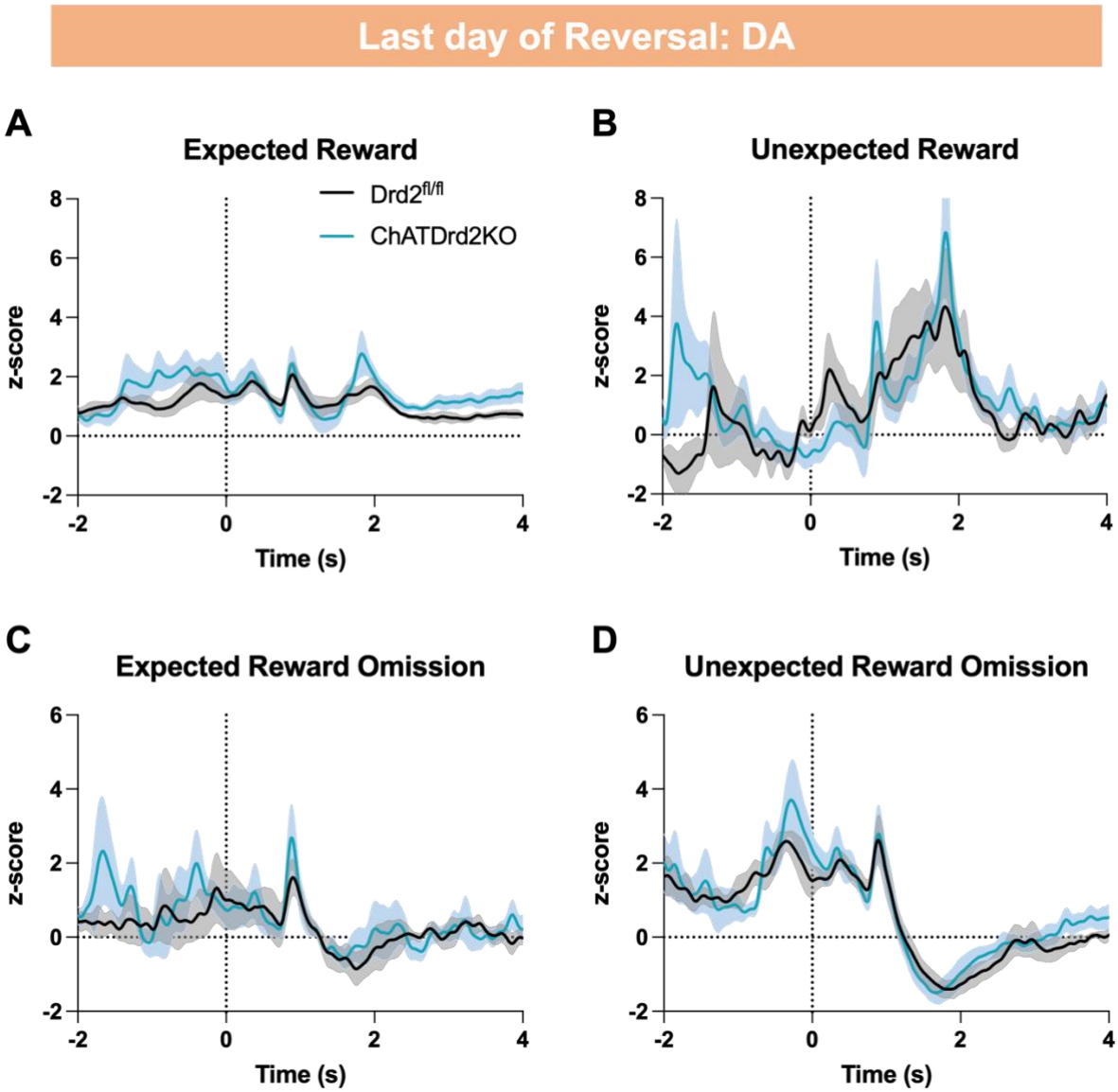

Figure 3.15. DA levels signal reward prediction errors on the last day of reversal. DA signals were aligned to the lever press (Time= 0 s) for different reward outcomes, **(A)** expected reward, **(B)** unexpected reward, **(C)** expected reward omission, and **(D)** unexpected reward omission. N= 3 for $Drd2^{fl/fl}$ control mice, N=4 for ChATDrd2KO mice.

3.4 Discussion

Here, we examined the role cholinergic D2Rs play in reward and flexible learning. We found that loss of CIN D2Rs did not affect reward learning in a classic Pavlovian task. In addition, we found that mice lacking CIN D2Rs were faster at reversal learning in a probabilistic choice task, while initial learning was unaffected. Moreover, we observed no difference in ACh or DA levels to reward outcome during the acquisition phase or early reversal learning between ChATDrd2KO and $Drd2^{fl/fl}$ mice. However, we observed a difference in ACh and DA levels to reward outcome on Day 3 of reversal learning between ChATDrd2KO and $Drd2^{fl/fl}$ mice, which may explain our behavioral result showing that mice lacking CIN D2Rs reversed faster than control mice.

3.4.1 Loss of CIN D2Rs does not affect reward learning

In Chapter 2, we reported that changes in DA and ACh levels that we recorded in a Pavlovian task are consistent with DA neurons and TANs/CINs encoding unexpected rewards and reward-predicting cues (Aosaki et al., 1994a; Joshua et al., 2008; Morris et al., 2004; Schultz et al., 1997). Here, we determined if loss of DA regulation of CINs via selective knockout of CIN D2Rs affected reward learning in a Pavlovian task. We found no difference in acquisition, extinction, or reinstatement learning between ChATDrd2KO and $Drd2^{fl/fl}$ mice. To determine learning, we quantified anticipatory responding during the conditioned and unconditioned stimuli. We found that both groups increased anticipatory behavior to the conditioned but not unconditioned stimuli over learning. This result suggest that CIN D2Rs are not important for initial learning and is consistent with previous studies showing that CINs are not important for initial

learning but rather necessary when task contingencies change (Aoki et al., 2015; Bradfield et al., 2013; Brown et al., 2010; Okada et al., 2017). Next, we determined if CIN D2Rs regulate cue-motivated behavior in a Pavlovian-to-Instrumental Transfer (PIT) task. We found no deficits in PIT learning in mice lacking CIN D2Rs, suggesting that CIN D2Rs do not affected cue-induced invigoration for a reward (Gallo et al., 2021).

3.4.2 Loss of CIN D2Rs enhances reversal but not acquisition learning in a probabilistic choice task

Striatal DA neurons and CINs have been shown to be important for guiding flexible learning (Adamantidis et al., 2011; Aoki et al., 2015; Bradfield et al., 2013; Brown et al., 2010; Clarke et al., 2011; Favier et al., 2020; Okada et al., 2014; Okada et al., 2017; Ragozzino et al., 2009; Rossi et al., 2013). Here, we determined if loss of CIN D2Rs affects flexible learning in a probabilistic choice task. We found that mice lacking CIN D2Rs were faster at reversal learning compared to controls, while initial learning was unaffected. This result is consistent with previous studies demonstrating that CINs are not necessary for initial learning but important when task contingencies change (Aoki et al., 2015; Bradfield et al., 2013; Brown et al., 2010; Okada et al., 2017). In Chapter 2, we reported that loss of CIN D2Rs shortens the cue induced ACh dip, which may explain this behavioral finding. This finding also supports a computational model, which predicts that shorter CIN pauses facilitates reversal learning compared to longer pauses (Franklin & Frank, 2015). However, future studies that can manipulate the duration of the CIN pause are necessary to confirm this theory.

3.4.3. Striatal ACh and DA levels signal reward expectation and outcomes in a probabilisitic choice task

DA neurons and TANs/CINs encoding unexpected rewards and reward-predicting cues (Aosaki et al., 1994a; Joshua et al., 2008; Morris et al., 2004; Schultz et al., 1997). Here, we observed changes in ACh and DA levels that signal reward outcomes in a probabilistic choice task. During acquisition, we observed enhanced rebound ACh levels in trials where a reward was expected in both ChATDrd2KO and $Drd2^{fl/fl}$ mice. In addition, we observed changes in DA levels that signal reward prediction errors in both groups. However, we observed no difference in ACh or DA levels to reward outcome between ChATDrd2KO and $Drd2^{fl/fl}$ mice during acquisition. These results are consistent with our behavioral finding showing no difference in initial learning between the two groups. Like acquisition, we found similar changes in ACh and DA levels to reward outcome between ChATDrd2KO and $Drd2^{fl/fl}$ mice, which supports no difference in performance on the first two days of reversal learning. However, on Day 3 of reversal, we observed differences in ACh levels to reward outcome between ChATDrd2KO and $Drd2^{fl/fl}$ mice. Notably, we observed higher ACh rebound levels in expected reward and unexpected reward omission trials in ChATDrd2KO mice compared to $Drd2^{fl/fl}$ mice. Enhanced rebound ACh levels in mice lacking D2Rs may facilitate reversal learning, which is consistent with our behavioral data.

In addition, we observed differences in DA levels to reward outcome between ChATDrd2KO and $Drd2^{fl/fl}$ mice. Notably, we observed a larger decrease in DA levels to unexpected reward omission in mice lacking CIN D2Rs compared to controls. Again, this difference in DA levels to reward outcome may be important for reversal learning. However, this data is preliminary, and further analyses and more mice are needed to confirm these results. Finally, we found that ACh and DA levels to reward outcome were comparable between

ChATDrd2KO and Drd2$^{fl/fl}$ mice on the final day of reversal learning. This finding is consistent with our behavioral data showing that at the end of reversal learning, both ChATDrd2KO and Drd2$^{fl/fl}$ mice reversed their behavior and learned the new rule change.

In conclusion, our data reveal a specific role for CIN D2Rs in flexible learning. We demonstrated that CIN D2Rs are not important initial learning, but rather inhibit learning when task contingencies change. Thus, further exploration striatal DA and ACh during reversal learning will provide a better understanding of each of their roles in regulating flexible learning.

3.5 Methods

Animals

Adult male and female C57BL/6J (JAX stock # 000664) mice (were bred in house. For control and KO animals: double-transgenic mice were generated by crossing heterozygous ChAT-IRES-Cre mice (Rossi et al., 2011) (JAX stock #031661) to homozygous *Drd2$^{fl/fl}$* (Drd2$^{fl/fl}$) mice (Bello et al., 2011). Control (Drd2$^{fl/fl}$) and ChATDrd2KO (ChAT-IRES-Cre x Drd2$^{fl/fl}$) mice are littermates, bred in house and back crossed onto C57BL/6J background. Mice were housed 1-4 per cage for most experiments on a 12-hr light/dark cycle, and all experiments were conducted in the light cycle. All experimental procedures were conducted following NIH guidelines and were approved by Institutional Animal Care and Use Committees by Columbia University and the New York State Psychiatric Institute.

Mice (≥ 8 weeks old) were induced with 4% isoflurane and maintained at 1-2% throughout the procedure. Mice were bilaterally injected with 450 nL/hemisphere with either AA5-hSYN-dLight1.2 (Addgene) (Patriarchi et al., 2018) or AA5-hSYN-ACh3.0 (Vigene) (Jing et al., 2020) (also known as GRAB-ACh3.0) into separate hemispheres of the dorsal medial striatum (DMS) using stereotactic Bregma-based coordinates: AP, +1.1mm; ML ± 1.4 mm; DV, -3.1 mm, -3.0 mm

and -2.9 mm (150 nL/DV site). Following virus injection, 400-µm fiber optic cannulas (Doric, Quebec, Canada) were carefully lowered to a depth of -3.0 mm and fixed in place to the skull with dental cement anchored to machine mini-screws. Groups of mice used for experiments were housed in a counterbalanced fashion that accounted for sex, age, and home cage origin. Cannula-implanted mice began behavioral training 4 weeks after surgery. At the end of experiments, animals were perfused, and brains were processed post-*hoc* to validate virus expression and optic fiber location as in (Gallo et al., 2021).

In vivo fiber photometry and optogenetics

Fiber photometry equipment was set up using two 4-channel LED Driver (Doric) connected to two sets of a 405-nm LED and a 465-nm LED (Doric, cLED_405 and cLED_465). The 405 nm LEDs were passed through 405-410 nm bandpass filters, while the 465 nm LEDs were passed through a 460-490 nm GFP excitation filters using two 6-port Doric minicubes. 405 and 465 LEDs were then coupled to a dichroic mirror to split excitation and emission lights. Two low-autofluorescence patch cords (400 µm/0.48NA, Doric) arising from the 2 minicubes were attached to the cannulas on the mouse's head and used to collect fluorescence emissions. These signals were filtered through 500-540 nm GFP emission filters via the same minicubes coupled to photodetectors (Doric, gain set to DC Low). Signals were sinusoidally modulated, using Synapse® software and RZ5P Multi I/O Processors (Tucker-Davis Technologies), at 210 Hz and 330 Hz (405 nm and 465 nm, respectively) to allow for low-pass filtering at 3 Hz via a lock-in amplification detector. 405 and 465 nm power at the patch cord were set to 30 µW or below.

Photometry data processing

All photometry and behavioral data utilized custom in-house MATLAB analysis scripts. Photometry signals were analyzed as time-locked events aligned to the lever extension (CRF) or

tone onset (Pavlovian) of each trial. The 405 nm channel was used to control for potential noise/movement artifacts and the 465 nm channel was used to detect the conformational modulation of either the GACh3.0 sensor by ACh or the dLight1.2 sensor by DA. Both demodulated signals were extracted as a 15 s window surrounding the event, which was denoted as time = 0. Both signals were down sampled by a factor of 10 using a moving window mean. The change in fluorescence, ΔF/F (%), was defined as (F-F0)/F0 x 100, where F represents the fluorescent signal (465 nm) at each time point. F0 was calculated by applying a least-squares linear fit to the 405 nm signal to align with the 465nm signal (Calipari et al., 2016). To normalize signals across animals and sessions, we calculated a local baseline fluorescence value for each trial using the average of the 5 s period preceding the event and subtracted that from the signal. The daily average GACh3.0 and dLight1.2 traces were calculated using session average traces from individual mice. ACh dip and DA peak amplitudes were calculated as the maximal change of the signal that was at least 1 or 2 STD below or above the local baseline, respectively. ACh dip duration was calculated using the last and the first zero crossings preceding and following the dip. Total AUC was calculated as the area of all three components of the ACh signal (initial peak, dip and rebound). Negative AUC was calculated as the area for only the negative component. Rebound AUC was calculated as the area of the positive component immediately following the dip. The AUC analysis was restricted to a 5 s time window following the task event. Individual CRF trial (ΔF/F (%)) traces were used for correlation analysis for CRF trials. For the ITI correlation, we examined any interaction between dLight1.2 and GACh3.0 regardless of event size during the variable 40 s of the ITI.

Operant apparatus

Four operant chambers (model Env-307w; Med-Associates, St. Albans, VT) equipped

with liquid dippers were used. Each chamber was in a light- and sound-attenuating cabinet equipped with an exhaust fan, which provided 72-dB background white noise in the chamber. The dimensions of the experimental chamber interior were $22 \times 18 \times 13$ cm, with flooring consisting of metal rods placed 0.87 cm apart. A feeder trough was centered on one wall of the chamber. An infrared photocell detector was used to record head entries into the trough. Raising of the dipper inside the trough delivered a drop of evaporated milk reward. A retractable lever was mounted on the same wall as the feeder trough, 5 cm away. A house light located on the wall opposite to trough illuminated the chamber throughout all sessions.

Dipper training

Four weeks after surgery, mice underwent operant training. Mice were weighed daily and food- restricted to 85–90% of baseline weight; water was available ad libitum. In the first training session, 20 dipper presentations were separated by a variable inter-trial interval (ITI) and ended after 20 rewards were earned or after 30 min had elapsed, whichever occurred first. Criterion consisted of the mouse making head entries during 20 dipper presentations in one session. In the second training session, criterion was achieved when mice made head entries during 30 of 30 dipper presentations.

Pavlovian conditioning, extinction, and reinstatement

Mice were trained for 16 consecutive days in a Pavlovian conditioning paradigm, which consisted of 12 conditioned stimulus-positive (CS+) trials and 12 unconditioned stimulus (CS-) trials occurring in a pseudorandom order. Each trial consisted of an 80-dB auditory cue presentation for 10 sec of an 8 kHz tone or white noise (counterbalanced between mice) and after cue offset a milk reward was delivered only in CS+ trials, whereas no reward was delivered in CS- trials. After conditioning, mice were trained for 4 days in Pavlovian extinction, which is similar

Pavlovian conditioning. However, in extinction the CS+ trials are not followed by a milk reward. After extinction, mice were subjected to 1 day of reinstatement learning where the CS+ was again paired with a milk reward. There was a 100 sec variable intertrial interval, drawn from an exponential distribution of times. Head entries in the food port were recorded throughout the session, and anticipatory head entries during the presentation of the cue were considered the conditioned response. Anticipatory responding was calculated as the difference in nose poking during the CS+ quintile with the maximum response (Q4 or 5) and the first quintile.

Continuous reinforcement schedule (CRF)

For lever press training, lever presses were reinforced on a continuous reinforcement (CRF) schedule. Levers were retracted after each reinforcer and were presented again after a variable ITI (average 40 s). The reward consisted of raising the dipper for 5 s. The session ended when the mouse earned 60 reinforcements, or one hour elapsed, whichever occurred first. Sessions were repeated daily until mice achieved 60 reinforcements.

Pavlovian- to-Instrumental Transfer (PIT)

Mice received 7 days of Pavlovian training in which an auditory CS+ (either a tone or white noise) was paired with a 20% sucrose-liquid reward. The CS+, which lasted 2 min, was presented 6 times with a variable ITI (mean of 5 min). Sucrose dippers were given on a random-time 30 sec schedule for 5 sec. After Pavlovian training, mice were trained to press a lever in a CRF schedule, as above, with the exception that the levers remained out once extended. The milk reward was delivered in a retractable dipper that was available for 5 sec. The session ended when the animal earned 30 rewards or 30 min elapsed, whichever occurred first. Sessions were repeated until the animals received 30 rewards. Mice then received 2-3 days each of random ratio 5 (RR5), RR10, and RR20 schedules in the absence of the CS+. After a Pavlovian “reminder” session, mice were

given a session where no rewards were given, and they were exposed to the CS that was not the initial CS+ ("$CS^{Ø}$"). Next, the mice received a 30 min session of lever press extinction, where no CS's were presented, and lever pressing was not rewarded. The next day, mice underwent a PIT test. The PIT test started with an 8 min extinction period, where lever pressing was not rewarded. The CS+ and the $CS^{Ø}$ were then presented four times each in the following order: (noise= n, tone= t: n-t-t-n-t-n-n-t). Each CS lasted 2 min followed by a 3 min fixed ITI, and no rewards were delivered.

Probabilistic choice

Mice received 5 days of probabilistic choice training, where each session consisted of 80 total trials: 40 forced and 40 choice trials. One of two levers was designated the "high payoff" lever and the other the "low payoff" lever. The "high payoff" lever if chosen would be followed by a 5 sec milk reward 80% of the time. The "low payoff lever if chosen would only be followed by a 5 sec milk reward 20% of the time. In forced trials, only one lever was presented at a time. In choice trials, both levers were presented. In both trial types, the lever(s) would be extended for 30 sec. If no lever was pressed within 30 sec, the lever(s) would retract, and a new trial would start after a variable ITI (mean of 30 sec). The criterion to determine learning was 80% choice of the "high payoff" lever in choice trials over 2 consecutive days. After 5 days of acquisition, the mice underwent reversal learning. Here, the "high payoff" and "low payoff" levers were switched. Everything else in the task remained the same. Again, criterion was achieved after 80% choice of the "high payoff" lever in choice trials over 2 consecutive days. Trial-by-trial analyses we conducted to determine any potential genotype differences in "win-stay or lose-shift" (WSLS) strategies. WSLS strategies are frequently used to examine decision-making tasks involving risk and reward, which can reveal changes in sensitivity to surprising outcomes and feedback learning

(Aguirre et al., 2020). A trial was classified as a "win" if the animal received a milk reward, and a "lose" if no reward was delivered. Decisions were classified as "win-stay" when the animal chose the same lever on the subsequent trial after a win, and a "lose-shift" if the animal switched to the other lever after a loss. In addition, we examined disadvantageous strategies. Decisions were classified as "win-shift" when the animal switched to the alternative lever after a reward, and a "lose-stay" when the animal stuck with the same lever after a loss. Using this, we compared the frequency of advantageous strategies and disadvantageous strategies by calculating an adaptive score (win-stay + lose-shit) – (win-shift + lose-stay).

Data analysis

Sample sizes were determined by performing statistical power analyses based on effect sizes observed in preliminary data or on similar work in the literature. Statistical analyses were performed using GraphPad Prism 9 (GraphPad), MATLAB (MathWorks). Data are generally expressed as mean ± standard error of the mean (SEM). Paired and unpaired two-tailed Student's t-tests were used to compare 2-group data, as appropriate. Multiple comparisons were evaluated by one- or two-way ANOVA and Bonferroni's post hoc test, when appropriate. In rare cases of values missing in repeated measures samples, the data were analyzed by fitting a mixed effects model, as implemented by Prism 9. Photometry correlation analyses were performed using Pearson's correlation coefficients. A p-value of < 0.05 was considered statistically significant. Behavioral findings were replicated with mice from different litters, ages, or sexes. Investigators were blinded to the genotype of mice during behavioral assays as well as throughout the data analysis.

Chapter 4: Conclusions and Future Directions

4.1 General conclusions

Here, I have described experiments conducted to explore the role of dopamine (DA) regulation on the cholinergic pause in the striatum during reinforced behaviors, reward learning and flexible learning. To do this, I first summarized the existing literature on this topic in Chapter 1. In Chapter 2, I detailed work done to develop a new imaging strategy to simultaneously monitor the activity of DA and acetylcholine (ACh) levels in the striatum within the same mouse during reinforced behaviors. In combination with this imaging strategy, we used pharmacology to systemically block DA D2 receptors (D2Rs) and mouse genetics to selectively ablate D2Rs on cholinergic interneurons (CINs) to determine the DA dependence of the acetylcholine (ACh) dip during reinforced behavior. In Chapter 3, we used this genetic inactivation of CIN D2Rs strategy to examine the consequences on reward learning. Using this strategy, we also determined the role of CIN D2Rs on flexible learning and examined changes in both DA and ACh levels during flexible learning.

In addition, we determined that the CIN D2Rs coordinate cue-evoked changes in striatal ACh levels. Blocking or ablating CIN D2Rs resulted in a shorter ACh dip, a larger rebound in ACh levels following the dip and decreased coordinated activity between striatal DA and ACh. Notably, this manipulation had no effect on reward learning tasks but did facilitate reversal learning in a probabilistic choice task. This work highlights the importance of DA regulation of CIN activity during natural behaviors. In this chapter, I will discuss the implications of these finding and describe future lines of research for this field.

4.2 Simultaneous monitoring of DA and ACh during operant based tasks

In Chapter 2, I described a new imaging approach to monitor changes in striatal dopamine (DA) and acetylcholine (ACh) at the same time in the same mouse during behavior, which provides many advantages over previous approaches. First, this approach allows us to simultaneously measure changes in DA and ACh levels using genetically encoded biosensors (Jing et al., 2020; Patriarchi et al., 2018). Measuring real-time changes in ACh has been particularly difficult (Jing et al., 2020; Jing et al., 2018; Kodama et al., 1992). Therefore, our understanding of changes in ACh levels is mostly based on electrophysiological recordings of "tonically active neurons" (TANs) in primates (Aosaki et al., 1994a). In addition, ACh release can be locally regulated independent of neuronal activity at the cell body (A. Mohebi et al., 2019). Finally, cholinergic interneurons (CINs) co-release other neurotransmitters and use of the genetically encoded sensor will enable selective measurement of ACh levels (El Mestikawy et al., 2011; Gras et al., 2008). Therefore, this new approach to simultaneously image the activity of DA and ACh in the same animal will allow us to examine the role each neuromodulator plays during behavior and determine if they mutually co-regulate each other to promote learning. Future work to monitor both DA and ACh not only in the same animal but in the same brain region will strengthen these experiments, and is now possible with the development of biosensors in different colors (Labouesse & Patriarchi, 2021).

Second, this new strategy does not rely on stimulation to evoke changes in striatal DA or ACh. Instead, we use operant based tasks to study naturally evoked changes in DA and ACh levels. Use of operant based tasks offers several advantages. First, we can align both DA and ACh signals to different task relevant events (i.e., lever extension, lever press, tone onset/offset, reward, etc.). Second, mice can complete many trials per day, which aids the quantification of DA and ACh

signals within mice per session and across days. Therefore, we can use this approach to determine if DA and ACh co-regulate each other during learning and behavior.

4.3 Co-development changes in dopamine and acetylcholine during learning

In Chapter 2, I found that changes in DA and ACh levels that were measured during Pavlovian conditioning (Figure 2.2B) are consistent with previous studies showing that DA neurons and TANs/CINs encode unexpected rewards and reward-predicting cues (Aosaki et al., 1994a; Apicella et al., 1991; Joshua et al., 2008; Morris et al., 2004; Schultz et al., 1997; Watanabe & Kimura, 1998). Like these previous studies, that recorded neuronal activity, we measured a robust increase in DA levels and a decrease in ACh levels to unexpected reward that diminishes as the reward becomes expected (Figure 2.2B). These results show that the neurotransmitter levels of both, DA and ACh, follow changes in neuronal activity of their respective neurons with sub-second kinetics. The fast induction of the ACh dip is especially striking as it indicates rapid degradation or diffusion of ACh.

In addition, we observed similar changes in both DA and ACh levels to the conditioned stimulus but not the unconditioned stimulus, which occur in parallel over learning (Figure 2.2B). This finding is consistent with previous studies showing that both DA neurons and CINs respond to salient and conditioned stimuli. We also reported that changes in DA and ACh levels correlated with behavioral responding in the Pavlovian task (Figure 2.2C). However, the correlation was not observed in all tested mice due to the nature of the Pavlovian task (Figure 2.2C). In this task, mice can learn the association between the conditioned stimulus and reward. As a result, some mice display anticipatory responding during the conditioned stimulus (head poking the reward port). However, because anticipatory responding is not required to obtain a reward, some mice did not exhibit any anticipatory behavior. Nevertheless, the development of a DA and ACh signal over

learning suggest that these mice do learn the stimulus-reward association. Together, these results demonstrate that task-evoked changes in DA and ACh levels support what has been previously described at the level of neuronal activity in primates (Aosaki et al., 1994a; Morris et al., 2004; Schultz et al., 1997).

4.4 Dopamine D2Rs coordinate cue-evoked changes in striatal acetylcholine levels

4.4.1 DA D2Rs modulate task-dependent decreases in ACh levels

In Chapter 2, we were determined to answer a long-standing question; is the cholinergic pause entirely dependent on DA? As previously described, salient and conditioned stimuli are known to induces pauses in CIN and TAN firing in rodents and primates, however, the dependence on striatal DA and cholinergic D2Rs for pause induction is widely debated (Aosaki et al., 1994b; Ding et al., 2010; Morris et al., 2004; Watanabe & Kimura, 1998; Zhang et al., 2018). Therefore, to determine the role CIN D2Rs play in modulating the stimulus induced dip in ACh levels, we pharmacologically blocked D2Rs or selectively ablated D2Rs from CINs and measured ACh and DA levels during an operant based reinforcement behavior (CRF). We found that blocking or ablating D2Rs shortened the stimulus induced ACh dip and enhanced rebound levels in ACh but did not prevent the induction of the dip (Figures 2.3, 2.5, 2.9). While the ACh dip was shortened in our ChATDrd2 KO mice, it was not further shortened and there was no enhanced rebound in ACh levels following D2R blockade (Figure 2.11G-L). Together, these results suggest that the induction of the ACh dip is not dependent on CIN D2Rs. Instead, CIN D2Rs are important for modulating the duration of the stimulus induced ACh dip and rebound levels following the dip.

Our data provide clarity on the controversial role that DA plays in the regulation of the ACh dip and suggest that the naturally evoked ACh dip is not entirely DA or CIN D2R-dependent as the literature suggested (Aosaki et al., 1994b; Ding et al., 2010; Watanabe & Kimura, 1998).

Therefore, the induction of the ACh dip is likely driven by non-DA inputs like glutamatergic cortical or thalamic inputs or long-range GABAergic inputs as suggested by previous stimulation studies (Brown et al., 2010; Brown et al., 2012; Cover et al., 2019; Doig et al., 2014; English et al., 2012b; Matsumoto et al., 2000; Zhang et al., 2018). Future studies that can inhibit or deactivate glutamatergic or GABAergic inputs to the striatum will uncover the mechanism driving each component of the ACh dip and help determine the behavioral significance of the ACh dip.

In addition to the effects on ACh dip durations, we found that CIN D2Rs also modulate rebound in ACh levels, which acts as a mechanism to limit enhanced ACh levels after the dip. Currently, it is unknown what roles the ACh dip and rebound play during behavior. As described in Chapter 1, CINs are thought to inhibit the activity of spiny projection neurons (SPNs) via nicotinic activation of local interneurons or through muscarinic M_2/M_4-mediated inhibition of corticostriatal projections. Thus, a larger ACh dip might lead to disinhibition of SPNs, while enhanced rebound ACh levels lead to a stronger inhibition of SPNs. We reported that D2R antagonism shortens the ACh dip and enhances ACh rebound levels after the dip, which might inhibit movement initiation leading to longer press latencies in lever pressing. In support of this, we found that the size of the stimulus induced ACh dip inversely correlated with lever press latency (the larger the ACh dip the shorter the press latency) (Figure 2.17C-F). Strikingly, this relationship was maintained for lever presses that occurred long after the stimulus induced ACh signal was gone. This result suggests that the stimulus induced ACh dips encode the motivational state of mice. Our data is consistent with a recent finding that demonstrated that inhibition of CINs in the ventral striatum (NAc) during a Pavlovian to Instrumental Transfer (PIT) task enhanced the ability of Pavlovian cues to invigorate motivated behavior (Collins et al., 2019).

In contrast, in Chapter 3, we found that selective D2R ablation from CINs did not affect performance in PIT (Figure 3.3) (Gallo et al., 2021). This result suggests a more nuanced impairment affecting press latencies rather than the vigor of responding. Therefore, future studies that can modulate the duration of the ACh dip and rebound levels *in vivo* will reveal the role of each component of the ACh in learning and behavior.

4.4.2 Dopamine D2Rs regulate the coordination between DA and ACh activity in vivo

In Chapter 2, we developed a new imaging approach to simultaneously measure DA and ACh levels in the same animal during behavior. This approach enabled us to explore the relationship between DA and ACh during task relevant events and during the intertrial interval (ITI). First during the CRF task, we identified a strong negative correlation with the stimulus induced increase in DA leading the ACh dip in our control mice, which was attenuated by D2R antagonism (Figures 2.7 & 2.13). Interestingly, this negative correlation between DA and ACh was significantly reduced in mice lacking CIN D2Rs compared to controls (Figure 2.12) and unaffected by D2R antagonism (Figure 2.15). These results suggest that CIN D2Rs are necessary for the coordinated activity between DA and ACh. In addition, in our control mice, we identified a weaker positive correlation with the cue induced increase in DA leading the ACh peak that was enhanced by D2R blockade (Figures 2.7 & 2.13). We believe this positive correlation represents the rebound in ACh levels following the dip that is typically blunted by activation of D2Rs at baseline. This result supports, as previously discussed above, that CIN D2Rs modulate both the ACh dip and rebound levels.

Second, in our control mice, we also observed negative and positive correlations between DA and ACh during the ITI, which were attenuated or enhanced by D2R antagonism, respectively (Figures 2.8 & 2.14). Again, we found that the correlation between DA and ACh signals during

the ITI was significantly reduced in ChATDrd2KO mice compared to controls (Figure 2.12). These results suggest that CIN D2Rs are not only necessary for coordinating the activity between DA and ACh during task events but also support a more general mechanism of coordination. However, it is unclear how the coordinated activity between DA and ACh affects learning and future studies that can address this will reveal the behavioral significance of this coordinated activity.

4.4.3 Loss of CIN D2Rs does not affect cue-induced changes in DA levels

Here, we focused on the DA regulation of ACh levels in the striatum. We made this decision to determine the mechanisms underlying the CIN pause. However, as outlined in Chapter 1, there is substantial evidence that suggests that CIN activity and ACh regulates local release of DA in the striatum (Cachope & Cheer, 2014; Cachope et al., 2012; David Sulzer et al., 2016; Threlfell et al., 2012). In Chapter 2, we described a new imaging technique to simultaneously monitor both striatal DA and ACh levels during behavior. Therefore, we also examined cue induced DA levels following pharmacological blockade of D2Rs and selective ablation of CIN D2Rs to determine how altered ACh levels affected DA signaling. We found that D2R antagonism significantly enhanced cue induced DA levels (Figure 2.6). However, this result was likely due to blocking D2-auto receptors on DA terminals, which prevents their activation. Surprisingly, we found no change in cue induced DA levels in ChATDrd2KO mice at baseline (Figure 2.10). However, cue induced DA levels in ChATDrd2KO mice were enhanced by D2R antagonism, indicating activation of D2-auto receptors (data not shown). We predicted that loss of CIN D2Rs, which shortens the ACh dip would enhance cue induced DA levels, however, we did not observe this. In addition, enhanced rebound ACh levels following D2R antagonism did not increase DA levels. One explanation is that under the circumstances measured changes in ACh levels may not regulate DA release. Another possibility is that increased ACh levels may desensitize nicotinic

receptors on DA neurons. Therefore, future studies that can selectively and transiently target CIN D2Rs will determine if an acute altered ACh signal affects cue induced changes in DA levels during behavior.

4.5 Behavioral implications of cholinergic dopamine D2Rs

In Chapter 2, we determined that DA via CIN D2Rs regulates the duration of the ACh dip and rebound ACh levels. What does an altered ACh signal mean for behavior? To address this, we tested mice in various striatal dependent functions, including reinforced behavior, reward learning and flexible learning.

4.5.1 Reinforced behavior

In Chapter 2, we used a continuous reinforcement (CRF) task to determine the role CIN D2Rs play during a reinforced behavior. In control mice but not ChATDrd2 KO mice, we found that antagonizing D2Rs increased lever press latency, in a dose-dependent manner consistent with previous findings (Figure 2.17A-B) (Kharkwal et al., 2016). This result demonstrates that acute inhibition of CIN D2Rs increases press latency, which we don't observe in ChATDrd2KO mice showing that the effect of the antagonist is mediated by CIN D2Rs. Surprisingly, we found no difference in press latency between control and ChATDrd2KO mice at baseline. We had predicted that loss of CIN D2Rs like D2R antagonism in control mice would enhance press latency in the CRF task, which we did not observe. We believe that this absence of effect in KO is due to developmental compensation. In support of a role of CIN ACh levels regulating press latency, we found a correlation between the size of the ACh dip and press latency in our control mice: the larger the dip the shorter the press latency (Figure 2.17C-F). This result suggests a role for ACh in regulating motivation. To directly test whether there is a relationship between ACh dip duration

and press latency, future work will need to manipulate the duration of the ACh dip to determine if ACh regulates press latencies and clarify our findings.

4.5.2 Reward learning

As previously described, both DA neurons and CINs respond to unexpected rewards and reward-predicting cues (Aosaki et al., 1994a; Apicella et al., 1991; Joshua et al., 2008; Morris et al., 2004; Schultz et al., 1997; Watanabe & Kimura, 1998).. Consistent with this, we reported in Chapter 2 that changes in both DA and ACh levels are associated with unexpected rewards and conditioned stimuli (Figure 2.2). Therefore, we sought out to determine the role CIN D2Rs play in reward learning. In Chapter 3, we reported that selective loss of CIN D2Rs does not affect Pavlovian conditioning, extinction, or reinstatement (Figures 3.1 & 3.2). However, these results were not surprising as CINs are not believed to play a role in initial learning but rather are important when task contingencies change (Aoki et al., 2015; Bradfield et al., 2013; Brown et al., 2010; Okada et al., 2017), which I will discuss next.

4.5.3 Flexible learning

In Chapter 1, I outlined the implications of striatal DA and ACh in behavioral flexibility (Adamantidis et al., 2011; Aoki et al., 2015; Bradfield et al., 2013; Brown et al., 2010; Clarke et al., 2011; Favier et al., 2020; Okada et al., 2014; Okada et al., 2017; Ragozzino et al., 2009; Rossi et al., 2013). However, it has yet to be determined if manipulating DA's regulation of CINs and vice versa affects flexible learning. Therefore, in Chapter 3, we tested how loss of CIN D2Rs affects acquisition and reversal learning in a probabilistic choice task. We found that mice lacking CIN D2Rs were faster at reversal learning compared to controls (Figure 3.6), with no difference in the initial acquisition between the two groups (Figure 3.5). In addition, we found that ChATDrd2KO mice developed advantageous strategies faster, as measured by enhanced win-stay

and lose-shift (WSLS) strategies, than Drd2$^{fl/fl}$ control mice during reversal learning, while they were comparable in the acquisition phase (Figure 3.7).

Next, we examined DA and ACh levels during acquisition and reversal learning in the probabilistic choice task. We found that both DA and ACh levels signal reward outcome in both phases of the task. During acquisition, rebound ACh levels are enhanced in trials with expected reward and unexpected reward omission, with no difference between ChATDrd2KO and Drd2$^{fl/fl}$ control mice (Figure 3.8). In addition, DA levels in both groups comparably signal reward prediction errors, with no difference between genotypes (Figure 3.9). These results support our behavioral finding demonstrating no difference in acquisition learning between ChATDrd2KO and Drd2$^{fl/fl}$ control mice. We observed similar changes in ACh and DA levels during the first two days of reversal learning, again with no difference between the two groups (Figures 3.10 & 3.11). However, on Day 3 of reversal, where we found the biggest difference in behavioral performance, we observed differences in ACh and DA levels to reward outcome between ChATDrd2KO and Drd2$^{fl/fl}$ control mice. In mice lacking CIN D2Rs, we observed a larger rebound in ACh levels in trials where a reward was expected, which may be important for reversal learning (Figure 3.12). Moreover, we observed a larger decrease in DA levels to unexpected reward omission in ChATDrd2KO mice compared to Drd2$^{fl/fl}$ mice (Figure 3.13). Together these results implicate a specific role for CIN D2Rs in limiting flexible behavior, which could be beneficial or harmful depending on the context. However, this data is preliminary and further analyses, and more mice are required to confirm these results. Below is a schematic depicting the overall conclusions of my thesis and future directions (Figure 4.1).

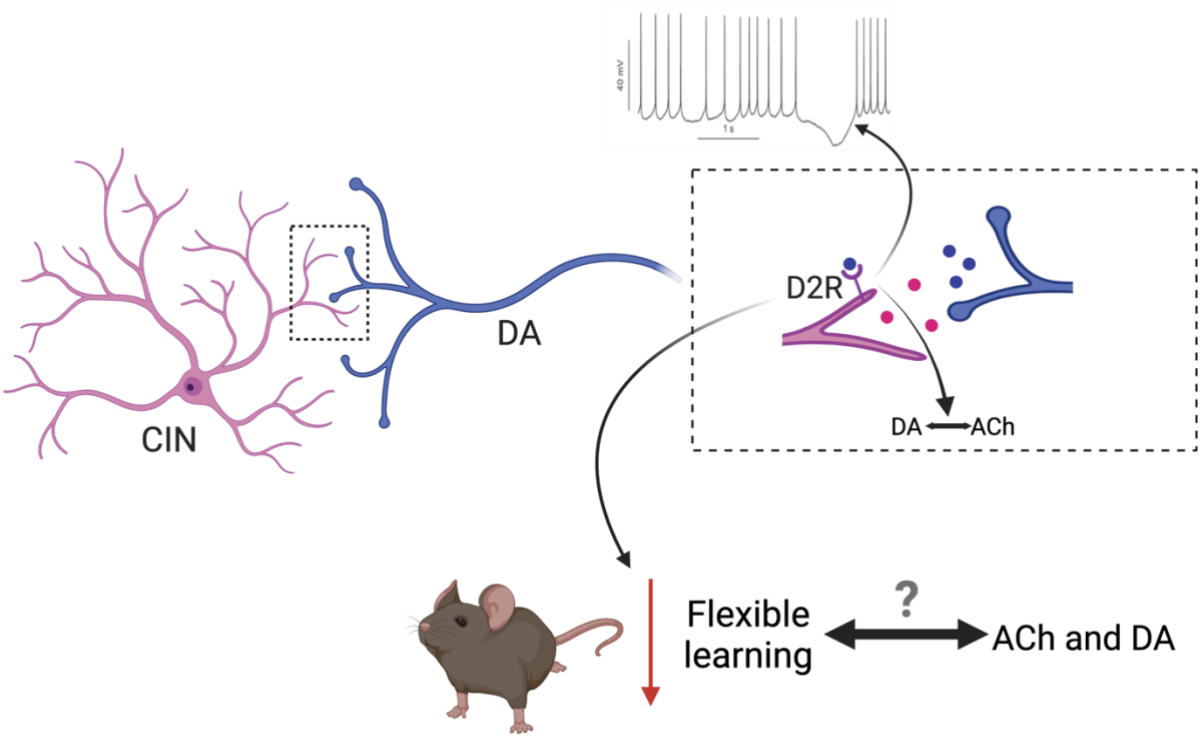

Figure 4.1. Overall conclusions and future directions. CIN D2Rs regulate the duration of the behaviorally induced CIN pause and rebound activity after the pause, the coordinated activity between DA and ACh and limit flexible learning in a reversal learning task. Future studies will determine how striatal ACh and DA regulate flexible learning.

References

Aberman, J. E., Ward, S. J., & Salamone, J. D. (1998, Dec). Effects of dopamine antagonists and accumbens dopamine depletions on time-constrained progressive-ratio performance. *Pharmacol Biochem Behav, 61*(4), 341-348. https://doi.org/10.1016/s0091-3057(98)00112-9

Adamantidis, A. R., Tsai, H. C., Boutrel, B., Zhang, F., Stuber, G. D., Budygin, E. A., Touriño, C., Bonci, A., Deisseroth, K., & de Lecea, L. (2011, Jul 27). Optogenetic interrogation of dopaminergic modulation of the multiple phases of reward-seeking behavior. *J Neurosci, 31*(30), 10829-10835. https://doi.org/10.1523/jneurosci.2246-11.2011

Aebischer, P., & Schultz, W. (1984, Sep 7). The activity of pars compacta neurons of the monkey substantia nigra is depressed by apomorphine. *Neurosci Lett, 50*(1-3), 25-29. https://doi.org/10.1016/0304-3940(84)90456-7

Aguirre, C. G., Stolyarova, A., Das, K., Kolli, S., Marty, V., Ray, L., Spigelman, I., & Izquierdo, A. (2020). Sex-dependent effects of chronic intermittent voluntary alcohol consumption on attentional, not motivational, measures during probabilistic learning and reversal. *PLoS One, 15*(6), e0234729. https://doi.org/10.1371/journal.pone.0234729

Akaike, A., Sasa, M., & Takaori, S. (1988, Sep). Muscarinic inhibition as a dominant role in cholinergic regulation of transmission in the caudate nucleus. *J Pharmacol Exp Ther, 246*(3), 1129-1136.

Albin, R. L., Young, A. B., & Penney, J. B. (1989, Oct). The functional anatomy of basal ganglia disorders. *Trends Neurosci, 12*(10), 366-375. https://doi.org/10.1016/0166-2236(89)90074-x

Aoki, S., Liu, A. W., Zucca, A., Zucca, S., & Wickens, J. R. (2015). Role of Striatal Cholinergic Interneurons in Set-Shifting in the Rat. *J Neurosci*. https://doi.org/10.1523/JNEUROSCI.0490-15.2015

Aosaki, T., Graybiel, A. M., & Kimura, M. (1994b). Effect of the Nigrostriatal Dopamine System on Acquired Neural Responses in the Striatum of Behaving Monkeys. *Science*, 265(5170): 5412-5415.

Aosaki T, T. H., Ishida A, Watanabe K, Graybiel AM, Kimura M. (1994b). Responses of Tonically Active Neurons in the Primate's Striatum Undergo Systematic Changes during Behavioral Sensorimotor Conditioning. *J Neurosci*. https://doi.org/14(6): 3969-3984

Aosaki, T., Tsubokawa, H., Ishida, A., Watanabe, K., Graybiel, A. M., & Kimura, M. (1994a). Responses of Tonically Active Neurons in the Primate's Striatum Undergo Systematic Changes during Behavioral Sensorimotor Conditioning. *J Neurosci*, 14(16): 3969-3984.

Apicella, P. (2007). Leading tonically active neurons of the striatum from reward detection to context recognition. *Trends Neurosci*, 30:299–306. https://doi.org/10.1016/j.tins.2007.03.011

Apicella, P., Deffains, M., Ravel, S., & Legallet, E. (2009). Tonically active neurons in the striatum differentiate between delivery and omission of expected reward in a probabilistic task context. *Eur. J. Neurosci*, 30:515–526. https://doi.org/10.1111/j.1460-9568.2009.06872.x

Apicella, P., Legallet, E., & Trouche, E. (1997, Oct). Responses of tonically discharging neurons in the monkey striatum to primary rewards delivered during different behavioral states. *Exp Brain Res*, *116*(3), 456-466. https://doi.org/10.1007/pl00005773

Apicella, P., Ravel, S., Deffains, M., & Legallet, E. (2011). The Role of Striatal Tonically Active Neurons in Reward Prediction Error Signaling during Instrumental Task Performance. *J Neurosci*. https://doi.org/10.1523/JNEUROSCI.4880-10.2011

Apicella, P., Scarnati, E., Ljungberg, T., & Schultz, W. (1992). Neuronal Activity in Monkey Striatum Related to the Expectation of Predictable Environmental Events *J Neurophysiol*.

Apicella, P., Scarnati, E., & Schultz, W. (1991). Tonically discharging neurons of monkey striatum respond to preparatory and rewarding stimuli. *Exp Brain Res*(84), 672–675.

Augustin, S. M., Chancey, J. H., & Lovinger, D. M. (2018). Dual Dopaminergic Regulation of Corticostriatal Plasticity by Cholinergic Interneurons and Indirect Pathway Medium Spiny Neurons. *Cell Rep*. https://doi.org/10.1016/j.celrep.2018.08.042.

Bailey, M. R., Chun, E., Schipani, E., Balsam, P. D., & Simpson, E. H. (2020, Apr). Dissociating the effects of dopamine D2 receptors on effort-based versus value-based decision making using a novel behavioral approach. *Behav Neurosci*, *134*(2), 101-118. https://doi.org/10.1037/bne0000361

Bailey, M. R., Goldman, O., Bello, E. P., Chohan, M. O., Jeong, N., Winiger, V., Chun, E., Schipani, E., Kalmbach, A., Cheer, J. F., Balsam, P. D., & Simpson, E. H. (2018, Feb 28). An Interaction between Serotonin Receptor Signaling and Dopamine Enhances Goal-Directed Vigor and Persistence in Mice. *J Neurosci*, *38*(9), 2149-2162. https://doi.org/10.1523/jneurosci.2088-17.2018

Bailey, M. R., Jensen, G., Taylor, K., Mezias, C., Williamson, C., Silver, R., Simpson, E. H., & Balsam, P. D. (2015, Jun). A novel strategy for dissecting goal-directed action and arousal components of motivated behavior with a progressive hold-down task. *Behav Neurosci*, *129*(3), 269-280. https://doi.org/10.1037/bne0000060

Bamford, N. S., Zhang, H., Schmitz, Y., Wu, N. P., Cepeda, C., Levine, M. S., Schmauss, C., Zakharenko, S. S., Zablow, L., & Sulzer, D. (2004, May 27). Heterosynaptic dopamine

neurotransmission selects sets of corticostriatal terminals. *Neuron, 42*(4), 653-663. https://doi.org/10.1016/s0896-6273(04)00265-x

Barral, J., Galarraga, E., & Bargas, J. (1999, Jul 1). Muscarinic presynaptic inhibition of neostriatal glutamatergic afferents is mediated by Q-type Ca2+ channels. *Brain Res Bull, 49*(4), 285-289. https://doi.org/10.1016/s0361-9230(99)00061-1

Bello, E., Mateo, Y., Gelman, D. M., Noaín, D., Shin, J. H., Low, M. J., Alvarez, V. A., Lovinger, D. M., & Rubinstein, M. (2011). Cocaine supersensitivity and enhanced motivation for reward in mice lacking dopamine D2 autoreceptors. *Nat Neurosci*, 14(18): 1033–1038. https://doi.org/10.1038/nn.2862.

Bennett, B. D., Callaway, J. C., & Wilson, C. J. (2000, Nov 15). Intrinsic membrane properties underlying spontaneous tonic firing in neostriatal cholinergic interneurons. *J Neurosci, 20*(22), 8493-8503. https://doi.org/10.1523/jneurosci.20-22-08493.2000

Berendse, H. W., Galis-de Graaf, Y., & Groenewegen, H. J. (1992, Feb 15). Topographical organization and relationship with ventral striatal compartments of prefrontal corticostriatal projections in the rat. *J Comp Neurol, 316*(3), 314-347. https://doi.org/10.1002/cne.903160305

Berridge, K. C. (2007, Apr). The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology (Berl), 191*(3), 391-431. https://doi.org/10.1007/s00213-006-0578-x

Bohnen, N. I., & Albin, R. L. (2009, Jul 28). Cholinergic denervation occurs early in Parkinson disease. *Neurology, 73*(4), 256-257. https://doi.org/10.1212/WNL.0b013e3181b0bd3d

Bolam, J. P., Hanley, J. J., Booth, P. A., & Bevan, M. D. (2000, May). Synaptic organisation of the basal ganglia. *J Anat, 196 (Pt 4)*(Pt 4), 527-542. https://doi.org/10.1046/j.1469-7580.2000.19640527.x

Bolam, J. P., Wainer, B. H., & Smith, A. D. (1984, Jul). Characterization of cholinergic neurons in the rat neostriatum. A combination of choline acetyltransferase immunocytochemistry, Golgi-impregnation and electron microscopy. *Neuroscience, 12*(3), 711-718. https://doi.org/10.1016/0306-4522(84)90165-9

Boulougouris, V., Castañé, A., & Robbins, T. W. (2009, Mar). Dopamine D2/D3 receptor agonist quinpirole impairs spatial reversal learning in rats: investigation of D3 receptor involvement in persistent behavior. *Psychopharmacology (Berl), 202*(4), 611-620. https://doi.org/10.1007/s00213-008-1341-2

Bradfield, L. A., Bertran-Gonzalez, J., Chieng, B., & Balleine, B. W. (2013). The thalamo-striatal pathway and cholinergic control of goal- directed action: Interlacing new and existing learning in the striatum. *Neuron*. https://doi.org/10.1016/j.neuron.2013.04.039.

Braun, S., & Hauber, W. (2011, Jul 7). The dorsomedial striatum mediates flexible choice behavior in spatial tasks. *Behav Brain Res, 220*(2), 288-293. https://doi.org/10.1016/j.bbr.2011.02.008

Briggs, R. S., Redgrave, P., & Nahorski, S. R. (1981, Feb 16). Effect of kainic acid lesions on muscarinic agonist receptor subtypes in rat striatum. *Brain Res, 206*(2), 451-456. https://doi.org/10.1016/0006-8993(81)90546-1

Bromberg-Martin, E. S., Matsumoto, M., & Hikosaka, O. (2010, Dec 9). Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron, 68*(5), 815-834. https://doi.org/10.1016/j.neuron.2010.11.022

Brown, H. D., Baker, P. M., & Ragozzino, M. E. (2010). The Parafascicular Thalamic Nucleus Concomitantly Influences Behavioral Flexibility and Dorsomedial Striatal Acetylcholine Output in Rats. *J Neurosci*, 30(43):14390 –14398. https://doi.org/10.1523/JNEUROSCI.2167-10.2010

Brown, M. T. C., Tan, K. R., O'Connor, E. C., Nikonenko, I., Muller, D., & Lüscher, C. (2012). Ventral tegmental area GABA projections pause accumbal cholinergic interneurons to enhance associative learning. *Nature*. https://doi.org/10.1038/nature11657

Cachope, R., & Cheer, J. F. (2014). Local control of striatal dopamine release. *Front Behav Neurosci*. https://doi.org/10.3389/fnbeh.2014.00188

Cachope, R., Mateo, Y., Mathur, B. N., Irving, J., Wang, H.-L., Morales, M., Lovinger, D. M., & Cheer, J. F. (2012). Selective activation of cholinergic interneurons enhances accumbal phasic dopamine release: setting the tone for reward processing. *Cell Rep*. https://doi.org/10.1016/j.celrep.2012.05.011.

Calipari, E. S., Bagot, R. C., Purushothaman, I., Davidson, T. J., Yorgason, J. T., Peña, C. J., Walker, D. M., Pirpinias, S. T., Guise, K. G., Ramakrishnan, C., Deisseroth, K., & Neslter, E. J. (2016). In vivo imaging identifies temporal signature of D1 and D2 medium spiny neurons in cocaine reward. *Proc. Natl. Acad. Sci. USA*. https://doi.org/10.1073/pnas.1521238113

Carvalho Poyraz, F., Holzner, E., Bailey, M. R., Meszaros, J., Kenney, L., Kheirbek, M. A., Balsam, P. D., & Kellendonk, C. (2016). Decreasing Striatopallidal Pathway Function Enhances Motivation by Energizing the Initiation of Goal-Directed Action. *The Journal of Neuroscience, 36*(22), 5988-6001. https://doi.org/10.1523/jneurosci.0444-16.2016

Castañé, A., Theobald, D. E., & Robbins, T. W. (2010, Jun 26). Selective lesions of the dorsomedial striatum impair serial spatial reversal learning in rats. *Behav Brain Res, 210*(1), 74-83. https://doi.org/10.1016/j.bbr.2010.02.017

Cazorla, M., de Carvalho, F. D., Chohan, M. O., Shegda, M., Chuhma, N., Rayport, S., Ahmari, S. E., Moore, H., & Kellendonk, C. (2014, Jan 8). Dopamine D2 receptors regulate the

anatomical and functional balance of basal ganglia circuitry. *Neuron, 81*(1), 153-164. https://doi.org/10.1016/j.neuron.2013.10.041

Centonze, D., Grande, C., Saulle, E., Martin, A. B., Gubellini, P., Pavón, N., Pisani, A., Bernardi, G., Moratalla, R., & Calabresi, P. (2003, Sep 17). Distinct roles of D1 and D5 dopamine receptors in motor activity and striatal synaptic plasticity. *J Neurosci, 23*(24), 8506-8512. https://doi.org/10.1523/jneurosci.23-24-08506.2003

Chuhma, N., Mingote, S., Moore, H., & Rayport, S. (2014). Dopamine neurons control striatal cholinergic neurons via regionally heterogeneous dopamine and glutamate signaling. *Neuron*. https://doi.org/10.1016/j.neuron.2013.12.027.

Clarke, H. F., Hill, G. J., Robbins, T. W., & Roberts, A. C. (2011, Mar 16). Dopamine, but not serotonin, regulates reversal learning in the marmoset caudate nucleus. *J Neurosci, 31*(11), 4290-4297. https://doi.org/10.1523/jneurosci.5066-10.2011

Clarke, H. F., Robbins, T. W., & Roberts, A. C. (2008, Oct 22). Lesions of the medial striatum in monkeys produce perseverative impairments during reversal learning similar to those produced by lesions of the orbitofrontal cortex. *J Neurosci, 28*(43), 10972-10982. https://doi.org/10.1523/jneurosci.1521-08.2008

Clatworthy, P. L., Lewis, S. J., Brichard, L., Hong, Y. T., Izquierdo, D., Clark, L., Cools, R., Aigbirhio, F. I., Baron, J. C., Fryer, T. D., & Robbins, T. W. (2009, Apr 15). Dopamine release in dissociable striatal subregions predicts the different effects of oral methylphenidate on reversal learning and spatial working memory. *J Neurosci, 29*(15), 4690-4696. https://doi.org/10.1523/jneurosci.3266-08.2009

Collins, A. L., Aitken, T. J., Greenfield, V. Y., Ostlund, S. B., & Wassum, K. M. (2016, Nov). Nucleus Accumbens Acetylcholine Receptors Modulate Dopamine and Motivation. *Neuropsychopharmacology, 41*(12), 2830-2838. https://doi.org/10.1038/npp.2016.81

Collins, A. L., Aitken, T. J., Huang, I.-W., Shieh, C., Greenfield, V. Y., Monbouquette, H. G., Ostlund, S. B., & Wassum, K. M. (2019). Nucleus accumbens cholinergic interneurons oppose cuemotivated behavior. *Biol Psychiatry*, 86(85): 388–396. https://doi.org/10.1016/j.biopsych.2019.02.014.

Contant, C., Umbriaco, D., Garcia, S., Watkins, K. C., & Descarries, L. (1996, Apr). Ultrastructural characterization of the acetylcholine innervation in adult rat neostriatum. *Neuroscience, 71*(4), 937-947. https://doi.org/10.1016/0306-4522(95)00507-2

Cools, R., Clark, L., Owen, A. M., & Robbins, T. W. (2002, Jun 1). Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. *J Neurosci, 22*(11), 4563-4567. https://doi.org/10.1523/jneurosci.22-11-04563.2002

Cooper, A. J., & Stanford, I. M. (2001, Jul). Dopamine D2 receptor mediated presynaptic inhibition of striatopallidal GABA(A) IPSCs in vitro. *Neuropharmacology*, *41*(1), 62-71. https://doi.org/10.1016/s0028-3908(01)00038-7

Corbit, L. H., Muir, J. L., & Balleine, B. W. (2001, May 1). The role of the nucleus accumbens in instrumental conditioning: Evidence of a functional dissociation between accumbens core and shell. *J Neurosci*, *21*(9), 3251-3260. https://doi.org/10.1523/jneurosci.21-09-03251.2001

Corvol, J. C., Studler, J. M., Schonn, J. S., Girault, J. A., & Hervé, D. (2001, Mar). Galpha(olf) is necessary for coupling D1 and A2a receptors to adenylyl cyclase in the striatum. *J Neurochem*, *76*(5), 1585-1588. https://doi.org/10.1046/j.1471-4159.2001.00201.x

Cover, K. K., Gyawali, U., Kerkhoff, W. G., Patton, M. H., Mu, C., White, M. G., Marquardt, A. E., Roberts, B. M., Cheer, J. F., & Mathur, B. N. (2019). Activation of the Rostral Intralaminar Thalamus Drives Reinforcement through Striatal Dopamine Release. *Cell Rep*. https://doi.org/10.1016/j.celrep.2019.01.044

Cragg, S. J. (2006). Meaningful silences: how dopamine listens to the ACh pause. *Trends Neurosci*. https://doi.org/10.1016/j.tins.2006.01.003

Dalley, J. W., Lääne, K., Theobald, D. E., Armstrong, H. C., Corlett, P. R., Chudasama, Y., & Robbins, T. W. (2005, Apr 26). Time-limited modulation of appetitive Pavlovian memory by D1 and NMDA receptors in the nucleus accumbens. *Proc Natl Acad Sci U S A*, *102*(17), 6189-6194. https://doi.org/10.1073/pnas.0502080102

Day, J. J., Roitman, M. F., Wightman, R. M., & Carelli, R. M. (2007, Aug). Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nat Neurosci*, *10*(8), 1020-1028. https://doi.org/10.1038/nn1923

Day, J. J., Roitman, M. F., Wightman, R. M., & Carelli, R. M. (2007). Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nat Neurosci*, 1020–1028. https://doi.org/10.1038/nn1923

Day, J. J., Wheeler, R. A., Roitman, M. F., & Carelli, R. M. (2006, Mar). Nucleus accumbens neurons encode Pavlovian approach behaviors: evidence from an autoshaping paradigm. *Eur J Neurosci*, *23*(5), 1341-1351. https://doi.org/10.1111/j.1460-9568.2006.04654.x

De Steno, D. A., & Schmauss, C. (2009, Aug 4). A role for dopamine D2 receptors in reversal learning. *Neuroscience*, *162*(1), 118-127. https://doi.org/10.1016/j.neuroscience.2009.04.052

DeLong, M. R. (1990, Jul). Primate models of movement disorders of basal ganglia origin. *Trends Neurosci*, *13*(7), 281-285. https://doi.org/10.1016/0166-2236(90)90110-v

Dias, R., Robbins, T. W., & Roberts, A. C. (1997, Dec 1). Dissociable forms of inhibitory control within prefrontal cortex with an analog of the Wisconsin Card Sort Test: restriction to novel situations and independence from "on-line" processing. *J Neurosci, 17*(23), 9285-9297. https://doi.org/10.1523/jneurosci.17-23-09285.1997

DiFiglia, M., Pasik, P., & Pasik, T. (1976, Sep 17). A Golgi study of neuronal types in the neostriatum of monkeys. *Brain Res, 114*(2), 245-256. https://doi.org/10.1016/0006-8993(76)90669-7

Ding, J., Peterson, J. D., & Surmeier, D. J. (2008, Jun 18). Corticostriatal and thalamostriatal synapses have distinctive properties. *J Neurosci, 28*(25), 6483-6492. https://doi.org/10.1523/jneurosci.0435-08.2008

Ding, J. B., Guzman, J. N., Peterson, J. D., Goldberg, J. A., & Surmeier, D. J. (2010). Thalamic gating of corticostriatal signaling by cholinergic interneurons. *Neuron*. https://doi.org/10.1016/j.neuron.2010.06.017.

Dobbs, L. K., Kaplan, A. R., Lemos, J. C., Matsui, A., Rubinstein, M., & Alvarez, V. A. (2016, Jun 1). Dopamine Regulation of Lateral Inhibition between Striatal Neurons Gates the Stimulant Actions of Cocaine. *Neuron, 90*(5), 1100-1113. https://doi.org/10.1016/j.neuron.2016.04.031

Doig, N. M., Magill, P. J., Apicella, P., Bolam, J. P., & Sharott, A. (2014). Cortical and Thalamic Excitation Mediate the Multiphasic Responses of Striatal Cholinergic Interneurons to Motivationally Salient Stimuli. *J Neurosci*. https://doi.org/10.1523/JNEUROSCI.4627-13.2014

Doig, N. M., Moss, J., & Bolam, J. P. (2010, Nov 3). Cortical and thalamic innervation of direct and indirect pathway medium-sized spiny neurons in mouse striatum. *J Neurosci, 30*(44), 14610-14618. https://doi.org/10.1523/jneurosci.1623-10.2010

Donthamsetti, P., Gallo, E. F., Buck, D. C., Stahl, E. L., Zhu, Y., Lane, J. R., Bohn, L. M., Neve, K. A., Kellendonk, C., & Javitch, J. A. (2020, Sep). Arrestin recruitment to dopamine D2 receptor mediates locomotion but not incentive motivation. *Mol Psychiatry, 25*(9), 2086-2100. https://doi.org/10.1038/s41380-018-0212-4

El Mestikawy, S., Wallén-Mackenzie, A., Fortin, G. M., Descarries, L., & Trudeau, L. E. (2011, Apr). From glutamate co-release to vesicular synergy: vesicular glutamate transporters. *Nat Rev Neurosci, 12*(4), 204-216. https://doi.org/10.1038/nrn2969

English, D. F., Ibanez-Sandoval, O., Stark, E., Tecuapetla, F., Buzsáki, G., Deisseroth, K., Tepper, J. M., & Koos, T. (2012a, 2012/01/01). GABAergic circuits mediate the reinforcement-related signals of striatal cholinergic interneurons. *Nature Neuroscience, 15*(1), 123-130. https://doi.org/10.1038/nn.2984

English, D. F., Ibanez-Sandoval, O., Stark, E., Tecuapetla, F., Buzsáki, G., Deisseroth, K., Tepper, J. M., & Koos, T. (2012b). GABAergic circuits mediate the reinforcement-related signals of striatal cholinergic interneurons. *Nat Neurosci*. https://doi.org/10.1038/nn.2984

Eyny, Y. S., & Horvitz, J. C. (2003, Mar 1). Opposing roles of D1 and D2 receptors in appetitive conditioning. *J Neurosci, 23*(5), 1584-1587. https://doi.org/10.1523/jneurosci.23-05-01584.2003

Faust, T. W., Assous, M., Shah, F., Tepper, J. M., & Koós, T. (2015, Jul). Novel fast adapting interneurons mediate cholinergic-induced fast GABAA inhibitory postsynaptic currents in striatal spiny neurons. *Eur J Neurosci, 42*(2), 1764-1774. https://doi.org/10.1111/ejn.12915

Faust, T. W., Assous, M., Shah, F., Tepper, J. M., & Koós, T. (2015). Novel Fast Adapting Interneurons Mediate Cholinergic-Induced Fast GABAA IPSCs In Striatal Spiny Neurons. *Eur. J. Neurosci*(42(2)), 1764-1774. https://doi.org/10.1111/ejn.12915.

Favier, M., Janickova, H., Justo, D., Kljakic, O., Runtz, L., Natsheh, J. Y., Pascoal, T. A., Germann, J., Gallino, D., Kang, J.-I., Meng, X. Q., Antinora, C., Raulic, S., Jacobsen, J. P. R., Moquin, L., Vigneuault, E., Gratton, A., Caron, M. G., Duriez, P., Brandon, M. P., Neto, P. R., Chakravarty, M. M., Herzallah, M. M., Gorwood, P., Prado, M. A., Prado, V. F., & Mestikawy, S. E. (2020). Cholinergic dysfunction in the dorsal striatum promotes habit formation and maladaptive eating. *Journal of Clinical Investigation*, 130(112):6616–6630. https://doi.org/10.1172/JCI138532.

Fields, H. L., Hjelmstad, G. O., Margolis, E. B., & Nicola, S. M. (2007). Ventral tegmental area neurons in learned appetitive behavior and positive reinforcement. *Annu Rev Neurosci, 30*, 289-316. https://doi.org/10.1146/annurev.neuro.30.051606.094341

Filla, I., Bailey, M. R., Schipani, E., Winiger, V., Mezias, C., Balsam, P. D., & Simpson, E. H. (2018, Oct). Striatal dopamine D2 receptors regulate effort but not value-based decision making and alter the dopaminergic encoding of cost. *Neuropsychopharmacology, 43*(11), 2180-2189. https://doi.org/10.1038/s41386-018-0159-9

Flagel, S. B., Clark, J. J., Robinson, T. E., Mayo, L., Czuj, A., Willuhn, I., Akers, C. A., Clinton, S. M., Phillips, P. E., & Akil, H. (2011, Jan 6). A selective role for dopamine in stimulus-reward learning. *Nature, 469*(7328), 53-57. https://doi.org/10.1038/nature09588

Floran, B., Floran, L., Sierra, A., & Aceves, J. (1997, Nov 14). D2 receptor-mediated inhibition of GABA release by endogenous dopamine in the rat globus pallidus. *Neurosci Lett, 237*(1), 1-4. https://doi.org/10.1016/s0304-3940(97)00784-2

Franklin, N. T., & Frank, M. J. (2015, Dec 25). A cholinergic feedback circuit to regulate striatal population uncertainty and optimize reinforcement learning. *eLife, 4*. https://doi.org/10.7554/eLife.12029

Gallo, E. F., Greenwald, J., Yeisley, J., Teboul, E., Martyniuk, K. M., Vallarin, J. M., Li, Y., Javith, J. A., Balsam, P. D., & Kellendonk, C. (2021). Dopamine D2 receptors modulate the cholinergic pause and inhibitory learning *Mol Psychiatry*. https://doi.org/10.1038/s41380-021-01364-y

Gallo, E. F., Meszaros, J., Sherman, J. D., Chohan, M. O., Teboul, E., Choi, C. S., Moore, H., Javitch, J. A., & Kellendonk, C. (2018, Mar 14). Accumbens dopamine D2 receptors increase motivation by decreasing inhibitory transmission to the ventral pallidum. *Nat Commun*, *9*(1), 1086. https://doi.org/10.1038/s41467-018-03272-2

Gallo, E. F., Salling, M. C., Feng, B., Morón, J. A., Harrison, N. L., Javitch, J. A., & Kellendonk, C. (2015, Jun). Upregulation of dopamine D2 receptors in the nucleus accumbens indirect pathway increases locomotion but does not reduce alcohol consumption. *Neuropsychopharmacology*, *40*(7), 1609-1618. https://doi.org/10.1038/npp.2015.11

Gerfen, C. R., Engber, T. M., Mahan, L. C., Susel, Z., Chase, T. N., Monsma, F. J., Jr., & Sibley, D. R. (1990, Dec 7). D1 and D2 dopamine receptor-regulated gene expression of striatonigral and striatopallidal neurons. *Science*, *250*(4986), 1429-1432. https://doi.org/10.1126/science.2147780

Gerfen, C. R., & Surmeier, D. J. (2011). Modulation of striatal projection systems by dopamine. *Annu Rev Neurosci*, *34*, 441-466. https://doi.org/10.1146/annurev-neuro-061010-113641

Goldberg, J. A., Ding, J. B., & Surmeier, D. J. (2012). Muscarinic modulation of striatal function and circuitry. *Handb Exp Pharmacol*(208), 223-241. https://doi.org/10.1007/978-3-642-23274-9_10

Goldberg, J. A., & Reynolds, J. N. (2011, Dec 15). Spontaneous firing and evoked pauses in the tonically active cholinergic interneurons of the striatum. *Neuroscience*, *198*, 27-43. https://doi.org/10.1016/j.neuroscience.2011.08.067

Goldberg, J. A., & Wilson, C. J. (2005, Nov 2). Control of spontaneous firing patterns by the selective coupling of calcium currents to calcium-activated potassium currents in striatal cholinergic interneurons. *J Neurosci*, *25*(44), 10230-10238. https://doi.org/10.1523/jneurosci.2734-05.2005

Gras, C., Amilhon, B., Lepicard, E. M., Poirel, O., Vinatier, J., Herbin, M., Dumas, S., Tzavara, E. T., Wade, M. R., Nomikos, G. G., Hanoun, N., Saurini, F., Kemel, M. L., Gasnier, B., Giros, B., & El Mestikawy, S. (2008, Mar). The vesicular glutamate transporter VGLUT3 synergizes striatal acetylcholine tone. *Nat Neurosci*, *11*(3), 292-300. https://doi.org/10.1038/nn2052

Graybiel, A. M. (2000, Jul 13). The basal ganglia. *Curr Biol*, *10*(14), R509-511. https://doi.org/10.1016/s0960-9822(00)00593-5

Graybiel, A. M. (2008). Habits, rituals, and the evaluative brain. *Annu Rev Neurosci, 31*, 359-387. https://doi.org/10.1146/annurev.neuro.29.051605.112851

Groman, S. M., Lee, B., London, E. D., Mandelkern, M. A., James, A. S., Feiler, K., Rivera, R., Dahlbom, M., Sossi, V., Vandervoort, E., & Jentsch, J. D. (2011, May 18). Dorsal striatal D2-like receptor availability covaries with sensitivity to positive reinforcement during discrimination learning. *J Neurosci, 31*(20), 7291-7299. https://doi.org/10.1523/jneurosci.0363-11.2011

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. (2016, Jan). Mesolimbic dopamine signals the value of work. *Nat Neurosci, 19*(1), 117-126. https://doi.org/10.1038/nn.4173

Helseth AR, H.-M. R., Hall VL, Oliver ML, Turner BD, Caffall ZF, Rittiner JE, Shipman MK, King CS, Gradinaru V, Gerfen C, Costa-Mattioli M, Calakos N. (2021). Cholinergic neurons constitutively engage the ISR for dopamine modulation and skill learning in mice. *Science, 372*(6540), eabe1931. https://doi.org/doi:10.1126/science.abe1931

Hervé, D., Le Moine, C., Corvol, J. C., Belluscio, L., Ledent, C., Fienberg, A. A., Jaber, M., Studler, J. M., & Girault, J. A. (2001, Jun 15). Galpha(olf) levels are regulated by receptor usage and control dopamine and adenosine action in the striatum. *J Neurosci, 21*(12), 4390-4399. https://doi.org/10.1523/jneurosci.21-12-04390.2001

Hervé, D., Lévi-Strauss, M., Marey-Semper, I., Verney, C., Tassin, J. P., Glowinski, J., & Girault, J. A. (1993, May). G(olf) and Gs in rat basal ganglia: possible involvement of G(olf) in the coupling of dopamine D1 receptor with adenylyl cyclase. *J Neurosci, 13*(5), 2237-2248. https://doi.org/10.1523/jneurosci.13-05-02237.1993

Heymann, G., Jo, Y. S., Reichard, K. L., McFarland, N., Chavkin, C., Palmiter, R. D., Soden, M. E., & Zweifel, L. S. (2020, Mar 4). Synergy of Distinct Dopamine Projection Populations in Behavioral Reinforcement. *Neuron, 105*(5), 909-920.e905. https://doi.org/10.1016/j.neuron.2019.11.024

Hikida, T., Kimura, K., Wada, N., Funabiki, K., & Nakanishi, S. (2010, Jun 24). Distinct roles of synaptic transmission in direct and indirect striatal pathways to reward and aversive behavior. *Neuron, 66*(6), 896-907. https://doi.org/10.1016/j.neuron.2010.05.011

Howes, O. D., Kambeitz, J., Kim, E., Stahl, D., Slifstein, M., Abi-Dargham, A., & Kapur, S. (2012, Aug). The nature of dopamine dysfunction in schizophrenia and what this means for treatment. *Arch Gen Psychiatry, 69*(8), 776-786. https://doi.org/10.1001/archgenpsychiatry.2012.169

Hyman, S. E., Malenka, R. C., & Nestler, E. J. (2006). Neural mechanisms of addiction: the role of reward-related learning and memory. *Annu Rev Neurosci, 29*, 565-598. https://doi.org/10.1146/annurev.neuro.29.051605.113009

Ikarashi, Y., Takahashi, A., Ishimaru, H., Arai, T., & Maruyama, Y. (1997, Feb). Suppression of cholinergic activity via the dopamine D2 receptor in the rat striatum. *Neurochem Int, 30*(2), 191-197. https://doi.org/10.1016/s0197-0186(96)00024-1

Izquierdo, A., Brigman, J. L., Radke, A. K., Rudebeck, P. H., & Holmes, A. (2017, Mar 14). The neural basis of reversal learning: An updated perspective. *Neuroscience, 345*, 12-26. https://doi.org/10.1016/j.neuroscience.2016.03.021

Jiang, Z. G., & North, R. A. (1991, Nov). Membrane properties and synaptic responses of rat striatal neurones in vitro. *J Physiol, 443*, 533-553. https://doi.org/10.1113/jphysiol.1991.sp018850

Jing, M., Li, Y., Zeng, J., Huang, P., Skirzewski, M., Kljakic, O., Peng, W., Qian, T., Tan, K., Zou, J., Trinh, S., Wu, R., Zhang, S., Pan, S., Hires, S. A., Xu, M., Li, H., Saksida, L. M., Prado, V. F., Bussey, T. J., Prado, M. A. M., Chen, L., Cheng, H., & Li, Y. (2020). An optimized acetylcholine sensor for monitoring in vivo cholinergic activity. *Nat Methods, 17*, 1139-1146. https://doi.org/10.1038/s41592-020-0953-2

Jing, M., Zhang, P., Wang, G., Feng, J., Mesik, L., Zeng, J., Jiang, H., Wang, S., Looby, J. C., Guagliardo, N. A., Langma, L. W., Lu, J., Zuo, Y., Talmage, D. A., Role, L. W., Barrett, P. Q., Zhang, L. I., Luo, M., Song, Y., Zhu, J. J., & Li, Y. (2018, 2018/09/01). A genetically encoded fluorescent acetylcholine indicator for in vitro and in vivo studies. *Nature Biotechnology, 36*(8), 726-737. https://doi.org/10.1038/nbt.4184

Jocham, G., Klein, T. A., Neumann, J., von Cramon, D. Y., Reuter, M., & Ullsperger, M. (2009, Mar 25). Dopamine DRD2 polymorphism alters reversal learning and associated neural activity. *J Neurosci, 29*(12), 3695-3704. https://doi.org/10.1523/jneurosci.5195-08.2009

Joel, D., & Weiner, I. (1994, Nov). The organization of the basal ganglia-thalamocortical circuits: open interconnected rather than closed segregated. *Neuroscience, 63*(2), 363-379. https://doi.org/10.1016/0306-4522(94)90536-3

Jones, B., & Mishkin, M. (1972, Aug). Limbic lesions and the problem of stimulus--reinforcement associations. *Exp Neurol, 36*(2), 362-377. https://doi.org/10.1016/0014-4886(72)90030-1

Joshua, M., Adler, A., Mitelman, R., Vaadia, E., & Bergman, H. (2008). Midbrain Dopaminergic Neurons and Striatal Cholinergic Interneurons Encode the Difference between Reward and Aversive Events at Different Epochs of Probabilistic Classical Conditioning Trials. *J Neurosci*. https://doi.org/10.1523/JNEUROSCI.3839-08.2008

Kawaguchi, Y. (1993, Nov). Physiological, morphological, and histochemical characterization of three classes of interneurons in rat neostriatum. *J Neurosci, 13*(11), 4908-4923. https://doi.org/10.1523/jneurosci.13-11-04908.1993

Kawaguchi, Y., Wilson, C. J., Augood, S. J., & Emson, P. C. (1995, Dec). Striatal interneurones: chemical, physiological and morphological characterization. *Trends Neurosci, 18*(12), 527-535. https://doi.org/10.1016/0166-2236(95)98374-8

Kebabian, J. W., & Calne, D. B. (1979, Jan 11). Multiple receptors for dopamine. *Nature, 277*(5692), 93-96. https://doi.org/10.1038/277093a0

Kellendonk, C., Simpson, E. H., Polan, H. J., Malleret, G., Vronskaya, S., Winiger, V., Moore, H., & Kandel, E. R. (2006, Feb 16). Transient and selective overexpression of dopamine D2 receptors in the striatum causes persistent abnormalities in prefrontal cortex functioning. *Neuron, 49*(4), 603-615. https://doi.org/10.1016/j.neuron.2006.01.023

Kelley, A. E., Baldo, B. A., Pratt, W. E., & Will, M. J. (2005, Dec 15). Corticostriatal-hypothalamic circuitry and food motivation: integration of energy, action and reward. *Physiol Behav, 86*(5), 773-795. https://doi.org/10.1016/j.physbeh.2005.08.066

Kemp, J. M., & Powell, T. P. (1971, Sep 30). The structure of the caudate nucleus of the cat: light and electron microscopy. *Philos Trans R Soc Lond B Biol Sci, 262*(845), 383-401. https://doi.org/10.1098/rstb.1971.0102

Kharkwal, G., Brami-Cherrier, K., Lizardi-Ortiz, J. E., Nelson, A. B., Ramos, M., Barrio, D. D., Sulzer, D., Kreitzer, A. C., & Borrelli, E. (2016). Parkinsonism driven by antipsychotics originates from dopaminergic control of striatal cholinergic interneurons. *Neuron*. https://doi.org/10.1016/j.neuron.2016.06.014.

Kimura, M. (1986, Jul). The role of primate putamen neurons in the association of sensory stimuli with movement. *Neurosci Res, 3*(5), 436-443. https://doi.org/10.1016/0168-0102(86)90035-0

Kimura, M. (1992, Apr 24). Behavioral modulation of sensory responses of primate putamen neurons. *Brain Res, 578*(1-2), 204-214. https://doi.org/10.1016/0006-8993(92)90249-9

Kimura, M., Rajkowski, J., & Evarts, E. (1984). Tonically discharging putamen neurons exhibit set-dependent responses. *Proc. Natl. Acad. Sci. USA*.

Kincaid, A. E., Zheng, T., & Wilson, C. J. (1998). Connectivity and Convergence of Single Corticostriatal Axons. *J Neurosci, 18*(12), 4722-4731. https://doi.org/10.1523/JNEUROSCI.18-12-04722.1998

Kita, H., & Kitai, S. T. (1988, May 3). Glutamate decarboxylase immunoreactive neurons in rat neostriatum: their morphological types and populations. *Brain Res, 447*(2), 346-352. https://doi.org/10.1016/0006-8993(88)91138-9

Klaus, A., Alves da Silva, J., & Costa, R. M. (2019, Jul 8). What, If, and When to Move: Basal Ganglia Circuits and Self-Paced Action Initiation. *Annu Rev Neurosci, 42*, 459-483. https://doi.org/10.1146/annurev-neuro-072116-031033

Kodama, T., Lai, Y. Y., & Siegel, J. M. (1992, May 15). Enhancement of acetylcholine release during REM sleep in the caudomedial medulla as measured by in vivo microdialysis. *Brain Res, 580*(1-2), 348-350. https://doi.org/10.1016/0006-8993(92)90967-e

Kohnomi, S., Koshikawa, N., & Kobayashi, M. (2012, Jan). D(2)-like dopamine receptors differentially regulate unitary IPSCs depending on presynaptic GABAergic neuron subtypes in rat nucleus accumbens shell. *J Neurophysiol, 107*(2), 692-703. https://doi.org/10.1152/jn.00281.2011

Kolb, B. (1977, Feb). Studies on the caudate-putamen and the dorsomedial thalamic nucleus of the rat: implications for mammalian frontal-lobe functions. *Physiol Behav, 18*(2), 237-244. https://doi.org/10.1016/0031-9384(77)90128-7

Koós, T., & Tepper, J. M. (2002, Jan 15). Dual cholinergic control of fast-spiking interneurons in the neostriatum. *J Neurosci, 22*(2), 529-535. https://doi.org/10.1523/jneurosci.22-02-00529.2002

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

Kreitzer, A. C. (2009). Physiology and pharmacology of striatal neurons. *Annu Rev Neurosci, 32*, 127-147. https://doi.org/10.1146/annurev.neuro.051508.135422

Kruzich, P. J., & Grandy, D. K. (2004, Apr 2). Dopamine D2 receptors mediate two-odor discrimination and reversal learning in C57BL/6 mice. *BMC Neurosci, 5*, 12. https://doi.org/10.1186/1471-2202-5-12

Kruzich, P. J., Mitchell, S. H., Younkin, A., & Grandy, D. K. (2006, Mar). Dopamine D2 receptors mediate reversal learning in male C57BL/6J mice. *Cogn Affect Behav Neurosci, 6*(1), 86-90. https://doi.org/10.3758/cabn.6.1.86

Labouesse, M. A., & Patriarchi, T. (2021). A versatile GPCR toolkit to track in vivo neuromodulation: not a one-size-fits-all sensor. *Neuropsychopharmacol, 46*, 2043-2047. https://doi.org/10.1038/s41386-021-00982-y

Laughlin, R. E., Grant, T. L., Williams, R. W., & Jentsch, J. D. (2011, Jun 1). Genetic dissection of behavioral flexibility: reversal learning in mice. *Biol Psychiatry, 69*(11), 1109-1116. https://doi.org/10.1016/j.biopsych.2011.01.014

Lee, B., Groman, S., London, E. D., & Jentsch, J. D. (2007, Oct). Dopamine D2/D3 receptors play a specific role in the reversal of a learned visual discrimination in monkeys. *Neuropsychopharmacology, 32*(10), 2125-2134. https://doi.org/10.1038/sj.npp.1301337

Lex, A., & Hauber, W. (2008, Jul). Dopamine D1 and D2 receptors in the nucleus accumbens core and shell mediate Pavlovian-instrumental transfer. *Learn Mem, 15*(7), 483-491. https://doi.org/10.1101/lm.978708

Lex, B., & Hauber, W. (2010, Feb). The role of nucleus accumbens dopamine in outcome encoding in instrumental and Pavlovian conditioning. *Neurobiol Learn Mem, 93*(2), 283-290. https://doi.org/10.1016/j.nlm.2009.11.002

Liles, S. L. (1974, Jan). Single-unit responses of caudate neurons to stimulation of frontal cortex, substantia nigra and entopeduncular nucleus in cats. *J Neurophysiol, 37*(1), 254-265. https://doi.org/10.1152/jn.1974.37.1.254

Mai, B., Sommer, S., & Hauber, W. (2012, Mar). Motivational states influence effort-based decision making in rats: the role of dopamine in the nucleus accumbens. *Cogn Affect Behav Neurosci, 12*(1), 74-84. https://doi.org/10.3758/s13415-011-0068-4

Malenka, R. C., & Kocsis, J. D. (1988, Oct). Presynaptic actions of carbachol and adenosine on corticostriatal synaptic transmission studied in vitro. *J Neurosci, 8*(10), 3750-3756. https://doi.org/10.1523/jneurosci.08-10-03750.1988

Martyniuk, K. M., Dandeneau, M., Balsam, P. D., & Kellendonk, C. (2021, Jun). Dopamine D2R upregulation in ventral striatopallidal neurons does not affect Pavlovian or go/no-go learning. *Behav Neurosci, 135*(3), 369-379. https://doi.org/10.1037/bne0000403

Matamales, M., Skrbis, Z., Hatch, R. J., Balleine, B. W., Götz, J., & Bertran-Gonzalez, J. (2016). Aging-Related Dysfunction of Striatal Cholinergic Interneurons Produces Conflict in Action Selection. *Neuron*. https://doi.org/10.1016/j.neuron.2016.03.006

Matsumoto, N., Minamimoto, T., Graybiel, A. M., & Kimura, M. (2000). Neurons in the Thalamic CM-Pf Complex Supply Striatal Neurons With Information About Behaviorally Significant Sensory Events. *J Neurophysiol*. https://doi.org/10.1152/jn.2001.85.2.960

Maurice, N., Liberge, M., Jaouen, F., Ztaou, S., Hanini, M., Camon, J., Deisseroth, K., Amalric, M., Goff, L. K.-L., & Beurrier, C. (2015). Striatal Cholinergic Interneurons Control Motor Behavior and Basal Ganglia Function in Experimental Parkinsonism. *Cell Rep*. https://doi.org/10.1016/j.celrep.2015.09.034

Mehta, M. A., Swainson, R., Ogilvie, A. D., Sahakian, J., & Robbins, T. W. (2001, Dec). Improved short-term spatial memory but impaired reversal learning following the dopamine D(2) agonist bromocriptine in human volunteers. *Psychopharmacology (Berl), 159*(1), 10-20. https://doi.org/10.1007/s002130100851

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*, 65–70. https://doi.org/10.1038/s41586-019-1235-y

Mohebi, A., Pettibone, J. R., Hamid, A. A., Wong, J. T., Vinson, L. T., Patriarchi, T., Tian, L., Kennedy, R. T., & Berke, J. D. (2019, Jun). Dissociable dopamine dynamics for learning and motivation. *Nature*, *570*(7759), 65-70. https://doi.org/10.1038/s41586-019-1235-y

Morris, G., Arkadir, D., Nevet, A., Vaadia, E., & Bergman, H. (2004). Coincident but Distinct Messages of Midbrain Dopamine and Striatal Tonically Active Neurons. *Neuron*. https://doi.org/2004;43:133–143.

Nasser, H., Calu, D. J., Schoenbaum, G., & Sharpe, M. J. (2017). The Dopamine Prediction Error: Contributions to Associative Models of Reward Learning. *Front Psychol*. https://doi.org/10.3389/fpsyg.2017.00244

Nicola, S. M. (2007, Apr). The nucleus accumbens as part of a basal ganglia action selection circuit. *Psychopharmacology (Berl), 191*(3), 521-550. https://doi.org/10.1007/s00213-006-0510-4

Nonneman, A. J., Voigt, J., & Kolb, B. E. (1974, Aug). Comparisons of behavioral effects of hippocampal and prefrontal cortex lesions in the rat. *J Comp Physiol Psychol, 87*(2), 249-260. https://doi.org/10.1037/h0036864

Nougaret, S., & Ravel, S. (2015). Modulation of Tonically Active Neurons of the Monkey Striatum by Events Carrying Different Force and Reward Information. *J Neurosci*. https://doi.org/10.1523/JNEUROSCI.0039-15.2015

Nowend, K. L., Arizzi, M., Carlson, B. B., & Salamone, J. D. (2001, Jul-Aug). D1 or D2 antagonism in nucleus accumbens core or dorsomedial shell suppresses lever pressing for food but leads to compensatory increases in chow consumption. *Pharmacol Biochem Behav, 69*(3-4), 373-382. https://doi.org/10.1016/s0091-3057(01)00524-x

Okada, K., Nishizawa, K., Fukabori, R., Kai, N., Shiota, A., Ueda, M., Tsutsui, Y., Sakata, S., Matsushita, N., & Kobayashi, K. (2014). Enhanced flexibility of place discrimination learning by targeting striatal cholinergic interneurons. https://doi.org/10.1038/ncomms4778

Okada, K., Nishizawa, K., Setogawa, S., Hashimoto, K., & Kobayashi, K. (2017). Task-dependent function of striatal cholinergic interneurons in behavioural flexibility. *Eur. J. Neurosci*. https://doi.org/10.1111/ejn.13768.

Ostlund, S. B., Wassum, K. M., Murphy, N. P., Balleine, B. W., & Maidment, N. T. (2011, Jan 5). Extracellular dopamine levels in striatal subregions track shifts in motivation and response cost during instrumental conditioning. *J Neurosci, 31*(1), 200-207. https://doi.org/10.1523/jneurosci.4759-10.2011

Pakhotin, P., & Bracci, E. (2007). Cholinergic Interneurons Control the Excitatory Input to the Striatum. *J Neurosci*(27(2)), 391-400.

Pakhotin, P., & Bracci, E. (2007, Jan 10). Cholinergic interneurons control the excitatory input to the striatum. *J Neurosci, 27*(2), 391-400. https://doi.org/10.1523/jneurosci.3709-06.2007

Parent, A., & Hazrati, L. N. (1995, Jan). Functional anatomy of the basal ganglia. I. The cortico-basal ganglia-thalamo-cortical loop. *Brain Res Brain Res Rev, 20*(1), 91-127. https://doi.org/10.1016/0165-0173(94)00007-c

Patriarchi, T., Cho, J. R., Merten, K., Howe, M. W., Marley, A., Xiong, W.-H., Folk, R. W., Broussard, G. J., Liang, R., Jang, M. J., Zhong, H., Dombeck, D., Zastrow, M. v., Nimmerjahn, A., Gradinaru, V., Williams, J. T., & Tian, L. (2018). Ultrafast neuronal imaging of dopamine dynamics with designed genetically encoded sensors. *Science*. https://doi.org/10.1126/science.aat4422

Phillips, P. E., Walton, M. E., & Jhou, T. C. (2007, Apr). Calculating utility: preclinical evidence for cost-benefit analysis by mesolimbic dopamine. *Psychopharmacology (Berl), 191*(3), 483-495. https://doi.org/10.1007/s00213-006-0626-6

Pisa, M., & Cyr, J. (1990, Mar 26). Regionally selective roles of the rat's striatum in modality-specific discrimination learning and forelimb reaching. *Behav Brain Res, 37*(3), 281-292. https://doi.org/10.1016/0166-4328(90)90140-a

Pisansky, M. T., Lefevre, E. M., Retzlaff, C. L., Trieu, B. H., Leipold, D. W., & Rothwell, P. E. (2019). Nucleus Accumbens Fast-Spiking Interneurons Constrain Impulsive Action. *Biol Psychiatry, 86*(11), 836-847. https://doi.org/10.1016/j.biopsych.2019.07.002

Pontieri, F. E., Tanda, G., Orzi, F., & Di Chiara, G. (1996, Jul 18). Effects of nicotine on the nucleus accumbens and similarity to those of addictive drugs. *Nature, 382*(6588), 255-257. https://doi.org/10.1038/382255a0

Precht, W., & Yoshida, M. (1971, Sep 10). Blockage of caudate-evoked inhibition of neurons in the substantia nigra by picrotoxin. *Brain Res, 32*(1), 229-233. https://doi.org/10.1016/0006-8993(71)90171-5

Prensa, L., & Parent, A. (2001, Sep 15). The nigrostriatal pathway in the rat: A single-axon study of the relationship between dorsal and ventral tier nigral neurons and the striosome/matrix striatal compartments. *J Neurosci, 21*(18), 7247-7260. https://doi.org/10.1523/jneurosci.21-18-07247.2001

Preston, R. J., Bishop, G. A., & Kitai, S. T. (1980, Feb 10). Medium spiny neuron projection from the rat striatum: an intracellular horseradish peroxidase study. *Brain Res, 183*(2), 253-263. https://doi.org/10.1016/0006-8993(80)90462-x

Ragozzino, M. E. (2007, Dec). The contribution of the medial prefrontal cortex, orbitofrontal cortex, and dorsomedial striatum to behavioral flexibility. *Ann N Y Acad Sci, 1121*, 355-375. https://doi.org/10.1196/annals.1401.013

Ragozzino, M. E., & Choi, D. (2004, Jan-Feb). Dynamic changes in acetylcholine output in the medial striatum during place reversal learning. *Learn Mem, 11*(1), 70-77. https://doi.org/10.1101/lm.65404

Ragozzino, M. E., Mohler, E. G., Prior, M., Palencia, C. A., & Rozman, S. (2009). Acetylcholine Activity in Selective Striatal Regions Supports Behavioral Flexibility. *Neurobiol Learn Mem*, 91(91): 13–22. https://doi.org/10.1016/j.nlm.2008.09.008.

Ragozzino, M. E., Ragozzino, K. E., Mizumori, S. J., & Kesner, R. P. (2002, Feb). Role of the dorsomedial striatum in behavioral flexibility for response and visual cue discrimination learning. *Behav Neurosci, 116*(1), 105-115. https://doi.org/10.1037//0735-7044.116.1.105

Ravel, S., Legallet, E., & Apicella, P. (2003). Responses of Tonically Active Neurons in the Monkey Striatum Discriminate between Motivationally Opposing Stimuli. *J Neurosci*(23), 8489–8497.

Reynolds, J. N., Hyland, B. I., & Wickens, J. R. (2004, Nov 3). Modulation of an afterhyperpolarization by the substantia nigra induces pauses in the tonic firing of striatal cholinergic interneurons. *J Neurosci, 24*(44), 9870-9877. https://doi.org/10.1523/jneurosci.3225-04.2004

Rice, M. E., & Cragg, S. J. (2004, Jun). Nicotine amplifies reward-related dopamine signals in striatum. *Nat Neurosci, 7*(6), 583-584. https://doi.org/10.1038/nn1244

Rivera, A., Alberti, I., Martín, A. B., Narváez, J. A., de la Calle, A., & Moratalla, R. (2002, Dec). Molecular phenotype of rat striatal neurons expressing the dopamine D5 receptor subtype. *Eur J Neurosci, 16*(11), 2049-2058. https://doi.org/10.1046/j.1460-9568.2002.02280.x

Rogers, R. D., Andrews, T. C., Grasby, P. M., Brooks, D. J., & Robbins, T. W. (2000, Jan). Contrasting cortical and subcortical activations produced by attentional-set shifting and reversal learning in humans. *J Cogn Neurosci, 12*(1), 142-162. https://doi.org/10.1162/089892900561931

Rossi, J., Balthasar, N., Olson, D., Scott, M., Berglund, E., Lee, C. E., Choi, M. J., Lauzon, D., Lowell, B. B., & Elmquist, J. K. (2011). Melanocortin-4 Receptors Expressed by Cholinergic Neurons Regulate Energy Balance and Glucose Homeostasis. *Cell Metab*, 13, 195–204. https://doi.org/10.1016/j.cmet.2011.01.010

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

Roughley, S., & Killcross, S. (2019, Jun). Differential involvement of dopamine receptor subtypes in the acquisition of Pavlovian sign-tracking and goal-tracking responses.

Psychopharmacology (Berl), 236(6), 1853-1862. https://doi.org/10.1007/s00213-019-5169-8

Salamone, J. D. (1998). Dopaminergic involvement in activational aspects of motivation: effects of haloperidol on schedule-induced activity, feeding, and foraging in rats. *Psychobiology, 16*, 196-206.

Salamone, J. D., Correa, M., Farrar, A., & Mingote, S. M. (2007, Apr). Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. *Psychopharmacology (Berl), 191*(3), 461-482. https://doi.org/10.1007/s00213-006-0668-9

Schultz, W. (1986, Nov). Responses of midbrain dopamine neurons to behavioral trigger stimuli in the monkey. *J Neurophysiol, 56*(5), 1439-1461. https://doi.org/10.1152/jn.1986.56.5.1439

Schultz, W. (1998, Jul). Predictive reward signal of dopamine neurons. *J Neurophysiol, 80*(1), 1-27. https://doi.org/10.1152/jn.1998.80.1.1

Schultz, W. (2007). Behavioral dopamine signals. *Trends Neurosci*. https://doi.org/10.1016/j.tins.2007.03.007

Schultz, W., Dayan, P., & Montague, P. R. (1997). A Neural Substrate of Prediction and Reward. *Science*. https://doi.org/10.1126/science.275.5306.1593.

Shimo, Y., & Hikosaka, O. (2001). Role of Tonically Active Neurons in Primate Caudate in Reward-Oriented Saccadic Eye Movement. *J Neurosci*(19), 7804-7814.

Shin, J. H., Adrover, M. F., Wess, J., & Alvarez, V. A. (2015, Jun 30). Muscarinic regulation of dopamine and glutamate transmission in the nucleus accumbens. *Proc Natl Acad Sci U S A, 112*(26), 8124-8129. https://doi.org/10.1073/pnas.1508846112

Simpson, E. H., & Balsam, P. D. (2016). The Behavioral Neuroscience of Motivation: An Overview of Concepts, Measures, and Translational Applications. *Curr Top Behav Neurosci, 27*, 1-12. https://doi.org/10.1007/7854_2015_402

Simpson, E. H., & Kellendonk, C. (2017, Jan 1). Insights About Striatal Circuit Function and Schizophrenia From a Mouse Model of Dopamine D(2) Receptor Upregulation. *Biol Psychiatry, 81*(1), 21-30. https://doi.org/10.1016/j.biopsych.2016.07.004

Smith, A. G., Neill, J. C., & Costall, B. (1999, Jun). The dopamine D3/D2 receptor agonist 7-OH-DPAT induces cognitive impairment in the marmoset. *Pharmacol Biochem Behav, 63*(2), 201-211. https://doi.org/10.1016/s0091-3057(98)00230-5

Smith, Y., Bennett, B. D., Bolam, J. P., Parent, A., & Sadikot, A. F. (1994, Jun). SYNAPTIC RELATIONSHIPS BETWEEN DOPAMINERGIC AFFERENTS AND CORTICAL OR THALAMIC INPUT IN THE SENSORIMOTOR TERRITORY OF THE STRIATUM

IN MONKEY. *Journal of Comparative Neurology*, *344*(1), 1-19. https://doi.org/10.1002/cne.903440102

Smith, Y., Bevan, M. D., Shink, E., & Bolam, J. P. (1998, Sep). Microcircuitry of the direct and indirect pathways of the basal ganglia. *Neuroscience*, *86*(2), 353-387. https://doi.org/10.1016/s0306-4522(98)00004-9

Smith, Y., Raju, D. V., Pare, J. F., & Sidibe, M. (2004, Sep). The thalamostriatal system: a highly specific network of the basal ganglia circuitry. *Trends Neurosci*, *27*(9), 520-527. https://doi.org/10.1016/j.tins.2004.07.004

Stalnaker, T. A., Berg, B., Aujla, N., & Schoenbaum, G. (2016, Jun 8). Cholinergic Interneurons Use Orbitofrontal Input to Track Beliefs about Current State. *J Neurosci*, *36*(23), 6242-6257. https://doi.org/10.1523/jneurosci.0157-16.2016

Steinberg, E. E., Keiflin, R., Boivin, J. R., Witten, I. B., Deisseroth, K., & Janak, P. H. (2013, Jul). A causal link between prediction errors, dopamine neurons and learning. *Nat Neurosci*, *16*(7), 966-973. https://doi.org/10.1038/nn.3413

Steinberg, E. E., Keiflin, R., Boivin, J. R., Witten, I. B., Deisseroth, K., & Janak, P. H. (2013). A causal link between prediction errors, dopamine neurons and learning. *Nat Neurosci*, 966–973 https://doi.org/10.1038/nn.3413

Straub, C., Tritsch, N. X., Hagan, N. A., Gu, C., & Sabatini, B. L. (2014). Multiphasic Modulation of Cholinergic Interneurons by Nigrostriatal Afferents. *J Neurosci*. https://doi.org/10.1523/JNEUROSCI.0589-14.2014

Sulzer, D., Cragg, S. J., & Rice, M. E. (2016). Striatal dopamine neurotransmission: Regulation of release and uptake. *Basal Ganglia*. https://doi.org/10.1016/j.baga.2016.02.001

Sulzer, D., Cragg, S. J., & Rice, M. E. (2016, Aug). Striatal dopamine neurotransmission: regulation of release and uptake. *Basal Ganglia*, *6*(3), 123-148. https://doi.org/10.1016/j.baga.2016.02.001

Surmeier, D. J., Song, W. J., & Yan, Z. (1996, Oct 15). Coordinated expression of dopamine receptors in neostriatal medium spiny neurons. *J Neurosci*, *16*(20), 6579-6591. https://doi.org/10.1523/jneurosci.16-20-06579.1996

Tai, L. H., Lee, A. M., Benavidez, N., Bonci, A., & Wilbrecht, L. (2012, Sep). Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value. *Nat Neurosci*, *15*(9), 1281-1289. https://doi.org/10.1038/nn.3188

Tecuapetla, F., Koós, T., Tepper, J. M., Kabbani, N., & Yeckel, M. F. (2009, Jul 15). Differential dopaminergic modulation of neostriatal synaptic connections of striatopallidal axon collaterals. *J Neurosci*, *29*(28), 8977-8990. https://doi.org/10.1523/jneurosci.6145-08.2009

Threlfell, S., Lalic, T., Platt, N. J., Jennings, K. A., Deisseroth, K., & Cragg, S. J. (2012). Striatal Dopamine Release Is Triggered by Synchronized Activity in Cholinergic Interneurons. *Neuron*. https://doi.org/10.1016/j.neuron.2012.04.038

Trifilieff, P., Feng, B., Urizar, E., Winiger, V., Ward, R. D., Taylor, K. M., Martinez, D., Moore, H., Balsam, P. D., Simpson, E. H., & Javitch, J. A. (2013, Sep). Increasing dopamine D2 receptor expression in the adult nucleus accumbens enhances motivation. *Mol Psychiatry*, *18*(9), 1025-1033. https://doi.org/10.1038/mp.2013.57

Tsutsui-Kimura, I., Natsubori, A., Mori, M., Kobayashi, K., Drew, M. R., de Kerchove d'Exaerde, A., Mimura, M., & Tanaka, K. F. (2017, Oct 9). Distinct Roles of Ventromedial versus Ventrolateral Striatal Medium Spiny Neurons in Reward-Oriented Behavior. *Curr Biol*, *27*(19), 3042-3048.e3044. https://doi.org/10.1016/j.cub.2017.08.061

Wanat, M. J., Kuhnen, C. M., & Phillips, P. E. (2010, Sep 8). Delays conferred by escalating costs modulate dopamine release to rewards but not their predictors. *J Neurosci*, *30*(36), 12020-12027. https://doi.org/10.1523/jneurosci.2691-10.2010

Ward, R. D., Simpson, E. H., Richards, V. L., Deo, G., Taylor, K., Glendinning, J. I., Kandel, E. R., & Balsam, P. D. (2012, Jun). Dissociation of hedonic reaction to reward and incentive motivation in an animal model of the negative symptoms of schizophrenia. *Neuropsychopharmacology*, *37*(7), 1699-1707. https://doi.org/10.1038/npp.2012.15

Watanabe, K., & Kimura, M. (1998). Dopamine Receptor–Mediated Mechanisms Involved in the Expression of Learned Activity of Primate Striatal Neurons. *J Neurophysiol*. https://doi.org/10.1152/jn.1998.79.5.2568

Weinstein, J. J., Chohan, M. O., Slifstein, M., Kegeles, L. S., Moore, H., & Abi-Dargham, A. (2017, Jan 1). Pathway-Specific Dopamine Abnormalities in Schizophrenia. *Biol Psychiatry*, *81*(1), 31-42. https://doi.org/10.1016/j.biopsych.2016.03.2104

Welch, A. C., Zhang, J., Lyu, J., McMurray, M. S., Javitch, J. A., Kellendonk, C., & Dulawa, S. C. (2021, Aug). Dopamine D2 receptor overexpression in the nucleus accumbens core induces robust weight loss during scheduled fasting selectively in female mice. *Mol Psychiatry*, *26*(8), 3765-3777. https://doi.org/10.1038/s41380-019-0633-8

Wieland, S., Du, D., Oswald, M. J., Parlato, R., Köhr, G., & Kelsch, W. (2014). Phasic Dopaminergic Activity Exerts Fast Control of Cholinergic Interneuron Firing via Sequential NMDA, D2, and D1 Receptor Activation. *J Neurosci*. https://doi.org/10.1523/JNEUROSCI.1175-14.2014

Wilson, C. J. (2005). The Mechanism of Intrinsic Amplification of Hyperpolarizations and Spontaneous Bursting in Striatal Cholinergic Interneurons. *Neuron*. https://doi.org/10.1016/j.neuron.2004.12.053

Wilson CJ, C. H., Kitai ST. (1990). Firing Patterns and Synaptic Potentials of Identified Giant Aspiny Interneurons in the Rat Neostriatum. *J Neurosci*. https://doi.org/10(2): 508-519

Wilson, C. J., & Goldberg, J. A. (2006, Jan). Origin of the slow afterhyperpolarization and slow rhythmic bursting in striatal cholinergic interneurons. *J Neurophysiol, 95*(1), 196-204. https://doi.org/10.1152/jn.00630.2005

Witten, I. B., Lin, S.-C., Brodsky, M., Prakash, R., Diester, I., Anikeeva, P., Gradinaru, V., Ramakrishnan, C., & Deisseroth, K. (2010). Cholinergic Interneurons Control Local Circuit Activity and Cocaine Conditioning. *Science*. https://doi.org/10.1126/science.1193771.

Yan, Z., Song, W. J., & Surmeier, J. (1997, Feb). D2 dopamine receptors reduce N-type Ca2+ currents in rat neostriatal cholinergic interneurons through a membrane-delimited, protein-kinase-C-insensitive pathway. *J Neurophysiol, 77*(2), 1003-1015. https://doi.org/10.1152/jn.1997.77.2.1003

Yan, Z., & Surmeier, D. J. (1991). D5 Dopamine Receptors Enhance Zn21-Sensitive GABAA Currents in Striatal Cholinergic Interneurons through a PKA/PP1 Cascade. *Neuron*. https://doi.org/10.1016/s0896-6273(00)80402-x.

Yan, Z., & Surmeier, D. J. (1997, Nov). D5 dopamine receptors enhance Zn2+-sensitive GABA(A) currents in striatal cholinergic interneurons through a PKA/PP1 cascade. *Neuron, 19*(5), 1115-1126. https://doi.org/10.1016/s0896-6273(00)80402-x

Yin, H. H., & Knowlton, B. J. (2006, Jun). The role of the basal ganglia in habit formation. *Nat Rev Neurosci, 7*(6), 464-476. https://doi.org/10.1038/nrn1919

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

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

Yoshida, M., & Precht, W. (1971, Sep 10). Monosynaptic inhibition of neurons of the substantia nigra by caudato-nigral fibers. *Brain Res, 32*(1), 225-228. https://doi.org/10.1016/0006-8993(71)90170-3

Zhang, H., & Sulzer, D. (2004, Jun). Frequency-dependent modulation of dopamine release by nicotine. *Nat Neurosci, 7*(6), 581-582. https://doi.org/10.1038/nn1243

Zhang, Y.-F., & Cragg, S. J. (2017). Pauses in Striatal Cholinergic Interneurons: What is Revealed by Their Common Themes and Variations? *Front Syst Neurosci*. https://doi.org/10.3389/fnsys.2017.00080

Zhang, Y.-F., Reynolds, J. N. J., & Cragg, S. J. (2018). Pauses in Cholinergic Interneuron Activity Are Driven by Excitatory Input and Delayed Rectification with Dopamine Modulation. *Neuron*. https://doi.org/10.1016/j.neuron.2018.04.027

Zhou, F. M., Liang, Y., & Dani, J. A. (2001, Dec). Endogenous nicotinic cholinergic activity regulates dopamine release in the striatum. *Nat Neurosci, 4*(12), 1224-1229. https://doi.org/10.1038/nn769

Zucca, S., Zucca, A., Nakano, T., Aoki, S., & Wickens, J. (2018, Mar 26). Pauses in cholinergic interneuron firing exert an inhibitory control on striatal output in vivo. *eLife, 7*. https://doi.org/10.7554/eLife.32510

Appendix A: Dopamine D2R Upregulation in Ventral Striatopallidal Neurons Does Not Affect Pavlovian or Go/No-Go Learning

A.1 Abstract

Ventral striatal dopamine is thought to be important for associative learning. Dopamine exerts its role via activation of dopamine D1 and D2 receptors in the ventral striatum. Upregulation of dopamine D2R in ventral striatopallidal neurons impairs incentive motivation via inhibiting synaptic transmission to the ventral pallidum. Here, we determined whether upregulation of D2Rs and the resulting impairment in ventral striatopallidal pathway function modulates associative learning in an auditory Pavlovian reward learning task as well as Go/No-Go learning in an operant based reward driven Go/No-Go task. We found that upregulation of D2Rs in indirect pathway neurons of the NAc did not affect Pavlovian learning or the extinction of Pavlovian responses, and neither did it alter No-Go learning. A delay in the Go component of the task however could indicate a deficit in learning though it may be attributed to locomotor hyperactivity of the mice. In combination with previously published findings our data suggest that D2Rs in ventral striatopallidal neurons play a specific role in regulating motivation by balancing cost/benefit computations but do not necessarily affect associative learning.

* This Appendix was published as an article in Behavioral Neuroscience in 2021. **Martyniuk KM,** Dandeneau M, Balsam PD, Kellendonk C. Dopamine D2R upregulation in ventral striatopallidal neurons does not affect Pavlovian or go/no-go learning. Behav Neurosci. 2021 Jun;135(3):369-379. doi: 10.1037/bne0000403. PMID: 34264690; PMCID: PMC8288487.

K.M.M. and C.K. designed the experiments. K.M.M performed the experiments and analyzed the data. M.D. assisted in the performance of the experiments. K.M.M., P.D.B. and C.K. interpreted the results and wrote the manuscript.

A.2 Introduction

The role of ventral striatal dopamine and its receptors in the regulation motivation and learning has been an intensive area of study for the last decades. Pharmacological studies have uncovered an important role for dopamine receptors in the nucleus accumbens (NAc) in the regulation of incentive motivation and the willingness to work for reward (Aberman et al., 1998; Berridge, 2007; Salamone et al., 2007). In this context dopamine is thought to regulate effort-related processes that are important to overcome work-related response costs rather than to adapt the animal's response to changes in reward value (Filla et al., 2018; Hamid et al., 2016; Kelley et al., 2005; Ostlund et al., 2011; Phillips et al., 2007; Salamone et al., 2007; Wanat et al., 2010). Upregulation of dopamine D2 receptors (D2Rs) in the adult NAc core enhances performance in a progressive ratio and a concurrent choice task that probe for incentive motivation and effort related decision making (Donthamsetti et al., 2020; Gallo et al., 2018; Trifilieff et al., 2013). Notably, cell specific upregulation of D2Rs in ventral striatopallidal projection neurons (D2R-OE_{NAcInd} mice) is sufficient to enhance motivation, whereas upregulation in cholinergic interneurons (D2R-OE_{ChAT} mice), which also express D2Rs had no effect on progressive ratio performance (Gallo et al., 2018).

D2Rs in ventral striatopallidal neurons are transported to axonal terminals, where they reduce inhibitory transmission at intra-striatal collaterals and striato-pallidal synapses (Cooper & Stanford, 2001; Dobbs et al., 2016; Floran et al., 1997; Kohnomi et al., 2012; Tecuapetla et al., 2009). Slice physiological recordings revealed that D2R upregulation in ventral striatopallidal neurons enhances this modulation by dopamine. Thus, D2R-OE_{NAcInd} mice display decreased baseline synaptic transmission and an enhanced inhibition of synaptic transmission by D2R

activation (Gallo et al., 2018). As you would expect this effect was recorded at intra-striatal collaterals to the direct pathway and the canonical projections to the ventral pallidum (Gallo et al., 2018). A follow up *in vivo* physiological analysis showed that the effects of disinhibition in D2R-OE_{NAcInd} mice are mostly measurable at the level of the striato-pallidal synapse (Gallo et al., 2018). Furthermore, selective inhibition of striato-pallidal synapses in the ventral pallidum is sufficient to enhance progressive ratio performance suggesting that ventral striatopallidal D2Rs promote incentive motivation via enhanced inhibition of striato-pallidal transmission (Gallo et al., 2018).

Ventral striatal dopamine has also been implicated in associative learning. Dopamine neurons have been shown to encode a reward prediction error providing a teaching signal that is required for learning and that is thought to be transmitted to the NAc via the release of dopamine (J. J. Day et al., 2007; Schultz et al., 1997; E. E. Steinberg et al., 2013). D2R-OE_{NAcInd} mice should be more sensitive to this signal so that dopamine released in response to a reward predicting cue leads to a stronger inhibition of synaptic transmission, which could affect associative learning. To address this hypothesis, we tested D2R-OE_{NAcInd} mice in Pavlovian conditioning, an associative learning task. In this task mice learn that an auditory stimulus (conditioned stimulus: CS+) predicts the delivery of a food reward, whereas a different auditory stimulus (CS-) is not reinforced. Importantly, CS+ presentation leads to the release of dopamine when mice are acquiring the task (Bailey et al., 2018). We then extinguished the importance of the CS+ by adding 5 days of extinction training in which animals were not rewarded.

As inhibition of the ventral striatopallidal pathway has been shown to increase response initiation, we further hypothesized that D2R upregulation impairs learning if actions must be suppressed (Carvalho Poyraz et al., 2016). We thus tested D2R-OE_{NAcInd} mice in an instrumental Go/No-Go learning task where in a first step mice learn to press a lever in the presence of a visual

stimulus. In a second step they then must learn to withhold from pressing the lever when the stimulus is absent. Last, we measured the activity of D2R-OE$_{NAcInd}$ mice in an open field to determine the functionality of the upregulated receptors in these new cohorts of mice.

We replicated previous findings showing hyperactivity in the open field (Donthamsetti et al., 2020; Gallo et al., 2015). In contrast to our expectations, D2R upregulation neither affected Pavlovian, extinction nor Go/No-Go learning suggesting that D2R upregulation in ventral striatopallidal neurons enhance motivation (Donthamsetti et al., 2020; Gallo et al., 2015) but does not affect Pavlovian or No-Go learning.

A.3 Results

To test the effects of D2R upregulation in ventral striatopallidal neurons on Pavlovian conditioning and Go/No-Go learning we generated two cohorts of mice. Cohort 1 was first tested in the Pavlovian conditioning task followed by the Go/No-Go task. Cohort 2 was first tested in the Pavlovian conditioning task followed by an extinction procedure. At the end of behavioral testing both cohorts were run in the open field as a positive control. Adult D2R-OE$_{NAcInd}$ mice were generated by injecting a Cre-dependent AAV1 expressing D2R-ires-mVenus or GFP into the NAc core of Drd2-Cre mice that express Cre recombinase in ventral striatopallidal neurons. This leads to a 3-fold increase in D2 receptor levels in the NAc (Gallo et al., 2018). Figure A.1A shows a

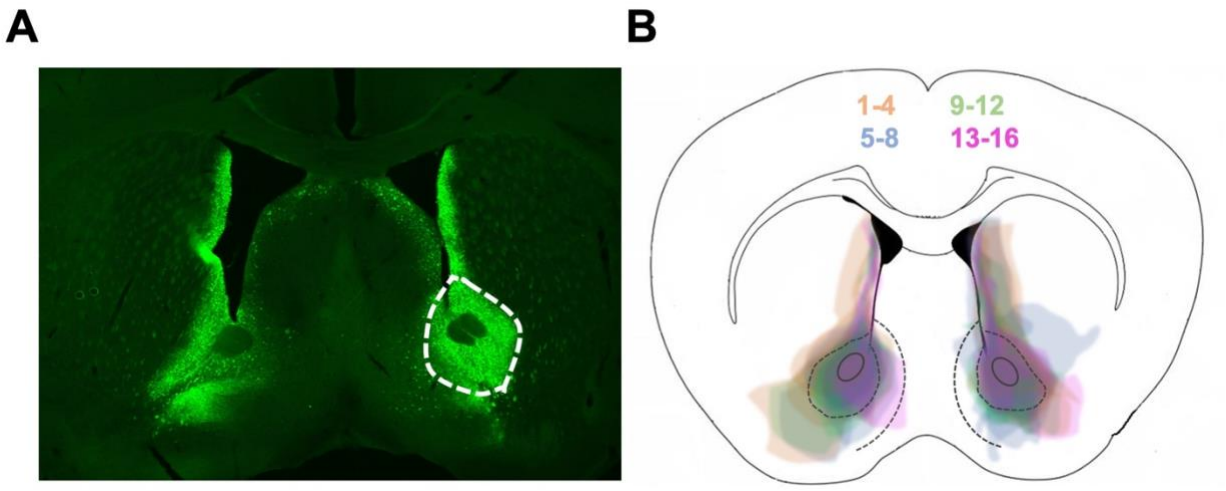

Figure A.1. Confirmation of viral spread. (A) Coronal section showing D2R-mVenus expression in the NAc of a *Drd2*-Cre mouse. **(B)** Superimposed traces of viral spread from coronal sections at ~1.0 mm anterior to bregma for all 16 D2R-OE$_{NAcInd}$ mice with the AAV1-hSyn-DIO-D2R-mVenus into the NAc. Mice are shown in 4 different colors (4 mice/color) for better visualization of viral spread.

representative image of a coronal section from a mouse with bilateral expression of the D2R-ires-mVenus virus in the NAc. Figure A.1B shows the spread of virus-mediated expression with injections from all mice superimposed on each other in different colors for better visualization. We see dense viral expression in the NAc core with some leakage into the lateral NAc shell and dorsal medial striatum (DMS).

Four weeks after AAV injections mice were tested in the Pavlovian conditioning task. First, mice were trained in an operant box to retrieve a food reward (evaporated milk) from an automatic dipper. Over the two training days, both groups learned that the dipper provided a reward (Figure A.2). There was no difference between the two groups, EGFP$_{NAcInd}$ and D2R-OE$_{NAcInd}$ mice decreased their latency to retrieve the milk reward (Figure A.2A, RM 2-way ANOVA: $F(1,15)= 123.1$, $p< 0.0001$, main effect of day) and increased the total number of rewards retrieved (Figure

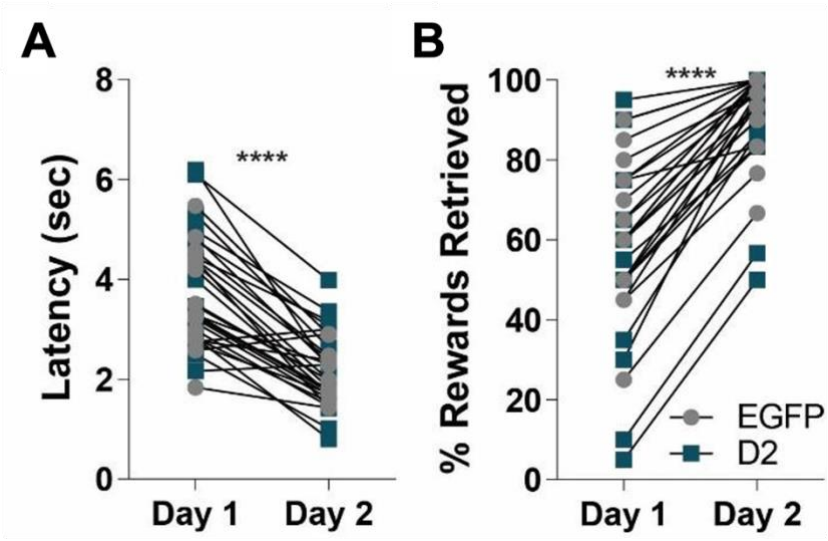

Figure A.2. D2R upregulation does not impair trough training. (A) Latency to retrieve an unexpected milk reward via head entry into a retractable dipper significantly decreased over 2 days of training for both D2R-OE$_{NAcInd}$ (dark squares) and EGFP$_{NAcInd}$ (light circles) mice. (**** $p<0.0001$). **(B)** Total number of rewards retrieved significantly increased from Day 1 to Day 2 for both groups (**** $p<0.0001$). Data from 16 animals/genotype was used to calculate all statistics reported.

A.2B, Mann Whitney, $p< 0.0001$, main effect of day). These results indicate that D2R upregulation in ventral striatopallidal neurons does not affect reward retrieval. During auditory Pavlovian conditioning, mice were presented with two 10 second auditory cues (tone verses white noise, counterbalanced between viral groups). After offset of the conditioned stimulus (CS+), a milk reward was delivered whereas no reward was delivered after the offset of the unconditioned stimulus (CS-). Figure A.3Aprovides a schematic of the task. Both groups of mice increased their head entries during the CS+ presentation over the 16 days of training (Figure A.3B). In contrast, during the CS-, head entries increased slightly during the first 4 days and then went down (Figure A.3B). Both groups were able to distinguish between the CS+ and CS- and demonstrated learning

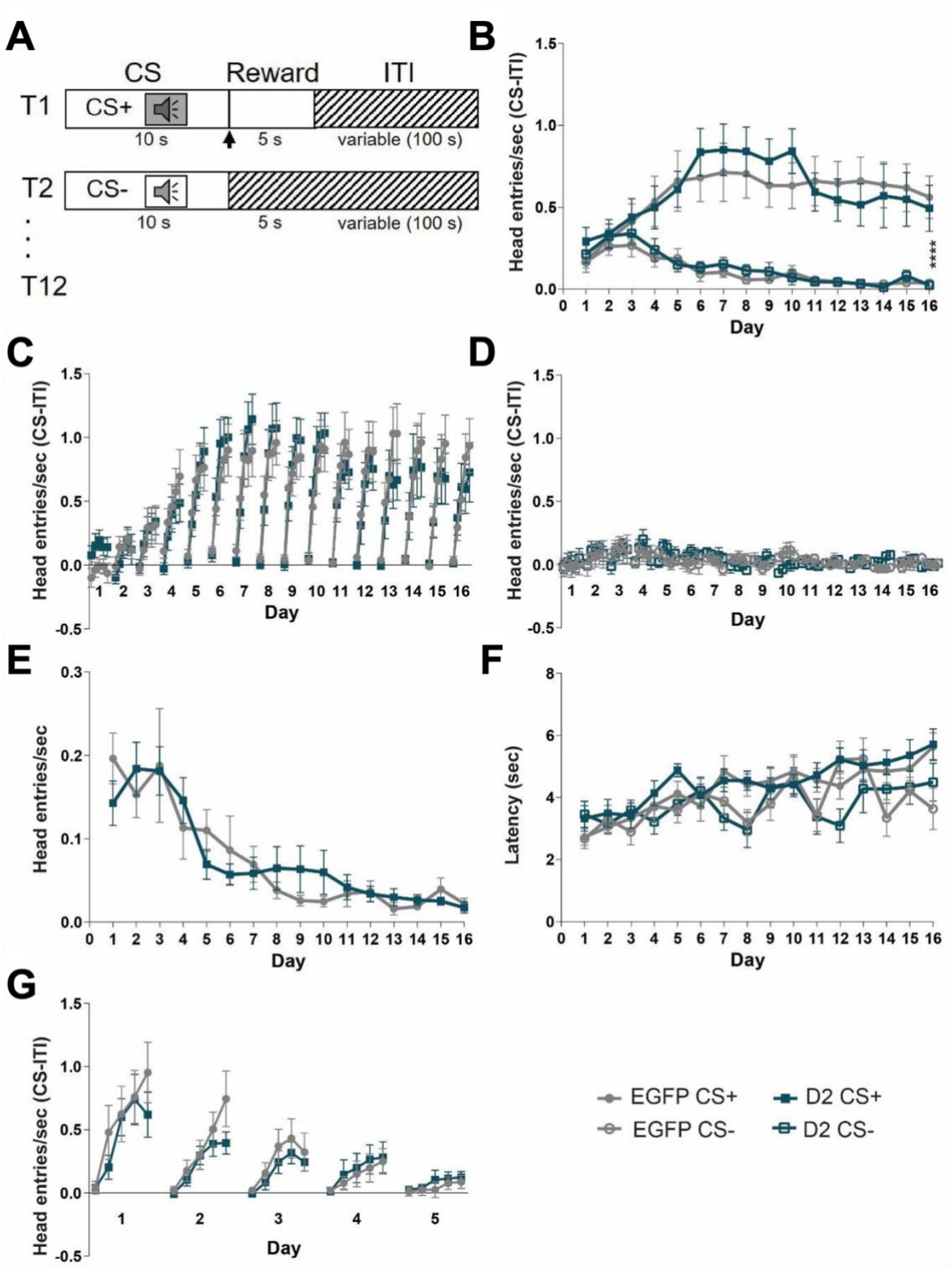
A
CS
Reward
ITI
T1
CS+
10 s
5 s
variable (100 s)
T2
CS-
10 s
5 s
variable (100 s)
T12
B
Head entries/sec (CS-ITI)
Day
****
C
Head entries/sec (CS-ITI)
Day
D
Head entries/sec (CS-ITI)
Day
E
Head entries/sec
Day
F
Latency (sec)
Day
G
Head entries/sec (CS-ITI)
Day
EGFP CS+
EGFP CS-
D2 CS+
D2 CS-

Figure A.3. D2R overexpression does not impair Pavlovian learning. (A) Task design. Mice are trained with 24 (12 CS+, 12 CS-) trials/day for 16 days. Each trial starts with a 10 sec tone (CS+ or CS-). At the end of the CS+ a dipper comes up presenting milk as a food reward for 5 sec. There is an intertrial interval (ITI) variable in length (100 sec). **(B)** To determine learning, anticipatory responses (head entries/sec (CS-ITI)) were measured. Anticipatory head entries during the CS+ increased for both $EGFP_{NAcInd}$ (filled light circles) and D2R-OE_{NAcInd} (filled dark squares) mice and decreased during the CS- for both $EGFP_{NAcInd}$ (open light circles) and D2R-OE_{NAcInd} (open dark squares) mice. Both groups learned to distinguish the CS+ and CS- (**** $p<0.0001$). There was no effect of genotype on learning over the 16 days of training ($p= 0.9977$). **(C)** Within each day, the 10 sec CS+ was split into 5 two-second time bins to better visualize the anticipatory response. $EGFP_{NAcInd}$ (filled light circles) and D2R-OE_{NAcInd} (filled dark squares) sharply increased head entries into the reward port across the 10 sec CS+ ($p< 0.0001$) and continued this anticipatory behavior across the 16 days of training ($p< 0.0001$). There was no effect of genotype on learning ($p= 0.9375$). **(D)** $EGFP_{NAcInd}$ (open light circles) and D2R-OE_{NAcInd} (open dark squares) showed no anticipatory response during the 10 sec CS-. There was no effect of genotype on learning ($p= 0.9988$). **(E)** Head entries during the ITI decreased over the 16 days of training comparably for both groups with no difference between groups ($p= 0.6173$). **(F)** Latency for the first head entry during the CS+ and CS- increased for the $EGFP_{NAcInd}$ (filled light circles/open light circles) and D2R-OE_{NAcInd} (filled dark squares/open dark squares). There was no difference between groups. **(G)** Anticipatory response to the CS+ attenuated over 5 days of extinction learning. Both groups similarly decreased anticipatory head entries during the CS+ ($p< 0.0001$). There was no difference in extinction learning between the two groups ($p= 0.5247$). Data from 8 or 16 animals per genotype was used to calculate all statistics reported.

across 16 days of training (RM 3-way ANOVA: $F(15, 15)= 4.744$, $p< 0.0001$, main effect of conditioned stimuli over 16 days of training). However, there were no significant difference in the interaction between rate of anticipatory head poking during the CS+ or CS- between D2R and EGFP expressing mice over the 16 days of training (Figure A.3B, RM 3-way ANOVA: $F(15,15)= 0.2672$, $p= 0.9977$, N=16/group).

To better visualize the pattern of anticipatory head entries, we plotted head entries during the CS+ and CS- in 2 second bins. As mice learn the fixed duration of the CS+, their anticipatory response sharply increased during the 10 second CS+ (Figure A.3C). We found that both D2R-OE_{NAcInd} and $EGFP_{NAcInd}$ mice comparably increased their head entries over the duration of the CS+ and continued this pattern of anticipatory behavior across the 16 days of training. However, there was no significant interaction between days, head entries during the CS+ and viral manipulation (Figure A.3C, RM 3-way ANOVA:0.5077, p=0.9375). Similarly, we observed no effect of day, head entries during CS- presentation and viral manipulation (Figure A.3D, RM 3-way ANOVA: $F(15, 15)= 0.2368$, $p= 0.9988$). In addition, we observed no effect of viral manipulation (D2R vs

EGFP) on the rate of head entries during the variable intertrial interval (ITI) (Figure A.3E, RM 2-way ANOVA: $F(15, 450)= 0.8535$, $p=0.6173$). To further quantify their ability to learn the duration of both cues, we measured the latency of first head entry with cue onset. For both groups, this latency increased only during the CS+ over the 16 days of conditioning, but no difference between groups (Figure A.3F). This result further indicates that the mice have learned both the cue that predicts the reward and the timing of the reward that follows.

To test the effects of D2R upregulation on eliminating a conditioned response, we ran our 2nd cohort of mice through 5 days of Pavlovian extinction. To reinstate the conditioned behavior, the mice received 3 days of Pavlovian conditioning as previously described (Figure A.3A). For extinction, the protocol was identical to Pavlovian conditioning except now the mice no longer received a reward following the CS+. Both D2R-OE_{NAcInd} and $EGFP_{NAcInd}$ mice attenuated their anticipatory head poking during the CS+ each day with near complete extinction by Day 5 (Figure A.3G). Both groups had a comparable decay in their anticipatory head entries during the 10 second CS+ over the 5 days of extinction testing (RM 2-way ANOVA: $F(4,4)= 6.536$, $p<0.0001$, main effect of day). However, there was no interaction between day, CS+ anticipatory head entries and viral manipulation (RM 3-way ANOVA: $F(4,4)= 0.8083$, $p= 0.5247$), suggesting that D2R upregulation has no effect on extinction learning.

Next, to determine if D2R-OE$_{NAcInd}$ mice have issues learning to withhold responses, we first trained mice to press a lever to earn a food reward and then tested them in a Go/No-Go task. Mice learned to press a lever to receive a milk reward using a continuous reinforcement (CRF) schedule. Over the two days of training, both D2R-OE$_{NAcInd}$ and EGFP$_{NAcInd}$ mice were able to complete all trials (30 or 60) during the allotted time (60 min) (Figure A.4A**)**. Furthermore, both

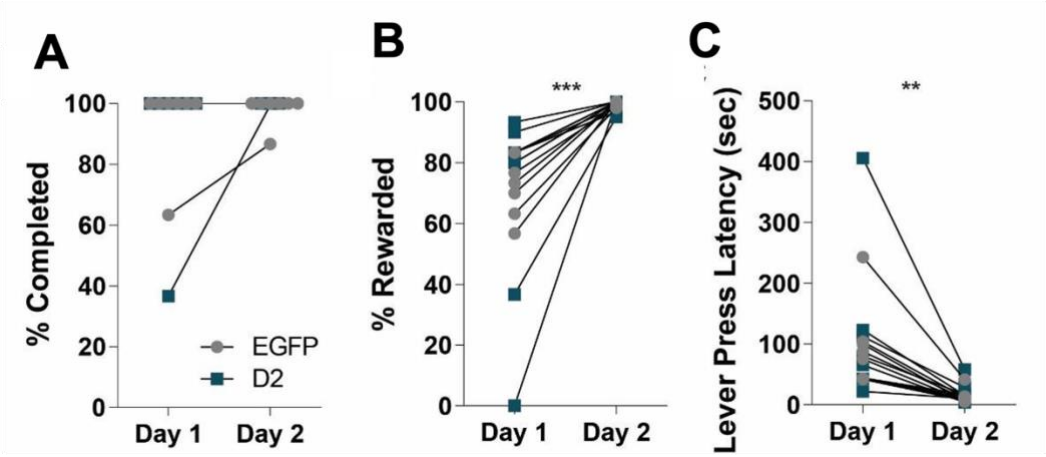

Figure A.4. D2R upregulation does not affect reinforcement learning. Mice were trained to press a lever to obtain a milk reward on a continuous reinforcement (CRF) schedule for two days. **(A)** EGFP$_{NAcInd}$ (light circles) and D2R-OE$_{NAcInd}$ (dark squares) were able to complete all CRF trials on both days of training. **(B)** The number of rewarded trials (% Rewarded) comparably increased for EGFP$_{NAcInd}$ (light circles) and D2R-OE$_{NAcInd}$ (dark squares) mice over the 2 days of training (*** p= 0.0003). There was no difference between groups (p= 0.6451). **(C)** Lever press latency decreased from Day 1 to Day 2 for both groups (** p= 0.0014). There was no difference between groups (p= 0.6422). Data from 8 animals per genotype was used to calculate all statistics reported.

groups improved their performance from Day 1 to Day 2. The number of successful trials (% Rewarded) increased (Figure A.4B) for both the EGFP$_{NAcInd}$ and D2R-OE$_{NAcInd}$ mice (RM 2-way ANOVA: F(1,14)= 22.12, *p*= 0.0003, main effect of day) while there was no group effect of the viral manipulation (RM 2-way ANOVA: F(1, 14)= 0.2216, *p*= 0.6451). Both EGFP$_{NAcInd}$ and D2R-OE$_{NAcInd}$ mice got faster at the task as measured by a decrease in lever press latency (Figure A.4C, RM 2-way ANOVA: F(1,14)= 15.76, *p*= 0.0014, main effect of day) with no difference between groups (RM 2-way ANOVA: F(1, 14)= 0.2255, *p*= 0.6422). We conclude that D2R upregulation

has no effect on learning to press a lever for a food reward as we observed no difference in performance between the two groups.

During the Go/No-Go task, mice use distinct visual cues to learn to either press a lever ("Go" trial) or withhold pressing of the same lever ("No-Go" trial) to receive a milk reward. In both trials, the lever is available for 5 seconds in which the animal must decide to press or not. A schematic of the task is shown in Figure A.5A. The mice were first trained exclusively on Go trials and learning was established once they reached a criterion of 75% accuracy over three consecutive days. Next, No-Go trials were randomly intermixed with Go trials (60 total trials). Surprisingly, D2R-OE_{NAcInd} mice showed a deficit in the acquisition of Go trials compared to the $EGFP_{NAcInd}$ mice (Figure A.5B, RM 2-way ANOVA: $F(15, 210)= 2.029$, $p= 0.0148$, interaction effect). During the Go/No-Go task, both D2R-OE_{NAcInd} and $EGFP_{NAcInd}$ control mice improved on withholding lever pressing during No-Go trials as measured by a decrease in incorrect responses (% incorrect) over the 35 days of testing (Figure A.5C, RM 2-way ANOVA: $F(34, 476)= 20.66$, $p<0.0001$, main effect of day). However, there was no significant difference between the two groups (RM 2-way ANOVA: $F(34, 476)= 0.4606$, $p=0.9965$). Both groups also improved their performance on Go trials as measured by an increase in correct responses (% correct) the 35 days of Go/No- Go testing (Figure A.5D, RM 2-way ANOVA: $F (34, 476)= 2.076$, $p=0.0005$, main effect of day). Again,

there was no difference in correct responses between the two groups during Go\No-Go testing (RM 2-way ANOVA: F(34, 476)= 1.2, *p*=0.2073).

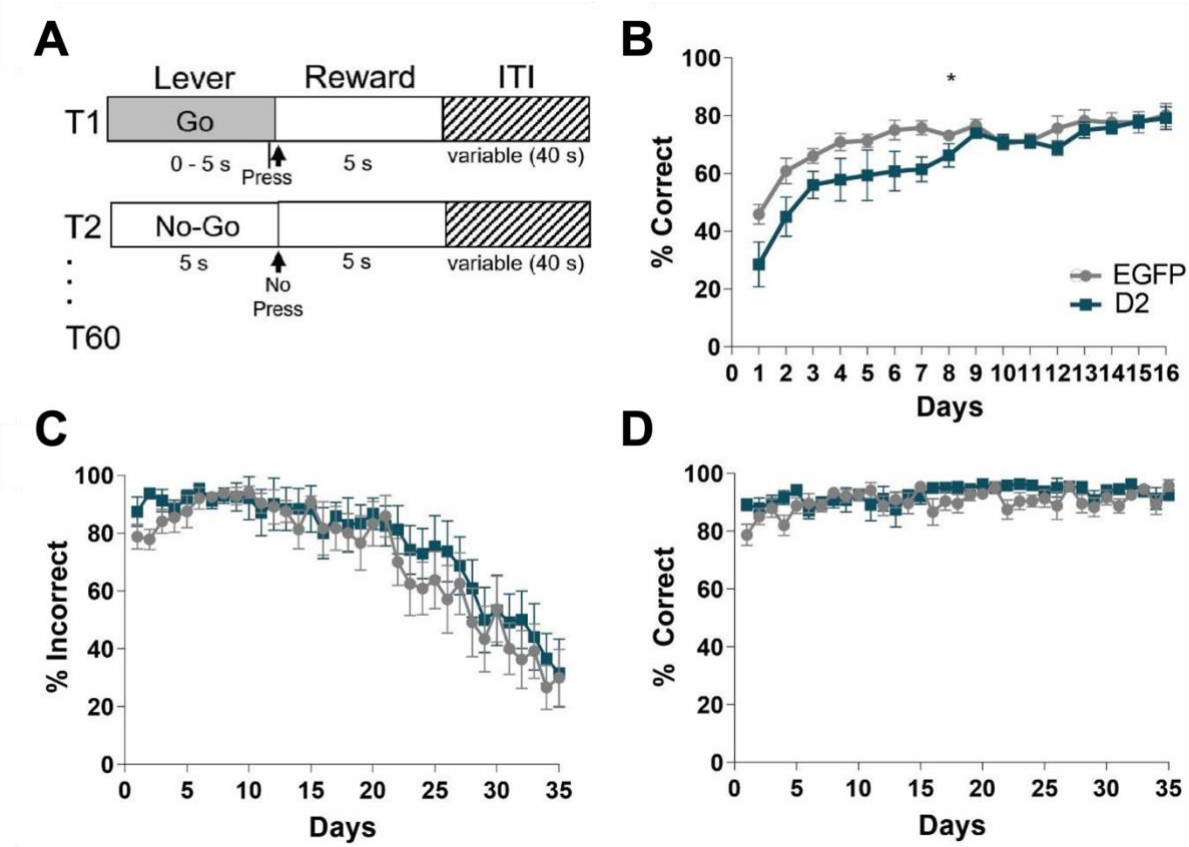

Figure A.5 D2R upregulation does not impair No-Go learning. (A) Task design. Go trials: mice were trained to press a lever within 5 seconds to obtain a milk reward. No-Go trials: mice learned to withhold pressing the lever for 5 seconds to receive a milk reward. Mice only received a reward if they made a correct choice. Each session consisted of 30 Go and 30 No-Go trials that were randomly mixed. **(B)** Acquisition of Go trials only. D2R-OE$_{NAcInd}$ (dark squares) showed a delay in the acquisition compared to EGFP$_{NAcInd}$ (light circles) (*p= 0.0148). **(C)** Performance on No-Go trials during Go/No-Go testing. Both EGFP$_{NAcInd}$ (light circles) and D2R-OE$_{NAcInd}$ (dark squares) improved on No-Go trials and learned to withhold responding (decrease in % incorrect) over the 35 days of testing (p< 0.0001). There was no difference in performance between the two groups (p= 0.9965). **(D)** Performance on Go trials during Go/No-Go testing. Both groups performed better on Go trials (increase in % correct) over the 35 days of testing (p= 0.0005). There was no difference in performance between groups (p= 0.2073). Data from 8 animals per genotype was used to calculate all statistics reported.

To determine whether D2R upregulation in these two cohorts leads to hyperlocomotion as has been described before (Donthamsetti et al., 2020; Gallo et al., 2018) we tested all mice in the open field. D2R-OE$_{NAcInd}$ mice showed an increase in locomotor activity in a standard 90-minute open field session replicating previous findings consistent with functional upregulation of D2Rs in ventral striatopallidal neurons (Figure A.6). D2R-OE$_{NAcInd}$ mice continued to traverse the

enclosure during the entire session while $EGFP_{NAcInd}$ control mice attenuated their locomotion (Figure A.6A, RM 2-way ANOVA: $F(17, 510)= 5.231$, $p<0.0001$, interaction effect). Furthermore, $D2R\text{-}OE_{NAcInd}$ mice traveled a significantly greater total distance compared to controls (Figure A.6B, 2-way ANOVA: $F(1,15)= 8.925$, $p= 0.0092$, main effect of genotype). These results confirmed that our viral manipulation of over-expressing D2Rs in ventral striatopallidal neurons was functional.

A.4 Discussion

Here, we examined enhanced D2R expression in ventral striatopallidal neurons and determined if the resulting deficit in ventral striatopallidal synaptic transmission is important for

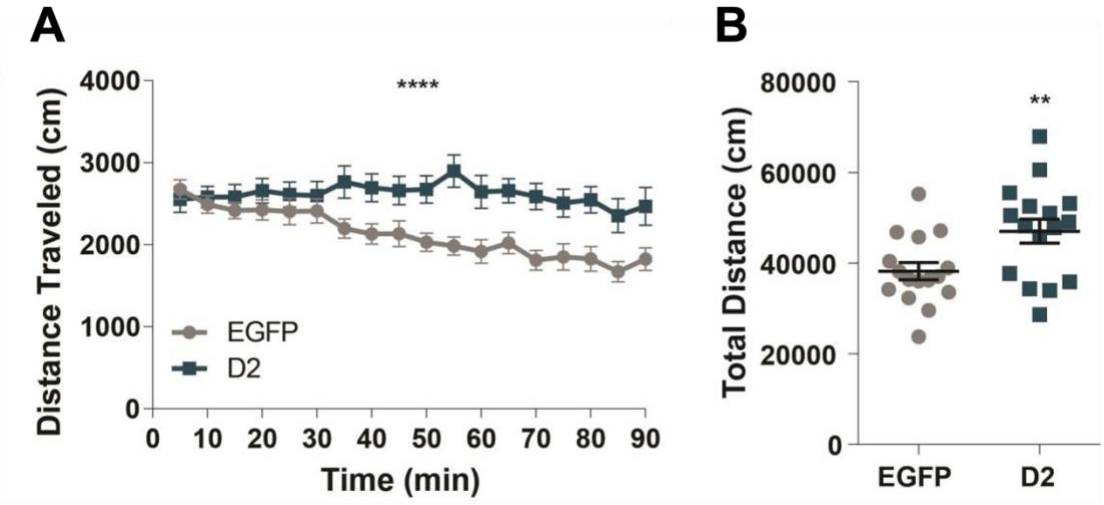

Figure A.6. D2R upregulation induces hyperlocomotion. (A) $D2R\text{-}OE_{NAcInd}$ (dark squares) mice traveled more distance in 5 min bins during a 90 min period than $EGFP_{NAcInd}$ (light circles) mice ($p< 0.0001$). **(B)** $D2R\text{-}OE_{NAcInd}$ mice traveled a greater total distance compared to EGFP mice ($p= 0.0092$). Data from 16 animals per genotype was used to calculate all statistics reported.

associative reward learning and cognitive function in mice. To determine this, we used a viral approach to selectively over-express D2Rs in ventral striatopallidal neurons and tested mice in an auditory Pavlovian conditioning task followed by a Go/No-Go paradigm. We found that upregulation of D2Rs in ventral striatopallidal neurons does not impair associative reward learning nor did it affect the extinction of response behavior when the Pavlovian cue was no longer paired

with a reward. In contrast to our initial hypothesis, decreased function of the ventral striatopallidal pathway did not cause deficits in No-Go learning, however, we did observe a slight delay in the acquisition of Go learning. Lastly, we replicated previous findings that D2R upregulation in ventral striatopallidal neurons enhances locomotion. These data suggest that while D2R upregulation in ventral striatopallidal neurons enhances locomotor activity and the motivation to work for food (Donthamsetti et al., 2020; Gallo et al., 2018) is does not affect reward learning, at least under the conditions tested.

A.4.1 D2R upregulation in ventral striatopallidal neurons does not impair Pavlovian conditioning and extinction learning

Accumulating evidence suggests that dopamine in the NAc core is important for associative reward learning. First, dopamine is released in response to reward predicting cues during a Pavlovian conditioning task (Bailey et al., 2018; Collins et al., 2016). Single-unit recordings from rat spiny projection neurons (SPNs) further showed that 75% of NAc neurons change their activity in response to reward-predicting cues. Of these neurons, half showed an increase in firing and half a decrease indicating a possible cell-type specific regulation consistent with D1R activation enhancing striatomesencephalic activity and D2R activation inhibiting striatopallidal activity (Day et al., 2006). However, ventral striatopallidal neurons are not always inhibited after reward predicting cues. Pathway specific Ca^{2+} imaging revealed that reward-predicting cues enhanced, ventral striatopallidal activity in the lateral part of the ventral striatum, whereas it decreased ventral striatopallidal activity in the ventro-medial striatum (Tsutsui-Kimura et al., 2017). This suggests that the regulation of ventral striatopallidal and striatomesencephalic activity in response to cue induced dopamine is more complicated than the dichotomous model would suggest.

Second, inhibition of NAc core projecting DA neurons disrupt Pavlovian reward learning (Heymann et al., 2020). Contrasting this systemic administration of a D2R antagonist was found to enhance approach behavior and promoted learning in rats during an auditory Pavlovian conditioning paradigm (Eyny & Horvitz, 2003). However, the latter finding is difficult to interpret due to the systemic actions of the antagonist. Local infusion of dopamine receptor antagonists into the NAc revealed that D1R antagonism impaired memory consolidation during appetitive Pavlovian learning, whereas D2R antagonism had no effect on learning (Dalley et al., 2005). Similarly, local NAc core infusion of a D1R antagonist have been shown to impair Pavlovian Instrumental Transfer, whereas the infusion of a D2R antagonist had only mild effects (Lex & Hauber, 2008).

Third, animals demonstrate a divergence in approach behavior during Pavlovian learning that may be directed either towards the CS itself (sign-tracking) or the location of the reward delivery (goal-tracking). Dopamine is differentially involved in these approach behaviors; it is necessary for the development and expression of sign-tracking behaviors but only required for the expression of goal-tracking behaviors (Flagel et al., 2011). Furthermore, dopamine D1 and D2 receptors (D1Rs and D2Rs) play a differential role in approach behaviors. Both D2Rs and D1Rs play an important role in sign-tracking behaviors while only the activity of D1Rs is necessary for the development of goal-tracking behaviors (Roughley & Killcross, 2019). Here, our behavioral paradigm limited our analysis to only measure goal-tracking behaviors, which may explain why D2R upregulation did not affect Pavlovian learning. Taken together, these results implicate a role for NAc dopamine in associative reward learning using Pavlovian cues, however, the underlying mechanism seems to involve D1Rs rather than D2Rs. In so far, our results that D2R upregulation in ventral striatopallidal neurons does not affect Pavlovian reward learning is not surprising.

Following the Pavlovian task, we tested D2R-OE_{NAcInd} mice in extinction learning. We rationalized that reward omission during extinction learning should result in a dip in dopamine release in accordance with the reward predicting error model (Bromberg-Martin et al., 2010). This dip in dopamine should decrease D2R mediated inhibition of ventral striatopallidal neurons, thereby facilitating the learning of avoiding the Pavlovian response (Hikida et al., 2010; Kravitz et al., 2012). Higher levels of D2Rs may make ventral striatopallidal neurons more responsive to this regulation if under wild-type conditions all receptors are occupied by dopamine at baseline, whereas in the condition of D2R upregulation additional receptors are occupied. However, we found that that both D2R- OE_{NAcInd} and $EGFP_{NAcInd}$ control mice attenuated their anticipatory head poking during the presentation of the previously reward predicting cue. Thus, D2R upregulation in ventral striatopallidal neurons neither enhanced nor impaired extinction learning.

A.4.2 D2R upregulation in ventral striatopallidal neurons does not affect "No-Go" learning

D2R-expressing ventral striatopallidal neurons have been proposed to suppress movement or action initiation by gating the output of the basal ganglia via connections through the globus pallidus external (GPe) (Albin et al., 1989). Furthermore, inhibition of these neurons using the G_i-coupled designer receptor $hM4D_{Gi}$ enhances response initiation in a progressive hold own task, where the mouse as to hold down a lever for an increasing amount of time to obtain a food reward (Bailey et al., 2015; Carvalho Poyraz et al., 2016). Thus, we hypothesized that upregulation of G_i-coupled D2Rs impairs learning if actions need to be suppressed. However, our Go/No-Go data did not reveal any impairment in No-Go learning and performance. During the Go/No-Go paradigm, both D2R-OE_{NAcInd} and control $EGFP_{NAcInd}$ mice improved their performance and plateaued at the same level, demonstrating their ability to withhold responding during No- Go trials. This result

shows that D2R upregulation in ventral striatopallidal neurons is not sufficient to disrupt learning when actions need to be withheld.

Prior to Go/No-Go testing, the mice were trained exclusively on Go trials. During 16 days of Go training, the D2R- OE_{NAcInd} mice showed a mild impairment during the first 9 days. We believe that this impairment was due to a newly implemented time restriction for lever availability during Go training. During the preceding two days of continuous reinforcement (CRF) training, the lever was extended until the animal pressed the lever. However, during Go training the lever was only available for 5 seconds, and if no press was made the lever would retract and a new trial would start. Previous work from our group showed that D2R-OE_{NAcInd} mice are hyperactive in the open field (Gallo et al., 2015; Welch et al., 2021), a finding we replicated here. One possibility is that increased exploration in the operant chamber distracts D2R-OE_{NAcInd} mice from lever responding within the 5 sec time limit. Alternatively, D2R-OE_{NAcInd} mice have a deficit in recognizing and adapting to the changing circumstance of the 5 sec time limit.

A.4.3 D2R upregulation in ventral striatopallidal neurons induces hyperlocomotion

D2R upregulation in ventral striatopallidal neurons increases locomotion (Gallo et al., 2015; Welch et al., 2021). To confirm that our manipulation was functional, we tested activity of D2R-OE_{NAcInd} and $EGFP_{NAcInd}$ control mice in a standard open field box. We replicated hyperactivity in D2R-OE_{NAcInd} mice. Furthermore, post-hoc immunohistochemistry staining confirmed that our viral expression was targeted to the ventral striatum. Since D2R-OE_{NAcInd} mice showed a bimodal distribution in the open field with 5 mice performing similar to $EGFP_{NAcInd}$ control mice, we re-analyzed the data only using the hyperactive D2R-OE_{NAcInd} mice. Still no alteration in behavior was observed in the Pavlovian or No-Go learning conditions (data not shown) but the deficit in the acquisition of the Go component became more significant (RM 2-

way ANOVA: F(15,180)=2.774, p=0.0008, interaction between genotype (EGFP and D2ROE) and day). Taken together, these results give us confidence that D2R upregulation was functional as in our previous publication where we established with slice and *in vivo* electrophysiological measures that D2R upregulation impairs synaptic transmission of ventral striatopallidal neurons collaterals within the NAc and the canonical projections to the ventral pallidum.

We set out to determine how altered D2R levels affects associative and Go/No-Go learning in mice. We found that selective D2R upregulation on ventral striatopallidal neurons does not impair associative reward learning or No-Go learning. In our previous published studies, we found that D2R upregulation enhances the willingness to work for food if response efforts are high (Donthamsetti et al., 2020; Gallo et al., 2018). In contrast, both the Pavlovian conditioning and the Go/No- Go task do not require much effort to perform the task. One explanation is that the behavioral assays we tested are not sensitive enough to detect any changes in learning and that our D2R-OE_{NAcInd} mice could have a deficit in Pavlovian learning with probabilistic outcomes, where task difficulty varies with reward predictability. Together with our previous published findings our data suggest that ventral striatopallidal D2Rs play a specific role in regulating motivation by balancing cost/benefit computations but does not affect associative learning (Bailey et al., 2020; Gallo et al., 2018; Simpson & Kellendonk, 2017; Ward et al., 2012).

A.5 Methods

Animals: Adult male and female *Drd2*-Cre BAC transgenic mice (ER44; GENSAT) backcrossed onto the C57BL/6J background were group housed under 12-h light/dark cycle. All experimental procedures were conducted following NIH guidelines and were approved by Institutional Animal Care and Use Committees by Columbia University and New York State Psychiatric Institute. We chose *Drd2*-Cre over A2A-Cre mice to recapitulate the conditions in which we saw enhanced

progressive ratio performance. Also, Cre levels are higher in *Drd2*-Cre mice as they can be visualized with anti-Cre immunohistochemistry (Cazorla et al., 2014; Gallo et al., 2018), whereas we cannot detect Cre expression in A2A-Cre mice using the same anti-serum. In our hands *Drd2-Cre* mice leads to recombination of AAV expression constructs in about 5% to 10% of ChAT neurons (Carvalho Poyraz et al., 2016; Gallo et al., 2018).

Stereotaxic Surgery: Mice (≥8 weeks old) were bilaterally injected with 450 nL/hemisphere of a previously characterized Cre-dependent double-inverted open reading frame (DIO) adeno-associated viruses (AAVs) encoding D2R-ires-Venus (5.1 X 10^{13} GC/mL) or EGFP (6.69 X 10^{13} GC/mL) (UNC Vector Core, Chapel Hill, NC) into the nucleus accumbens (NAc) using stereotactic Bregma-based coordinates: AP, + 1.70 mm; ML, ± 1.20 mm; DV, –4.1 mm (from dura). Mice were induced with 4% isoflurane anesthetic and maintained at 1-2% throughout the stereotaxic surgery. Following induction, mice were placed on the stereotaxic setup and a midline incision was made using a sterile scalpel. A high-speed rotary micromotor kit (Foredom, Bethel, CT) was used to make holes in the skull and a glass pipette was lowered into the brain to deliver the AAVs. Mice were given 4 weeks to recover from the surgery. Groups of mice used for experiments were first assigned their AAV-genotype in a counterbalanced fashion that accounted for sex, age, home cage origin.

Operant Apparatus: Eight operant chambers (model Env-307w; Med-Associates, St. Albans, VT) equipped with liquid dippers were used. Each chamber was in a light- and sound-attenuating cabinet equipped with an exhaust fan, which provided 72-dB background white noise in the chamber. The dimensions of the experimental chamber interior were 22 × 18 × 13 cm, with flooring consisting of metal rods placed 0.87 cm apart. A feeder trough was centered on one wall of the chamber. An infrared photocell detector was used to record head entries into the trough.

Raising of the dipper inside the trough delivered a drop of evaporated milk reward. A retractable lever was mounted on the same wall as the feeder trough, 5 cm away. A house light located on wall opposite to trough illuminated the chamber throughout all sessions.

Dipper Training: Four weeks after AAV surgery, mice underwent operant training. Mice were weighed daily and food- restricted to 85–90% of baseline weight; water was available ad libitum. In the first training session, 20 dipper presentations were separated by a variable inter-trial interval (ITI) and ended after 20 rewards were earned or after 30 min had elapsed, whichever occurred first. Criterion consisted of the mouse making head entries during 20 dipper presentations in one session. In the second training session, criterion was achieved when mice had made head entries during 30 of 30 dipper presentations.

Pavlovian Conditioning: Mice were trained for 16 consecutive days in a Pavlovian conditioning paradigm, which consisted of 12 conditioned stimulus-positive (CS+) trials and 12 CS- trials occurring in a pseudorandom order. Each trial consisted of an 80-dB auditory cue presentation for 10 sec of either a 3 kHz (Cohort 1) or 8 kHz (Cohort 2) tone or white noise (counterbalanced between mice) and after cue offset a milk reward was delivered only in CS+ trials, whereas no reward was delivered in CS- trials. There was a 100 sec variable intertrial interval, drawn from an exponential distribution of times. Head entries in the food port were recorded throughout the session, and anticipatory head entries during the presentation of the cue were considered the conditioned response. The differential score (Head entries/sec (CS – (ITI))) was calculated using either the 2 sec or 10 sec of ITI proceeding the cue. We used 10 sec of ITI to calculate the differential score during the entire 10 sec cue (Figure A.3B) and 2 sec ITI for the 2 sec binning (Figure A.3C & D).

Continuous Reinforcement schedule (CRF): For lever press training, lever presses were reinforced on a continuous reinforcement (CRF) schedule. Levers were retracted after each reinforcer and were presented again after a variable ITI (average 40 sec). The reward consisted of raising the dipper for 5 sec. The session ended when the mouse earned 60 reinforcements, or one hour elapsed, whichever occurred first. Sessions were repeated daily until mice achieved 60 reinforcements.

Go/No- Go schedule: Mice were first trained on Go trials in which they were required to press a lever within 5 sec of its presentation to receive a reward. If the 5 sec elapsed with no response, the lever would retract, no reward would be presented, and a new ITI (average 40 sec) would begin. Mice were trained on these 5 sec Go-only trials until they reached 75% accuracy over three consecutive days. Once this criterion was achieved, No- Go trials were added in which the lever was presented simultaneously with two cues (the house lights turning off, and a small LED light above the lever turning on). A lack of any lever press within 5 sec, resulted in a reward. A lever press during this period caused the lever to retract, the house lights to turn on, the LED light to turn off, and a new ITI to begin without any reward for that trial. In each session, 30 Go trials were interspersed with 30 No-Go trials presented pseudo-randomly such that there were an equal number of both kinds of trials in every block of 10 trials. Mice were tested for 35 days, and false alarm rate and hit rate were analyzed.

Locomotor Activity: *D2*-Cre mice injected with D2R- or EGFP-expressing AAVs were tested in open field boxes equipped with infrared photobeams to measure locomotor activity (Med Associates, St. Albans, VT). Data were acquired using Kinder Scientific Motor Monitor software (Poway, CA) and expressed as total distance traveled (cm) over 90 min.

Histology: Mice were transcardially perfused with ice-cold 4% paraformaldehyde (Sigma, St. Louis, MO) in PBS under deep anesthesia. Brains were harvested, post-fixed overnight and washed in PBS. Free-floating 50-µm coronal sections were obtained using a Leica VT2000 vibratome (Richmond, VA). After incubation in blocking solution (5% horse serum, 0.5% bovine serum albumin in 0.5% PBS-Triton X-100) for 2 hr at room temperature, sections were labeled overnight at 4 °C with primary antibodies against GFP (chicken; 1:1000; AB13970 Abcam, Cambridge, MA). Sections were incubated with fluorescent secondary antibodies (Goat anti-chicken, A488, A11039, ThermoFisher) for 2 hr at RT. Sections were then mounted on slides and cover slipped with Vectashield containing DAPI (Vector, Burlingame, CA). Digital images were acquired using a Zeiss epifluorescence microscope.

Statistical analysis: Data are expressed as mean ± SEM. Students' *t*-tests were used to compare between two groups. Analyses involving multiple conditions were evaluated by one-way, two-way, or three-way repeated measures ANOVA, using GraphPad Prism software. Statistical significance was considered for $p < 0.05$.