ABSTRACT

Title of Dissertation:	NEURAL MECHANISMS OF APPROACH AND AVOIDANCE
	Ronny Nicole Gentry, Doctor of Philosophy, 2018
Dissertation directed by:	Dr. Matthew R. Roesch, Associate Professor, Department of Psychology, Neuroscience and Cognitive Sciences Program

Using environmental cues to acquire good things and avoid harmful things is critical for survival. Rewards and punishments both drive behavior through reinforcement learning mechanisms and sometimes occur together in the environment, but it remains unclear how these signals are encoded within the brain and if signals for positive and negative reinforcement are encoded similarly. The dopaminergic system and, more broadly, the corticomesolimbic circuit are known to be involved in the processing of positive and negative reinforcement. Here, I investigated neural correlates of decision-making and associated behavioral patterns within two key corticomesolimbic regions: the ventromedial prefrontal cortex (vmPFC), which is thought to generate contextually appropriate responses, and the nucleus accumbens (NAc), which is thought to use dopamine (DA) prediction error signals to motivate behavior.

The goal of this work was to uncover the underlying brain mechanisms encoding positive and negative reinforcement signals and to explore individual differences in neural and behavioral patterns that arise during learning and performance. To achieve this, I recorded from single neurons within vmPFC and measured DA release within NAc core during two behavioral tasks examining distinct aspects of learning: initial Pavlovian responses, as well as more complex combined positive and negative reinforcement. I found that, within the vmPFC, cell firing was modulated more often and more robustly by cues predicting reward than by cues preceding avoidable shock; overall, we found very few cells that responded to shock cues, and responses to shock avoidance and reward cues were not colocalized within the same cells. Alternatively, I found that DA release within the NAc increased to both reward and shock avoidance cues compared to neutral cues, and these changes occurred within the same microdomain of the NAc. Additionally, we uncovered intriguing individual differences in NAc DA release and behavioral responses during both our combined approach avoidance and autoshaping tasks and, in the final chapter, shifted these responses by manipulating task parameters and inhibiting VTA-NAc DA neurons. Together, these results help further our understanding of how differences in vmPFC activity and accumbal DA release influence cue-driven learning and behavioral performance across various contexts.

NEURAL MECHANISMS OF APPROACH AND AVOIDANCE

by

Ronny Nicole Gentry

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2018

Advisory Committee:
Professor Matthew R. Roesch, Chair
Professor Jens Herberholz
Professor Luiz Pessoa
Professor Elizabeth Quinlan, Dean's Representative
Professor Alexander Shackman

Dedication and Acknowledgements

I would like to dedicate this dissertation to my mom, Mary Gentry, and to the memory of my dad, Chris Gentry. Growing up, my dad's passion for nature and science constantly inspired me, and I thank him daily for passing on those traits to me. To my mom, you have shown me what it is to be a strong woman, and you never cease to encourage me in everything I do. I can't thank you enough for your love and support along the way.

The work presented in this dissertation would not have been possible without the support of many people. To my fiancé, Jon Jaffe, thank you for everything that you do. You always have a hug, a thoughtful gesture, or a kind word for me when I need it most. You're the best. I would also like to thank the rest of my friends and family for always lending an ear or a shoulder when I needed it. I love all of y'all.

An enormous thank you to Dr. Matthew R. Roesch, my PhD advisor. During my time in your lab, you have helped me grow so much as a scientist and a person. Thank you for your ceaseless patience, understanding, and guidance every step of the way. It was a pleasure learning from you, and I am honored to have been a part of your group.

An additional thank you to the members of the Roesch lab, past and present, who were always there for me, inside and outside of the lab. Finally, I would like to thank the Coutureau lab for welcoming me into your group and making Bordeaux feel like home.

Table of Contents

Chapter 1: Introduction

General Introduction to reinforcement learning

The ability to make beneficial choices is critical for normal, everyday behavior and, as a consequence, decision-making has remained a fundamental executive function across species (Calvert, Green, & Myerson, 2011; Kalenscher & van Wingerden, 2011). Making good decisions is not always easy: it is a multifaceted process. To make a decision, an animal must weigh many possible outcomes against short- and long- term goals before deciding a course of action. By successfully monitoring environmental cues and predicting consequences, animals can select behaviors that facilitate the attainment of a chosen goal. Behaviors that lead to the acquisition of a positive outcome or the evasion of a negative outcome are reinforcing and will increase the probability the same behavior will be selected again in the future. In this way, behavior can be shaped toward the good and away from the bad in one's environment.

Possible beneficial outcomes can be grouped into the probability of obtaining something rewarding or avoiding an outcome that is negative or punishing. However, in many cases, the expectation of emotionally charged outcomes also alters other functions related to motivation, salience, arousal and attention that serve to facilitate response mechanisms to approach or avoid; thus, to better understand how these associations are formed within the brain and how appropriate behaviors are selected and executed, we must try to dissociate these factors. For example, two odors could be equally salient but indicate outcomes with different valences: an odor that predicts

the potential presence of a predator and an odor that predicts the presence of a ripe fruit would both be highly salient cues in one's natural environment, but the former would likely lead to a negative association, while the latter would likely lead to a positive association. While many studies have examined appetitive and aversive stimuli separately, some tasks are beginning to vary appetitive and aversive stimuli within the same task. In studies that combine appetitive and aversive stimuli within a task, there are typically three basic trial types, such that: (1) a conditioned stimulus (CS) predicts a rewarding outcome; (2) another CS predicts a neutral condition or a smaller reward; and (3) a third CS predicts a smaller reward (or no reward) with the threat of an aversive outcome. In animal studies, the aversive outcome is usually a concrete punisher, which may range from a time-out, a bolus of a bitter quinine solution, an electric shock, or an air-puff to the eye, while human studies may also employ abstract punishers, such as loss of money, in addition to concrete punishers (Anderson et al., 2003; Anstrom, Miczek, & Budygin, 2009; Bissonette, Burton, et al., 2013; Brischoux, Chakraborty, Brierley, & Ungless, 2009; Calu, Roesch, Haney, Holland, & Schoenbaum, 2010; Carter, MacInnes, Huettel, & Adcock, 2009; J. M. Choi, Padmala, & Pessoa, 2012; J. M. Choi, Padmala, Spechler, & Pessoa, 2014; J. C. Cooper & Knutson, 2008; M R Delgado, Nystrom, Fissell, Noll, & Fiez, 2000; Lammel, Ion, Roeper, & Malenka, 2011; Litt, Plassmann, Shiv, & Rangel, 2011; Matsumoto & Hikosaka, 2009; Matthew R Roesch & Olson, 2004; Rolls, Sienkiewicz, & Yaxley, 1989; Small et al., 2003). These aversive outcomes and their predictive cues may produce a variety of response behaviors: freezing/helplessness, escape, or avoidance, depending on the state of the animal. Rewarding outcomes, in

contrast, must be valued, wanted and must drive approach behavior; rewards used in animal and human studies can also be concrete or abstract, such as delivery of food or water, access to a mate/conspecific, access to addictive drugs, or gain of money. Importantly, all of these stimuli and their predictors can drive behavior either toward or away from specific outcomes with continued experience.

Approach and avoidance behaviors are driven by positive and negative reinforcement learning strategies (Bromberg-Martin, Matsumoto, & Hikosaka, 2010). In short, a reinforcer can be defined as any behavioral consequence that will increase the probability that a certain behavior (i.e., the one performed to receive the reinforcer) will be repeated in the future whenever a specific environmental stimulus is presented. In this way, both rewarding and aversive consequences and the cues that predict them can drive behavioral shifts with learning, toward the acquisition of more positive outcomes and the avoidance of more negative outcomes in the future.

During positive reinforcement, a desirable unconditioned stimulus (US; e.g., sucrose pellet reward) is added to the environment contingent upon the subject's behavior; the acquisition of the reward then acts to increase the subject's performance of this behavior in the future. For example, if the delivery of a sucrose pellet is signaled by a cue (CS; e.g., cue light, auditory tone), an animal will learn to associate the cue preceding sucrose delivery with reward, and the cue itself will start to drive the appropriate behavioral response. This behavioral response would include an action or sequence of actions that must be performed in order to obtain the reward, such as pressing a lever and/or visiting a food receptacle. Generally, positive reinforcement is more often and more easily studied, as this is a fairly straightforward

behavior for the animal to learn and perform. With that said, positive reinforcement paradigms evoke several different associative mechanisms (e.g., stimulus-outcome, stimulus-response, response-outcome, attention, arousal, motivation, etc.)

During negative reinforcement, a noxious US (e.g., foot shock) is removed from the environment contingent upon the subject's behavior; the removal of this threat also acts to increase performance of the subject's performance of this behavior in the future. For example, if the delivery of the foot shock is first signaled by a cue CS (e.g., light or auditory tone), the animal will learn to associate this cue with the shock and the cue will start to drive the appropriate behavioral response to avoid or escape the shock threat, if possible. In negative reinforcement scenarios, it is thought that, with learning, the cue CS becomes aversive in itself and the animal will work to turn it off (Budygin et al., 2012; Mowrer & Lamoreaux, 1946; Wenzel, Rauscher, Cheer, & Oleson, 2015b); this behavior becomes the reinforcing step in avoidance paradigms and is likely driven by habit-like stimulus-response (S-R) encoding systems within the brain. In contrast with positive reinforcement, negative reinforcement can be very difficult for the animal to learn and perform, and, thus, its underlying mechanisms have not been thoroughly studied.

Negative reinforcement is thought to be more difficult to learn due to the formation of an initial Pavlovian freezing response to the threat, which must be overcome in order to perform an avoidance action. Classically, Mowrer's two-factor theory proposed that avoidance learning comprises two stages that are in direct competition with one another (Mowrer & Lamoreaux, 1946; Rescorla & Solomon, 1967). In the first stage, a previously neutral stimulus comes to predict an aversive

unconditioned stimulus (US) via Pavlovian conditioning and induces freezing to the CS+; in the second stage, this Pavlovian association then motivates the acquisition of an instrumental escape action to the CS+ in order to consequently avoid delivery of the unpleasant US, which ultimately leads to reduction of threat and hence induces "relief" (Mowrer & Lamoreaux, 1946). Importantly, these stages create opposing behavioral endpoints that must be reconciled to avoid the noxious US in the future, namely by somehow suppressing freezing induced by Pavlovian associations established during the first stage (Moscarello & LeDoux, 2013). In brief, the two-factor theory suggests that acquisition of the avoidance response could develop as a result of the drive to terminate the learned CS, which had been paired with threat, rather than from thoughts of the outcome of the potential threat itself (the US) (Lovibond, Saunders, Weidemann, & Mitchell, 2008; Maia, 2010; Maia & Frank, 2011). This theory takes reference from Hull's drive reduction theory, suggesting that avoiding a threat would reduce a drive (e.g., "fear") in a rewarding way, just as acquiring a food treat would reduce a drive such as hunger in a rewarding way (Hull, 1943).

Recently, interesting parallels have been drawn between Mowrer's two-factor theory and a prominent theory within the dopamine literature, the actor-critic model (Maia, 2010). In the actor-critic model, "the critic" learns about the values of states (by predicting the reinforcement that is signaled in these states) and calculates a prediction error, which is used by "the actor" to assign values to actions, learn preferred actions, and select appropriate responses accordingly (Barto, 1995; Maia, 2010; Niv & Schoenbaum, 2008; Redish, 2004; Sutton & Barto, 1998). Thinking

about these theories in parallel, the critic oversees learning the value of states, which would correspond to the drive acquired by the animal in a certain environment (i.e., hunger or fear), and the actor learns which behaviors to promote based on these states (i.e., approach or avoidance) (Maia, 2010). In the context of negative reinforcement, a reduction in the threat-related fear drive produced by the predictive cue would lead to a positive prediction error within the critic, which would go on to inform future behavior toward the production of more avoidance responses by the actor. In short, the critic essentially would implement the classical conditioning component of two-factor theory, while the actor would decide on and activate the instrumental avoidance component based on the critic's input. This system is thought to be embodied by activity of the striatum, which will be discussed in depth in the following section.

While both positive and negative reinforcement strategies increase behavior, it is possible they are driven by opposing neural mechanisms. By investigating neural correlates during cued positive and negative reinforcement, we can better understand how the brain connects environmental cues to actions and how we are able to make beneficial decisions based on these associations during healthy decision-making. Although an established body of literature has extensively studied neural systems involved in both functions, very few have set out to explicitly study how these neural systems directly reconcile both appetitive and aversive neural signals in a single task. Even fewer have addressed questions related to how anticipated appetitive and aversive outcomes interact to alter neural signals related to expected value, motivation, and salience.

If appetitive and aversive stimuli are encoded by independent neural populations, then activity should be modulated by either appetitive or aversive stimuli but not both. If appetitive and aversive stimuli are indeed encoded by the same populations, they may exhibit several patterns. If activity is modulated by factors that encode the value of appetitive and aversive stimuli, then activity should respond differently for appetitive and aversive trials compared to neutral trials (e.g., increase to reward and decrease to punishment). However, if activity is modulated by factors that vary with the strength of both appetitive and aversive stimuli, signaling salience, then activity should respond similarly for appetitive and aversive trials compared to neutral trials. A final possibility is that cues predicting reward or avoidable punishers (e.g., foot shock) could both be construed as high value in brain areas that combine information from both positive and negative reinforcers within single units or a particular microdomain. Here, the avoidability of the punisher is highly relevant: cues that predict unavoidable punishment would be salient, attention-grabbing, and arousing, but they would not have a high value since no action can be taken; conversely, cues signaling reward or avoidable punishment would also have value, since action could be taken to acquire the reward or avoid the punishment, respectively.

Overview of the corticomesolimbic circuit

Many key brain areas have been independently implicated in various aspects of learning and decision-making, such as learning about environmental cues and encoding expectations surrounding cues and their predicted outcomes, imbuing cues with motivational properties based on predicted outcomes, and integrating these

signals to invigorate behavioral responses toward or away from cues that help obtain good and avoid bad outcomes (Balleine & O'Doherty, 2010; Gentry, Lee, & Roesch, 2016a; Wolfram Schultz, 2006). While numerous brain regions are involved, activity within specific areas of the mesolimbic pathway, basal ganglia and prefrontal cortex have been shown to be modulated by appetitive and aversive stimuli. Some of the many brain regions that make up this reinforcement circuit include the orbitofrontal cortex (OFC), the anterior cingulate cortex (ACC), the basolateral (BLA) and central amygdala (CeA), the medial prefrontal cortex (mPFC), the ventral tegmental area (VTA) and substantia nigra pars compacta (SNc) dopamine (DA) neurons, and the dorsal (DMS, DLS) and ventral (NAc core and shell) striatum. A subset of this larger circuit (**Fig. 1**) shows how the specific brain regions we target within this dissertation may interact with one another. Individual brain regions within this circuit are often described in terms of their isolated function; however, it is important to note that these areas are not independent of each other but act in concert to process and encode rewarding and punishing events and their predictive environmental cues to produce goal-directed behaviors. Since encoding of outcomes and related cues relies on the combined feedback between midbrain and cortical regions and directionality is often difficult to determine, it is important to consider the circuit as a whole.

The orbitofrontal cortex

The orbitofrontal cortex (OFC) has been shown to encode expectations about future appetitive and aversive outcomes that are critical for guiding learning and decision-making (Morrison, Saez, Lau, & Salzman, 2011; Morrison & Salzman, 2011; Plassmann, O'Doherty, & Rangel, 2010; Roesch & Olson, 2004; Schoenbaum,

Chiba, & Gallagher, 1998; Schoenbaum & Roesch, 2005). For example, neurons in the OFC are modulated by cues that predict different appetitive outcomes, such as different types of food and magnitudes of reward; other OFC neurons signal when an aversive stimulus is anticipated, such as quinine or air- puff. However, since motivation and value are difficult to disentangle in many experiments, it long remained unknown whether neural signals genuinely represented the value of the predicted outcome, or the motivational level associated with obtaining reward or avoiding aversive outcomes. For example, neurons in OFC fire strongly when an animal anticipates a desirable outcome (Schoenbaum et al., 1998; Schoenbaum, Chiba, & Gallagher, 1999), but if that outcome is paired with a chance for another, preferable outcome (Wallis & Miller, 2003), or is devalued through satiation (Rolls et al., 1989), then the rate of firing decreases. This activity modulation might reflect the decrease in value, but it might also reflect changes in motivation. A similar situation holds true for OFC neurons that predict aversive outcomes; activity might reflect how aversive the stimulus is or how motivated the animal is to avoid it.

A study by Roesch and Olson was able to dissociate value from motivation by varying both reward and punishment within the same task: monkeys would receive a big or small reward for correct responses and would receive a big or small time-out penalty for failed responses (Roesch & Olson, 2004). They found that OFC neurons fired most for cues that predicted large reward/small penalty and least for cues that predicted small reward/large penalty, relative to neutral conditions (small reward/small penalty); hence, the strength of responding in OFC reflected the value conveyed by the combination of reward and penalty cues. These cells have been

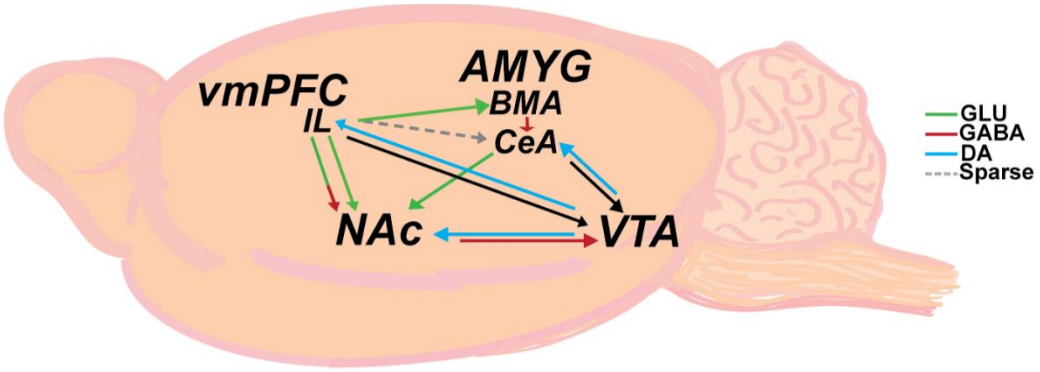


Figure 1. Simplified corticostriatal circuit within the rat brain. Recent work suggests that the medial prefrontal cortex (mPFC), specifically the ventromedial (infralimbic; IL) prefrontal cortex (vmPFC) exhibits top-down control over Pavlovian fear responses produced in the central amygdala (CeA) via inhibitory projections from basomedial amygdala (BMA) (Adhikari et al., 2015). The vmPFC also sends excitatory projections to the nucleus accumbens (NAc), which is mainly composed of inhibitory medium-spiny neurons (MSNs); however, some of these IL-NAc projections synapse onto inhibitory interneurons within the NAc causing overall disinhibition. The NAc receives dopaminergic (DA) inputs from the ventral tegmental area (VTA), which also sends DA projections to vmPFC and amygdala.

found to respond to reward or punishment, anticipation of reward or punishment, and cues predicting reward or punishment; however, they are different from VTA dopamine neurons in that they do not respond differently based on expectations (Schoenbaum, Roesch, Stalnaker, & Takahashi, 2009). It has recently been shown, however, that VTA DA neuronal PE responses depend on signals from the OFC (Takahashi, Stalnaker, Roesch, & Schoenbaum, 2017). Therefore, it is thought that the OFC signals how good or bad an expected outcome should be by integrating associative information with reference to the animal's internal state and goal representations (Schoenbaum & Roesch, 2005; Schoenbaum et al., 2009). Other studies have further shown that other populations of OFC neurons do not represent the overall value associated with a given situation but instead reflect the offers combined with the option that will be chosen (Hosokawa, Kato, Inoue, & Mikami, 2007; Morrison et al., 2011; Morrison & Salzman, 2011; Padoa-Schioppa & Assad, 2006, 2008).

The anterior cingulate cortex

The anterior cingulate cortex (ACC) has also been discussed in terms of its involvement in reinforcement learning. Work in macaque monkeys has uncovered neural correlates related to unsigned PEs (Hayden, Heilbronner, Pearson, & Platt, 2011), potentially signaling the necessity for additional resources in the face of signaling a need for behavioral modification. Using a variable size/delay task in rats, Bryden et al. demonstrated that ACC can signal errors and recruit additional attentional resources during unexpected shifts in value (Bryden, Johnson, Tobia, Kashtelyan, & Roesch, 2011). Unlike activity in BLA, firing in the ACC was

significantly stronger after both unexpected appetitive and aversive events during and before sampling of cues on subsequent trials, likely reflecting salience or attention being drawn to conditioned stimuli to update contingencies.

Alternatively, work in rhesus monkeys has demonstrated that the ACC encodes value as it relates to the integration of previous outcomes with current choices (Kennerley, Behrens, & Wallis, 2011). Additional research suggests that ACC may signal both positive and negative PEs of action values (Matsumoto et al., 2007). A recent study used a Pavlovian task in monkeys to measure ACC response to cues predicting certain or uncertain reward and punishment across blocks, finding that some ACC neurons represented expected value and uncertainty in a valence-specific manner, while other neurons were excited by both cued reward and punishment in a salience-like pattern (Ilya E Monosov, 2017).

The amygdala

Although for many years it has been hypothesized that the amygdala is important for acquiring and storing associative information related to both appetitive and aversive outcomes, work has also emerged suggesting that amygdala also supports other functions related to associative learning, such as the signaling of attention, uncertainty, and intensity (Belova, Paton, Morrison, & Salzman, 2007; LeDoux, 2000; Morrison et al., 2011; Saddoris, Gallagher, & Schoenbaum, 2005; Tye, Cone, Schairer, & Janak, 2010; Tye & Janak, 2007). In addition to this, sub-regions of the amygdala have been heavily implicated in various aspects of reinforcement processing.

For instance, it is thought that the basolateral amygdala (BLA) integrates information about appetitive and aversive events and their intensity or salience to modify behavior via signaling errors in predictions or recruitment of attentional/ executive functions. In this sense, BLA would be critical for reporting attentional need, arousal or intensity during sampling of unconditioned stimuli in the service of learning to predict the appetitive and aversive nature of the outcomes during sampling of conditioned stimuli. This hypothesis, however, is drawn from a vast literature of recordings during outcomes and does not address the role of BLA during predictive cues. At the single neuron level, BLA is modulated by the predictability of both appetitive and aversive events, specifically when expectancies are violated (Belova et al., 2007; Calu et al., 2010; M. R. Roesch, Calu, Esber, & Schoenbaum, 2010; Tye et al., 2010). In other words, BLA neurons increase firing when outcomes are unexpectedly delivered or omitted, events that are highly salient. Other studies have shown that lesions to basal and lateral amygdala using NMDA infusions impairs avoidance performance even after extensive training, confirming that BLA is involved in mediating avoidance conditioned responses (Choi, Cain, & LeDoux, 2010).

It has recently been shown that BLA interference disrupts development of cue selectivity in other areas, such as OFC and ventral striatum (VS) (Hatfield, Han, Conley, Gallagher, & Holland, 1996; Lucantonio, Stalnaker, Shaham, Niv, & Schoenbaum, 2012; Stalnaker et al., 2007; Stalnaker, España, & Berridge, 2009). This is demonstrated by experiments showing that rats with OFC lesions exhibit perseverative conditioned behavior after successful food devaluation, continuing to

respond to devalued reward-predictive cues though they do not consume their earned reward; this behavior is mirrored by decreased cue-associative firing in BLA (Pickens et al., 2003; Pickens, Saddoris, Gallagher, & Holland, 2005; Zeeb & Winstanley, 2013). Alternatively, lesions to BLA cause a complete lack of cue association to outcomes. In fact, a recent study showed that single units in BLA reflect signals that use information from prior learning to inform novel outcome estimates, and these signals were abolished with ipsilateral OFC lesions (Lucantonio et al., 2015). Additionally, inactivation of BLA to OFC, but not OFC to BLA, projections using inhibitory DREADDs has been shown to disrupt cue-triggered reward representations during conditioned approach and PIT (Lichtenberg et al., 2017).

This suggests that OFC activity facilitates rapid associative learning via BLA, while this associative learning in BLA in turn helps OFC store these associations so they can be used to guide future behaviors via the VTA/Striatum. Together, these results may indicate that, within these reciprocal connections between BLA and OFC, the BLA maybe critical for acquiring information about outcomes but not for using it to make predictions, and vice versa for OFC.

The central amygdala (CeA) is well known for its role in threat encoding. Where it was once thought that CeA activation only occurred in response to immediate threat, new data has caused a revision in this hypothesis to include CeA contributions to long-lasting, sustained threat (Davis, Walker, Miles, & Grillon, 2010; Shackman & Fox, 2016). It has been shown that Pavlovian freezing behaviors induced by punishment or cues predicting punishment originate from activation of the central amygdala (CeA) (Bolles, 1970; Justin M Moscarello & LeDoux, 2013;

Purgert, Wheeler, McDannald, & Holland, 2012); indeed, after rats were exposed to food-omission and shock-paired cues, neuronal activity in the CeA, as measured by increased mRNA presence of *Arc* and *Homer1a* (two immediate-early genes used as synaptic activity markers), was elevated compared to rats exposed to control cues (Purgert et al., 2012). Additionally, a recent study demonstrated that, in animals that failed to acquire successful avoidance behavior, lesioning CeA led to the immediate rescue of avoidance learning, suggesting that CeA activity was inhibiting the acquisition of the instrumental response in these animals (Choi et al., 2010). Rats with lesions to the CeA also show impaired freezing responses to cues paired with shock or omission of expected food, compared with control cues (Purgert et al., 2012).

While the CeA is most often thought of in terms of its role in Pavlovian freezing responses related to aversive cues and outcomes, it has also been implicated in reward and appetitive behaviors. Early work has shown that the CeA is involved in appetitive behaviors (Gallagher, Graham, & Holland, 1990; Kim, Zhang, Muralidhar, LeBlanc, & Tonegawa, 2017; Parkinson, Robbins, & Everitt, 2000; Robinson, Warlow, & Berridge, 2014; Seo et al., 2016), increasing reward saliency (Mahler & Berridge, 2009; Robinson et al., 2014; Seo et al., 2016), and modulating food consumption (Cai, Haubensak, Anthony, & Anderson, 2014; Mahler & Berridge, 2009). Recently, optogenetic and chemogenetic techniques have revealed that specific GABAergic cell types expressing serotonin 2a receptors within the CeA circuit play a significant role in modulating positive reinforcement and food consumption using a positive valence mechanism (Douglass et al., 2017). Genetically distinct BLA to CeA connections have also been discovered that either promote or suppress appetitive

behavior, which has been compared to the direct/indirect pathways within the basal ganglia pathways (Kim et al., 2017).

The ventromedial prefrontal cortex

Another critical cortical component within this circuit is the ventromedial prefrontal cortex (vmPFC). In humans, the vmPFC is presumed to be the functional homolog of the infralimbic (IL) and dorsopeduncular (DP) regions of the mPFC in non-human primates and rodents (Milad et al., 2007; Öngür, Ferry, & Price, 2003; Peters, Pattij, & De Vries, 2013). The vmPFC is well-connected to many other areas involved in appetitive and aversive processing, with IL receiving inputs from BLA, hippocampus, thalamus and sending projections to basomedial amygdala (BMA), the lateral CeA, and the striatum (McDonald, 1998; Mcdonald, Mascagni, & Guo, 1996; Vertes, 2004). Optogenetic studies have also revealed intra-mPFC connections between IL and its dorsal neighbor prelimbic mPFC (PL), with IL activation shown to have inhibitory effects on PL activity (Ji & Neugebauer, 2012). Taking these connections into account, the vmFPC has been implicated in a variety of seemingly contradictory functions, such as fear suppression, habit, and extinction; these varied results have recently been reconciled into an overarching IL theory of behavioral flexibility, which can be applied across appetitive and aversive contexts.

It is thought that the CeA-driven freezing response mentioned previously, defined as a complete pause in movement except for what is necessary to breathe, can be overridden by the activation of the ventral subdivision of the mPFC (vmPFC). Originally, it was proposed that this inhibition occurred through prefrontal connections with GABAergic intercalated cells (ITC) within the amygdala that relay

information from BLA to CeA, since lesioning these cells was shown to impair the expression of fear extinction (Asede, Bosch, Lüthi, Ferraguti, & Ehrlich, 2015; Likhtik, Popa, Apergis-Schoute, Fidacaro, & Paré, 2008). However, there has been some conflicting reports on the level of connection between vmPFC and ITCs, with some studies reporting moderate to heavy connections while other more recent work has shown sparse innervation (Berretta, Pantazopoulos, Caldera, Pantazopoulos, & Paré, 2005; Pinard, Mascagni, & McDonald, 2012; Quirk & Mueller, 2008; Vertes, 2004). However, it is possible that sparse connections between ITCs and vmPFC could still play a role in fear extinction and may even help explain phenomena like spontaneous recovery, renewal, and reinstatement of freezing responses (Giustino & Maren, 2015). Recent evidence, however, has also proposed that vmPFC may strongly inhibit the CeA via basomedial amydala (BMA) connections, which are able to differentiate between safe and aversive contexts and decrease freezing related to shock cues when activated (Adhikari et al., 2015).

Through various lesion and inactivation studies, we see that failure to activate infralimbic portions of the vmPFC leads to failure to avoid shock during shock escape-avoidance procedures (Maren & Quirk, 2004; Morgan & LeDoux, 1995; Moscarello & LeDoux, 2013). Another study by Sangha et al. further demonstrated that IL inactivation caused failed discrimination between shock and safety cues in a discriminative conditioning task including reward, shock, and compound shock/safety cues (Sangha, Robinson, Greba, Davies, & Howland, 2014a). However, few have recorded from vmPFC neurons during avoidance behavior, especially in combination with approach in the same task. To better understand the function of vmPFC within

the overall circuit, it is important to ask if its activity increases during successful avoidance and if it shows reduced firing during poor avoidance behavior.

Related to its role in the suppression of freezing, the vmPFC has also traditionally been associated with fear extinction. In a foundational study, Milad and Quirk showed that IL neurons increased firing during recall of fear extinction, but not during fear conditioning or early extinction (Milad & Quirk, 2002). A human fMRI study found that, while the amygdala and striatum showed increased BOLD activation to cues predicting shock, vmPFC seemed to track stimuli that were not paired with shock; specifically, vmPFC was activated most strongly by cues that recently switched from signaling shock to signaling no-shock (Schiller & Delgado, 2010). This suggests that the vmPFC might signal important positive switches in cue valence (Delgado, Li, Schiller, & Phelps, 2008).

More recently, important work has shown that vmPFC is also involved in extinction in appetitive environments, not only in aversive contexts. Inactivation of vmPFC resulted in increased responding during extinction of an appetitive food reinforcer (Eddy, Todd, Bouton, & Green, 2016). A study using a rat homolog to fMRI BOLD responses, in vivo oxygen amperometry, found that IL showed increased response during early extinction for reward cues; thus, early in learning, IL might suppress contextually inappropriate action toward cues that no longer signal reward (Francois et al., 2014).

While it has been shown that vmPFC activation is critical for the acquisition and expression of extinction, vmPFC has also been implicated in the expression of habit. A study by Coutureau and Killcross showed that inactivation of IL could

reinstate flexible, goal-directed behavior even after habits had been formed after extensive training on a task; in fact, these animals could no longer express stimulus-response (S-R) habit-like behaviors (Coutureau & Killcross, 2003). Further, D1 receptor blockade and D2 receptor activation in IL also reversed habit formation and encouraged flexible goal-oriented responding for reward (Barker, Torregrossa, & Taylor, 2013).

While results suggesting vmPFC involvement in extinction and habit in both aversive and appetitive contexts seem to be contradictory at first glance, taken together they may suggest a broader role for vmPFC in the regulation of contextually appropriate behaviors. Single unit recordings in IL during appetitive tasks where rats had to press a lever for sucrose revealed delayed and prolonged IL activation in 25% of neurons to the collection of sucrose, delayed sucrose collection when IL was inactivated, and further encoding of contextually appropriate behavioral initiation during reinforced and extinction blocks (i.e., seeking when reward was present, withholding during extinction when reward was no longer signaled) (Burgos-Robles, Bravo-Rivera, & Quirk, 2013; Moorman & Aston-Jones, 2015a).

In studies that alternated or combined rewarding and aversive components, populations within vmPFC were activated for both. In monkeys, separate populations of vmPFC single units were found to process of reward and punishment outcomes when they were presented in alternating blocked trials; specifically, ventral vmPFC was more active during blocks where reward was delivered, while dorsal vmPFC was more active during blocks where punishment was delivered (Monosov & Hikosaka, 2012). In a task that presented conflict between approach cues and an interfering

pain-predictive cue, NAc to vmPFC connections were found to mediate this conflict (Schwartz, Miller, & Fields, 2017).

Finally, increased BOLD responses were found in vmFPC when human participants were forced to decide between options to choose the 'better' of two good choices; additionally, vmPFC seemed to encode and track outcome expectations surrounding both possible outcomes (Blair et al., 2006). This suggests that vmPFC responses are shaped by contextual information about specific outcomes. Thus, taken together, the role of vmPFC might deal more with contextually appropriate behavioral selection than inhibition, extinction, or habit exclusively (Gourley & Taylor, 2016; O'Doherty, 2011; Roy, Shohamy, & Wager, 2012; Schiller & Delgado, 2010).

Dopamine neurons and dopamine release

Within the midbrain lie the two major sources of dopamine neurons, the VTA and SNc, which send broad projections within the basal ganglia (striatum, pallidum, substantia nigra pars reticulata, subthalamic nucleus) and prefrontal cortex (Björklund & Dunnett, 2007; Bromberg-Martin et al., 2010; Cragg, Baufreton, Xue, Bolam, & Bevan, 2004; Lavoie, Smith, & Parent, 1989; Matsumoto & Hikosaka, 2009; Parker et al., 2016). These DA neurons are anatomically distributed across a topographical gradient, with the ventromedial SNc and lateral parts of the VTA containing mostly cells that encode motivational value signals using prediction errors, while cells that encode motivational salience signals are more densely packed within the dorsolateral SNc (Bromberg-Martin et al., 2010; Matsumoto & Hikosaka, 2009).

Historically, signals from midbrain DA neurons have been shown to play a critical role in reinforcement learning by providing a physiological correlate to the well-studied prediction error (PE). Phasic bursts or pauses in neuronal activity, together with resulting neurotransmitter release, encodes this PE signal, which guides goal-directed behavior by informing the system which aspects of the environment are appetitive or aversive and initiating actions to obtain the good and avoid the bad (Schultz, Dayan, & Montague, 1997). The PE signal measures the difference between an expected outcome and the actual outcome to inform future behavior. A better-than-expected outcome activates dopaminergic neurons resulting in phasic neurotransmitter release (positive PE), while a worse-than-expected outcome induces a pause in dopaminergic firing and neurotransmitter release downstream (negative PE). A fully predictable outcome elicits no change in firing of DA neurons or baseline DA concentrations.

Based on the mismatch of expectation and consequence, the DA signal acts as a teaching mechanism, updating expectations and potential behavioral responses based on feedback received from the environment. With continued exposure to outcomes and environmental cues that precede them, the same firing pattern is then applied to sensory cues that come to predict or give information about future outcomes. As the system continues to learn about the environment, the primary reinforcer (i.e., the unconditioned stimulus; the outcome) will start to become more easily predictable and the DA response to the outcome will diminish over time; simultaneously, the DA signal will gradually shift to the outcome's first predictor (e.g., cue light, drug paraphernalia, etc.) (Schultz et al., 1997). Most of the value

signaling described in the brain areas above likely relies on DA to form associations between stimuli and outcomes during learning and decision-making. Importantly, with continued cue-outcome learning, DA firing and release tends to shift away from the delivery of primary rewards as they come to be predicted by cues during learning, resulting in more or less firing for cues that predict appetitive and aversive outcomes, respectively (Day, Roitman, Wightman, & Carelli, 2007; Schultz, 1998; Schultz et al., 1997).

Although DA PE signaling is often studied in the context of rewarding or appetitive stimuli, it also applies to aversive stimuli, such as air puff and shock. In primates, neurons that encode reward PEs are depressed by unexpected air puff and visual cues that predict them (Bromberg-Martin et al., 2010). Furthermore, DA firing increases when an expected air puff is omitted, an event that is more appetitive or better than expected, signaling a positive PE. Similarly, in rats performing an instrumental escape-avoidance paradigm, phasic DA activity to the predictive cue presentation can predict if rats will successfully avoid an upcoming foot shock (Oleson, Gentry, Chioma, & Cheer, 2012). Successful avoidance behavior was contingent upon DA release time-locked to the warning cue; dopamine release was also seen during the safety period, once shock had been successfully avoided. Importantly, dopamine release was not present to cues predicting escape responses during the avoidance paradigm or unavoidable shock in a fear conditioning paradigm; here, the dopamine was uninformative in that the animal could do nothing to avoid the shocks. Thus, as during appetitive paradigms, DA release signals still seem to adhere to the general rule of firing more or less strongly for cues and outcomes that

are better or worse than expected, respectively. However, dopamine release to aversive outcomes has not been as well-studied as dopamine release during appetitive cues and outcomes, and it is beginning to become clear that the PE mechanism may be more complicated when it comes to aversive outcome avoidance or combined approach and avoidance.

It is important to note, however, that not all DA neurons transmit reward PE signals. Other, anatomically discrete, DA neurons appear to be more concerned about the motivational salience of appetitive and aversive stimuli (Matsumoto & Hikosaka, 2009). These DA neurons fire similarly for both appetitive and aversive outcomes and the cues that predict them. In experiments where visual stimuli predict delivery of either reward or an aversive air-puff, these motivational salience DA neurons fire more strongly for these outcomes and the cues that predict them compared to neutral trials where no reward or air-puff is delivered. As mentioned briefly above, these two types of value- and salience-encoding DA neurons are somewhat segregated within VTA and SNc, with value encoding cells mostly located in VTA and motivational salience DA neurons mostly located in SNc; further evidence shows that salience-encoding DA neurons might preferentially project to the prefrontal cortex (PFC) and the core of nucleus accumbens (NAc), while reward-preferring or PE DA neurons project preferentially to the ventromedial PFC and the shell of NAc (Bromberg-Martin et al., 2010). Given this data, it seems likely that such a signal would be critical for driving attention/motivation toward salient and important (appetitive or aversive) events promoting learning in regions that these neurons project to, whereas

PEs signal might be critical for specifically updating representations of associations between events and their respective outcomes.

Terminal dopamine release can be thought of in terms of either tonic or phasic release patterns. Tonic release results from the baseline, rhythmic spontaneous firing of DA neurons regulated by prefrontal afferents, while phasic release results from transient DA neuron burst firing (Grace, 1991). Tonic release of DA provides a sustained, but low, extracellular DA concentration (i.e., 5-10 nM), activating mainly high affinity D2-type receptors which activates the indirect pathway; activation of this pathway initiates the G_i-protein cascade, decreasing cAMP activity and promoting LTD at the synapse (Grace, 1991, 1995; Rice, Patel, & Cragg, 2011). Interestingly, this pattern of activity also acts as a homeostatic mechanism by lowering the threshold for future synaptic potentiation within these cells, promoting higher probability for activation in the future. Burst firing of DA neurons and phasic DA release, on the other hand, dramatically increases DA released into the synapse well above tonic levels (i.e., 150-400 nM) for a short period of time, activating low affinity D1-type receptors (which are not already activated by tonic DA) and initiating the direct pathway; activation of this pathway initiates the G_s-protein cascade, increasing cAMP activity and promoting LTP at the synapse (Grace, 1991, 1995; Rice et al., 2011). Again, this also acts as a homeostatic mechanism by heightening the threshold for future synaptic potentiation, making it more difficult to induce cell firing.

However, DA-induced LTP or LTD at the synapse merely sets the stage for long-term changes in activation and behavior but does not guarantee it. Pre- and post-

synaptic modifications will only occur in the presence of coincident phasic DA release, presynaptic, and postsynaptic depolarization, as in response to a rewarding or salient event; this is the synaptic basis of reinforcement learning. Thus, if a stimulus has caused these value-encoding neurons to fire in the past, this system tells us we should approach these cues, as the outcome is of high value; alternatively, if a stimulus has caused these value-encoding neurons to be inhibited in the past, this system tells us we should avoid this cue, as the following outcome is aversive (Bromberg-Martin et al., 2010). DA neurons that fire synchronously and release DA as a result, are reinforced and are more likely to be activated in the future, promoting paired behaviors. The synchronized firing of dopaminergic neurons follows Hebb's law that "neurons that fire together, wire together", but DA must be released in order for reinforcement learning to occur and the synaptic connection between neurons to be strengthened (Bromberg-Martin et al., 2010; Montague, Dayan, & Sejnowski, 1996; Wolfram Schultz, 1998). If the outcome is not salient, but already completely predictable, synapses will not undergo changes and no neuronal error signal is produced.

The striatum

The striatum contains many dopamine release sites (Heien, Johnson, & Wightman, 2004). DA release within the striatum, driven by the activation of midbrain dopamine neuronal projections, maintains PE signals to drive behavioral selection toward action or suppression of action (Graybiel, Aosaki, Flaherty, & Kimura, 1994; Kreitzer & Malenka, 2008). Along with dopaminergic inputs received from VTA and SNc, striatal neurons also receive glutamatergic inputs from cortex,

amygdala, hippocampus, and thalamus (Britt et al., 2012). These glutamatergic and dopaminergic inputs have different target sites onto striatal neurons, which allows for complex modulation of their activity: glutamatergic inputs synapse onto dendritic spine heads, while dopaminergic inputs synapse on the spine necks (Freund, Powell, & Smith, 1984; Yager, Garcia, Wunsch, & Ferguson, 2015).

Importantly, 95% of all neurons within the striatum are inhibitory GABAergic medium spiny neurons (MSNs); there also exists a minority of cholinergic and GABAergic interneurons within the striatum that express a number of identifying proteins (e.g., parvalbumin [fast-spiking], calretinin, somatostatin, etc.) (Kemp & Powell, 1971). MSNs, have two primary phenotypes based on the type of dopaminergic G-protein coupled receptor found on the neuron and its overall action within the basal ganglia; D1-type MSN neurons (containing D1, D5 receptors) and D2-type MSN neurons (containing D2, D3, D4 receptors). D1-type receptors (D1R) have a low affinity for dopamine and are largely considered part of the direct motor pathway in that they stimulate adenylyl cyclase (AC) activity via the $G_{s/olf}$ subunit; D2-type receptors (D2r) have a high affinity for dopamine and are largely considered part of the indirect motor pathway in that they inhibit AC activity via the G_i subunit. In general, the direct pathway facilitates the initiation of body movements, while the indirect pathway suppresses body movements (Bromberg-Martin et al., 2010; Hikida, Kimura, Wada, Funabiki, & Nakanishi, 2010).

Consequently, with their low and high affinities and opposing pathways, it is hypothesized that this receptor dichotomy is at the basis of reward prediction error execution. For instance, fully predicted events would result in the maintenance of

low, tonic levels of DA, so D2Rs, but not D1Rs, and the indirect pathway would be activated; in contrast, large bursts of dopamine release in response to a reward-predictive cue would increase baseline levels of DA and activate low-affinity D1Rs and the direct pathway to promote action to obtain the predicted reward (Bromberg-Martin et al., 2010; Hikosaka, 2007). However, though it was long thought that these two types of receptor were mutually exclusive across neurons, it has recently been shown that almost 40% of striatal neurons express mRNA for both D1 and D2-like receptors (Nishi, Kuroiwa, & Shuto, 2011; Surmeier, Song, & Yan, 1996; Yager et al., 2015).

The striatum itself can be divided into two major sub-regions, the dorsal striatum and the ventral striatum. The dorsal striatum is usually broken down into its dorsomedial and dorsolateral components, while the ventral striatum can further be broken down into the nucleus accumbens (NAc) core and shell. These divisions are based on a number of factors, including anatomical projections, morphological differences between cells types, and proposed function (Cragg, 2003; Haber, Fudge, & McFarland, 2000).

The dorsomedial striatum (DMS) receives inputs from the mPFC and sends reciprocal projections to the ventral SNc and some sparse projections to dorsal SNc. The dorsolateral striatum (DLS) receives in puts from secondary and primary motor cortices and sends reciprocal projections to dorsal SNc. Additionally, both sub-regions of dorsal striatum receive information from ventral striatum; these combined inputs from cortical regions, midbrain dopamine, and ventral striatum promote encoding of response-outcome (R-O) goal and stimulus-response (S-R) habit

contingencies in dorsomedial and dorsolateral sub-regions, respectively (Balleine & O'Doherty, 2010; Kim, Lee, & Jung, 2013). Importantly, the spiraling midbrain-striatum-midbrain interactions formed by the reciprocal connections between dorsoventral striatum and VTA/SNc projections allows information to be propagated forward and is thought to underlie the gradual transformation of goal-oriented behaviors to habit.

Within the ventral striatum, the NAc shell receives inputs from OFC and infralimbic vmPFC and sends reciprocal projections to the medial VTA and some sparse projections to the lateral VTA (Salgado & Kaplitt, 2015; Yin, Ostlund, & Balleine, 2008). The NAc core receives inputs from the PL, insula, ACC and BLA and sends reciprocal projections to the lateral VTA and some sparse projections to the ventral SNc. Morphologically, in rats, cells within the NAc core tend to be larger cells with more dendrites and dendritic spines than those in the NAc shell; interestingly, the opposite is true in humans with the shell containing more dendritic arborization (Heimer et al., 1997; Salgado & Kaplitt, 2015).

Pharmacological and lesion studies have shown that DA in the ventral striatum seems to be involved in salience and motivation encoding, as well as reward prediction errors and value-encoding (Berridge, 2007; Burton, Bissonette, Lichtenberg, Kashtelyan, & Roesch, 2014; Lex & Hauber, 2010; Salamone & Correa, 2012). Previous single unit work in rodents and monkeys has clearly demonstrated that firing in VS is modulated by the value associated with cues that predict reward (Carelli & Deadwyler, 1994; Janak, Chen, & Caulder, 2004; Schultz, Apicella, Scarnati, & Ljungberg, 1992; Setlow, Schoenbaum, & Gallagher, 2003). Further, a set

of experiments using combined fast-scan cyclic voltammetry and electrophysiology in the NAc core and shell demonstrated that increased DA release preceded goal-directed behavior and occurred simultaneously with changes in NAc firing rates; these neural changes, as well as approach behavior, were inhibited using D1 receptor antagonists within NAc shell and $GABA_a$ receptor antagonists within NAc core, showing the behavioral importance of both DA and GABA receptors on MSN cells (Cheer, Heien, Garris, Carelli, & Wightman, 2005; Cheer et al., 2007). In humans, value-type signals have also been found in dorsal and ventral striatum, such that BOLD responses were greatest for reward, weaker for neutral and weakest for punishment trials, and cue-related activity in both NAc and VTA increased for both gain and loss trials, providing evidence for salience signals in these regions (Breiter, Aharon, Kahneman, Dale, & Shizgal, 2001; Cooper & Knutson, 2008; Delgado et al., 2000). Thus, it is possible that the ventral striatum might be required for integrating both value and salience types of information that are central to actor-critic models, as well frameworks that view the VS as a "limbic-motor" interface, in order to drive flexible approach responses in both appetitive and avoidance contexts (Bissonette, Burton, et al., 2013; Bissonette, Gentry, Padmala, Pessoa, & Roesch, 2014; Ikemoto & Panksepp, 1999).

Often within the literature, the functions of striatal regions are described in isolation, though interregional connection is still assumed. For instance, the NAc shell has been most often implicated in the reinforcing properties of novel unconditioned stimuli, reward and appetitive behaviors and drug relapse, while the core is more often associated with encoding conditioned cues and driving conditioned responses

toward these motivational incentive stimuli (Cardinal & Cheung, 2005; Corbit, Muir, & Balleine, 2001; Everitt et al., 1999; Roesch, Singh, Brown, Mullins, & Schoenbaum, 2009; Takahashi, Roesch, Stalnaker, & Schoenbaum, 2007). Together, the NAc uses information about outcomes to motivate behavior toward particular cues predicting and actions acquiring these outcomes (Roesch et al., 2009; Takahashi et al., 2007). Cue information from the NAc is then used by the dorsal striatum and combined with DMS response-outcome signals and DLS S-R signals to drive actions or inhibit action to obtain desirable outcomes (Balleine & O'Doherty, 2010; Everitt & Robbins, 2005; Graybiel et al., 1994; Ikemoto & Panksepp, 1999). In the actor-critic model, previously described above, the ventral portions of the striatum are thought to act as the "critic," assigning value information to states and maintaining predictions errors or DA neurons, and the dorsal portions of the striatum are thought to embody the "actor," using PEs from ventral striatum to drive appropriate and beneficial actions towards goals (Barto, 1995; O'Doherty et al., 2004; Sutton & Barto, 1998).

Importantly, it is also hypothesized that the spiraling connections between VTA and SNc DA neurons and the striatal sub-regions is responsible for the experience-based transition from goal-directed (R-O) behaviors to more habit-driven (S-R) behaviors (Haber et al., 2000; Haber, 2003; Haber & Knutson, 2010; Yin et al., 2008). As information moves more dorsolaterally on this spectrum, behaviors become more habit-driven. Thus, this spiraling interaction between regions emphasizes the importance of examining the circuit in totality, as well as its individual regional components. We can examine how interactions among striatal and dopaminergic brain regions affect goal-directed behavior through studies examining the effects of

ventral striatum lesions during a devaluation task (Singh, McDannald, Haney, Cerri, & Schoenbaum, 2010). Animals using goal-directed behavior will halt responding to a lever that produces previously devalued food (e.g., food that has made them sick); however, animals whose responding is habitual will continue to work for food regardless of its value. Animals with ventral striatum neurotoxic lesion as if they are under the control of habit and continue to press for devalued rewards due to an increased proportion of DS:VS activation (Singh et al., 2010).

Circuit summary

When we step back and look at how all of these regions interact with one another, a dynamic and complex circuit emerges. In this circuit, the ACC likely increases attentional control to ensure that neural processes are prioritized depending on expected actions and unsigned errors in reward prediction. The OFC represents value expectancies necessary for guiding decision-making and learning. These signals depend on BLA, which not only encodes associative appetitive and aversive information during sampling of conditioned stimuli and across states, but integrates value and intensity/salience during delivery of appetitive and aversive outcomes. OFC and BLA both broadcast this information to VS and DA neurons, which carry both evaluative (VTA) and motivational salience (SNc) signals in separate populations of neurons. PE signals generated by VTA DA neurons provide feed-forward information to more dorsal- medial and dorsal-lateral regions in striatum, which are critical for goal-directed and habitual behaviors, respectively. In this way, the NAc is often thought of as the limbic-motor interface (Mogenson, Jones, & Yim, 1980). Last, but not least, vmPFC might use integrated value and salience signals

from BLA to send information to CeA and striatum to inform flexible, contextually relevant behavioral choice. Therefore, the signals carried by cortical areas not only inform deeper mesolimbic brain regions (VTA/SNc DA projections, VS/DS, BLA/CeA) but are simultaneously dependent upon their feedback and the feed-forward outputs to motor-oriented regions (globus pallidus, subthalamic nucleus, thalamus, motor cortices) in order to drive behavioral responses.

Individual differences in positive and negative reinforcement

Good avoidance versus Poor avoidance

While DA is well-known for its role in reward processing, much less is known about its role in the avoidance of something aversive. Both reward and punishment can promote instrumental responding, but respective approach and avoidance behaviors are differentially governed by positive and negative reinforcement strategies (Dayan, 2012; Dayan & Niv, 2008). Additionally, the processes that guide punishment avoidance are thought to be much more complex than those governing reward approach, since punishment avoidance involves both an initial Pavlovian freezing response, which must be overcome in order to perform a subsequent instrumental response (Bolles, 1970; Morris, 1974; Rescorla & Solomon, 1967). Because of this initial fear response, rats tend to freeze to cues that predict an aversive outcome and are unable to produce behaviors necessary to avoid the negative consequence; in this stage, the cue itself becomes aversive (Moscarello & LeDoux, 2013). Aversive cues and cues leading to freezing behavior have been shown to inhibit dopamine release in the nucleus accumbens (Oleson et al., 2012). However,

with experience, most rats are able to learn that behaviors that terminate the cue also lead to the escape from or avoidance of the aversive outcome; when the desired behavior is fully learned, the cue that predicts shock loses its aversive qualities and rats are able to successfully avoid punishment (Oleson et al., 2012; Rescorla & Solomon, 1967). It has been shown that accumbal dopamine release during shock cues predicts successful avoidance; specifically, in rats, phasic DA release was seen in the NAc core to cues predicting successful instrumental shock avoidance and during subsequent safety from shock but not when rats escaped or failed to avoid (Oleson et al., 2012).

However, in most studies, a subset of rats consistently freeze to the shock cue and, hence, maintain poor avoidance behavior throughout the experiment. Unfortunately, animals exhibiting poor avoidance behavior are usually excluded from these studies under the assumption that they are not learning the task; recently, this behavioral difference has been shown to be due to performance deficits and not a failure to learn (Brush, 2003; Lazaro-Munoz, LeDoux, & Cain, 2010; Martinez et al., 2013). It is thought that these behavioral differences likely arise from differences in the prefrontal control of freezing responses and DA reinforcement signals. This important behavioral difference could provide insight into the neural mechanisms underlying individual differences in cue processing and could give insight into disorders such as addiction and PTSD, whose symptoms are highly linked to maladaptive cue processing. In *Chapter 3*, we took advantage of these behavioral differences to uncover underlying variance in accumbal dopamine release between rats expressing good or poor shock avoidance.

It has been shown that the freezing response induced by punishment cues originates from activation of the CeA (Bolles, 1970; Moscarello & LeDoux, 2013; Purgert et al., 2012). Recently, studies have shown that this Pavlovian freezing response can be overridden by the activation of the ventral subdivision of the mPFC (vmPFC), which is thought to inhibit the CeA via basomedial amydala (BMA) connections (Adhikari et al., 2015).Through various lesion and inactivation studies, we see that failure to activate vmPFC, specifically the infralimbic sub-region, leads to failure to avoid shock during shock escape-avoidance procedures (Maren & Quirk, 2004; Morgan & LeDoux, 1995; Moscarello & LeDoux, 2013). However, no one has recorded from vmPFC during avoidance. To address these gaps, we conducted single-unit recordings in the vmPFC during combined a combined approach and avoidance task; in *Chapter 2*, this project will be discussed in depth. These results will help us determine if: 1) vmPFC does indeed increase firing during successful avoidance, and 2) vmPFC shows reduced firing during poor avoidance behavior.

Sign-tracking versus Goal-tracking

As mentioned previously, avoidance behavior consists of both an initial Pavlovian component and a secondary instrumental action component. A Pavlovian freezing response to cues predicting shock can sometimes persist during shock avoidance training, which is thought to underlie performance differences across animals. Animals who freeze more to cues predicting shock are less likely to generate actions needed to avoid the oncoming shock and are, thus, usually worse at the task (i.e., poor avoiders). Still, it is unclear what drives an animal to focus on and respond to these Pavlovian-conditioned stimuli as opposed to expected outcomes, and vice

versa. It is also currently unknown if avoidance is encoded in an analogous way to reward approach within the brain, but studying the Pavlovian aspect of reinforcement learning could be a simple way to ask and answer these questions.

Pavlovian autoshaping procedures could lend some insight. In these types of tasks, two distinct behavioral phenotypes tend to emerge along a gradient: rats can be categorized as either sign- or goal- trackers, or somewhere in between, depending the conditioned response they form after training and whether they respond most vigorously to cues or outcome locations, respectively (Flagel et al., 2011; Kaveri & Nakahara, 2014; Lesaint, Sigaud, Flagel, Robinson, & Khamassi, 2014; Saunders & Robinson, 2013b). During an autoshaping task, an animal is presented with a neutral cue (e.g., a lever) which is then paired with a primary reward (e.g., food pellet). With continued pairings, the previously neutral cue acquires predictive power and motivational value, becoming a conditioned stimulus. Once this association is learned, animals will develop a conditioned response, which varies between individuals; this conditioned response usually resembles a consummatory behavior related to the modality of the delivered reward (e.g., biting and chewing if the US is a food reward).

For some animals, the conditioned cue becomes incredibly motivationally salient, and they will work for it even in the absence of reward. These animals, termed “sign-trackers” (ST), may not even approach the location of reward delivery during the cue presentation and, instead, will engage solely with the lever cue. Goal-trackers, however, pay little attention to the conditioned cue and immediately begin to interact with the location of reward delivery upon cue presentation. Intermediate

animals perform a more even mix of both sign-tracking and goal-tracking behaviors. Some suggest that differences between sign- and goal-trackers arise from the way the animals learn reinforcement tasks, via model free versus model based systems (Cardinal et al., 2002; Clark, Hollon, & Phillips, 2012; Doll, Simon, & Daw, 2011).

According to the literature, there are at least two ways animals can learn from reinforcers: model-based or model-free learning systems. In a model-based system, animals use representations of their environment and related expectations to make predictions about the future; this type of system is usually thought to aid in making goal-directed behaviors, such as those exhibited by goal-trackers (Doll et al., 2011; Flagel et al., 2011; Lesaint et al., 2014; A. Solway & Botvinick, 2011; Yin, Ostlund, Knowlton, & Balleine, 2005). By using an internal approximation of possible consequences, it is possible to infer potential outcomes using this system and adapt rapidly to novel or changing environments. In a model-free system, animals do not use an internal model and instead use cached estimates of context and action values gathered from previous experiences; this type of system is thought to be seen most often during reflexive or habitual behavior, such as stimulus-response behaviors exhibited by sign-trackers (Doll et al., 2011; Graybiel, 2008; Yin et al., 2005). Pavlovian learning has also typically been presumed to be model-free in nature.

However, recent work has called into question the role of model-free learning during Pavlovian conditioning experiments, suggesting the development of sign- and goal- tracking behavioral phenotypes is likely a complex process involving both model-free and model-based learning mechanisms. Flagel et al. used an auto-shaping task to identify sign- and goal-trackers and examine phasic DA release within the

NAc core of each group (Flagel et al., 2011). They found that DA release matched RPE signaling and was required for the acquisition and performance of sign-tracking behaviors, while goal-tracking behaviors could be developed and performed in the absence of an RPE-like DA signals, including under the systemic influence of a DA antagonist, flupenthixol (Danna & Elmer, 2011; Flagel et al., 2011). Additionally, dopamine release during goal-tracking behaviors did not seem to represent a typical RPE-like signal, with a transfer of release from the primary reward to the cue across time; instead, dopamine release to the primary reward was maintained throughout the sessions, as in early learning. Dayan and Berridge also suggest that Pavlovian training can involve a form of model-based learning in that a novel state may change the value of the primary reward and enhance incentive salience of that US before new learning can occur (Dayan & Berridge, 2014).

A revision to the traditional model-based and model-free dichotomy is described by the Lesaint-Khamassi model, which we test in *Chapter 4* (Lesaint, Sigaud, Clark, Flagel, & Khamassi, 2015; Lesaint et al., 2014). Instead of viewing these systems separately, the Lesaint-Khamassi model combines model-based and model-free systems and weighs the contribution of each system in its estimation of conditioned responses. By doing this, the Lesaint-Khamassi model computes a spectrum of potential observed behaviors, with sign-trackers and goal-trackers at the extremes and intermediates in between, instead of simply separating animals into two discrete categories (Meyer et al., 2012). We believe viewing these behaviors on along a gradient is a more accurate representation of true behavioral development and performance than the more simplified dichotomy that has been previously tested. This

model has the potential to help us understand why the cue (e.g. light, lever) and reward location (food cup) might acquire different motivational values in different individuals, even when these individuals are trained in the same task (Robinson & Flagel, 2010). In addition, this model could help us explain why dopaminergic responses vary so widely between sign-trackers and goal-trackers, specifically why sign-tracking is DA dependent and goal-tracking is not (Flagel et al., 2011).

We also believe interesting comparisons may come to light between good and poor avoiders and sign- and goal- trackers. We propose that rats demonstrating poor avoidance are highly focused on the outcome, which we believe to be similar to behavior seen in goal trackers; however, whether this behavior is beneficial or harmful seems to be depending on the context (appetitive or aversive). On the other hand, we propose that rats demonstrating good avoidance are not focused on the potential aversive outcome but are driven by the cues, similar to what is seen in sign-tracking behavior. We will examine this idea further in *Chapter 5*, the general discussion.

Individual differences and the formation of psychiatric disorder

Individual differences that arise during these tasks may also provide insight into the development of psychiatric disorder, specifically why some individuals demonstrate vulnerabilities while others do not. A broad range of psychiatric disorders (e.g., addiction, anxiety disorders, obsessive-compulsive disorder [OCD], Tourette's syndrome, attention-deficit hyperactivity disorder [ADHD]) seem to hijack positive and negative reinforcement systems to produce maladaptive behaviors. These vulnerabilities can also be linked to disturbances in the dopaminergic system and

corticomesolimbic circuit (Maia & Frank, 2011). By understanding how these reinforcement and neural systems differ in individuals presenting with each of these disorders, we may be able to better understand possible links across tasks and disorders and potentially inform better, more targeted treatments.

During the development of addiction, both positive and negative reinforcement systems become dysfunctional. In the beginning, fast phasic DA release in the NAc signals a positive prediction error when a drug (or other rewarding stimulus) is received; with time, cues that predict drug receipt also induce phasic DA release to drive drug-seeking. Stimulation of the direct or indirect dopaminergic pathway has been shown to either increase or decrease this positive reinforcement effect, respectively (Lobo et al., 2010). These recurring cue-triggered motivations may become quite powerful and maladaptive, resulting in compulsive drug use over time, compounded by a decrease in sensitivity to rewards (Dezfouli et al., 2009; Maia & Frank, 2011; Redish, 2004). However, with time, a transition from positive to negative reinforcement mechanisms is thought to take place, driven by an attempt to avoid symptoms of withdrawal and associated negative affect (Ahmed & Koob, 2005; Koob, 2013). Neural correlates of these effects can be seen in the overall decreased activation of the corticomesolimbic system, specifically through decreased D2 receptors in the striatum and hypoactivity of OFC & vmPFC (Volkow, Fowler, & Wang, 2003).

Anxiety disorders are also thought to result from a conflict between approach-related and avoidance-related drives, and treatments usually focus on reducing avoidance-related drives to ultimately reduce anxiety (Aupperle & Paulus, 2010). In

patients with general anxiety, warning signals that predict an aversive stimulus become aversive in themselves, and these individuals seem to be more sensitive to uncertain and unexpected negative events on the whole (Borkovec, Alcaine, & Behar, 2004; Mowrer & Lamoreaux, 1946). Traditionally, the formation of anxiety disorders has been associated with deficiencies in extinction learning and the reduced activation of the vmFPC (Milad et al., 2007, 2009; Stein & Paulus, 2009). However, positive reinforcement systems might also be affected. Specifically, patients presenting with post-traumatic stress disorder (PTSD) show decreased expectancy and satisfaction for rewarding events, will not work as hard for reward, and show a decreased ability to learn reward contingencies (Aupperle & Paulus, 2010). In accordance with the idea that both reinforcement systems are affected, patients with anxiety disorders show increased neural activity in the amygdala and insula, while activity in the striatum is decreased (Aupperle & Paulus, 2010; Sailer et al., 2008).

Disorders of compulsion and impulsivity, such as OCD, Tourette's syndrome, and ADHD, are also thought to result from irregularities within corticomesolimbic reinforcement systems. For instance, in OCD, it is thought that a relationship between anxiety and accidental non-contingent reinforcement produces the hallmark superstitious behaviors (Bloom, Venard, Harden, & Seetharaman, 2007; Catania & Currs, 1963). These compulsive behaviors help relieve anxiety and intrusive obsessive thoughts. Tourette's syndrome is marked by compulsive motor and vocal stereotypies, which are also thought to be negatively reinforced by the removal of invasive, aversive premonitory urges (Capriotti, Brandt, Turkel, Lee, & Woods, 2014; Maia & Frank, 2011). Patients suffering from Tourette's syndrome show abnormal

activation of the D1 direct pathway and increased numbers of D2 receptors within the striatum; administration of D2 blockers has been shown to reduce related compulsive behaviors (Maia & Frank, 2011; Mink, 2001). Alternatively, patients showing increased hyperactivity and impulsivity related to ADHD, have abnormally low tonic DA release within the prefrontal cortex, which parallels deficits that have been shown to occur in addiction (Frank, Santamaria, O'Reilly, & Willcutt, 2007; Kheramin et al., 2004; Maia & Frank, 2011; Smith, Becker, & Kapur, 2005; Volkow et al., 2003). Interestingly, patients across these disorders have been shown to learn better during tasks employing rewards compared to tasks employing punishment, so it seems both positive and negative reinforcement systems may be affected (Palminteri et al., 2009).

Behavioral tasks and measurements

My combined approach and avoidance task

As mentioned previously, most studies that have investigated the neural mechanisms of approach or avoidance have done so separately. Recently, more work has aimed to address gaps in the value and motivation literature by employing tasks that combine reward approach and punishment avoidance. The task I created for use in *Chapters 2-3*, recording activity from single units in vmPFC and phasic dopamine release in NAc, also aims to help bridge this gap by using combined approach and avoidance within the same session. In order to investigate positive and negative reinforcement simultaneously, we designed a behavioral task that manipulates both the promise of reward and the threat of punishment within the session.

In brief, each session consisted of three trial types that were pseudo-randomly interleaved: reward, neutral, and shock trials. Each trial consisted of a distinct audiovisual cue presentation, signaling whether the trial was reward, shock, or neutral, followed by access to an operant lever; animals could press the lever to gain a sucrose pellet reward (reward trial, positive reinforcement), avoid a shock punishment (shock trial, negative reinforcement), or encounter no consequence (neutral trial, non-reinforced). Animals were progressively trained on this task (e.g., shaping toward the lever, shock escape paradigm, avoidance paradigm, addition of reward and neutral trial types). All animals were well-trained on this task (approximately 2 months of behavioral training; at least 60% avoidance success) prior to recordings, therefore our neural results measure neural correlates during performance rather than learning. This task was used for both *Chapter 2* and *Chapter 3* and will be elaborated upon further within the *Materials and Methods* section of my *Chapter 2*. The basic design of each trial type used in our combined approach and avoidance task is illustrated in **Figure 2a**.

This combined approach and avoidance task has a few key characteristics that are important to note. First, reward and punishment trials are pseudo-randomly

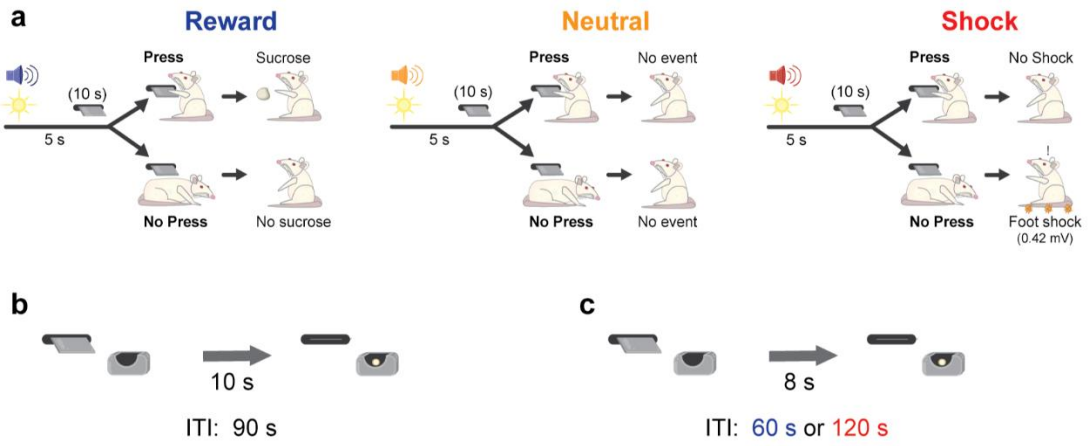


Figure 2. Task designs for experiments outlined in *Chapters 2-4*. (**A**) Combined approach and avoidance task used in *Chapters 2-3*, which consists of three trial types: Reward (positive reinforcement), Neutral (unreinforced), and Shock (negative reinforcement), pseudorandomly interleaved within each session. After cue onset, animals are required to press a lever in order to receive sucrose in Reward trials, encounter no consequence in Neutral trials, and avoid foot shock in Shock trials. (**B-C**) Two variations of a standard autoshaping task that were used in *Chapter 4* to measure and manipulate the development and expression of sign- and goal- tracking behaviors. In each task, a lever extended into the test chamber for a certain amount of time (B: 10s, C: 8s), after which the lever was retracted, and a sucrose pellet was delivered to an adjacent food cup. Time between trials (ITI) was either 90s (B), 60 or 120s (C) ± 30s.

interleaved within the same session. Therefore, each rat will receive the same number of trial types and should not develop any effects due to sequence. Many previous studies that combine reward and punishment trials do so in a blocked format, which can open the door to effects of timing, sequence, learning, and tissue drift during electrophysiological studies. Due to the high temporal specificity of our techniques, timing of neural data acquisition during interleaved trials was not a critical factor in deciding between blocked and interleaved trials, as it may be for fMRI studies. We were also not interested in sequence effects, such as reversals, which would benefit from blocked trials. Interleaving trials has also been shown to boost learning, which may be beneficial for such a complex task (Carvalho & Goldstone, 2015; Ethridge, Brahmbhatt, Gao, McDowell, & Clementz, 2009).

Second, our task also includes neutral, non-reinforced trial type as a control. Some combined studies have employed a neutral trial type, and we find this extremely important to the comparison of effects across reward and punishment trial types. Importantly, the same behavioral sequence is required across all three trial types; the animal must approach the lever and press it within the given time period (10s) in order to gain reward, avoid punishment, or obtain nothing on neutral trials. This controls for activity related to the action alone and also allows us to compare activity across all three trial types.

Lastly, the 5 second time delay between audiovisual cue onset and lever access allowed us to temporally isolate neural activity to the cue, the action, and the outcome delivery. This ability to easily dissociate between trial events allowed for precise control during analysis.

Additionally, the inclusion of a simultaneous audiovisual cue provides the animal with both a localizable (visual) and a potentially more attention-grabbing, specific (auditory) cue elements. Many studies use the presentation of a cue light, but if the animal is not facing the cue light, this could result in more unintentionally failed trials and could exacerbate shock-related learning difficulties; however, this is localizable, unlike the auditory cue. The inclusion of the auditory stimulus allowed the animal to discriminate between trial types. By using a task employing all of these unique features, we were able to observe and record individual differences in neural activity and behavior that may not have been discernible using a more simplified task.

The behavioral measures that we examined from this task included percent correct, reaction times, and correlations between these two measures for each trial type across sessions and across rats. Percent correct was calculated based on the number of lever presses versus total number of trials for each trial type. Reaction times were calculated based on the time difference between auditory cue offset/lever extension and when the animal pressed the lever; therefore, reaction times were only calculated for successful trials when the animal pressed the lever.

Pavlovian autoshaping task

As mentioned above, most studies investigating sign- and goal- tracking behavioral differences employ an Pavlovian autoshaping task. This is a very simple task where an animal is presented with an initially neutral cue (e.g., a lever) which is followed by a primary reward (e.g., food pellet) after some delay. With continued pairings of the cue and reward, the previously neutral cue can acquire predictive and motivational power, becoming a conditioned stimulus. Once this association is

learned, animals will develop a conditioned response to the cue presentation, which varies between individuals. This conditioned response usually presents as a consummatory behavior based on the modality of the reward (e.g., biting and chewing the CS if the US is a food reward). During autoshaping procedures, two distinct behavioral phenotypes tend to reliably emerge: rats can be categorized as either sign- or goal- trackers, or somewhere in between, depending the conditioned response they form after training. Importantly, this behavior does not emerge as a dichotomy but animals usually fall along a sign- and goal- tracking spectrum. Generally, sign-trackers respond most vigorously to cues, while goal-trackers respond most vigorousy to outcome locations (Flagel et al., 2011; Kaveri & Nakahara, 2014; Lesaint et al., 2014; Saunders & Robinson, 2013b).

In the fourth chapter of my dissertation work, we employed a Pavlovian autoshaping task to attempt to manipulate the development and expression of sign- and goal- tracking behaviors by varying the length of the ITI and inhibiting VTA to NAC DA neurons during learning and performance of the task. There were two notable task differences between the two experiments outlined in *Chapter 4*, though the general sequence remained the same. First, the length of the ITI and, secondly, timing between cue presentation and reward delivery varied between experiments. This task was used for both experiments within *Chapter 4* and will be elaborated upon further within the *Materials and Methods* section of that chapter. The basic design of each trial type for both experiments is outlined in Fig. 2.

In our DREADD autoshaping experiment (**Fig. 2b**), we used the standard 90 second variable interval ITI, which is used in the majority of prior studies examining

sign- and goal- tracking during autoshaping (Robinson & Flagel, 2009). For our FSCV autoshaping experiment (**Fig. 2c**), we wanted to test the effect of shifting this ITI to provide less time (60 s ITI) or more time (120 s ITI) to revise food cup values on the acquisition and performance of sign- and goal-tracking behaviors. Additionally, cue presentation varied slightly between our FSCV and DREADD autoshaping procedures; our DREADD experiment used an 8 s delay between lever extension and reward delivery, which has been previously used by Flagel et al. (Robinson & Flagel, 2009), while our FSCV experiment used a 10 s delay between lever extension and reward delivery. Importantly, lever presentation was not further signaled using a cue light or sound in either experiment, which has been shown to produce different results (Kohler et al., unpublished).

The behavioral measures that we examined from this task included PCA score and its three components (Probability, Response Bias, and Latency). Each of these measures is represented as a score from -1.0 to +1.0, with scores closer to -1.0 signifying more goal-tracking behaviors while scores closer to +1.0 signify more sign-tracking behavior and scores from -0.5 to 0.5 signifying intermediate animals. Probability difference was calculated as $(P_{lever}-P_{receptacle})$/total trials, where the total number of trials containing either a lever press or a food cup entry were divided by the total number of trials; thus, the probability score gives the proportion of lever presses or magazine visits across a session. Response Bias was calculated as (#Lever Presses – #Food Cup Entries) / (#Lever Presses + #Food Cup Entries), which focuses instead on the actual number of times each behavior was performed during a session as opposed to the ratio. Latency was calculated as ($\overline{x}$ Cup Entry Latency – $\overline{x}$ Lever

Press Latency) / cue length, giving the average latency to act toward food cup or lever during a session. The PCA index is an average of these three measures. We also measured the frequency of food cup visits and the frequency of lever press across each trial, across each session.

Chapter 2: Neural activity in ventral medial prefrontal cortex is modulated more during approach than avoidance

This chapter is currently under review for publication.

Authors: Ronny N. Gentry[1,2] and Matthew R. Roesch[1,2]

1. Department of Psychology

2. Program in Neuroscience and Cognitive Science
1147 Biology-Psychology Building
University of Maryland
College Park, MD 20742

Abstract

Ventromedial prefrontal cortex (vmPFC) is thought to provide regulatory control over Pavlovian fear responses and has recently been implicated in appetitive approach behavior, but much less is known about its role in contexts where appetitive and aversive outcomes can be obtained and avoided, respectively. To address this issue, we recorded from single neurons within vmPFC while rats performed our combined approach and avoidance task under reinforced and non-reinforced (extinction) conditions. Surprisingly, we found that cues predicting reward modulated cell firing in vmPFC more often and more robustly than cues preceding avoidable shock; additionally, firing of vmPFC neurons was both response (press or no-press) and outcome (reinforced or extinction) selective. These results suggest a complex role for vmPFC in regulating behavior and supports its role in appetitive contexts during both reinforced and non-reinforced conditions.

Significance Statement

Selecting context-appropriate behaviors to gain reward or avoid punishment is critical for survival. While the role of ventromedial prefrontal cortex (vmPFC) in mediating fear responses is well-established, vmPFC has also recently been implicated in the regulation of reward-guided approach and extinction. Many studies have used indirect methods and simple behavioral paradigms to study vmPFC, which leaves the literature incomplete. We measured vmFPC neural activity during a complex cue-driven combined approach and avoidance task and during extinction. Surprisingly, we found very little vmPFC modulation to cues predicting avoidable shock, while cues predicting reward approach robustly modulated vmPFC firing in a response- and outcome-selective manner. This suggests a more complex role for vmPFC than current theories suggest, specifically regarding context-specific behavioral optimization.

Introduction

This chapter is currently under review for publication.

The medial prefrontal cortex (mPFC) is thought to exhibit control over appetitive behavior and Pavlovian fear responses. Broadly, dorsal mPFC is implicated in reward processing and goal-directed behaviors, whereas ventral prefrontal cortex (vmPFC) is associated with aversive processing and the formation and expression of extinction behaviors and habit (Ostlund & Balleine, 2005; Senn et al., 2014; Sotres-Bayon & Quirk, 2010). This dissociation is supported by divergent anatomical projections, with the dmPFC connecting to nucleus accumbens (NAc) core and dorsomedial striatum, while the vmPFC projects to NAc shell, amygdala, and bed nucleus of the stria terminalis (Haber et al., 2000; Killcross & Coutureau, 2003).

Though several recording studies have examined the role of dorsal mPFC in behaving rats, fewer have explored the function of single neurons in vmPFC. Generally, vmPFC activity is thought to regulate fear-related behaviors. For example, trace fear conditioning has been shown to transiently increase intrinsic vmPFC excitability in basolateral amygdala (BLA) projections, which was positively correlated with freezing behavior (Song, Ehlers, & Moyer, 2015). Others have shown that fear-induced freezing can be overcome by the activation of vmPFC to basomedial amygdala (BMA) projections, which can differentiate between aversive and safe environments (Adhikari et al., 2015). While these studies show the importance of vmPFC in fear conditioning, others have supported its role in avoidance behavior (Giustino, Fitzgerald, & Maren, 2016; Schwartz et al., 2017; Soler-Cedeño, Cruz, Criado-Marrero, & Porter, 2016); lesion and inactivation of vmPFC leads to failed

shock avoidance and muddles discrimination between fear and safety cues (Adhikari et al., 2015; Sangha, Robinson, Greba, Davies, & Howland, 2014b).

Emerging evidence from Moorman et al. suggests that single neurons in vmPFC may also be critically involved in appetitive behavior (Moorman & Aston-Jones, 2015b). In line with these findings, recent vmPFC lesion and inactivation studies show suppressed reward seeking behavior in contexts that were previously associated with gaining reward and suggest these connections to NAc are necessary for the expression of reward-predictive cue-driven behavior (Bossert et al., 2011; Keistler, Barker, & Taylor, 2015; Zeeb, Baarendse, Vanderschuren, & Winstanley, 2015). However, the opposite outcome has also been reported— that *activating* vmPFC projection neurons suppresses cue-induced drug seeking behavior (LaLumiere, Smith, & Kalivas, 2012; Peters, LaLumiere, & Kalivas, 2008). Such variance could arise from the simultaneous existence of separate, but intermingled, neural ensembles within vmPFC that selectively encode opposing environmental actions for learned cue-driven responses, like reward or extinction (Suto et al., 2016; Warren et al., 2016).

In addition to the regulation of fear and appetitive behaviors, extensive evidence also implicates the vmPFC in extinction learning and expression. Several studies have found that the degree to which extinction memories are retrieved scales with firing and burst activity within vmPFC (Burgos-Robles, Vidal-Gonzalez, Santini, & Quirk, 2007; Maroun, Kavushansky, Holmes, Wellman, & Motanis, 2012; Milad & Quirk, 2002; Wilber et al., 2011). In line with this, inactivation of the vmPFC has been shown to increase responding to food-predictive cues during extinction and impair

extinction retrieval, while stimulation of the vmPFC promoted extinction to fear-related cues (Eddy et al., 2016; Milad & Quirk, 2002). This is consistent with the finding that vmPFC both encodes contextually appropriate behavioral initiation during reward seeking and withholding during extinction (Moorman & Aston-Jones, 2015b).

Thus, it is clear that vmPFC is involved in reward seeking and fear processing during both conditioning and extinction. However, it remains unknown how vmPFC encodes avoidance, if single neurons are modulated during both approach and avoidance, and how these correlates change when task-relevant cues become non-reinforced during extinction. To address these issues, we recorded from single neurons in vmPFC while rats performed a combined approach and avoidance task when behavior was reinforced and not reinforced (i.e., extinction). We observed distinct correlates within vmPFC during approach that were response (press; no press) and block (reinforced; extinction) selective but found very few cells that were modulated during avoidance.

Materials and Methods

Animals. Eight male Sprague-Dawley rats were obtained from Charles River Labs (Wilmington, Massachusetts) at 250-300g (7-8 weeks old). Animals were individually-housed in a temperature- and humidity-controlled environment and kept on a 12-h light/dark cycle (0700-1900 in light); all tests were run during the light phase. Animals had access to water *ad libitum* and body weight was maintained at 85% of baseline weight by food restriction (15g standard rat chow provided daily, in

addition to approximately 1g sucrose pellets during experimental trials). All procedures were performed in accordance with National Institutes of Health guidelines and the University of Maryland, College Park Institutional Animal Care and Use Committee (IACUC) protocols.

Combined positive and negative reinforcement behavioral task. Eight rats were run on a combined positive and negative reinforcement behavioral task. We first trained animals on positive reinforcement trials and, once this behavior was learned, proceeded to train them on shock trials. Animals were first trained daily on a 45 min FR1 reward shaping program to establish the lever response reward contingency. Once the reward contingency was learned (~3 days), animals were then trained daily on a 45-min foot shock (0.42mA) escape procedure to establish the response shock termination contingency. Foot shock intensity was selected based on the conditioned foot shock intensity optimization protocol for avoidance behavior outlined in Oleson *et al*.(E. B. Oleson et al., 2012) and previous success in shock avoidance paradigms (Gentry, Lee, & Roesch, 2016b); we used the moderately aversive stimulus strength of 0.42 mV in order to balance aversiveness with response probability for shock trials. During each shock escape training session, subjects were presented with a lever paired with a cue light and an auditory cue; a response on the lever at any point during the session resulted in the retraction of the lever and termination of the cue light and foot shock, as well as progression to the ITI (20 s). Subjects were gradually shaped toward the lever (safe side, quadrant with lever, orientation toward the lever, rearing, pressing) by the experimenter as needed until escape behavior acquisition.

Once subjects acquired consistent escape behavior, trials were altered to allow for shock avoidance; once shock avoidance was introduced, positive reinforcement (as described above) and neutral contingencies were also added to the program. At trial onset, a cue light and one of three discriminatory auditory cues (tone, white noise, or clicker) were activated; house lights remained on at all times. Five seconds after the onset of the auditory cue and cue light, the lever could be pressed in order to produce a response; the 5 s delay was implemented to reduce compulsory pressing and to allow for separate epoch analysis around cue and lever press. Lever pressing after this 5 s delay would produce one of three outcomes (dependent on the auditory cue identity): delivery of a food reward (a sucrose pellet; positive reinforcement behavior), prevention of foot shock (negative reinforcement behavior), or no consequence. If the animal failed to press the lever within a 10 s period, no food reward was delivered, foot shock commenced (2 s duration with automatic termination), or there was no consequence. After response or termination of the trial, an ITI (20 s) was initiated. Auditory cue identities were counterbalanced across rats. Animals were very well trained on this task, completing >30 sessions and displaying >60% avoidance responses for 3 consecutive sessions

Each regular reinforced session consisted of an average of 32.6 ± 5.9 trials per trial type. Once animals were well trained, electrode surgeries were performed and electrophysiological recording began 2 weeks after. During each session, rats were also run on an extinction program after the regular reinforced session, where no consequences occurred whether or not the animal pressed the lever during reward and shock trials. Extinction sessions consisted of an average of 21.5 ± 4.6 trials per trial

type. Combined, each daily session lasted approximately 75 minutes (45 minutes reinforced, 30 minutes extinction). Behavioral sessions in combination with single-unit recordings were run for approximately two months.

Intracranial surgical procedures. All surgical procedures were performed after rats were initially trained on the task. All animals were anesthetized using isoflurane in O_2 (5% induction, 1% maintenance) and each of the eight rats were chronically implanted with a drivable bundle of 10, 25μm diameter FeNiCr (iron, nickel, chromium) wires in the left or right hemisphere in mPFC just dorsal to the infralimbic cortex (+3.0 AP, ±0.6 ML, -4.0 DV from brain). The recording electrode and anchoring screws were stabilized using dental cement (Dentsply), and animals then received post-operative care: subcutaneous injection of 5 mL saline containing 0.04 mL carprofen (Rimadyl), topical application of lidocaine cream to the surgical area, and placement on a heating pad until full consciousness was regained. Animals were also given antibiotic treatment with Cephlexin orally one day prior to surgery and twice daily for a week post-surgery to prevent infection of the surgical site. All subjects were allowed at least a week for full recovery before experimentation.

Data acquisition. Experiments were performed in a plexiglass behavioral chamber (MedAssociates). The behavioral chamber was fitted with shock-grid flooring, with a retractable lever, cue light (above lever portal) and food cup on the left side. Auditory cue sounds were recorded and played back to the rat via an Arduino system.

Electrodes were screened daily to monitor active wires, and the electrode assembly was advanced by 40-80μm per day at the end of the recording session,

which allowed us to record from a different neuronal population each day. Neural activity was recorded using Plexon Multichannel Acquisition Processor systems (Dallas, TX). Signals from the electrode channels were amplified 20 times by an op-amp headstage (Plexon, HST/8o50-G20-GR), located on the electrode array. Immediately outside the chamber, signals were passed through a differential pre-amplifier (Plexon, PBX2/16sp-r-G50/16fp-G50), where the single-unit signals were amplified 50 times and filtered at 150–9000 Hz. The single-unit signals were then sent to the Multichannel Acquisition Processor box, where they were further filtered at 250–8000 Hz, digitized at 40 kHz and amplified at 1–32 times. Waveforms >2.5:1 signal-to-noise were extracted from active channels and recorded to disk. Neurons were sorted using Offline Sorter and Neuroexplorer and exported for further analysis in Matlab (Bissonette, Powell, & Roesch, 2013; Burton et al., 2014).

Experimental design and statistical analysis. Our analysis epochs (cue and baseline) were computed by dividing the total number of spikes by time. The cue epoch consisted of average spikes across time during the 5 s period after cue onset, and the baseline epoch consisted of average spikes across time during a 1 s period 2 s before cue onset. Neurons were characterized by comparing firing rate during baseline to firing rate during the cue epoch, or firing rate during the cue epoch of reward or shock trials to firing rate during the cue epoch of neutral trials, averaged over all trial types (Wilcoxon; $p < 0.05$). We also computed a reward index (reward – neutral) and shock index (shock – neutral) to normalize firing to neutral trials and determine if firing was significantly shifted across the population (Wilcoxon; $p < 0.05$ and to determine correlations between firing on shock and reward trials. Chi-square

tests were performed to assess differences in the counts of neurons showing significant modulation across groups of interest.

Behavior during performance of the task was evaluated by computing percent press and reaction times for each trial type. A two-factor ANOVA ($p < 0.05$) was performed on these behavioral measures to determine if activity was modulated by trial-type (reward, neutral, and shock) and block (reinforced, extinction) or if there were any interactions between these factors.

Histology. Following the completion of the study, animals were terminally anesthetized with an overdose of isoflurane (5%) and transcardially-perfused with saline and buffered 4% paraformaldehyde. Brain tissue was removed and post-fixed with paraformaldehyde at 4°C. Brains were then placed in 30% sucrose solution for 72 hr and sectioned coronally (50µm) using a freezing microtome. Tissues slices were mounted onto slides and stained with thionin for histological reconstruction. Electrode placement was verified under light microscope and drawn onto plates adapted from the rat brain atlas(Paxinos & Watson, 2007).

Results

Behavior during combined approach avoidance

Rats ($n = 8$) were trained on a combined approach and avoidance task (**Fig. 1a-c**). At the start of each session, a lever was extended into the behavioral chamber and remained extended until session completion. At the start of each trial, one of three

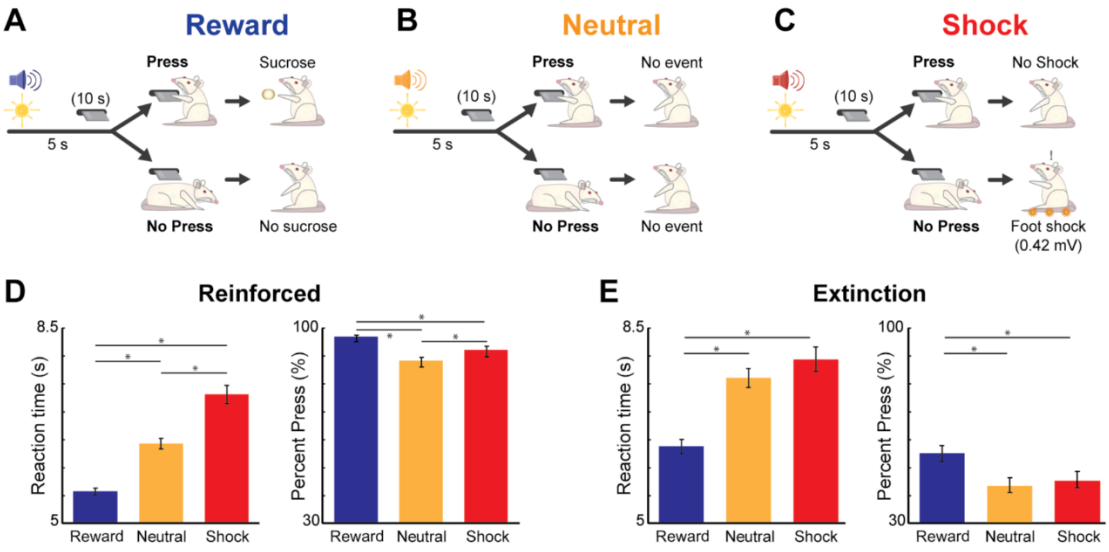


Figure 1: Task design. **A-C**, Sessions consisted of 3 trial types: reward (**A**), neutral (**B**), and shock (**C**), which could be identified by a unique auditory cue. A lever was introduced into the chamber at the start of each session and remained extended for the duration of the session. At the beginning of each trial, rats were presented with a light cue and trial-specific sound cue for 5s and then had a maximum of 10s to press the lever. If rats pressed the lever during this 10s interval, they could receive a sucrose pellet reward, avoid an impending foot shock (0.42 mV), or experience no consequence, depending on the identity of the sound cue. If rats failed to press the lever within 10s, they would alternatively receive no sucrose reward, receive a foot shock (0.42 mV; 2s duration with automatic termination), or experience no consequence depending on the identity of the sound cue. After each consequence, the trial progressed into a 20s ITI. Trial types were pseudo-randomly interleaved within each session (~45 min) and sound cue identity was counterbalanced across rat. During extinction sessions (~30 min), cues produced no outcome regardless of previous association with reward, neutral, or shock. **D-E**, Reaction time and percent press computed across reinforced (**D**) and during extinction (**E**) sessions (n = 84). Bars with asterisks represent significance (T-test; $p < 0.05$). Error bars represent S.E.M.

distinct discriminatory auditory cues and a cue light were presented to the animal for 5 s, signaling whether the current trial would be a reward, shock or neutral trial. After termination of the auditory and visual cues, a lever press produced one of three outcomes (dependent upon auditory cue identity): delivery of a food reward (positive reinforcement behavior, i.e. reward trials), prevention of foot shock (negative reinforcement behavior, i.e. shock trials), or no consequence (i.e. neutral trials). If the animal failed to press the lever within 10 s of cue termination, no food reward was delivered on reward trials, foot shock (0.42mA, 2 s with automatic termination) commenced on shock trials, or there was no consequence on neutral trials. These three trial types were pseudo-randomly interleaved (i.e., random without replacement) within each session. For the remainder of the manuscript, we will refer to this first block of trials as the 'reinforced' trial block. Each reinforced session consisted of an average of 32.6 ± 5.9 (SD) trials per trial type. During sessions where single neurons were recorded, each animal immediately went through extinction (i.e., no shocks or rewards were administered) after completion of the regular reinforced block of trials. Extinction (i.e., non-reinforced) sessions consisted of an average of 21.5 ± 4.6 (SD) trials per trial type. For the remainder of the manuscript we will refer to this second block of trials as the 'extinction' trial block.

Figure 1D-E illustrates behavioral measures across recording sessions (N = 84). During reinforced sessions (**Fig. 1d**), rats produced the most responses and were fastest to respond on reward trials compared with neutral trials (% Press (%P): $t(83) = 7.64$, $p < 0.0001$; Reaction time (RT): $t(83) = 12.29$, $p < 0.0001$); rats were slowest to press for shock trials (Shk vs. Neu RT: $t(83) = 7.17$, $p < 0.0001$; Shk vs. Rew RT:

$t(83) = 11.94$, $p < 0.0001$) but pressed significantly more often during shock trials than during neutral trials (%P: $t(83) = 2.74$, $p < 0.01$). During extinction sessions (**Fig. 1e**), rats were still faster to press and pressed more often during reward trials compared with neutral trials (%P: $t(83) = 7.60$, $p < 0.0001$; RT: $t(83) = 9.24$, $p < 0.0001$), but reaction times and percent press were no longer different between shock and neutral (%P: $t(83) = 1.46$, $p = 0.15$; RT: $t(83) = 1.72$, $p = 0.09$).

To further assess behavior across reinforced and extinction trial blocks, we performed a 2-factor ANOVA with trial-type (reward, neutral, and shock) and block (reinforced and extinction) as factors. Consistent with the above analysis, we found a main effect of trial-type, (RT: $F_{2, 498} = 62.4$, $p < 0.05$; %P: $F_{2, 498} = 10.55$, $p < 0.05$). We also found a main effect of trial block, demonstrating that rats were slower (RT: $F_{1, 498} = 47.53$, $p < 0.05$) and pressed less often (%P: $F_{1, 498} = 579.97$, $p < 0.05$) during extinction compared to reinforced trials blocks. There were no interactions between trial-type and extinction for either RT or percent lever press, indicating that the pattern of behavior observed on reward, neutral, and shock trials was similar across trial blocks (RT: $F_{2, 498} = 1.76$, $p = 0.17$; %P: $F_{2, 498} = 0.58$, $p = 0.56$). Overall, these behavioral results suggest that rats can discriminate between the three trial types during both reinforced sessions and extinction, and, as expected, behavior declined during extinction when outcomes were omitted.

Activity in vmPFC was strongly and weakly modulated during reward and shock trials, respectively

To further understand the role of vmPFC in our combined approach and avoidance task, we recorded from a total of 289 neurons within the ventromedial prefrontal cortex (vmPFC) of rats (N = 6 rats). In our initial analysis, we broadly determined if neurons increased (i.e., increasing-type cell) or decreased (i.e., decreasing-type cell) firing during the cue epoch (5 s after cue onset) compared to baseline (1 s epoch taken 2 s before cue onset; Wilcoxon; $p < 0.05$). We found that 60 (22.5%; $\chi 2 = 495.0$, $p < 0.05$) and 15 (5.6%; $\chi 2 = 13.6$, $p < 0.05$) neurons significantly increased or decreased firing rate during the cue epoch, respectively. **Figure 2a-b** illustrates the average firing rate of increasing and decreasing-type neurons over trial-time, broken down by trial type (blue = reward, yellow = neutral, red = shock). In both populations of cells, there were clear differences in firing rate between reward and neutral trials (as defined) but no difference between shock and neutral trials.

To quantify these results, we computed a reward (reward minus neutral) and shock (shock minus neutral) index for each cell by subtracting average firing rates during the cue epoch on neutrals press trials from reward-press and shock-press (i.e., avoid) trials. These indices are plotted against each other in **Fig. 2c**. We found no correlation between firing rates for cells that were modulated by reward and shock trials compared with neutral (**Fig. 2c**; $r^2 = 0.02$, $p = 0.26$), indicating that the same cells were not significantly modulated by both. For increasing-type neurons, the reward index was significantly shifted in the positive direction (Wilcoxon; Z = 2.22;

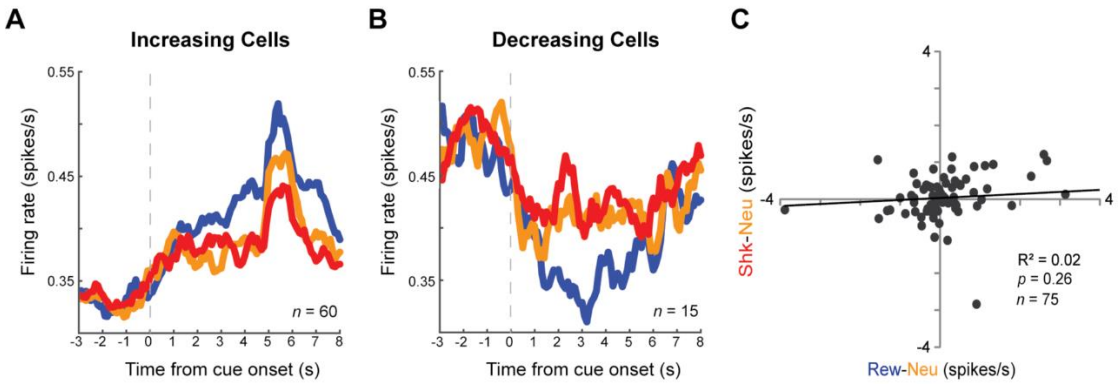
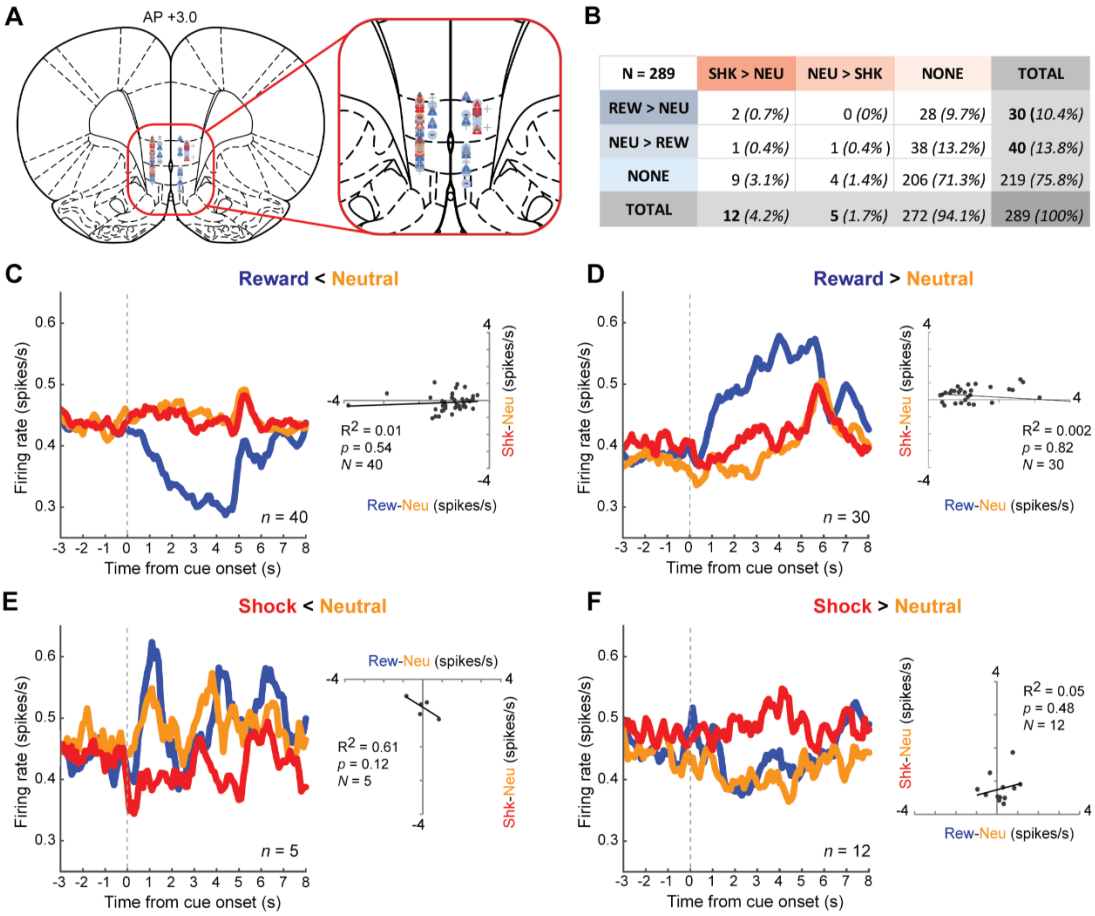


Figure 2: Increasing- and decreasing- type cells in vmPFC were modulated by cues that predict reward. **A, B,** Histograms depicting average firing rate (spikes/second) for cells increasing (n = 60) or decreasing (n = 15) within the overall population (N = 289 cells) across trial time for reward (blue), neutral (orange), and shock (red) trial types. Cue onset is depicted with a gray dashed line aligned to Time = 0. **C**, Scatter plot depicting combined increasing and decreasing cells (n = 75) along computed reward (reward minus neutral; X-axis) and shock (shock minus neutral; Y-axis) indices for each cell. Indices were calculated by subtracting average firing rates during the cue epoch on neutral press trials from reward-press and shock-press (i.e., avoid) trials.

μ = 0.43; $p < 0.05$), while the reward index for decreasing-type neurons was shifted in the negative direction (Wilcoxon; μ = -0.63; $p < 0.05$). Thus, both increasing and decreasing populations were significantly modulated by reward expectation. This was not true, however, for cues that predicted shock; the distribution of shock indices was not significantly shifted from zero in either increasing or decreasing populations (Wilcoxons; Increasing: Z = 0.86, μ = 0.04, $p = 0.39$; Decreasing: μ = 0.09, $p = 0.64$). Further, there was no correlation between reward and shock indices individually for increasing ($r^2 = 0.013$, $p = 0.144$) or decreasing ($r^2 = 0.157$, $p = 0.391$) cells when analyzed separately, nor when combined ($r^2 = 0.017$; $p = 0.259$). Thus, we conclude that average firing rates in vmPFC were modulated by cues that predict reward but not shock and, furthermore, that these reward effects were not linked to parallel signals reflecting the value or the motivational level associated with avoiding shock.

From the analysis above, it appears that firing in vmPFC was significantly modulated by reward expectancy during the cue epoch, with little to no modulation during shock trials when cells were divided into increasing- and decreasing-type neurons. However, it is possible that, by broadly dividing neurons in this way, we overlooked neurons that were shock-selective independent from modulation on reward or neutral trial-types. To address this issue, we asked which neurons within the vmPFC were modulated more or less on reward and shock trials compared with neutral trials. **Figure 3a** shows the distribution of recording locations for these cells within vmPFC and **Table 3b** further quantifies the breakdown of these cells, showing counts and percentages of cells that fired significantly differently between reward and shock trials relative to neutral (Wilcoxon; $p < 0.05$) indicating how many and what



N = 289	SHK > NEU	NEU > SHK	NONE	TOTAL
REW > NEU	2 (0.7%)	0 (0%)	28 (9.7%)	**30** (10.4%)
NEU > REW	1 (0.4%)	1 (0.4%)	38 (13.2%)	**40** (13.8%)
NONE	9 (3.1%)	4 (1.4%)	206 (71.3%)	219 (75.8%)
TOTAL	**12** (4.2%)	**5** (1.7%)	272 (94.1%)	289 (100%)

Figure 3: Neurons in vmPFC are strongly and weakly modulated during reward and shock cues, respectively, and very few are modulated by both. **A**, Location of recording sites based on histology (Paxinos & Watson, 2007). Each symbol represents the location of neurons that showed differential firing (Wilcoxon; $p < 0.05$) in the analyses described in the text (See Results) and shown in the table in **B**. Dark blue = Reward > Neutral; Light blue = Neutral > Reward; Dark red = Shock > Neutral; Light red = Neutral > Shock; '-' = decreasing-cells; '+' = increasing cells. **B**, Table quantifying numbers and percentages of cells that were Reward > Neutral, Neutral > Reward, Shock > Neutral, Neutral > Reward, or none of the above. **C**-**F**, Histograms depicting average firing rate (spikes/second) for cells where Reward < Neutral (n = 40; **C**), Reward > Neutral (n = 30; **D**), Shock < Neutral (n = 5; **E**), and Shock > Neutral (n = 12; **F**) within the overall population (N = 289 cells) across trial time for reward (blue), neutral (orange), and shock (red) trial types. Cue onset is depicted with a gray dashed line aligned to Time = 0. Insets show scatter plots depicting each cell within each sub-population (Reward < Neutral, Reward > Neutral, Shock > Neutral, Shock > Neutral) along computed reward (reward minus neutral; X-axis) and shock (shock minus neutral; Y-axis) indices. Indices were calculated by subtracting average firing rates during the cue epoch on neutral press trials from reward-press and shock-press (i.e., avoid) trials.

percentage of individual cells increased or decreased to both shock and reward, were modulated only by shock or only by reward, or were not significantly modulated by either.

From the total population of neurons (n = 289), 70 (24.2%) cells fired significantly differently to cues on reward compared to neutral trials ($\chi2 = 224.39$, $p < 0.0001$), whereas only 17 (5.9%) cells fired significantly differently to cues on shock trials, which was not significantly more than expected by chance alone ($\chi2 = 0.46$, $p = 0.49$). Further, the frequency of neurons modulated during reward trials significantly outnumbered those modulated during shock trials ($\chi2 = 27.03$, $p < 0.0001$).

The average firing rate for reward- and shock modulated neurons is plotted across trial-time in **Figures 3c-f**, along with inset correlations between the reward and shock indices (reward – neutral, shock – neutral) for these cells. Although, by definition, reward-modulated cells exhibited differential firing rates on reward versus neutral trials, they did not also show differential firing during shock trials; the same is true for shock-modulated neurons, showing differential firing rates on shock versus neutral trials but no difference during reward trials. Of the 40 (13.8%) cells that fired significantly less on reward trials relative to neutral, only 2 of these cells were also modulated by shock cues ($\chi2 = 0.001$, $p = 1.0$), and of the 30 (10.4%) that fired significantly more on reward trials relative to neutral, again only 2 cells were also modulated by shock cues ($\chi2 = 0.14$, $p = 0.68$). Correlation insets show no correlation between reward and shock indices for either neutral greater than reward **(Fig 3c**; $r^2 = 0.01$, $p = 0.54$) or reward greater than neutral (**Fig. 3d**; $r^2 = 0.002$, $p = 0.82$) cells. Of the 5 (1.9%) cells that fired less on shock trials relative to neutral, only 1 of these

cells was also modulated by reward cues ($\chi 2 = 2.06$, $p = 0.12$), and of the 12 (4.2%) cells that tired significantly more on shock trials relative to neutral, 3 cells were also modulated by reward ($\chi 2 = 9.69$, $p < 0.01$). Correlation insets show no correlation between reward and shock indices for either neutral greater than shock (**Fig. 3e**; $r^2 = 0.61$, $p = 0.12$) or shock greater than neutral (**Fig. 3f**; $r^2 = 0.05$, $p = 048$) shock-modulated cells. Overall, only 1.4% of all neurons were modulated on both reward and shock trials. These single unit results are consistent with our population findings in **Figure 2c**, showing no correlation between reward and shock indices. Taken together, these results demonstrate that neurons in vmPFC are strongly and weakly modulated during reward and shock cues, respectively, and that very few neurons were modulated during both reward-seeking and shock-avoidance behaviors.

Neurons selective for outcome during conditioning became non-selective during extinction and vice versa

Ventromedial prefrontal cortex is thought to play a key role in extinction (Burgos-Robles et al., 2007; Giustino et al., 2016; Hefner et al., 2008; Holmes et al., 2012; Maroun et al., 2012; Milad & Quirk, 2002; Wilber et al., 2011). To determine how neural signals encoding the promise of reward and the threat of shock were modulated when outcomes were no longer delivered, we next determined how many neurons that were selective during reinforced trials blocks became non-selective during extinction, and vice versa. Since extinction sessions naturally have fewer press trials, we only examined sessions where there were at least 2 press trials per trial-type during both reinforced and extinction sessions ($n = 241$) and asked whether neurons that were selective for press during reinforced trials (i.e., reward > neutral, neutral >

reward, shock > neutral, neutral > shock) were also selective for press trials during extinction. Due to the low overall number of neurons that were significantly modulated during shock trials, the following figures are restricted to cells that were modulated by reward, though data and statistics for cells modulated by shock trials will still be reported in the text using parallel analyses.

Figures 4a-b illustrate the average firing rate over trial-time of an extinction-matched subgroup of neurons that showed significantly different firing rates to reward relative to neutral trials during reinforced sessions. This group of neurons did not exhibit differential firing between reward and neutral trial-types during extinction (**Fig 4c-d**). Of the 32 (13.3%; $\chi2 = 114.4$, $p < 0.05$) and 25 (10.4%; $\chi2 = 61.0$, $p < 0.05$) neurons that showed lower (**Fig. 4a**) and higher **(Fig. 4b**) firing on reward trials during reinforced sessions, respectively, only 4 ($\chi2 = 3.63$, $p = 0.05$) and 2 ($\chi2 = 0.41$, $p = 0.49$) of these cells were also selective during extinction. Similarly, of the 5 (2.1%; $\chi2 = 0.16$, $p = 0.69$) and 8 (3.3%; $\chi2 = 0.63$, $p = 0.43$) neurons that showed lower and higher firing on shock trials during conditioning, respectively, none of these cells were selective during extinction. Thus, we conclude that neurons selective during reinforced sessions were not also selective during extinction when outcomes were omitted.

Interestingly, another sub-population of neurons that were not selective during conditioning became selective during extinction, showing significantly different firing rates to reward relative to neutral during extinction (**Fig. 4e-h**). There were 24 (10%) neurons in total that were significantly modulated by expected reward during extinction ($\chi2 = 12.4$, $p < 0.06$). Of these, 18 (7.5%; $\chi2 = 24.2$, $p < 0.05$) and 6 (2.5%;

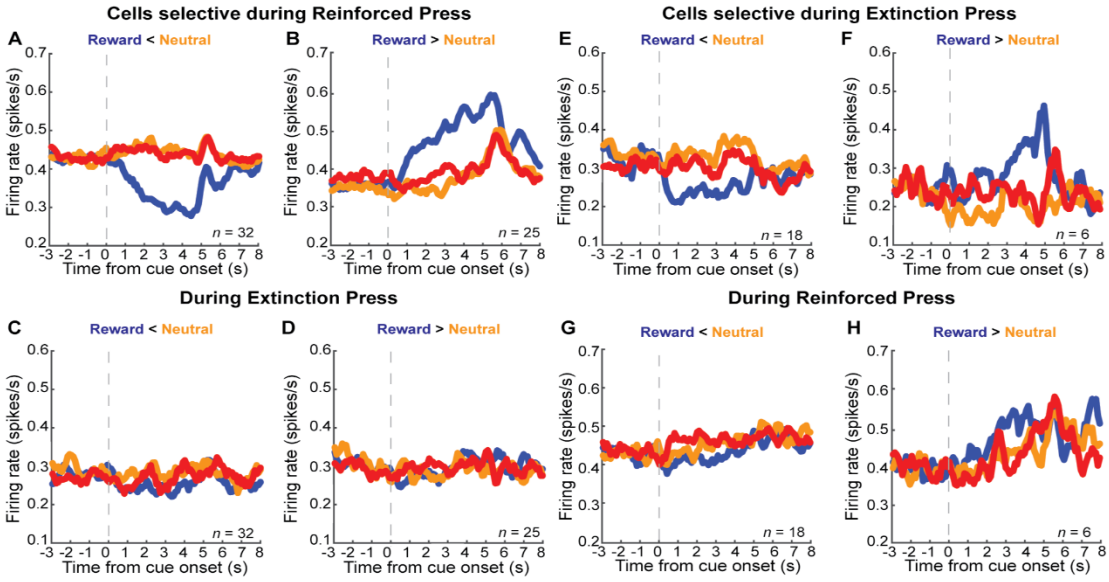


Figure 4: Cells were selective for outcome during either conditioning or extinction, not during both contexts. **A-D**, Histograms depicting average firing rate (spikes/second) for Reward < Neutral (n = 32; **A, C**) and Reward > Neutral (n = 25; **B, D**) cells that are modulated when rats press the lever during Reinforced cues (**A, B**) but not during Extinction (**C, D**) for reward (blue), neutral (orange), and shock (red) trial types. **E-H**, Histograms depicting average firing rate (spikes/second) for Reward < Neutral (n = 18; **E, G**) and Reward > Neutral (n = 6; **F, H**) cells that are modulated when rats press the lever during Extinction cues (**E, F**) but not during Reinforced cues (**G, H**) for reward (blue), neutral (orange), and shock (red) trial types. Cue onset is depicted with a gray dashed line aligned to Time = 0. Cells are drawn from the total population and were behaviorally matched across Reinforced and Extinction (N = 241).

$\chi 2 > 0.001$, $p = 0.99$) neurons exhibited lower (**Fig. 4e**) and higher (**Fig. 4f**) firing on reward trials during extinction, respectively. Out of these 24 neurons that were reward-selective during extinction, only 6 were also selective during conditioning, which is not significantly greater than chance alone (6 out of 241; 2.49%; $\chi 2 < 0.001$, $p = 0.99$). Similarly, of the 5 (2.1%; $\chi 2 = 0.16$, $p = 0.69$) and 6 (2.5%; $\chi 2 < 0.001$, $p = 0.99$) neurons that showed lower and higher firing on shock trials during extinction, respectively, none of these cells were also selective during the reinforced trial block. Thus, we conclude that neurons selective during extinction were not selective during trials when outcomes were present.

Outcome selectivity was response selective during extinction

In the previous section, we show that a sub-population of neurons were outcome selective during either reinforced trials or extinction but not in both. Next, we wanted to test if this outcome selectivity seen during extinction was dependent upon the behavioral response (i.e., if the rat pressed or failed to press). For these analyses, we focused on extinction trials from sessions with an adequate number of press and non-press trials for both conditions (i.e., we excluded sessions with too few non-press trials during reinforced trials), and we only included sessions were there were at least 2 press and non-press trials for each trial-type ($n = 244$).

Indeed, we found that outcome selectivity during extinction was also response-dependent. **Figures 5a-d** demonstrate average firing activity over trial-time for neurons that fired significantly less (**Fig. 5a**) or more (**Fig. 5b**) on reward press trials compared with neutral press trials during extinction. This same population of neurons did not, however, show outcome selectivity on non-press trials during

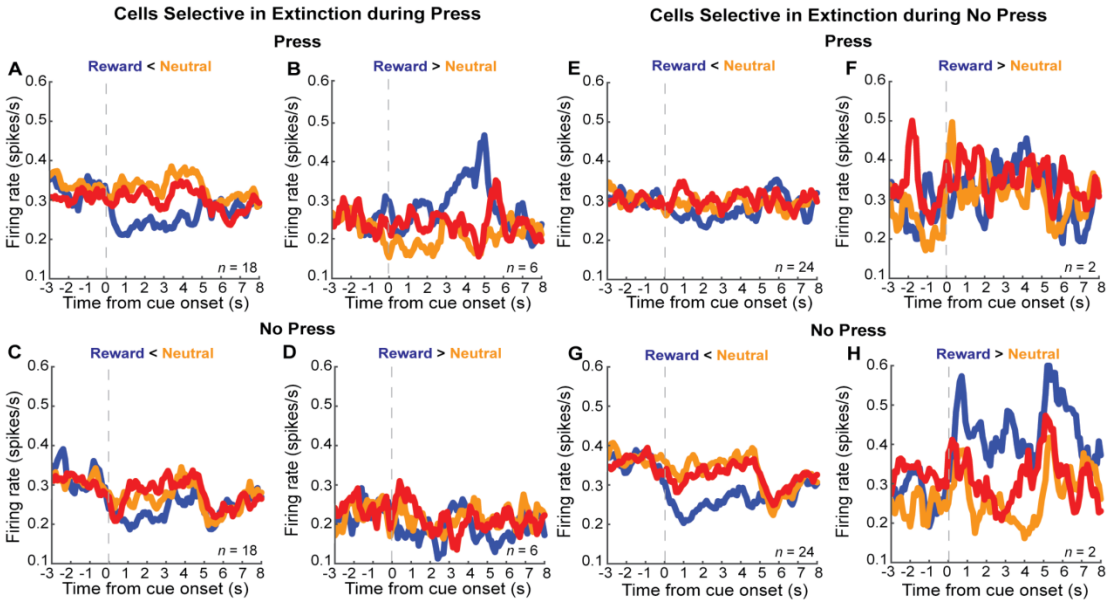


Figure 5: Outcome selectivity during extinction was also response selective. **A-D**, Histograms depicting average firing rate (spikes/second) for Reward < Neutral (*n* = 18; **A, C**) and Reward > Neutral (*n* = 6; **B, D**) cells that are modulated when rats press the lever during Extinction (**A, B**) but not when they fail to press (**C, D**) for reward (blue), neutral (orange), and shock (red) trial types. **E-H**, Histograms depicting average firing rate (spikes/second) for Reward < Neutral (*n* = 24; **E, G**) and Reward > Neutral (*n* = 2; **F, H**) cells that are modulated when rats fail to press the lever during Extinction cues (**E, F**) but not when they press (**G, H**) for reward (blue), neutral (orange), and shock (red) trial types. Cue onset is depicted with a gray dashed line aligned to Time = 0. Cells are drawn from the total population and were behaviorally matched across Reinforced and Extinction (*N* = 244).

extinction. A total of 24 neurons exhibited significantly different firing on reward versus neutrals trials (9.8%; $\chi2 = 11.9.6$, $p < 0.05$) during extinction press trials. Of these, 18 (7.4%; $\chi2 = 23.6$, $p < 0.05$) and 6 (2.5%; $\chi2 < 0.001$, $p = 0.98$) neurons exhibited significantly lower (**Fig. 5a**) and higher (**Fig. 5b**) firing on reward press trials compared to neutral press trials in extinction. Of the 24 neurons selective for reward on press trials, only 4 neurons were also significantly selective during no press trials in extinction, which is significantly fewer than expected by chance alone (4 out of 244; 1.6%; $\chi2 = 5.7$, $p < 0.05$). Thus, during extinction, neurons that were reward selective on press trials were not also selective on no-press trials.

Interestingly, other cells showed the opposite pattern. **Figure 5e-h** illustrates the average firing activity over trial-time for neurons that exhibited significantly less (**Fig. 5e**) or more (**Fig. 5f**) firing on reward non-press trials compared with press trials during extinction. This same population of neurons did not, however, show outcome selectivity on press trials during extinction. A total of 26 neurons showed significantly different firing on reward versus neutrals trials (10.7%; $\chi2 = 16.3$, $p < 0.05$) during extinction non-press trials. Of these, 24 (9.8%; $\chi2 = 53.6$, $p < 0.05$) and 2 (1%; $\chi2 = 8.89$, $p = 0.10$) neurons exhibited significantly lower (**Fig. 5g**) and higher (**Fig. 5h**) firing on reward no-press trials compared with neutral. Of the 26 neurons that showed significant reward modulation on no-press trials, only 4 also showed selectivity during press trials (1.6%; $\chi2 = 0.71$, $p = 0.40$). Thus, during extinction, neurons that were reward selective on no-press trials were not also selective on press trials.

Overall, these results suggest that vmPFC neurons were both outcome and response selective, in that sub-populations of vmPFC neurons showed differential firing on reward or shock trials relative to neutral on either press or non-press trials in either conditioning or extinction sessions, but not in opposing contexts.

Chapter discussion

While vmPFC activity is often associated with fear attenuation and extinction learning, little is known about how it processes complex environments that present opportunities for both punishment and reward. Historically, most vmPFC studies have used fear extinction paradigms to measure activity in an aversive context, but to date no one has measured activity during punishment avoidance. Here, we recorded from neurons within vmPFC while rats performed a cued combined approach and avoidance task and then during extinction of cues. We found that neurons within the vmPFC were both outcome and response selective, in that sub-populations of vmPFC neurons fired differently on reward or shock trials relative to neutral on either press or non-press trials in either conditioning or extinction sessions, but not in conflicting conditions. This effect was more robust for reward trials than shock trials.

Most strikingly, firing rates were significantly modulated in response to cues signaling subsequent reward approach, consistent with many emerging studies reporting a role for vmPFC in reward seeking behavior. One recent electrophysiology study showed that, in addition to extinction, single units in vmPFC are modulated during cue-evoked approach responses in a simple discriminative stimulus sucrose-seeking task (Moorman & Aston-Jones, 2015b). Interestingly, recent evidence also demonstrates the existence of specific inhibitory projections from CeA to vmPFC that

may influence reward-related behaviors, as opposed to fear. Seo et al. showed that specifically activating a subset of GABAergic neurons projecting from CeA to vmPFC in mice amplified external reward valuation, increasing nose poke behavior for sucrose reward in an operant conditioning paradigm, while producing no effect on internal motivation or value states or overall reward consumption (Seo et al., 2016). Our data also generally suggest that vmPFC activity reflects external reward approach, not the value or motivational drive of cues.

Though an abundance of literature emphasizes the role of vmPFC in the suppression of amygdala-driven fear responses, suggesting it may be a critical player in inhibiting freezing and allowing behaviors that lead to avoidance, we saw little vmPFC modulation during cues predicting successful shock avoidance. This result was surprising to us, since previous studies have shown that vmPFC lesion or inactivation disrupts discrimination between fear and safety cues and, hence, disturbs avoidance behavior (Adhikari et al., 2015; Sangha et al., 2014b). Still, our data do not necessarily contradict current literature regarding the role of vmPFC in fear suppression. Fear conditioning has been shown to reduce vmPFC excitability and low vmPFC activity may contribute to the encoding of contextual fear, while extinction of fear has been shown to increase vmPFC firing rates that were previously low during fearful cues (Cruz, López, & Porter, 2014; Giustino et al., 2016; Soler-Cedeño et al., 2016). However, these studies did not examine the role of vmPFC during active avoidance and, to date, few have measured vmPFC activity directly in this type of behavioral paradigm. Emerging perspectives are beginning to emphasize the need to study these regions in a broader context than fear conditioning, by using more

naturalistic approach and avoidance paradigms (Bravo-Rivera, Roman-Ortiz, Brignoni-Perez, Sotres-Bayon, & Quirk, 2014; Christian Bravo-Rivera, Roman-Ortiz, Montesinos-Cartagena, & Quirk, 2015; Delgado et al., 2016).

One recent study by Schwartz, Miller and Fields found that vmPFC to NAc connections are recruited when animals are required to make choices involving conflicting actions to promote reward-directed behavior; they concluded that activation of this pathway drives the animal to choose the action with the most rewarding outcome while simultaneously inhibiting actions that may interfere with this choice (Schwartz et al., 2017). In this study, it was necessary for animals to learn to suppress the drive to avoid a risky pain-predictive cue in order to gain a reward. By using designer receptors exclusively activated by designer drugs (DREADDs) and micro-injections of a GABA agonist, Schwartz et al. were able to temporarily inactivate infralimbic cortex (IL) during performance of their approach-avoidance task and reinstate avoidance of the pain predictive cue in rats. They concluded that, after learning, IL function is needed to overcome the drive to avoid punishment in order to gain valuable food reward, which differs from fear-conditioning studies showing that IL activity is only critical during training. Thus, it is likely that the role of vmFPC is more complex than proposed thus far by fear conditioning and extinction studies, and theories may need to be revised to account for data from avoidance and approach studies.

There are several possible explanations for why we saw little vmFPC modulation by shock in our task. First, our animals were very highly trained on our combined approach and avoidance task, having completed over 30 sessions before

vmPFC recordings were collected for an additional two months. Though the literature is conflicted regarding the role of vmFPC during and after learning, it might be argued that animals in our study were no longer using predicted outcomes (i.e., promise of reward and threat of foot shock) to guide behavior but were responding habitually. It is possible, then, that responding on avoidance trials could have initially been governed by vmFPC but control had been transferred to more habit-oriented regions, like dorsal striatum, by the time of recording. Though this is consistent with our behavioral findings that rats responded at a high rate for all trial types, it seems unlikely that this is the case, since we would also expect reward responses to become habitual over time. In contrast, we still saw significant modulation to reward cues in vmPFC.

Another possibility is that the vmPFC may be more important during early avoidance learning; others have shown that sub-regions of vmPFC display CS-evoked responses during early and late extinction, but these responses decrease in magnitude with training (Chang, Berke, & Maren, 2010). Alternatively, it is possible that when rats are well-trained for shock avoidance, as in our study, cues that predict successful avoidance with certainty may no longer elicit a fear response but are instead interpreted as safety cues; as stated previously, vmPFC has been shown to be important in discriminating between safety and fear cues (Sangha et al., 2014b). Additionally, it is possible that cues predicting reward and neutral outcomes may also be interpreted as safety cues during late training, since shock is not possible on these trial types. However, we found that rats were slower to respond during avoidance

trials compared with reward or neutral trials, suggesting that they were indeed still behaviorally impacted by the potential for shock.

We also saw context-dependent firing related to block type (reinforced or extinction) and response type (press or non-press). This finding is consistent with recent work by Moorman et al. showing context-dependent firing in vmPFC to optimize behavioral output for reward-seeking and extinction contexts; for instance, neurons fired more strongly for reward approach in reinforced contexts but also fired more during behavioral inhibition in extinction (Moorman & Aston-Jones, 2015b). This study is one of many implicating the vmFPC in extinction and context-dependent behavioral control (Burgos-Robles et al., 2007; Camp et al., 2009; Hefner et al., 2008; Holmes et al., 2012; Milad & Quirk, 2002; Wilber et al., 2011). Many studies have also hypothesized that separate neural ensembles within the vmPFC encode reward-seeking and extinction, and there is likely some intermingling within the same cortical region, since inactivation of food-seeking and extinction-related ensembles decreased and increased food seeking, respectively(Warren et al., 2016). These results are in line with our current findings, in that separate individual neurons were modulated by either press or no press and either reinforcement or extinction contexts, for both reward and shock, within the same region of vmPFC.

The data presented here provide evidence that the vmPFC is involved in cue-driven, reward-guided behavioral optimization. This finding is of great interest, since this region of the prefrontal cortex has commonly been linked to fear extinction and the processing of aversive outcomes. Here, we found distinct correlates within the vmPFC for reward-modulated cues that were response (press; no press) and block

(reinforced; extinction) specific. Surprisingly, we found little vmPFC modulation related to shock avoidance cues in our task. This work provides new insights into the neurobiological underpinnings of approach and avoidance behaviors and extinction learning.

Chapter 3: Phasic dopamine release in the rat nucleus accumbens predicts approach and avoidance performance

Published manuscript

Nature Communications. 2016 Oct 27; 7:13154. doi: 10.1038/ncomms13154.

Authors: Ronny N. Gentry[1,2], Brian Lee[1], and Matthew R. Roesch[1,2]

1. Department of Psychology

2. Program in Neuroscience and Cognitive Science
1147 Biology-Psychology Building
University of Maryland
College Park, MD 20742

Abstract

Dopamine (DA) is critical for reward processing, but significantly less is known about its role in punishment avoidance. Using a combined approach-avoidance task, we measured phasic DA release in the nucleus accumbens (NAc) of rats during presentation of cues that predicted reward, punishment or neutral outcomes and investigated individual differences based on avoidance performance. Here we show that DA release within a single microenvironment is higher for reward and avoidance cues compared with neutral cues and positively correlated with poor avoidance behavior. We found that DA release delineates trial-type during sessions with good avoidance but is non-selective during poor avoidance, with high release correlating with poor performance. These data demonstrate that phasic DA is released during cued approach and avoidance within the same microenvironment and abnormal processing of value signals is correlated with poor performance.

Introduction

While a breadth of literature has examined the role of phasic dopamine (DA) release within the context of unexpected rewards and the cues that come to predict them (Schultz, Dayan, & Montague, 1997; Wise, 2004), fewer studies have explored the function of DA signaling in aversive situations. Both reward-seeking and punishment-avoidance paradigms promote instrumental responding (Bromberg-Martin et al., 2010; Dayan, 2012; Ikemoto & Panksepp, 1999), but these behaviors are differentially governed by positive and negative reinforcement learning strategies, respectively. It is still unknown how conditioned stimuli promote avoidance behaviors, how these behaviors are modified by DA release, or if these effects are analogous to those seen during appetitive tasks. These questions have spurred discussion regarding the further heterogeneity of the dopamine response and a recent surge of models aiming to describe negative reinforcement using DA-like prediction error signaling (Bromberg-Martin et al., 2010; Dayan, 2012; Dayan & Berridge, 2014; Dayan & Niv, 2008; Oleson et al., 2012; Wolfram Schultz, 2000). To date, these issues have not been adequately addressed, largely because few studies have examined DA signals in the context of both positive and negative reinforcement.

Since the mechanisms governing punishment avoidance have been studied considerably less than those of reward seeking, the circuit underlying avoidance behavior remains poorly understood. The behavioral processes that guide punishment avoidance are complex, involving both an initial Pavlovian response and a secondary instrumental component (Bolles, 1970; Morris, 1974; Moscarello & LeDoux, 2013;

Rescorla & Solomon, 1967). Cues that predict the possibility of shock also produce fear, often leading to freezing behavior and inaction that reduces the likelihood of avoidance (Bolles, 1969, 1970; Kumar, Bhat, & Kumar, 2013). The transition from freezing to successfully pressing a lever or shuttling to avoid foot shock requires overcoming this initial fear response in order to initiate action (Bolles, 1969; van Meurs, Wiggert, Wicker, & Lissek, 2014). This is very different from behavior driven by reward; cues that predict reward generally arouse animals, promoting action and increasing the probability of responding (Bromberg-Martin et al., 2010; Ikemoto & Panksepp, 1999; Salamone & Correa, 2002; Schultz et al., 1997).

Given these complications, it is no surprise that punishment avoidance tasks are generally more difficult to learn than reward-seeking tasks. This distinction is greatly influenced by the mode of response (nose poke, lever press, shuttle response) employed within the task, as well as whether this behavioral response is in conflict or concert with the underlying Pavlovian response. Though the majority of animals are able to learn to avoid a noxious stimulus, many fail to perform at high levels even after training (Bolles & Popp, 1964; Brush, 2003; Martinez et al., 2013). Most studies exclude poor avoiders from analyses due to difficulty in determining whether these animals are suffering from a learning or performance deficiency. This is unfortunate, since these individuals may provide insight into specific neural impairments present in psychiatric disorders involving negative reinforcement deficits, such as addiction and anxiety disorders; these connections will be discussed in further detail in the "Future Directions" section of *Chapter V: General Discussion* (Lissek & van Meurs, 2014; Nestler, 2005). Studies that have examined this subgroup suggest that the

breakdown in behavior does not reflect a learning deficit, but rather one of performance (Martinez et al., 2013); these studies reveal extensive freezing during conditioned stimuli that predict shock, which reduces the likelihood the animal will react to avoid punishment.

One way to overcome fear associated with potential shock is to adopt a habitual responding pattern, driven by stimulus-response associations instead of the anticipated negative outcome. This strategy could increase successful avoidance performance during tasks that involve punishment. Indeed, it has been suggested that stress can prompt a transition from goal-directed to habitual responding; specifically, it has been shown that stress makes instrumental responding insensitive to changes in reinforcement value and reduces explicit knowledge of action-outcome contingencies (Dias-Ferreira et al., 2009; Schwabe & Wolf, 2009, 2011; Soares et al., 2012; Taylor et al., 2014).

Recent work has begun to address these issues by recording DA release during avoidance-only procedures (Oleson et al., 2012); however, it is still unclear whether DA correlates seen during avoidance behavior are similar to those observed during appetitive scenarios. Further, very few studies have examined differences between good and poor avoiders to determine how behavior and its neural underpinnings vary among individuals (Brush, 2003; Martinez et al., 2013). This information could help explain why some individuals are able to overcome anxiety in stressful situations, while others are not. Here, in order to address these concerns, we recorded sub-second DA release within the nucleus accumbens core (NAc) using fast-scan cyclic voltammetry as rats performed a combined positive and negative reinforcement

procedure. We show that DA release delineates trial type and is higher for both reward and avoidance cues compared with neutral cues only during good avoidance performance, while indiscriminately high DA release is correlated with poor avoidance performance. These results suggest that reward approach and punishment avoidance is signaled within the same microenvironment of the NAc and abnormal processing of these cues may disrupt successful avoidance.

Materials and Methods

Animals. Sixteen male Sprague-Dawley rats were obtained from Charles River Labs at 300-350g (90-120 days old). Animals were individually-housed in a temperature- and humidity-controlled environment and kept on a 12-h light/dark cycle (0700-1900 in light); all tests were run during the light phase. Animals had access to water *ad libitum* and body weight was maintained at 85% of baseline weight by food restriction (15g standard rat chow provided daily, in addition to approximately 1g sucrose pellets during experimental trials). Of the 16 animals entering the study, 10 animals provided reliable cyclic voltammograms. All procedures were performed in concordance with the University of Maryland, College Park Institutional Animal Care and Use Committee (IACUC) protocols.

Chronic microelectrode fabrication. Electrodes were constructed according to the methods of Clark *et al*. (Clark et al., 2010). A single carbon fiber (Goodfellow Corporation) was inserted into a 15 mm cut segment of fused silica (Polymicro Technologies) while submerged in isopropyl alcohol. One end of the silica tubing was sealed with a two-part epoxy (T-QS12 Epoxy, Super Glue) and left to dry overnight,

leaving untouched carbon fiber extending past the seal. The protruding carbon fiber was cut to a length of 150 μm. A silver connector (Newark) was secured to the carbon fiber at the opposing end of the silica tubing using silver epoxy (MG Chemicals) and was allowed to dry. A final coat of two-part epoxy was then applied to the pin connection to provide insulation and structural support for the electrode and was allowed to dry overnight.

Intra-cranial surgical procedures. All animals were anesthetized using isoflurane in O_2 (5% induction, 1% maintenance) and implanted with a chronic voltammetry microelectrode aimed at the NAc core (+1.3 AP, +1.4 ML, -6.9 DV), an ipsilateral bipolar stimulating electrode (Plastics One) in the medial forebrain bundle (-2.8 AP, +1.7 ML, -8.8 DV), and a contralateral Ag/AgCl reference electrode (Sigma-Aldrich). The reference electrode and anchoring screws were stabilized using a thin layer of dental cement (Dentsply), leaving the holes for the stimulating and recording electrodes unobstructed. The stimulating and recording electrodes were attached to a constant current isolator (A-M Systems) and voltammetric amplifier, respectively, and lowered to the most dorsal point of the target region (-6.6 DV for the working electrode and -8.5 DV for the stimulating electrode). At this depth, a triangular voltammetric input waveform (-0.4 to +1.3 V vs. Ag/AgCl, 400 V/s; Heien et al., 2003) was applied to the recording electrode at 60 Hz for 30 minutes and then reduced to 10 Hz for the remainder of the surgery. Electrical stimulation (24 biphasic pulses, 60 Hz, 120 μA) was applied to the stimulating electrode in order to evoke dopamine release, which was monitored at increasing depths by the recording electrode. If neither an evoked change in DA nor a physical response (whisker

movement or blinking) was observed, the stimulating electrode was lowered by 0.05mm until a response was achieved or to a maximum depth of 8.8mm. The working electrode was then lowered by 0.05mm until DA release was observed or to a maximum depth of 6.9mm. Once electrically-evoked DA release was detected in the NAc core, a thin layer of dental cement was used to secure the stimulating and recording electrodes in place. A Ginder implant (Ginder Scientific; constructed in house) was connected to the reference, stimulating, and recording electrodes and fully insulated using dental cement, leaving only the screw-top connector exposed, in order to reduce noise and prevent loss of connectivity during behavioral training. Animals then received post-operative care: subcutaneous injection of 5 mL saline containing 0.04 mL carprofen (Rimadyl), topical application of lidocaine cream to the surgical area, and placement on a heating pad until full consciousness was regained. Animals were also given antibiotic treatment with Cephlexin orally twice daily post-surgery for two weeks to prevent infection of the surgical site. All subjects were allowed a month for full recovery and stabilization of the electrode before experimentation.

Combined positive and negative reinforcement behavioral task. Animals were first trained daily on a 45 min foot shock (0.42 mA) escape procedure to establish the response-shock termination contingency. Foot shock intensity was selected based on the conditioned foot shock intensity optimization protocol for avoidance behavior outlined in Oleson *et al*. (Oleson et al., 2012). For behavioral sessions accompanied with FSCV recording, we used the moderately aversive stimulus strength of 0.42 mV in order to balance aversiveness with response probability; however, our task employed continuous shock for punishment as opposed to intermittently spaced

shock, as used in Oleson *et al* (Oleson et al., 2012). During each session, subjects were presented with a lever paired with a cue light and an auditory cue; a response on the lever at any point during the session resulted in the retraction of the lever and termination of the cue light and foot shock, as well as progression to the ITI (20 s). Subjects were gradually shaped toward the lever (safe side, quadrant with lever, orientation toward the lever, rearing, pressing) by the experimenter as needed until escape behavior acquisition.

Once subjects acquired consistent escape behavior, trials were altered to allow for shock avoidance; positive reinforcement and neutral contingencies were also added. At trial onset, a cue light and one of three discriminatory auditory cues (tone, white noise, or clicker) were activated; house lights remained on at all times. After 5 s, the lever was extended into the chamber; the 5 s delay was implemented to reduce compulsory pressing and to allow for separate epoch analysis around cue and lever press. Once extended, the lever could be pressed to produce one of three outcomes (dependent on the auditory cue identity): delivery of a food reward (a sucrose pellet; positive reinforcement behavior), prevention of foot shock (0.42 mV; negative reinforcement behavior), or no consequence. If the animal failed to press the lever within a 10 s period, no food reward was delivered, foot shock commenced, or no there was no consequence. Similar to the previous protocol, rats were able to press the lever at any time to escape the foot shock once it commenced; if rats failed to press the lever, foot shock automatically terminated after 15 s. After response or termination of the trial, an ITI of (20 s) was initiated. Auditory cue identities were counterbalanced across rats. Animals were very well trained on this task, completing

>30 sessions and displaying >60% avoidance responses for 3 consecutive sessions. Session duration during FSCV recording was 60 minutes.

Fast-scan cyclic voltammetry. For recordings, animals were connected to a head-mounted voltammetric amplifier (current-to voltage converter) and a commutator (Crist Instruments) mounted above the recording chamber. During each session, an electrical potential was applied to the recording electrode in the same manner as described above (see *Intra-cranial surgical procedures*). In order to detect changes in dopaminergic concentration over time, the current at its peak oxidation potential was plotted for successive voltammetric scans and background signal was subtracted. Two PC-based systems, fitted with PCI multifunction data acquisition cards and software written in LabVIEW (National Instruments), were used for waveform generation, data collection, and analysis. The signal was low-pass filtered at 2,000Hz. Event timestamps from Med Associates were recorded, in order to analyze behaviorally relevant changes in dopamine release.

Dopamine was identified by its stereotypical and specific cyclic voltammogram signature. Behaviorally-evoked DA signals met electrochemical criterion if the cyclic voltammogram was highly correlated to that of the DA templates produced during the training set. The training set is a template containing six each of background-subtracted cyclic voltammograms and corresponding calibrated concentrations for both dopamine and pH extracted from data pooled across animals acquired during electrical stimulations that are known to evoke DA release (stimulation at 1V: 30 Hz, 6 pulses; 30 Hz, 12 pulses; 30 Hz, 24 pulses; 60 Hz, 6 pulses; 60 Hz, 12 pulses; 60 Hz, 24 pulses) . The data collected during a session

were not analyzed if reward trials did not elicit DA release that satisfied these chemical verification criteria. Voltammetric data was analyzed using software written in LabView and Matlab. A principal component regression (Tar Heel CV chemometrics software) was used to extract the DA component from the raw voltammetric data (Heien et al., 2004; Keithley, Heien, & Wightman, 2009). Eigenvalues (principal components) are calculated that describe relevant components of our training set, and we perform multivariate regression analysis to determine a correlation coefficient to describe our recorded behavioral data versus the training set. The number of factors we select to keep in our PCA analysis accounts for >99% of the variance (at least 3, but usually 4-5 factors are kept). Factor selection is a very important step, as retaining more factors than we need would add noise to our data but retaining too few could mean discarding potentially meaningful information (Kramer, 1998). FSCV results may be influenced by the way in which the variance is apportioned to the components. Importantly, the exact same method was applied to each trial-type (neutral, reward, and shock) allowing for fair comparison between conditions.

We also use the residual to examine the quality of the fit. In general, the residual is the difference between the experimental observation and the predicted value derived from a model/template (our regression values) and is a measure of the unknown portion of the signal that is not accounted for by the principal components of the regression. This is important when considering the accuracy and the applicability of the model and is important for identifying possible interfering molecules or noise (such as drift). The sum of squares of the difference between the

template and the experimental data is the residual value (Q) and the threshold Qa establishes whether the retained principal components provide a satisfactory description of the experimental data; the discarded principal components should provide a measure of noise (Heien et al., 2005; Keithley et al., 2009). We use this Qa measure in combination with our regression analysis to establish our concentration corrections.

Chemometrics is a widely-used analytical method that separates changes in current that are caused by DA release from those caused by pH shift or other electrochemical 'noise' by comparing eigenvalues derived from stimulated DA release and changes in pH to those derived from behavioral release (Joseph F. Cheer et al., 2007; Heien et al., 2005; Phillips, Robinson, Stuber, Carelli, & Wightman, 2003; Wightman et al., 2007)

Histology. Following the completion of the study, animals were terminally anesthetized with an overdose of isoflurane (5%) and transcardially-perfused with saline and 4% paraformaldehyde. Brain tissue was removed and post-fixed with paraformaldehyde. Brains were then placed in 30% sucrose solution for 72 hr and sectioned coronally (50µm) using a microtome. Tissues slices were mounted onto slides and stained with thionin for histological reconstruction.

Data analysis and statistics. Behavioral videos from the combined positive and negative reinforcement task were scored for measures of fear (freezing, rearing, orienting to the lever) during the cue presentation epoch (cue onset to lever extension) for all trial types. For behavioral analysis, this epoch was divided into 2 sub-epochs (first half and last half) and separate binary (0 or 1) scores were recorded for each

behavioral measure during each sub-epoch. These behavioral analyses were scored blindly.

As described above, all voltammetric data was analyzed using software written in LabView and then further analyzed in Matlab (Mathworks). The dopamine component of our signal was first isolated from the raw voltammetric signal using principal component regression and calibration to a CV/concentration matrix. Analysis was centered on various epochs: cue epoch (cue onset to lever extension), lever epoch (1 s after lever extension), and baseline epoch (5 s before cue onset). Behavioral measures were correlated to dopamine release using linear regression ($p < 0.05$).

Data availability. All data that support the findings of this study are available from the corresponding author upon request.

Results

Behavior during combined approach-avoidance

Rats ($n = 10$) were trained on a combined approach-avoidance task (**Fig. 1a-c**). At the start of each trial, one of three discriminatory auditory cues and a cue light were presented. Auditory cues signaled whether the current trial would be a reward, shock or neutral trial. Five seconds after cue presentation, a lever was extended into the chamber where it could be pressed to produce one of three outcomes (dependent upon auditory cue identity): delivery of a food reward (positive reinforcement behavior, *i.e.* reward trials), prevention of foot shock (negative reinforcement behavior, *i.e.* shock trials), or no consequence (*i.e.* neutral trials). If the animal failed

to press the lever within a 10 s period, no food reward was delivered on reward trials, foot shock commenced on shock trials, or there was no consequence on neutral trials.

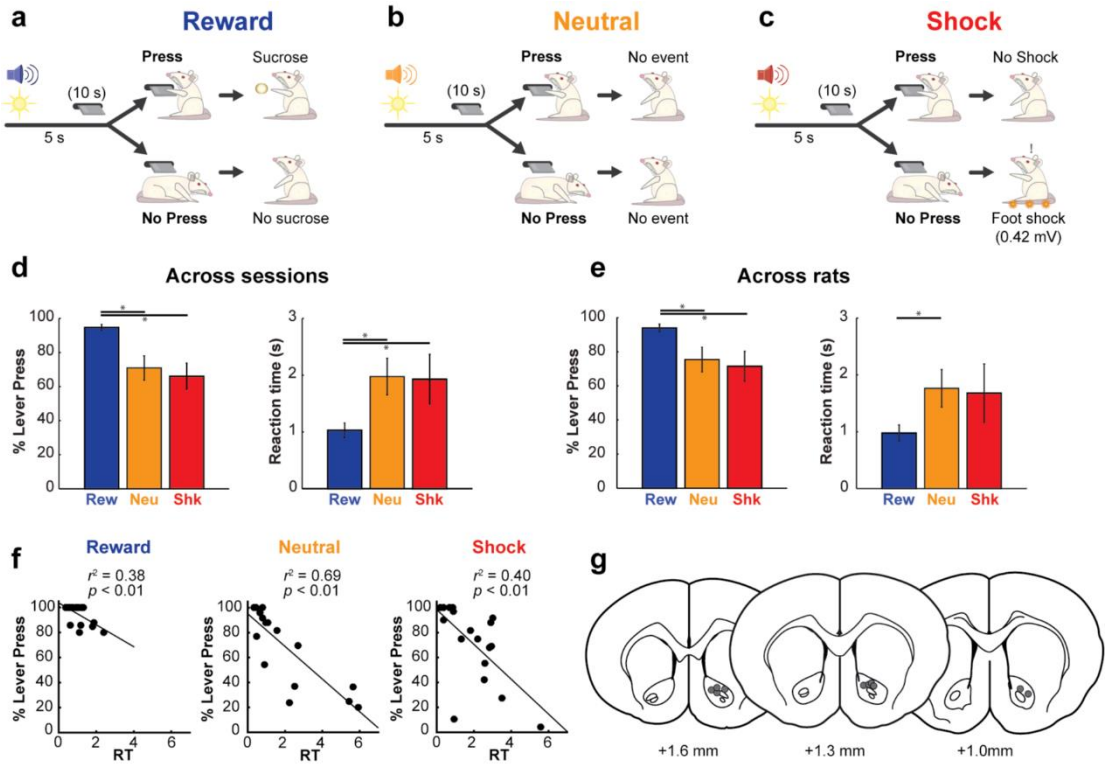


Figure 1. Task design and population behavioral results (N = 10 rats; 18 sessions). Sessions consisted of 3 trial types: reward (**a**), neutral (**b**), and shock (**c**), which could be identified by their unique auditory cue. (**a-c**) At the beginning of each trial, rats were presented with a light cue and trial-specific sound cue 5s before lever extension and then had a maximum of 10s to press the lever before it was retracted. If rats pressed the lever, they could receive a sucrose pellet reward, avoid an impending foot shock (0.42 mV), or experience no consequence, depending on the identity of the sound cue. If rats failed to press the lever within 10s after its extension into the chamber, they would alternatively receive no sucrose reward, receive continuous foot shock (0.42 mV), or experience no consequence depending on the identity of the sound cue. Once shock commenced, it could be terminated by lever press. After each consequence, the trial progressed into a 20s ITI. Trial types were pseudo-randomly interleaved within each session (~60 min) and sound cue identity was counterbalanced across rats. (**d-e**) Percent lever press and reaction time computed across each session (**d**) and across rats (**e**). Bars with asterisks represent significance (T-test; $p < 0.05$; n = 18 for (**d**) and n = 10 for (**e**)). Error bars represent S.E.M. (**f**)

Correlation between percentage lever press and reaction time to press for each trial type (reward, neutral, and shock) across all sessions (**g**) Placement of chronic recording electrodes within the NAc core based on histology (Paxinos & Watson, 2007).

These three trial types were pseudo-randomly interleaved (i.e., random without replacement) within each session. The average number of trials per session was 78 (26 per trial type).

The data described below were collected during 18 different behavioral sessions (i.e., 3 sessions from 1 rat, 2 sessions per rat from 6 rats and 1 session per rat from 3 rats, to equal 10 rats total) performed in combination with fast-scan cyclic voltammetry (FSCV) recording within the NAc (**Fig. 1g**). During these sessions, rats produced the most responses and were the fastest to respond on reward trials compared to neutral (% P: $t(17) = 3.67$, $p < 0.01$; RT: $t(17) = 3.71$, $p < 0.01$) and shock trials (% P: $t(17) = 3.88$, $p < 0.01$; RT: $t(16) = 1.97$, $p = 0.07$); there was no significant difference between neutral and shock trials for either behavioral measure (% P: $t(17) = 1.42$, $p = 0.17$; RT: $t(16) = 0.33$, $p = 0.74$; **Fig. 1d**). During sessions where at least one shock was delivered (*i.e.* rat failed to avoid shock on at least one shock trial; 15 out of 18 sessions), rats escaped shock on 56% of non-avoid trials. Lastly, there was a significant negative correlation between reaction time and percent lever press for all trial types when examining data across sessions (Reward $r^2 = 0.38$, Neutral $r^2 = 0.69$, Shock $r^2 = 0.40$, all $p < 0.01$; **Fig. 1f**).

Similar results were obtained when we averaged across sessions within a rat and then averaged across rats (i.e., one data point for each rat; $n = 10$; **Fig. 1e**). Across rats, percent lever pressing was higher for reward trials relative to neutral

(%P: $t(9) = 2.52$, $p < 0.05$) and shock trials (%P: $t(9) = 2.46$, $p < 0.05$); there was no difference in lever pressing between neutral and shock trials (%P: $t(9) = 0.88$; $p = 0.40$). Rats were also slower to respond on neutral and shock trials relative to reward trials; however, this comparison was only significant for neutral versus reward (Rew vs Neu: $t(9) = 2.86$, $p < 0.05$; Rew vs Shk: $t(9) = 1.22$, $p = 0.25$). There was no significant difference in reaction times between neutral and shock trials (Neu vs Shk: $t(9) = 0.18$, $p = 0.86$).

Overall, these behavioral measures demonstrate that rats indeed dissociated reward from the other two trial types; variability in behavioral performance observed across recording sessions will be discussed below. Further, we will show that rats also understood the difference between neutral and shock trials, as illustrated by significantly increased freezing behavior during the presentation of the shock cue relative to cues predicting reward or neutral trials.

Phasic DA release is high for approach and avoidance cues

As a first step in understanding the role of DA in task performance, we examined changes in phasic DA release across all animals ($n = 10$) when rats pressed ('Press'; **Fig. 2a-c**) or did not press ('Non-Press'; **Fig. 2d-e**) the lever. Average DA release across time is displayed for each of the three trial types in **Fig. 2a**. Increases in DA release were observed shortly after cue onset and were higher for reward (blue) and shock (red) cues compared to neutral (yellow).

We focused our following analyses on two behaviorally-relevant epochs, a cue epoch (5 s after cue onset) and a lever epoch (1s after lever extension). Both analysis epochs precede shock and reward delivery, and all data shown are taken prior

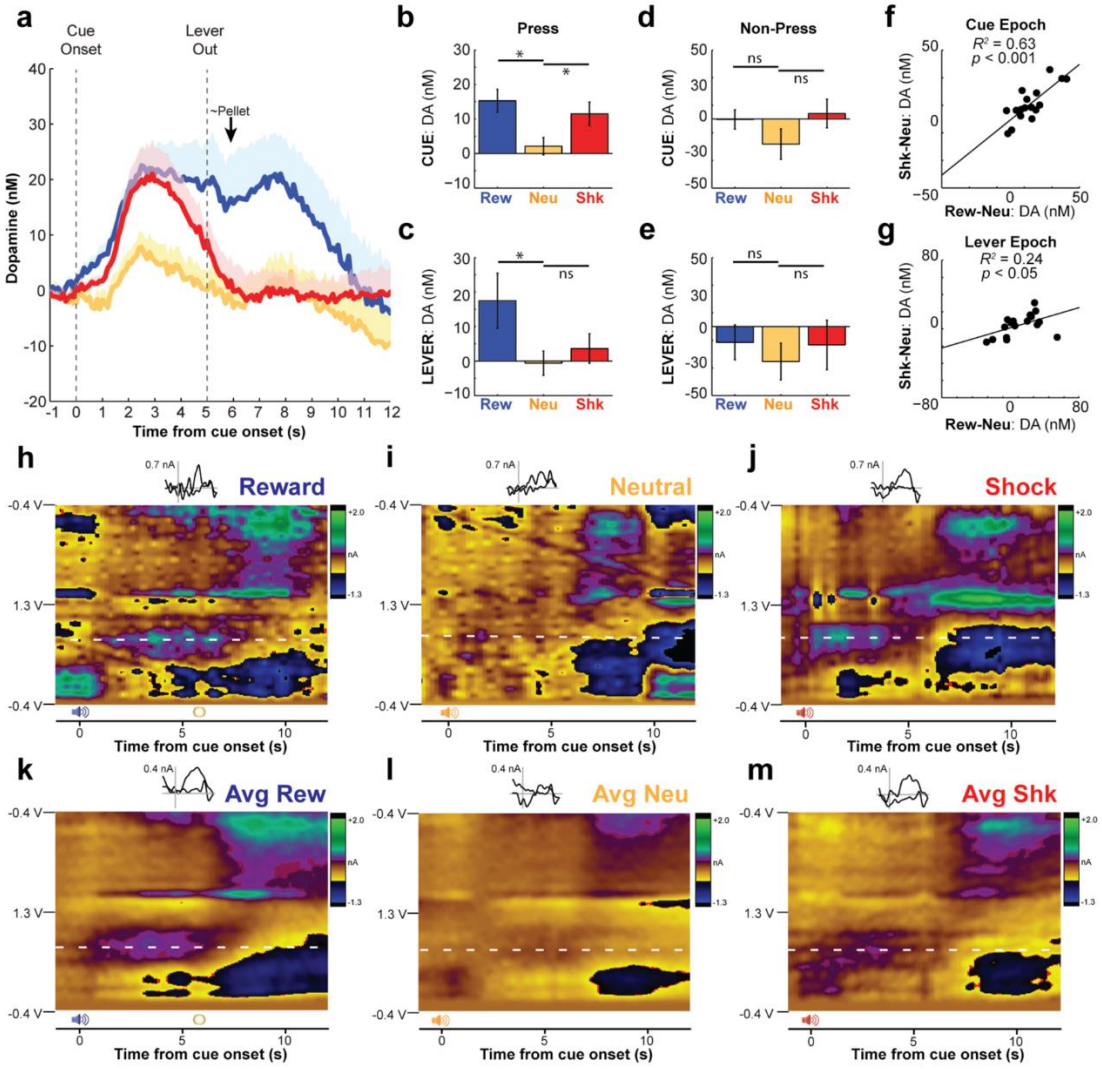


Figure 2. Average dopamine release (N = 10 rats) during cue and lever epochs for each trial type. (**a**) Dopamine release (nM) across time for reward (blue), neutral (yellow), and shock (red) trials. Dopamine release is baseline (5s before light onset to light onset) subtracted. (**b-e**) Quantification of DA release for press and non-press responses during the cue epoch (cue onset to lever extension; 5s) and lever epoch (lever extension plus 1s). Bars with asterisks represent significance (T-test; $p < 0.05$). Error bars represent S.E.M. (**f-g**) Correlation between shock and reward trials normalized over neutral trials (shock minus neutral; reward minus neutral) for both cue epoch and lever epoch. (**h-m**) False-color plots indicate voltammetric current (z-axis) plotted against applied scan potential (y-axis) and time (x-axis) for representative press trials aligned to cue onset for each of the 3 trial types **(h-j**; Reward, Neutral, and Shock), as well as averaged press trials aligned to cue onset for each of the 3 trial types (**k-m**; Average Reward, Average Neutral, Average Shock). Insets show cyclic voltammogram for dopamine; scale bars are set to 0.7 nA for individual examples and 0.4 nA for averages.

to shock delivery to exclude shock artifact. A one-factor ANOVA during the cue epoch revealed a significant main effect of trial type during lever press trials ($F(2,27) = 4.3$, $p < 0.05$; $n = 10$). Mean DA release during both reward and shock cues was significantly elevated compared to neutral trials when rats pressed the lever (**Fig. 2a-b**; Rew vs. Neu: $t(9) = 3.96$, $p < 0.01$; Shk vs. Neu: $t(9) = 2.57$, $p < 0.05$; $n = 10$). DA release during the cue epoch was not significantly different between reward and shock trials (Rew vs. Shk: $t(9) = 1.85$, $p = 0.10$; $n = 10$). During the lever epoch, the main effect of 'trial type' was not significant on lever press trials ($F(2,27) = 2.56$, $p = 0.096$; $n = 10$); DA release was only significantly elevated during the lever epoch of reward trials, relative to shock and neutral (**Fig. 2a**, **2c**; Rew vs. Neu: $t(9) = 2.75$, $p < 0.05$; Rew vs. Shk: $t(9) = 2.83$, $p < 0.05$; Shk vs. Neu: $t(9) = 0.98$, $p = 0.35$; $n = 10$). False-color plots shown in figure 2 indicate voltammetric current (z-axis) plotted against applied scan potential (y-axis) and time (x-axis) for representative press trials aligned to cue onset for each of the 3 trial types **(Fig. 2h-j**; Reward, Neutral, and Shock), as well as averaged press trials aligned to cue onset for each of the three trial types for one session (**Fig. 2k-m**; Average Reward, Average Neutral, Average Shock). Additional examples of stimulated and behaviorally-evoked DA release can be found within the supplementary materials (**Supplementary Fig. 1**, **Supplementary Fig. 2**). We conclude that, on press trials, DA release was significantly increased for reward and shock trials compared to neutral trials during the cue epoch, but it was only significantly increased for reward trials during the lever epoch. Notably, when rats did not press the lever, there was not a significant main effect of 'trial type' for either epoch (Cue Epoch: $F(2,18) = 0.93$, $p = 0.41$; Lever

Epoch: $F(2,18) = 0.15$, $p = 0.86$). DA release did not significantly differ between any of the trial types during the cue epoch (**Fig. 2d;** Rew vs Neu: $t(4) = 2.60$, $p = 0.06$, $n = 5$ and Rew vs Shk: $t(4) = 1.35$, $p = 0.25$, $n = 5$; Shk vs Neu: $t(7) = 1.09$, $p = 0.31$, $n = 8$) or the lever epoch (**Fig. 2e;** Rew vs Neu: $t(4) = 1.97$, $p = 0.12$, $n = 5$ and Rew vs Neu: $t(4) = 0.68$, $p = 0.53$, $n = 5$; Shk vs Neu: $t(7) = 0.38$, $p = 0.72$, $n = 8$). Note that the degrees of freedom were fewer for the analysis of 'non-press trials' due to sessions where rats pressed for all trials within a trial type (i.e., two rats pressed for all trials across all trial types and two rats pressed for all reward trials but failed to press for some neutral and shock trials).

We see increases in NAc DA during cues that predict potential reward or shock during successful acquisition or avoidance behavior, respectively. Since this data is averaged across all sessions, it is possible that these increases in DA release to reward and shock cues may have occurred in different microdomains (Wightman et al., 2007). That is, DA release might be high during reward cues and low during shock and neutral cues in some sessions but high during shock cues and low during reward and neutral cues in other sessions. To address this issue, we computed a reward index (reward - neutral/reward + neutral) and a shock index (shock - neutral/shock + neutral) for each session during the cue and lever epochs. We found significant positive correlations between DA release on reward and shock trials relative to neutral trials during the cue epoch and lever epoch, indicating that increases in DA release to reward cues occurred in the same session and, hence, the same microdomain as increases in DA release to shock cues during avoidance trials (**Fig. 2f-g**; $r^2 = 0.63$ and $r^2 = 0.24$, respectively; $p < 0.05$ for both; $n = 10$ rats).

DA release is negatively correlated with avoidance

Next, we examined the relationship between DA release and behavior during both cue and lever epochs separately for each trial type by plotting percent lever press and reaction time against DA release for all sessions (**Fig. 3**). For reward trials, correlations were not significant, suggesting that increased dopamine release during the cue or lever epoch does not predict performance or there was not enough variance to capture the relationship between the two (**Fig. 3a-d**). However, when DA release was high during the cue or lever epoch for neutral trials, reaction times tended to be slower (Cue: $r^2 = 0.491$, $p < 0.01$; Lever: $r^2 = 0.487$, $p < 0.01$; $n = 10$ rats) and there were fewer responses on the lever (Cue: $r^2 = 0.333$, $p < 0.05$; Lever: $r^2 = 0.366$, $p < 0.01$; $n = 10$ rats; **Fig. 3e-h**). This pattern was conserved for shock trials, but only significant during the cue epoch (% P: $r^2 = 0.327$, $p < 0.05$; RT: $r^2 = 0.213$, $p = 0.06$; $n = 10$ rats; **Fig 3i-l**). Thus, increased DA release during the shock cue was positively correlated with worse performance on the task. This is an intriguing finding, since prior studies predict increased DA during the cue or lever epoch results in more and faster lever pressing for both reward and avoidable shock(E. B. Oleson et al., 2012). Instead, here we find excessive DA at the cue is associated with poor performance during shock avoidance.

Distinct DA patterns for good and poor avoidance behavior

When rats are anxious, they tend to perform poorly in active shock avoidance paradigms due to the perseveration of freezing behavior, which inhibits the initiation of voluntary actions needed to avoid shock (Bolles, 1970; Brush, 2003). In contrast, other rats are able to overcome this Pavlovian response in order to avoid shock

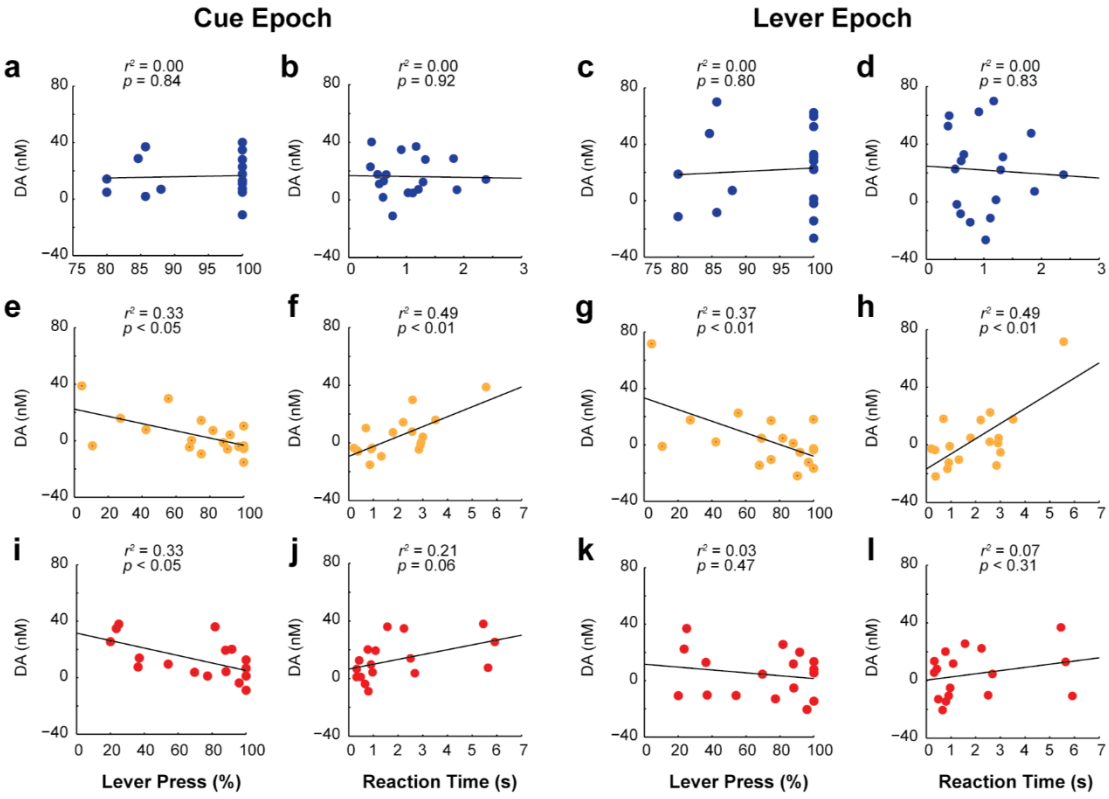


Figure 3. Correlation of DA release with behavioral measures. Each dot represents an individual session; all recording sessions are represented. (**a-d**) DA release is not significantly correlated with lever press or reaction time for reward trials (blue). (**e-f, i-j**) DA release is negatively correlated with lever press and positively correlated with reaction time for both neutral (yellow) and shock (red) trials during the cue epoch. (**g-h**) DA release is negatively correlated with lever press and positively correlated with reaction time for neutral trials in the lever epoch; (**k-l**) DA is not significantly correlated with behavioral measures for shock trials during the lever epoch.

successfully. Based on these findings, we predicted that some rats would press the lever less frequently during cues that predict shock compared to cues that predict neutral trials. As in previous studies, this would enable us to divide our sessions into those displaying good or poor avoidance performance (Martinez et al., 2013). Indeed, we found a subset of sessions (n = 9) contained pressing behavior that differed significantly on shock trials compared with neutral trials. Lever pressing during these sessions showed a significant main effect of trial type in a one-way ANOVA ($F(2,24) = 8.91$, $p < 0.01$; n = 5 rats). During these sessions, response rates were significantly higher and lower for reward and shock trials, respectively, relative to neutral trials (Rew vs. Neu: $t(8) = 3.65$, $p < 0.01$; Rew vs. Shk: $t(8) = 5.28$, $p < 0.001$; Shk vs. Neu: $t(8) = 3.04$, $p < 0.05$; n = 5 rats), and reaction times were slower for neutral and shock trials relative to reward trials (Rew vs. Neu: $t(8) = 4.23$, $p < 0.01$; Rew vs. Shk: $t(8) = 2.35$, $p = 0.05$; n = 5 rats; **Fig. 4a-b**). Thus, in these sessions, rat pressed significantly less on shock trials compared to reward and neutral trials. We will refer to these sessions as poor avoidance sessions.

The remainder of sessions (n = 9) showed no significant main effect of trial type on lever pressing ($F(2,24) = 0.55$, $p = 0.58$) or reaction time ($F(2,24) = 0.26$, $p = 0.77$). Instead, during these sessions, rats pressed at a high rate for all trial types (Rew vs. Neu: $t(8) = 1.42$, $p = 0.19$; Rew vs. Shk: $t(8) = 0.63$; $p = 0.55$; Neu vs. Shk: $t(8) = 1.80$, $p = 0.11$; n = 6 rats) and were equally fast on neutral and shock trials as reward trials (Rew vs. Neu; $t(8) = 1.37$, $p = 0.21$; Rew vs Shk: $t(8) = 0.64$, $p = 0.54$; n = 6 rats; **Fig. 4c-d**). We will refer to these sessions as good avoidance sessions. When demonstrating good avoidance, rats only received shock on 4.6% of total trials (i.e.,

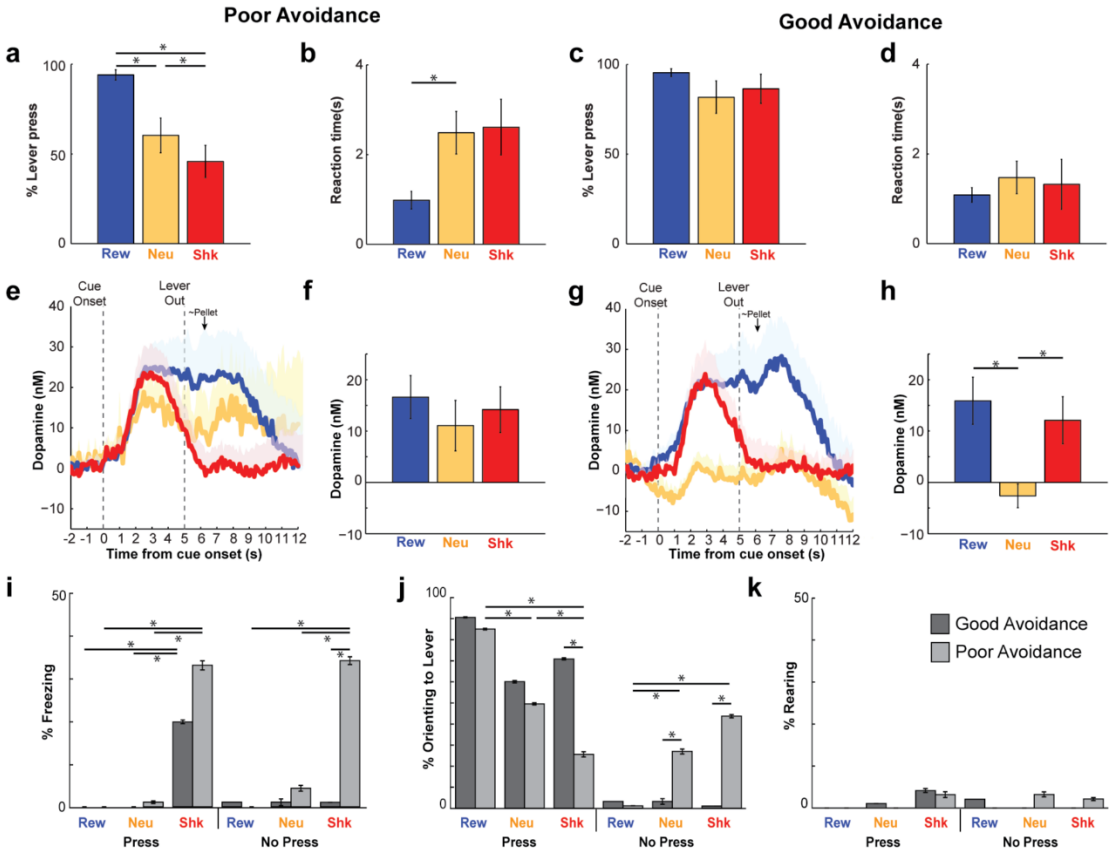


Figure 4. Poor and good avoidance exhibit differences in behavior and dopamine release across trial types. Reward, neutral, and shock trial types are represented by blue, yellow, and red, respectively. Behavioral differences are shown using percent lever press (**a**) and reaction times (**b**) for poor avoidance and percent lever press (**c**) and reaction times (**d**) for good avoiders. Bars with asterisks represent significance (T-test; $p < 0.05$; $n = 6$ rats for good avoidance, 5 rats for poor avoidance). Error bars represent S.E.M. (**e-h**) Dopamine quantification for poor and good avoiders. Dopamine release (nM) for each trial type is shown across time with cue and lever epochs indicated for poor (**e**) and good avoiders (**g**), respectively. Dopamine release is quantified during the cue epoch for each trial type for poor (**f**) and good (**h**) avoiders. Error bars represent S.E.M. (**i-k** Analysis of stress-related behaviors during press and failed press. Percent freezing (**i**), orienting to the lever (**j**), and rearing (**k**) during poor and good avoidance. Asterisks indicate $p < 0.05$ in chi-square; $n = 6$ rats.

sum of all 3 trial types), which was significantly less than 19% received during poor avoidance sessions ($t(16) = 3.14$, $p < 0.01$; $n = 6$ rats). There was no significant difference between the number of rewards received between groups; during good and poor avoidance sessions rats received reward on 32% and 31% of the total trials (i.e., sum of all 3 trial types), respectively ($t(16) = 0.86$, $p = 0.40$; $n = 6$ rats).

Thus, overall, 6 different rats contributed to sessions demonstrating good avoidance ($n = 9$ sessions: 2 sessions per rat for 3 rats and 1 session per rat for 3 rats) and 5 rats contributed sessions demonstrating poor avoidance ($n = 9$ sessions: 2 sessions per rat for 4 rats and 1 session from 1 rat). Note, only 1 of the 10 recorded rats contributed sessions to both categories (1 and 2 sessions to good and poor avoidance, respectively).

As suggested above, poor avoidance behavior is thought to result from unmanaged fear-evoked defensive reactions. To determine if this holds true for our dataset, we asked if freezing, lever orienting, and rearing behaviors were different between good and poor avoiders (**Fig. 4i-k**). Though both groups exhibited increased freezing behavior during shock trials, poor avoiders froze more than good avoiders (**Fig. 4i**). Good avoiders (Rew vs. Shk: $X^2 = 15.31$, $p < 0.001$, Shk vs. Neu: $X^2 = 15.31$, $p < 0.001$; $n = 3$ rats) and poor avoiders (Rew vs. Shk: $X^2 = 24.89$, $p < 0.0001$, Shk vs. Neu: $X^2 = 22.36$, $p < 0.0001$; $n = 3$ rats) exhibited increased freezing behavior during shock trials when the lever was pressed, compared with freezing during reward or neutral trials. Freezing on shock trials when rats failed to press the lever was significantly increased relative to reward and neutral trials (Shk vs. Rew: $X^2 = 25.66$, $p < 0.0001$; Shk vs. Neu: $X^2 = 16.78$, $p < 0.0001$; $n = 6$ rats) and relative to

shock trials when good avoiders did not press the lever (Poor Shk Non-press vs. Good Shk Non-press: $X^2 = 23.91$, $p < 0.0001$; $n = 6$ rats). Thus, we found that poor avoiders froze more on both press and non-press shock trials when compared to good avoiders. Notably, good avoiders still expressed fear responses during shock trials, demonstrating that they clearly understood task contingencies.

Rats generally oriented toward the lever more often when they were successful in pressing (**Fig. 4j**). Poor avoiders showed significant step-wise decreases in orienting behavior following the same pattern as their lever pressing behavior (Rew vs. Neu: $X^2 = 4.75$, $p < 0.05$; Rew vs. Shk: $X^2 = 18.91$, $p < 0.0001$; Shk vs. Neu: $X^2 = 4.61$, $p < 0.05$; $n = 3$ rats); orienting behavior was not significantly different between trial types when good avoiders pressed the lever (Rew vs. Neu: $X^2 = 2.98$, $p = 0.08$; $n = 3$ rats). Poor avoiders oriented toward the lever more often during failed shock trials and neutral trials than during failed reward trials (Rew vs. Neu: $X^2 = 29.85$, $p < 0.0001$; Rew vs. Shk: $X^2 = 17.86$, $p < 0.0001$; $n = 3$ rats), unlike good avoiders who failed to orient on non-press trials regardless of trial type (Good vs. Poor for Neu: $X^2 = 14.02$, $p < 0.001$; Good vs. Poor for Shk: $X^2 = 30.82$, $p < 0.0001$; $n = 6$ rats). There were no significant differences in rearing behavior, a measure of general motor activity, attention, and environmental engagement(Bailey & Crawley, 2009), between good and poor avoiders across any trial type (**Fig. 4k**).

To determine how DA release patterns differ among good and poor avoidance behaviors, we performed a three-factor ANOVA across trial type (reward, shock, neutral), group (‘good’ or ‘poor’ avoidance), and response type (press or non-press). This revealed a main effect of response type ($F(1,75) = 7.07$, $p < 0.01$; $n = 10$), trial

type ($F(2,75) = 5.83$, $p < 0.01$; $n = 10$), and group ($F(1,75) = 6.92$, $p < 0.05$; $n = 10$). In addition, there was a significant two-way interaction between trial type and group ($F(2,75) = 3.38$, $p < 0.05$; $n = 10$). Interactions between response type and group ($F(1,75)=1.29$, $p = 0.28$; $n = 10$), between response type and trial type ($F(2,75) = 0.77$, $p = 0.47$; $n = 10$), and between all three factors did not achieve significance ($F(2,75) = 2.49$, $p = 0.09$; $n = 10$).

Next, we examined average DA release over time for good and poor avoidance sessions (**Fig. 4e-h**). When rats performed poorly on avoidance trials, DA release was nonselective during the cue epoch (**Fig. 4f**); there was no main effect of trial type in the one-factor ANOVA ($F(2,23) = 0.33$, $p = 0.72$; $n = 5$ rats) and no comparisons between trial types were significant (Rew vs. Neu: $t(8) = 1.22$, $p = 0.25$; Rew vs. Shk $t(7) = 1.04$, $p = 0.33$; Shk vs. Neu: $t(7) = 2.09$, $p = 0.074$; $n = 5$ rats). To the contrary, when rats that responded at a high rate for all trial-types (i.e., good avoidance), DA release clearly delineated reward, shock, or neutral cues (**Fig. 4g**). During the cue epoch, we found a significant main effect of trial type (**Fig. 4h**; $F(2,24) = 5.37$, $p < 0.05$; $n = 6$ rats) and DA release during both reward and shock cues differed from release seen during neutral cues (Rew vs. Neu: $t(8) = 3.81$, $p < 0.01$; Shk vs. Neu: $t(8) = 3.01$, $p < 0.05$; Rew vs. Shk: $t(8) = 1.27$; $p = 0.24$; $n = 6$ rats; **Fig. 4h**).

With these group distinctions in mind, we re-examined the correlation between DA release and behavior. During good avoidance, DA release was not correlated with behavior (% P or RT) for any trial type or analysis epoch (**Supplementary Table 1**). However, during poor avoidance, DA release was

a

Poor Avoiders

		Cue Epoch			Lever Epoch		
		Reward	Neutral	Shock	Reward	Neutral	Shock
PC	R^2	0.25(+)	**0.60(-)**	**0.66(-)**	0.28(+)	**0.66(-)**	0.46(-)
	p	0.1667	**0.0145**	**0.0149**	0.1451	**0.008**	0.0646
RT	R^2	0.08(-)	**0.62(+)**	0.21(+)	0.05(-)	**0.76(+)**	**0.58(+)**
	p	0.4604	**0.0120**	0.2557	0.5515	**0.0023**	**0.0276**

b

Good Avoiders

		Cue Epoch			Lever Epoch		
		Reward	Neutral	Shock	Reward	Neutral	Shock
PC	R^2	0.23(-)	0.00	0.28(-)	0.29(-)	0.00	0.04(+)
	p	0.1889	0.8862	0.1419	0.1350	0.9502	0.5897
RT	R^2	0.07(+)	0.08(+)	0.23(+)	0.03(+)	0.00	0.07(-)
	p	0.4855	0.4490	0.1954	0.6815	0.8589	0.4813

Supplementary Table 1. Correlation between DA release and behavior in poor and good avoiders. R2 values (top) and p values (bottom) for correlations between DA release during each trial type and behavior during cue and lever epochs for poor (**A**) and good (**B**) avoiders.

negatively correlated with % P during both neutral (Cue: $r^2 = 0.60$, $p < 0.05$; Lever: $r^2 = 0.66$, $p < 0.01$; $n = 5$ rats) and shock trials (Cue: $r^2 = 0.66$, $p < 0.05$; $n = 5$ rats; **Supplementary Table 1**). In these sessions, DA release was also positively correlated with reaction times during both neutral (Cue: $r^2 = 0.62$, $p < 0.05$; Lever: $r^2 = 0.76$, $p < 0.01$; $n = 5$ rats) and shock trials (Lever: $r^2 = 0.58$, $p < 0.05$; $n = 5$ rats). Together, these data suggest that increased cue-evoked dopamine release in poor avoiders promotes maladaptive behavior such as slower and fewer lever presses during avoidance, but not during reward-seeking.

Differences in DA release between press and no-press responses during good and poor avoidance

Finally, we asked if there were differences in DA release during trials when rats press the lever versus trials when rats did not (**Fig. 5**). Dopamine release was reduced in trials containing neutral cues during the lever epoch during non-press relative to press trials for both good and poor avoiders (**Fig. 5a-b**). A two-factor ANOVA with response type (press and non-press) and performance (good and poor avoidance) as factors during neutrals trials produced significant main effects of response type ($F(1, 28) = 6.81$, $p < 0.05$) and group ($F(1,28) = 5.67$, $p < 0.05$), but there was no interaction between response type and group ($F(1,28) = 2.29$, $p = 0.14$).

Although DA release was reduced on neutral non-press relative to neutral press trials for both groups, DA release was only significantly reduced on non-press shock trials versus press shock trials during good avoidance sessions (**Fig. 5d**). A two-factor ANOVA with response type (press and non-press) and performance (good and poor avoidance) as factors during shock trials produced significant main effects

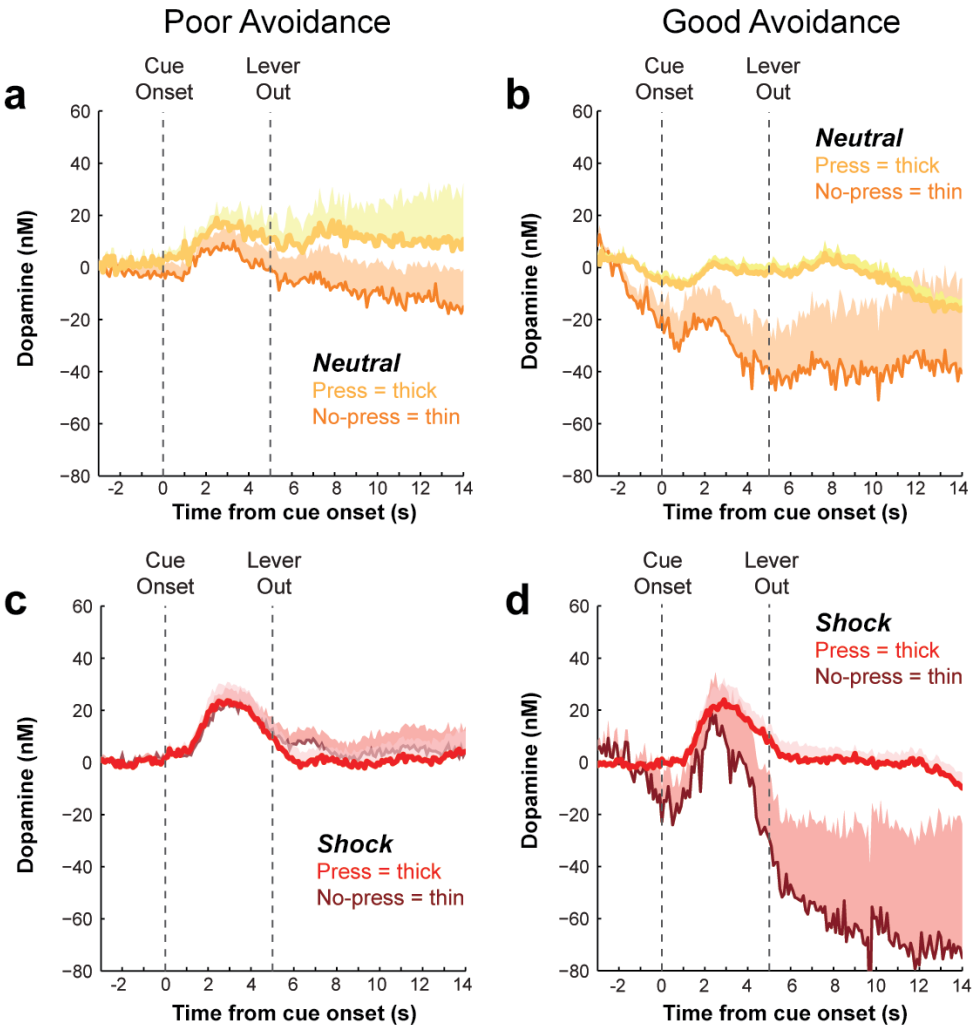


Figure 5. Dopamine release during poor and good avoidance for neutral and shock trials when rats pressed or did not press the lever. (**a-b**) Press and non-press during neutral trials in poor and good avoidance. (**c-d**) Press and non-press during shock trials in poor and good avoidance. DA release during lever press is represented by a thick line, non-press with a thin line. Reward trials are not shown here, as pressing for reward was at ceiling for all animals. Error bars represent S.E.M.

of response type ($F(1, 27) = 4.40$, $p < 0.05$) and group ($F(1,27) = 5.22$, $p < 0.05$), and a significant interaction between response type and group ($F(1,27) = 6.35$, $p < 0.05$). Thus, during good avoidance sessions, there was a reduction in DA release when rats failed to press the lever relative to when the lever was pressed and shock was avoided.

Chapter Discussion

While dopaminergic activity within the mesolimbic pathway has been widely implicated in the construction of reward expectations, a growing literature has recently emerged investigating its role during punishment and avoidance. Recent studies suggest increased cue-evoked DA release in the NAc predicts punishment avoidance, whereas a pause in DA transients occurs during unavoidable punishment across modalities (Badrinarayan et al., 2012; Darvas, Fadok, & Palmiter, 2011; Oleson et al., 2012; Oleson & Cheer, 2013; Roitman, Wheeler, Wightman, & Carelli, 2008; Volman et al., 2013). Yet, activation of DA neurons and D1 receptors is necessary for the formation of fear memories, and increases in DA release in the NAc core occurs in direct response to punishments, such as tail pinch (Budygin et al., 2012; Ikegami, Uemura, Kishioka, Sakimura, & Mishina, 2014; Kishioka et al., 2009). These seemingly contradictory findings have made it difficult to pinpoint the exact role of DA during punishment and negative reinforcement.

Here, we show that phasic increases in DA release can signal the need for approach or avoidance behavior within the same microenvironment. Our group data reveals higher cue-evoked DA release during shock and reward cues compared with

neutral cues, when the cue promotes lever press. By temporally dissociating the onset of the cue and the extension of the lever, we also found that the increase in DA release seen during shock avoidance is to the cue, not the action. Importantly, increased DA release to cues predicting shock and reward do not appear to reflect salience, since cues that predict unavoidable shock – although salient – inhibit DA release (Katoh et al., 1996; Oleson et al., 2012; Wenzel, Rauscher, Cheer, & Oleson, 2015). Taken together, these results suggest that increased DA release to cues predicting successful avoidance and reward-seeking report the predicted value associated with each.

Our results are consistent with a previous report from Oleson et al. showing increased DA release to cues that predict successful avoidance (Oleson et al., 2012); however, their study also found increases in DA release during a cued safety period, when shock would have been delivered had the animal not successfully pressed the lever to avoid it. This increase in DA release was interpreted as a reinforcement signal similar to those seen during reward delivery in appetitive tasks. It is worth noting that our current study did not overtly signal entry into the safety period, and, in turn, we did not witness an increase in DA release during this time point in our data set. There are several possible explanations as to why we did not replicate this effect. Firstly, Oleson et al. presented a safety cue that turned on after rats successfully avoided foot shock. In our task, there was no cue to explicitly signal safety from shock. It is possible, then, that an external safety cue is necessary to elicit a DA response during the safety period and these increases will not occur simply to the absence of predicted shock. Secondly, our rats may have been more thoroughly

trained in our task than rats were in the previous report, and, thus, DA release could have completely transferred to the avoidance cue; however, this was not the case for reward trials. Lastly, not getting shocked when a potential shock was predicted is an outcome that is better than expected; this is true in both behavioral paradigms. However, in our task, the shock trial type also implicitly signifies that food reward will not be delivered. It is possible that any increases in DA release we would have seen during the safety period were attenuated by a simultaneous pause in DA release that occurs in the absence of a food reward. Further research will be necessary to rule out these interpretations; however, it was clear in both studies that DA release was high during cues that predicted successful shock avoidance.

We only observed increases in DA release during reward and shock cues relative to neutral cues when rats demonstrated good avoidance behavior. These animals responded reliably and at comparably high speeds for all three predictive cues. Compared to poor avoiders, good avoiders also froze less to cues predicting shock and responded quickly on shock trials. Thus, this group seems to be responding without being deterred by the potential negative outcome of shock trials, as if they were responding habitually. The development of a habit-like strategy is supported by previous research showing that stress can lead to an insensitivity to changes in reinforcement value and a reduction in explicit knowledge of action-outcome contingencies (Dias-Ferreira et al., 2009; Schwabe & Wolf, 2009, 2011; Soares et al., 2012; Taylor et al., 2014). Both goal-directed and habitual processes are thought to be involved in successful avoidance learning, and the behavioral pattern of good avoiders could reflect the utilization of a proactive habitual strategy under the control

of dorsal lateral striatum (DLS; habit center) in order to maximally obtain reward and avoid punishment (Seymour, Singer, & Dolan, 2007). However, note that in our task this remains speculation, since our current data set does not allow us to prove that our rats were acting habitually in response to all three trial type cues. Recent work has shown that rats well-trained on avoidance paradigms still show sensitivity to the devaluation of the shock outcome, which suggests that they remain goal-directed with respect to this action-outcome contingency (Fernando, Urcelay, Mar, Dickinson, & Robbins, 2014). This could suggest, then, that the NAc is monitoring predictions but does not directly initiate action in this task unless there are changes in action-outcome contingencies. Indeed, we found that DA release during good avoidance was not correlated with behavioral output; despite this, DA release clearly and correctly reflected the value of the predictive cues. Such signals are likely critical to maintaining appropriate responding behavior during our task, consistent with previous studies demonstrating that NAc lesions (6-hydroxydopamine, quinolinic acid, electrolytic) and D1 receptor antagonists disrupt avoidance behavior (Beninger, Mason, Phillips, & Fibiger, 1980; B. R. Cooper, Howard, Grant, Smith, & Breese, 1974; McCullough, Sokolowski, & Salamone, 1993; Wendler et al., 2013).

Based on the existing literature, it would be logical to conclude that poor avoidance behavior likely reflects low phasic DA release in NAc to shock cues. However, with few exceptions (Brush, 2003; Martinez et al., 2013), current animal research on avoidance behavior has focused on subjects who avoid at high rates. Animals that perform poorly on avoidance tasks are often omitted under the assumption that they fail to learn task contingencies; however, it has been shown that

poor avoiders do learn and instead suffer from performance deficits that arise from persistent species-specific defense reactions (Bolles, 1970; Brush, 2003; Choi, Cain, & LeDoux, 2010; Martinez et al., 2013; Moscarello & LeDoux, 2013). For example, poor avoiders tend to demonstrate higher baseline levels of anxiety and exhibit persistent freezing behavior (Brush, 2003; Choi et al., 2010; Moscarello & LeDoux, 2013). For these reasons, poor avoiders might better represent human populations with psychiatric disease.

During poor avoidance sessions in our task, when rats responded most for reward and least for shock trials, DA release during the cue was indiscriminately high across all trial types. Thus, DA release failed to properly reflect the value of cues, including cues predicting failed shock avoidance and neutral trials, when an animal's behavior was ruled by the fear of an expected aversive outcome. Such a signal could confuse processing in downstream areas, where the predictive value of future action or inaction would be indistinguishable. Increased lever pressing during reward trials versus neutral or shock trials might reflect higher overall value associated with the combined promise of reward and relief of avoiding shock; however, we do not feel that this is a complete explanation, since rats do not press more for reward than shock during good avoidance and rats also press more for neutral than shock during poor avoidance. We also found that decreased responding on shock versus neutral trials corresponded with increases in freezing to the cue, reflecting a species-specific defense reaction described previously (Bolles, 1970; Brush, 2003; Martinez et al., 2013); high DA release preceding failed avoidance might also reinforce these inappropriate freezing behaviors during avoidance trials. Recent studies have

suggested that misappropriated increases in DA release to irrelevant or misinterpreted stimuli, like our neutral cues or failed shock avoidance cues, could be critically linked to dysfunctional salience attribution in many psychological disorders (Boehme et al., 2015; Heinz, 2002; Kapur, 2003; Mishara & Fusar-Poli, 2013; Robbins & Sahakian, 1983; Winton-Brown, Fusar-Poli, Ungless, & Howes, 2014). In contrast, accumbal DA release during good avoidance clearly assigned value to cues based on their predictive valence, namely exhibiting high DA release for lever press trials during which reward was obtained or punishment was avoided.

Altogether, these data suggest that abnormal processing of value signals in NAc hinders adaptive behavior during active avoidance. That is, when rats are intractably focused on the outcome, avoidance performance is poor and is correlated with higher overall DA release in NAc. Though reliance on expected outcomes is adaptive for behavior driven by rewards and their predictive cues, this is maladaptive during punishment avoidance. These results should provide insight into the underlying neural mechanisms involved in psychiatric disorders such as addiction, anxiety disorders, and psychosis.

Chapter 4: The role of nucleus accumbens dopamine in the development and execution of sign- and goal- tracking behaviors

Portions of this chapter are currently under review for publication.

Authors: Ronny N. Gentry[1,2,5], Brian Lee[1], Gregory B. Bissonette[1], Rae J. Herman[1], John J. Mallon[1], Daniel W. Bryden[1,2], Donna J. Calu[3], Geoffrey Schoenbaum[4], Etienne Coutureau[5], Alain R. Marchand[5], Mehdi Khamassi[6], and Matthew R. Roesch[1,2]

1. Department of Psychology,
2. Program in Neuroscience and Cognitive Science
   1147 Biology-Psychology Building
   University of Maryland
   College Park, MD 20742 USA
3. Department of Anatomy and Neurobiology
   University of Maryland
   Baltimore, MD USA
4. NIDA Intramural Research Program
   Baltimore, Maryland, USA
5. Institut de Neurosciences Cognitives et Intégratives d'Aquitaine (INCIA), CNRS
   Université de Bordeaux
   Bordeaux, France
6. Sorbonne Universités
   UPMC Univ Paris 06
   Institut des Systèmes Intelligents et de Robotique, CNRS
   Paris, France

Abstract

Recent work showing that dopamine (DA) is not necessary for all forms of learning has challenged the hypothesis that phasic DA corresponds to a reward prediction error signal. A recent computational model has accounted for these observations and has provided a set of predictions to further its validity by manipulating the inter-trial interval (ITI) during autoshaping. We found that lengthening the ITI increased behavioral engagement with conditioned stimuli (CS; i.e., sign-tracking) and cue-induced phasic DA release. Importantly, DA release was also present at the time of reward delivery, even after learning. During conditioning with shorter ITIs, goal-tracking was prominent (i.e., engagement with food cup), and DA release to the CS was weaker and absent at the time of reward delivery after learning. We also found that inhibiting DA activity in VTA-NAc neurons during autoshaping using an inhibitory DREADD might lead to increased goal-tracking behavior during learning and potentiation of sign-tracking behavior once inhibition is lifted; however, this research is complicated by recent debate over the mechanism and specificity of the small molecule actuator of DREADD receptors. Overall, these results validate recent proposed computational hypotheses, opening new perspectives on the understanding of inter-individual differences in Pavlovian conditioning and DA signaling.

Introduction

All experiments discussed thus far have employed a complex combined approach and avoidance behavioral task to uncover neural correlates of reinforcement learning, where rats were required to make an operant behavioral response to avoid punishment or gain reward. One of the difficulties in uncovering the neural underpinnings of avoidance behavior is that the behavioral sequence itself is complex and multifaceted, involving both Pavlovian and instrumental conditioning components. A stimulus that predicts something aversive, like shock, elicits a Pavlovian defense response in the animal, causing the animal to freeze, which by its nature disrupts instrumental avoidance. To successfully avoid punishment, the animal must overcome this initial Pavlovian response to perform an action, like shuttling to a safe zone or pressing a lever. Individual differences in avoidance performance, like those that emerge between good and poor avoiders, might stem from how the Pavlovian stimulus is processed within the brain. While our combined approach and avoidance task is useful for studying the development individual differences that arise during learning in a more complex environment, it does not allow us to easily parse or manipulate potential causal mechanisms.

To address this issue, the final two experiments presented here employed a simple and well-understood Pavlovian autoshaping task to better understand individual differences in learning in a more controlled system. Pavlovian autoshaping is a simple procedure where a cue (in this case, the extension of a lever into the test chamber) predicts that reward will soon be delivered to an adjacent food cup. Recently, a breadth of studies have started employing this task to uncover the role of

the dopamine system in learning and how individuals differ in the way they process conditioned stimuli.

During Pavlovian autoshaping tasks, different behavioral responses tend to develop with learning, despite identical training parameters. For instance, in a typical Pavlovian autoshaping task, a cue (e.g., a lever) is presented to the animal followed by a reward (e.g., a food pellet) after some delay; interaction with the cue has no effect on the outcome and food is delivered regardless of the animal's behavior. However, despite identical training in identical environments, rats undergoing this task can learn to exhibit disparate behavioral patterns: some rats, known as sign-trackers (STs), learn to rapidly approach and engage the CS lever, whereas other rats, known as goal-trackers (GTs), learn to approach and enter the food cup upon presentation of the CS lever.

It is thought that these behavioral differences may arise due to differences in dopamine transmission. Khamassi and colleagues (Lesaint et al., 2014) recently proposed a new computational model – the "STGT model" (for Sign-Tracking and Goal-Tracking) – which accounts for a large set of behavioral, physiological and pharmacological data obtained from studies investigating individual variation in Pavlovian conditioned approach behavior (DiFeliceantonio & Berridge, 2012; Flagel et al., 2011; Flagel, Akil, & Robinson, 2009; Flagel, Watson, Robinson, & Akil, 2007; Mahler & Berridge, 2009; T. E. Robinson & Flagel, 2009; Saunders & Robinson, 2012). Most notably, the model can account for recent work by Flagel et al. (2011) that has called into question the classic hypothesis that phasic dopamine

release corresponds to a reward prediction error signal arising from a classical model-free system (Flagel et al., 2011).

In their experiments, Flagel and colleagues trained rats on a classical autoshaping procedure where the presentation of a retractable-lever conditioned stimulus (CS; 8 seconds) was followed immediately by delivery of a food pellet (unconditioned stimulus; US) into an adjacent food cup. Although both sign- and goal-trackers learn the CS-US relationship equally well, it has been shown that phasic dopamine release in the nucleus accumbens core (NAc) matches reward prediction error (RPE) signals only in STs (Flagel et al., 2011). Specifically, during learning in ST rats, DA release to unexpected reward decreases while DA release to the CS increases. In contrast, even though GTs acquire a Pavlovian conditioned approach response, DA release to reward does not decline, and CS-evoked DA is much weaker. Further, systemic administration of a DA antagonist flupenthixol (also known as flupentixol) blocked the acquisition and performance of sign-tracking behaviors, but had no effect on the acquisition or performance of goal-tracking behaviors (Danna & Elmer, 2010; Flagel et al., 2011).

The STGT computational model accounts for these data by attributing different weights to model-free (MF) and model-based (MB) reinforcement systems during conditioning (Lesaint et al., 2014). Further, this model suggests that GTs revise the value of the food cup multiple times during trials and during the 90 s inter-trial interval (ITI). During the trial, the food cup gains value since reward is delivered to this location; however, visits to the food cup during the ITI would not result in reward and, thus, would progressively reduce the value assigned to the food cup.

Importantly, all animals visit the food cup more often during the ITI, since the CS is not present at this time. However, only goal-trackers will also visit the food cup during the trial when the CS is present, which counteracts this downward revision, since the CS is maintaining the expectation of reward at this time. This mechanism is thought to prevent the gradual transfer of reward value signals from the US to the CS and, hence, could explain the absence of a normal DA RPE pattern in goal-trackers. Alternatively, visiting the food cup only during the ITI and not during the trial would lead to the downward revision of food cup value, and, as a result, the animal should visit the food cup less in subsequent trials and exhibit more sign-tracking behaviors in the future.

This model also predicts that changing the length of the ITI would systematically shift ST/GT behavioral responses and corresponding DA responses. Specifically, decreasing the ITI should reduce the amplitude of US DA bursts (i.e., rats have less time to negatively revise the value of the food cup and, thus, the size of the RPE would decrease); resultant higher food cup value should lead to an increase in the tendency to GT in the overall population. Alternatively, lengthening the ITI should increase the amplitude of US DA bursts (i.e., rats have more time to negatively revise the value of the food cup and, thus, the size of the RPE would increase); resultant lower food cup value should lead to an increased tendency to ST, accompanied by a large phasic DA response to the highly salient lever-CS. Here, we tested these predictions by recording DA release in NAc core using fast-scan cyclic voltammetry (FSCV) during 10 days of Pavlovian conditioning in rats that either had a short ITI of 60 s or a long ITI of 120 s.

This model predicts that DA is indeed critical for both sign- and goal-tracking, but possibly to varying degrees. Thus, altering DA release within this system should modify the development and expression of sign- and goal- tracking behaviors. However, as mentioned previously, Flagel et al. demonstrated that systemically inhibiting DA activity using a dopamine antagonist flupenthixol blocked the acquisition of sign-tracking, but not goal-tracking, behavior (Flagel et al., 2011). Flupenthixol can also inhibit serotonin, specifically via the $5HT_{2A}$ receptor subtype, as well as adrenergic and m-type acetylcholine receptors, which could potentially complicate these results ("Flupentixol," 2005). Further, since the administration of flupenthixol in this experiment was systemic, it could have impacted several downstream systems in addition to DA release in NAc. Importantly, this experiment formed the crux of their argument that DA is critical for sign-tracking but not goal tracking, in that they only found shifts in DA release from reward delivery to cue, typical of a traditional RPE signal, in sign-trackers but not goal- trackers. Thus, in the next experiment, we aimed to reproduce their results using a more targeted approach.

Based on these combined data, we predicted that inhibiting mesolimbic DA activity would slow down or stop the development of sign-tracking behaviors, enhance goal-tracking behaviors, or some combination of these effects. To test this prediction, we used a recently developed pharmacogenetic technique, Designer Receptors Exclusively Activated by Designer Drugs (DREADDs). This technique utilizes a viral vector to insert either an excitatory or inhibitory receptor and fluorescent marker into a brain area of interest; vectors can contain retrograde or anterograde tracers and can also be linked to certain neural pathways by utilizing

floxed segments. Once the virus has been fully expressed, these receptors can only be activated using an otherwise biologically inert pharmaceutical drug (e.g., clozapine-n-oxide or CNO). Theoretically, this puzzle-piece combination allows for the specificity of the technique. The onset and degradation of CNO is also fairly well-characterized, making this a non-permanent inhibition with some behaviorally relevant temporal specificity; after CNO administration, CNO remains active within the system for approximately 6 hours. Here, we inserted a retrograde viral vector (CAV-2) containing an inhibitory hM4Di receptor and mCherry fluorescent tag into the nucleus accumbens core of wild-type or transgenic Th-Cre rats, in order to specifically and temporarily inhibit DA neurons projecting from VTA to NAc core during the acquisition and expression of sign- and goal- tracking behaviors in an autoshaping task.

Materials and Methods

Animals. Animals used for the fast-scan cyclic voltammetry (FSCV) experiments included twenty-nine male Sprague-Dawley rats that were obtained from Charles River Labs at 250-275g (90-120 days old). Animals were individually-housed in a temperature- and humidity-controlled environment and kept on a 12-h light/dark cycle (0700-1900 in light); all tests were run during the light phase. Animals had access to water *ad libitum* and body weight was maintained at 85% of baseline weight by food restriction (15g standard rat chow provided daily, in addition to approximately 1g sucrose pellets during experimental trials). All procedures were performed in concordance with the University of Maryland, College Park Institutional Animal Care and Use Committee (IACUC) protocols.

Animals for experiments using Designer Receptors Exclusively Activated by Designer Drugs (DREADD) included 46 (10 wild-type, 36 TH-Cre+) male Long Evans rats (300-350g) that were housed in pairs in a temperature- and humidity-controlled environment. All subjects were genotyped and microchipped prior to surgery and study initiation. All animals were kept on a 12-h light/dark cycle (0700-1900 in light) and all tests were run during the light phase. Animals had access to water *ad libitum* and were food restricted three days prior to behavioral procedures; weight was maintained at approximately 90% of baseline weight throughout behavioral training (15g standard rat chow provided daily, in addition to approximately 1g grain pellets during experimental trials). All animals were handled for five days prior to the beginning of experimental procedures. All procedures were performed in accordance with National Institutes of Health guidelines and the University of Maryland, College Park Institutional Animal Care and Use Committee (IACUC) protocols; experiments were also in agreement with French (council directive 2013-118, 1 February 2013) and international (directive 2010-63, 22 September 2010, European Community) legislations and received approval # 5012053-A from the local Ethics Committee.

Viral vector. A CRE-recombinase expressing canine adenovirus (CAV-2) vector carrying the inhibitory hM4Di designer receptor exclusively activated by designer drugs (DREADDs) was obtained from University of North Carolina Vector Core (Chapel Hill, NC), specifically CAV-2-DIO-hM$_4$D(G$_i$)-mCherry (Armbruster, Li, Pausch, Herlitze, & Roth, 2007; Rogan & Roth, 2011). CAV-2 retrogradely infects projection neurons, which allowed us to specifically express hM$_4$D(G$_i$)-

mCherry in neurons that project from the ventral tegmental area (VTA) to the nucleus accumbens (NAc) core. Virus stored at -80°C in 20µL aliquots prior to use. The exogenous ligand, clozapine-N-oxide (CNO; Enzo Life Sciences) was dissolved in 0.9% saline containing 0.5% of dimethyl sulfoxide (DMSO; Sigma) to obtain a final concentration of 1 mg/ml. CNO was injected intraperitoneally (1 mg/kg) at least 45 min before behavioral training in a room separate from the animal facility and behavioral test room.

Chronic FSCV microelectrode fabrication. FSCV electrodes were constructed according to the methods of Clark *et al*. (Clark et al., 2010). A single carbon fiber (Goodfellow Corporation) was inserted into a 15 mm cut segment of fused silica (Polymicro Technologies) while submerged in isopropyl alcohol. One end of the silica tubing was sealed with a two-part epoxy (T-QS12 Epoxy, Super Glue) and left to dry overnight, leaving untouched carbon fiber extending past the seal. The protruding carbon fiber was cut to a length of 150 µm. A silver connector (Newark) was secured to the carbon fiber at the opposing end of the silica tubing using silver epoxy (MG Chemicals) and was allowed to dry. A final coat of two-part epoxy was then applied to the pin connection to provide insulation and structural support for the electrode and was allowed to dry overnight.

Intra-cranial surgical procedures. For FSCV experiments, all animals were anesthetized using isoflurane in O_2 (5% induction, 1% maintenance) and implanted with a chronic voltammetry microelectrode aimed at the NAc core (+1.3 AP, +1.8 ML, -6.6 DV), an ipsilateral bipolar stimulating electrode (Plastics One) in the medial forebrain bundle (-2.8 AP, +1.7 ML, -8.5 DV), and a contralateral Ag/AgCl reference

electrode (Sigma-Aldrich). The reference electrode and anchoring screws were stabilized using a thin layer of dental cement (Dentsply), leaving the holes for the stimulating and recording electrodes unobstructed. The stimulating and recording electrodes were attached to a constant current isolator (A-M Systems) and voltammetric amplifier, respectively, and lowered to the most dorsal point of the target region (-6.6 DV for the working electrode and -8.5 DV for the stimulating electrode). At this depth, a triangular voltammetric input waveform (-0.4 to +1.3 V vs. Ag/AgCl, 400 V/s; Heien et al., 2003) was applied to the recording electrode at 60 Hz for 30 minutes and then reduced to 10 Hz for the remainder of the surgery. Electrical stimulation (24 biphasic pulses, 60 Hz, 120 μA) was applied to the stimulating electrode in order to evoke dopamine release, which was monitored at increasing depths by the recording electrode. If neither an evoked change in DA nor a physical response (whisker movement or blinking) was observed, the stimulating electrode was lowered by 0.05mm until a response was achieved or to a maximum depth of 8.8mm. The working electrode was then lowered by 0.05mm until DA release was observed or to a maximum depth of 6.9mm. Once electrically-evoked DA release was detected in the NAc core, a thin layer of dental cement was used to secure the stimulating and recording electrodes in place. A Ginder implant (Ginder Scientific; constructed in house) was connected to the reference, stimulating, and recording electrodes and fully insulated using dental cement, leaving only the screw-top connector exposed, in order to reduce noise and prevent loss of connectivity during behavioral training. Animals then received post-operative care: subcutaneous injection of 5 mL saline containing 0.04 mL carprofen (Rimadyl), topical application

of lidocaine cream to the surgical area, and placement on a heating pad until full consciousness was regained. Animals were also given antibiotic treatment of Cephlexin orally twice daily post-surgery for one week to prevent infection of the surgical site. All subjects were allowed a month for full recovery and stabilization of the electrode before experimentation.

For DREADD microinfusion surgeries, all animals were anaesthetized using isoflurane in air (5% induction; 1.5% maintenance) and underwent microinfusion of CAV-2-DIO-hM4D(Gi)-mCherry viral vector bilaterally within the NAc core (+1.4 AP, ±1.7 ML, -6.8 DV) using a stereotaxic apparatus (Kopf). All coordinates are given in millimeters from bregma (Paxinos & Watson, 2007). Rats were subcutaneously injected with 0.05 mg/kg buprenorphine (Buprècare) and the incision site was treated with the local anesthetic xylocaine. The viral vector was infused using microinfusion pump and 10µL syringe (Micropump4 controller and UMP3 UltraMicroPump, World Precision Instruments) at a rate of 200nL/minute for a total of 5 minutes (1uL total) for each hemisphere. The syringe tip of the microinfusion pump remained at DV depth for an additional 5 minutes after infusion was completed to allow full diffusion of the virus and to prevent backflow. This equates to 10 minutes total for each hemispheric injection. Patency of the pump syringe was tested before and after each infusion. After infusions were completed, the scalp was sutured and sprayed with Aluspray (Vétoquinol) to aid in healing. All animals were allowed at least 4 weeks to recover before the start of the behavioral procedures, which allowed sufficient time for viral infection, during which time they were monitored daily and weighed.

Behavioral Tasks. For FSCV experiments, all behavioral procedures were conducted in Med Associates test chambers equipped with a food-tray and a retractable lever located to the left or right of the food-tray (counterbalanced). Head entries into the food tray were timestamped during disruption of the photobeam located inside the receptacle. Similarly, timestamps were generated during downward deflection of the lever.

Three pre-training sessions were conducted that consisted of the delivery of 25 sucrose pellets, which were randomly delivered on a variable-interval 30 +/- 15 s schedule. Following pre-training, rats began Pavlovian training sessions which consisted of the presentation of the lever (CS) for 8 s, which was immediately followed with delivery of a sucrose pellet upon its retraction. The CS was presented on a random-interval of either 60 +/- 30 s (n = 7 rats) or 120 +/- 30 s (*n* = 11 rats) and each Pavlovian session consisted of 25 trials. Pavlovian training continued for 10 sessions which were accompanied with FSCV recording.

For DREADD experiments, we had three groups of interest that were gathered across two separate experiments; each group was labeled according to whether the rat was transgenic (TH+) or wild-type (WT) and whether they received CNO or vehicle (VEH) before autoshaping training. For our first round of experiments, we used all TH+ animals with approximately half receiving CNO (*n* = 12) and half receiving vehicle (*n* = 10) prior to training; for the second round of experiments, we used approximately half TH+ animals (*n* = 14) and half WT animals (*n* = 10) with all animals receiving CNO prior to training. This resulted in three groups of interest: an experimental group of TH+ animals that received CNO during training (*N* = 26), a

control group of TH+ animals that received vehicle during training (N = 10), and a control group of wild-type animals that received CNO during training (N = 10). At the start of behavioral training, all animals received 1 session of food cup training, in which they were confined to the operant chamber and a 45mg grain pellet was delivered to the food cup on a VT90 schedule until 25 pellets were received (39 minutes). The number of grain pellets consumed were recorded.

Once animals were familiar with the grain pellets used in the task, rats were trained on a Pavlovian autoshaping task for 16 days. On days 1-10, rats were injected intraperitoneally (i.p.) with either vehicle or CNO (1 mg/kg) at least 45 minutes prior to the start of each session; on days 11-16, no vehicle or CNO injection was administered. During each autoshaping session, a lever (left) was introduced into the operant chamber for 10 seconds prior to pellet delivery; after lever retraction, a 45mg grain pellet was delivered to the food cup. Acting on the lever had no consequence and did not affect pellet delivery or inter-trial interval (ITI) in any way. The ITI range was varied slightly every other day to prevent animals from tracking time. Each session consisted of 25 lever (CS+) and pellet pairings and lasted for 39 minutes.

Fast-scan cyclic voltammetry. For recordings, animals were connected to a head-mounted voltammetric amplifier (current-to voltage converter) and a commutator (Crist Instruments) mounted above the recording chamber. During each session, an electrical potential was applied to the recording electrode in the same manner as described above (see *Intra-cranial surgical procedures*). In order to detect changes in dopaminergic concentration over time, the current at its peak oxidation potential was plotted for successive voltammetric scans and background signal was

subtracted. Two PC-based systems, fitted with PCI multifunction data acquisition cards and software written in LabVIEW (National Instruments), were used for waveform generation, data collection, and analysis. The signal was low-pass filtered at 2,000Hz. Event timestamps from Med Associates were recorded, in order to analyze behaviorally relevant changes in dopamine release.

Dopamine was identified by its stereotypical and specific cyclic voltammogram signature. Behaviorally-evoked DA signals met electrochemical criterion if the cyclic voltammogram was highly correlated to that of the DA templates produced during the training set. The training set is a template extracted from each individual animal that contained six each of background-subtracted cyclic voltammograms and corresponding calibrated concentrations for both dopamine and pH acquired during electrical stimulations that are known to evoke DA release (stimulation at 1V: 30 Hz, 6 pulses; 30 Hz, 12 pulses; 30 Hz, 24 pulses; 60 Hz, 6 pulses; 60 Hz, 12 pulses; 60 Hz, 24 pulses). The data collected during a session were not analyzed if DA release did not satisfy these chemical verification criteria (e.g., Fig. 1D and E). Voltammetric data were analyzed using software written in LabView and Matlab. A principal component regression (Tar Heel CV chemometrics software) was used to extract the DA component from the raw voltammetric data (Heien et al., 2004, 2003; Keithley et al., 2009). Eigenvalues (principal components) are calculated that describe relevant components of our training set, and we perform multivariate regression analysis to determine a correlation coefficient to describe our recorded behavioral data versus the training set. The number of factors we select to keep in our PCA analysis accounts for >99% of the variance (at least 3, but usually 4-5 factors are

kept). Factor selection is a very important step, as retaining more factors than we need would add noise to our data but retaining too few could mean discarding potentially meaningful information (Kramer, 1998). Importantly, the exact same method was applied to both groups allowing for fair comparisons.

We also use the residual to examine the quality of the fit. In general, the residual is the difference between the experimental observation and the predicted value derived from a model/template (our regression values) and is a measure of the unknown portion of the signal that is not accounted for by the principal components of the regression. This is important when considering the accuracy and the applicability of the model and is important for identifying possible interfering molecules or noise (such as drift). The sum of squares of the difference between the template and the experimental data is the residual value (Q) and the threshold Qa establishes whether the retained principal components provide a satisfactory description of the experimental data; the discarded principal components should provide a measure of noise (Heien et al., 2003; Keithley et al., 2009; Kramer, 1998). We use this Qa measure in combination with our regression analysis to establish our concentration corrections.

Chemometrics is a widely-used analytical method that separates changes in current that are caused by DA release from those caused by pH shift or other electrochemical 'noise' by comparing eigenvalues derived from stimulated DA release and changes in pH to those derived from behavioral release (Joseph F. Cheer et al., 2007; Flagel et al., 2011; Heien et al., 2003; Keithley et al., 2009; Kramer, 1998; Phillips et al., 2003; Wightman et al., 2007). Once converted to concentrations,

DA release was examined over 3 analysis epochs (1) Baseline = 3 s before CS onset, (2) CS epoch = 3 s starting 1 s after CS onset, and (3) US epoch = 3 s starting 1 s after reward delivery (i.e., lever in).

Histology and Immunohistochemistry. Following the completion of the study, animals were terminally anesthetized with an overdose of isoflurane (5%) and transcardially-perfused with saline and 4% paraformaldehyde. Brain tissue was removed and post-fixed with paraformaldehyde. For FSCV experiments, brains were post-fixed using 4% paraformaldehyde solution at 4°C overnight then placed in 30% sucrose solution for 72 hr and sectioned coronally (50μm) using a freezing microtome; issue slices were then mounted onto slides and stained with thionin for histological reconstruction.

For DREADD experiments, brain tissue was post-fixed in 4% paraformaldehyde at 4°C overnight, and then 50 μm coronal sections were cut using a VT1200S Vibratome (Leica Microsystems) and stored in wells filled with phosphate buffer until further processing. Every fourth section was collected to form a series, and a double immunoreactivity experiment was performed for TH and mCherry labeling. Free-floating sections were prepared by rinsing four times in 0.1 M phosphate buffered saline (PBS) for 5 minutes each, incubated in a blocking solution (1 h, PBS 0.1 M, 0.2% Triton-X, 4% goat serum), and then placed in wells containing primary antibodies 1:1000 rabbit anti-RFP (red fluorescent protein; PM005, MBL International Corporation) and 1:5000 monoclonal mouse anti-TH (tyrosine-hydroxylase; Merck Millipore, MAB318) in blocking solution at 4°C for 48 h. After incubation with primary antibodies, sections were then rinsed four times in 0.1M PBS

for 5 minutes each and then incubated in 1:200 rhodamine TRITC goat anti-rabbit (Jackson Immunoresearch, 111-025-003) and 1:200 fluorescein FITC goat anti-mouse (Jackson Immunoresearch, 115-095-003) diluted in 0.1M PBS for 2 h at room temperature; plates were wrapped in aluminum foil to protect from the light. Sections were then washed 4 times for 5 minutes each in PBS, and then incubated with hoescht solution (1:5000 in 0.1M PBS) for 15 minutes. Sections were again rinsed four times for 5 minutes each with PBS 0.1M, and were then mounted, and cover-slipped with Fluoromount-G (SouthernBiotech). Sections were stored at 4°C away from the light and imaged using a Nanozoomer slide scanner (Hamamatsu Photonics) and analyzed with the NDP.view 2 freeware (Hamamatsu Photonics).

Statistical analyses. Behavior during performance of the autoshaping task was evaluated by measuring multiple behavioral parameters during the cue: average response bias, probability difference, latency, frequency of press, and frequency of food cup visits. Average response bias is calculated as (Lever Presses – Food Cup Entries) / (Lever Presses + Food Cup Entries). The probability difference is calculated as ($P_{lever} - P_{receptacle}$). Latency index is calculated as ($\overline{x}$ Cup Entry Latency – $\overline{x}$ Lever Press Latency) / 10 sec. The PCA index is an average calculated from the response bias, probability difference and latency difference indices described above. The ratios derived from these 4 equations range from -1.0 to +1.0, with more positive and negative for animals being categorized as sign trackers and goal trackers, respectively, and animals between -0.5 and 0.5 categorized as intermediate. The frequency of lever presses and frequency of food cup entries during the cue was also computed (# lever press/10 sec and # food cup visits/10 sec). Two-factor ANOVAs (*p*

< 0.05) were performed on these behavioral measures to determine if activity was modulated by group (experimental, control) and day block (learning, expression, no injection) or if there were any interactions between these factors. Post-hoc t-tests (p < 0.05) were then performed in order to uncover the nature of these effects.

Results

Longer ITIs increased sign tracking and DA release to the CS and US

Dopamine (DA) release was recorded from nucleus accumbens core (**Fig. 1B-E**) during a standard Pavlovian conditioned-approach behavior task (**Fig. 1A**) for 10 days. Each trial began with the presentation of a lever (CS) located to the left or right side of a food cup (counterbalanced) for 8 s. Upon the lever's retraction, a 45-mg sucrose pellet was delivered into the food cup, independent of any interaction with the lever. Each behavioral session consisted of 25 trials presented at a random time-interval of either 60 s (VT60; n = 7 rats) or 120 s (VT120; n = 11 rats). To quantify the degree to which rats engaged in sign- vs goal- tracking behavior, we used the Pavlovian Conditioned Approach (PCA) index (Meyer et al., 2012), which comprised the average of three ratios: (1) the response bias, which is (Lever Presses – Food Cup Entries) / (Lever Presses + Food Cup Entries), (2) the probability difference, which is ($P_{lever} - P_{receptacle}$), and (3) the latency index, which is ($\overline{x}$ Cup Entry Latency – $\overline{x}$ Lever Press Latency) / 8. All of these ratios range from -1.0 to +1.0 (similarly for PCA index) and are more positive or negative for animals that sign track or goal track, respectively. All behavioral indices are derived from sessions during which DA was recorded (**Fig. 1B-E**). For the initial analysis described in this section, behavior and

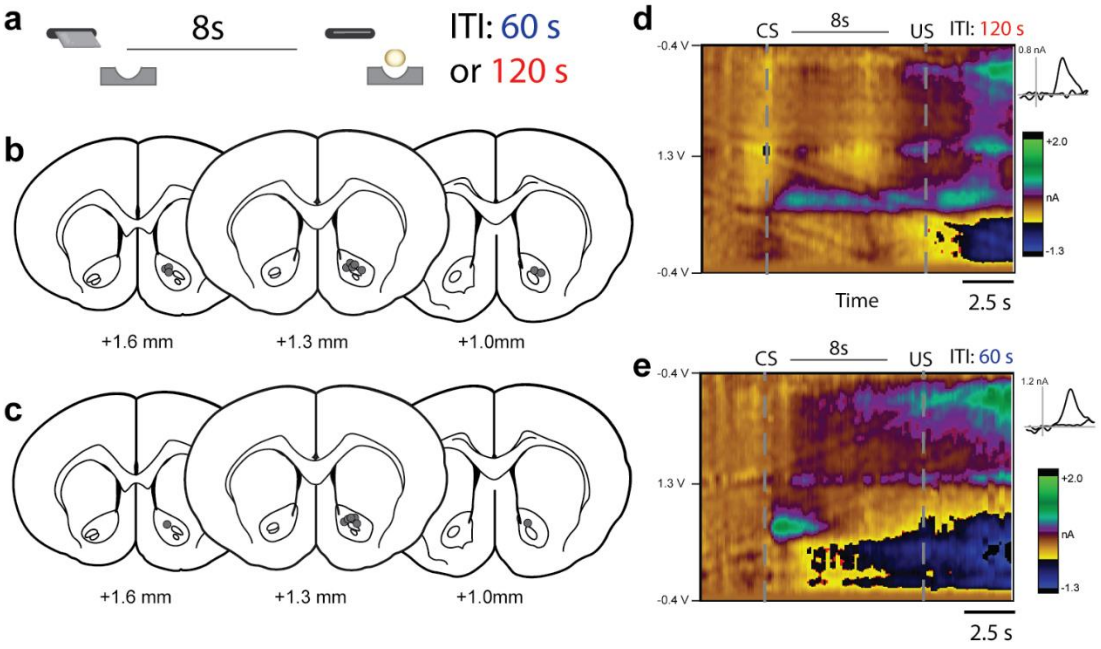


Figure 1. (**A**) DA release was recorded during a standard Pavlovian conditional approach behavior task for 10 days. Each behavioral session consisted of 25 trials presented at a random time-interval of either 60 s (+/-30; n = 7 rats) or 120 s (+/-30); n = 11 rats). (**B-C**) Placement of chronic recording electrodes within the NAc core based on histology for the 60 s (B) and 120 s (C) groups. (**D-E**) False-color plots indicate voltammetric current (z-axis) plotted against applied scan potential (y-axis) and time (x-axis) for example 120 s (D) and 60 s ITI trials.

DA were examined across all sessions; the development of behavior and DA over training is examined in later sections.

The distributions of behavioral session scores are shown in **Figure 2A-D** for each group. As predicted, rats with the 120 s ITI tended to sign track more, whereas rats with the 60 s ITI tended to goal track more. Across all behavioral indices (i.e., response bias, probability, latency, PCA), the mean distributions were significantly positively-skewed (biased toward sign-tracking) for rats in the 120 s ITI group (**Fig. 2A-D**, Left; Wilcoxon; μ's > 0.17, *p*'s < 0.05). An opposite trend was observed in the 60 s ITI group, in that all distributions were negatively shifted from zero (**Fig. 2A-D**, Right; Wilcoxon; response bias: μ = -0.06, *p* = 0.06; lever probability: μ = -0.03, *p* = 0.58; PCA index: μ = -0.11, *p* = 0.097); however, only the shift in the latency difference distribution reached significance (**Fig. 2C**, right; Wilcoxon; μ = -0.10; *p* < 0.05). Direct comparisons between 60 and 120s ITI groups produced significant differences across all four measures (Wilcoxons; *p*'s < 0.01). Thus, we conclude that lengthening the ITI indeed increased sign tracking behavior, as predicted by the STGT model (Lesaint et al., 2015, 2014).

Notably, the degree of sign- or goal- tracking within the 60 s ITI group was highly dependent upon when behavior was examined during the 8 s CS period. This is illustrated in **Figures 2G and H**, which show percent beam breaks in the food cup (solid lines) and lever pressing (dashed lines) across trial time. Consistent with the ratio analysis described above (**Fig. 2A-D**), rats in the 120 s ITI group (red) showed sustained pressing (red dashed) that started shortly after lever extension and persisted

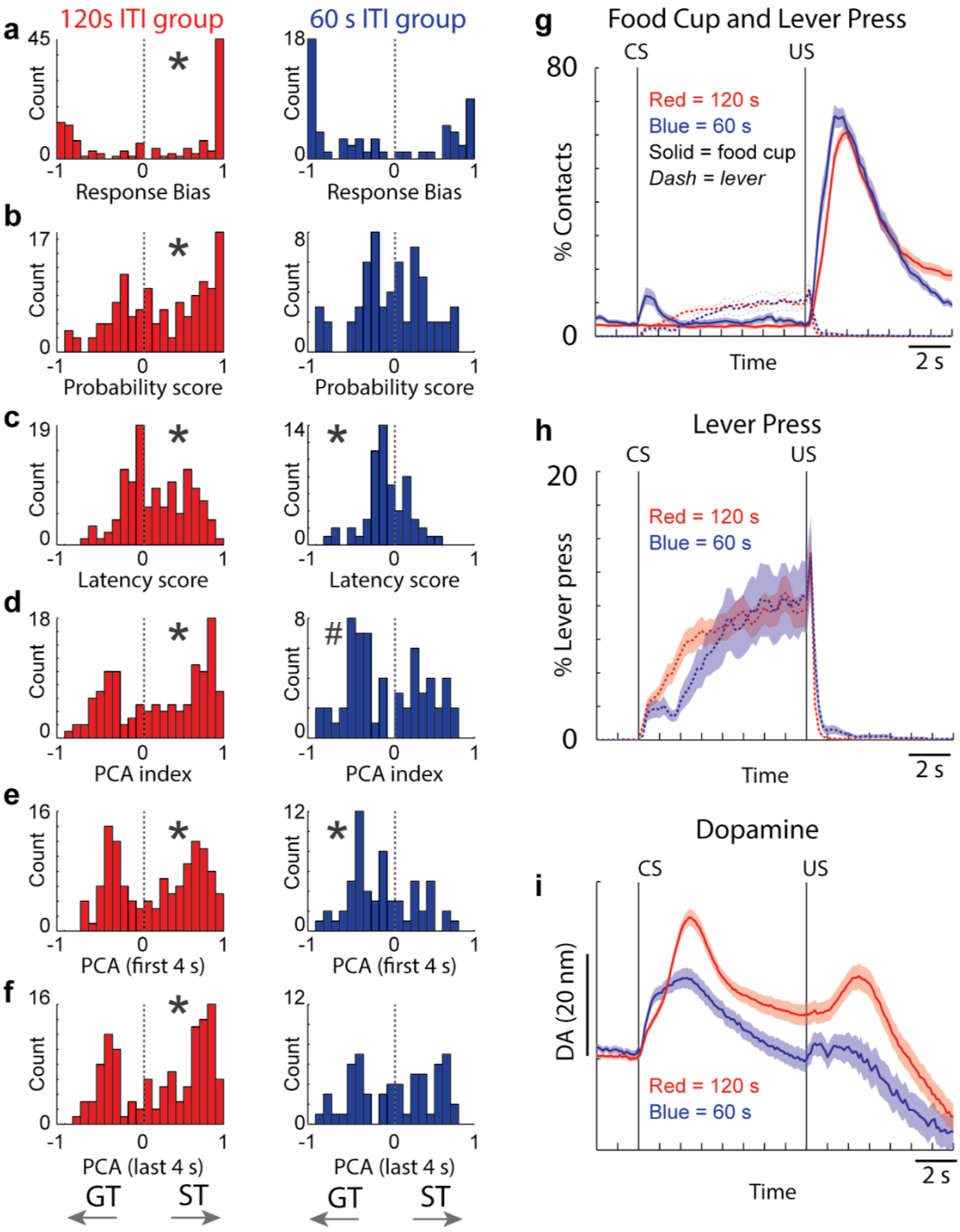


Figure 2. (**A**) Response bias, which is (Lever Presses – Food Cup Entries) / (Lever Presses + Food Cup Entries). (**B**) the probability difference, which is (Plever – Preceptacle). (**C**) Latency index, which is ($\bar{x}$ Cup Entry Latency – $\bar{x}$ Lever Press Latency) / 8. (**D**) PCA index = average of response bias, probability difference and latency difference indices described in A-C. All of these ratios range from -1.0 to +1.0 and are more positive and negative for animals that sign track and goal track, respectively. All behavioral indices are derived from sessions during which DA was recorded (60 s ITI groups = 7 rats; 120 s ITI group = 11 rats) and used behavior

during the entire 8 s CS epoch. Each of the above ratios was computed by session. (**E-F**) PCA index computed using just the first 4s (E) and last 4 s (F) of the CS period. 120 s ITI group = red; 60 s ITI group = blue. (**G**) Average beam break (solid) and lever press (dashed) rate for 60 s (red) and 120 s (blue) ITI sessions. (**H**) Average lever press rate for 60 s (red) and 120 s (blue) ITI sessions. Data is the same as in G but with a smaller scale so that differences and timing can be better visualized. (**I**) Average DA release over time for 60 s (red) and 120 s (blue) ITI sessions. Error bars represent S.E.M.

throughout the 8 s CS period, while showing no increase in food cup entries (red solid) after CS presentation (**Fig. 2G**, red solid vs. dashed).

Although it is clear that rats in the 120 s ITI group sign track more than goal track during the CS period, the relationship between lever pressing and food cup entry was far more dynamic during sessions with 60 s ITIs (**Fig. 2G**; blue). During 60 s ITI sessions, rats would briefly enter the food cup for approximately 2 s immediately upon CS presentation (**Fig. 2G**, solid blue) before engaging with the lever (**Fig. 2G**, dashed blue). As a result, lever pressing was delayed in the 60 s ITI groups relative to the 120 s ITI group (**Fig. 2G and H**; blue versus red dashed). This suggests that the goal-tracking tendencies described above during the entire 8 s CS period were largely due to the distribution of behaviors observed early in the CS period. To quantify this observation, we recomputed the PCA index using either the first or the last 4 seconds of the 8 s CS period. For the 120 s ITI group, the PCA index was significantly shifted in the positive direction during both the first and last 4 seconds of the cue period (i.e., more sign tracking; **Fig. 2E and F**, Left; Wilcoxon; μ's > 0.16; p's < 0.05). For the 60 s ITI group, the PCA index was significantly shifted in the negative direction during the first 4 seconds (i.e., more goal-tracking; **Fig. 2E**, Right; Wilcoxon; $\mu = -0.16$; $p < 0.05$), but not during the last 4 seconds (**Fig. 2F**, Wilcoxon; $\mu = 0.01$; $p = 0.81$).

The behavioral data described above support the model predictions that increasing and decreasing the ITI would produce more and less sign-tracking, respectively. Next, we tested the prediction that longer ITIs would elevate DA release to the CS and the US. **Figure 2I** shows average DA release across all sessions for the

60 s and 120 s groups. Rats in the 120 s ITI group exhibited significantly higher DA release to the CS and the US relative to rats in the 60 s ITI group (CS: $t = 2.99$, df = 178, $p < 0.05$; US: $t = 3.07$, df = 178, $p < 0.05$). In the 120 s ITI group, DA release to both the CS and the US was significantly higher than baseline (CS: $t = 14.77$, df = 119, $p < 0.05$; US: $t = 4.79$, df = 119, $p < 0.05$); in the 60 s ITI group, this was only true during CS presentation ($t = 7.34$; df = 59; $p < 0.05$) but not at US delivery ($t = 0.99$; df = 59; $p = 0.33$). These results are in line with the STGT model, which predicts that reducing ITI duration would prevent the downward revision of the food cup value during the ITI and, hence, the high predictive value associated with the food cup would drive a high amplitude DA response to the CS, consistent with the dopamine reward prediction error hypothesis (W Schultz et al., 1997). Conversely, increasing ITI duration resulted in higher DA release during both CS presentation and US delivery; these results are also consistent with model predictions, which suggests that more positive reward prediction errors may result from positive surprise associated with receiving reward in a low-valued food cup. Together, these results suggest that the absence of a DA RPE pattern in GTs observed by Flagel et al. does not necessarily mean that these animals do not use such model-free reinforcement learning (MFRL) mechanisms (Flagel et al., 2011). Instead, in the STGT model, both STs and GTs have MFRL mechanisms for computing RPE signals and such a RPE signal may be restored in GTs by manipulating the ITI to reduce the downward revision of food cup value, as performed here.

Development of sign tracking and DA signals over training

In the analysis above, we averaged DA release and behavior from all recording sessions. Next, we asked how behavior and DA release patterns evolved with training. As a first step to addressing this issue, we recomputed the PCA analysis for the first and last 5 days of training. For the 60 s ITI group, the PCA index distribution was significantly shifted in the negative direction (i.e., goal-tracking) during the first 5 sessions (Wilcoxon; $\mu = -0.38$; $p < 0.05$) but not in the last 5 sessions (Wilcoxon; $\mu = 0.15$; $p = 0.07$). Thus, early in training, rats with the 60 s ITI exhibited goal-tracking more than sign-tracking but did not fully transition to sign-tracking, at least when looking over the last 5 sessions. For the 120 s ITI group, the PCA index was significantly shifted in the positive direction (i.e., sign-tracking) during the last 5 sessions (Wilcoxon; $\mu = 0.28$; $p < 0.05$) but was not shifted during the first 5 sessions (Wilcoxon; $\mu = 0.10$; $p = 0.11$). Thus, when the ITI was long (120 s), rats exhibited sign- and goal- tracking in roughly equal proportions during the first 5 sessions, but tended to sign track significantly more during later sessions.

To more accurately pinpoint when during training rats in the 120 s group shift toward sign-tracking, we examined the four distributions individually for each session. Sign-tracking became apparent during session 4, when the latency and lever probability distributions first became significant (Wilcoxon; latency: $\mu = 0.28$, $p < 0.05$; lever probability: $\mu = 0.40$, $p < 0.05$). To visualize changes in behavior and dopamine release around this timepoint, we plotted food cup beam breaks, lever pressing, and DA release averaged across the first 3 days of training and across days 4-10 (**Fig. 3**). Consistent with the behavioral distributions described above (first 5 and

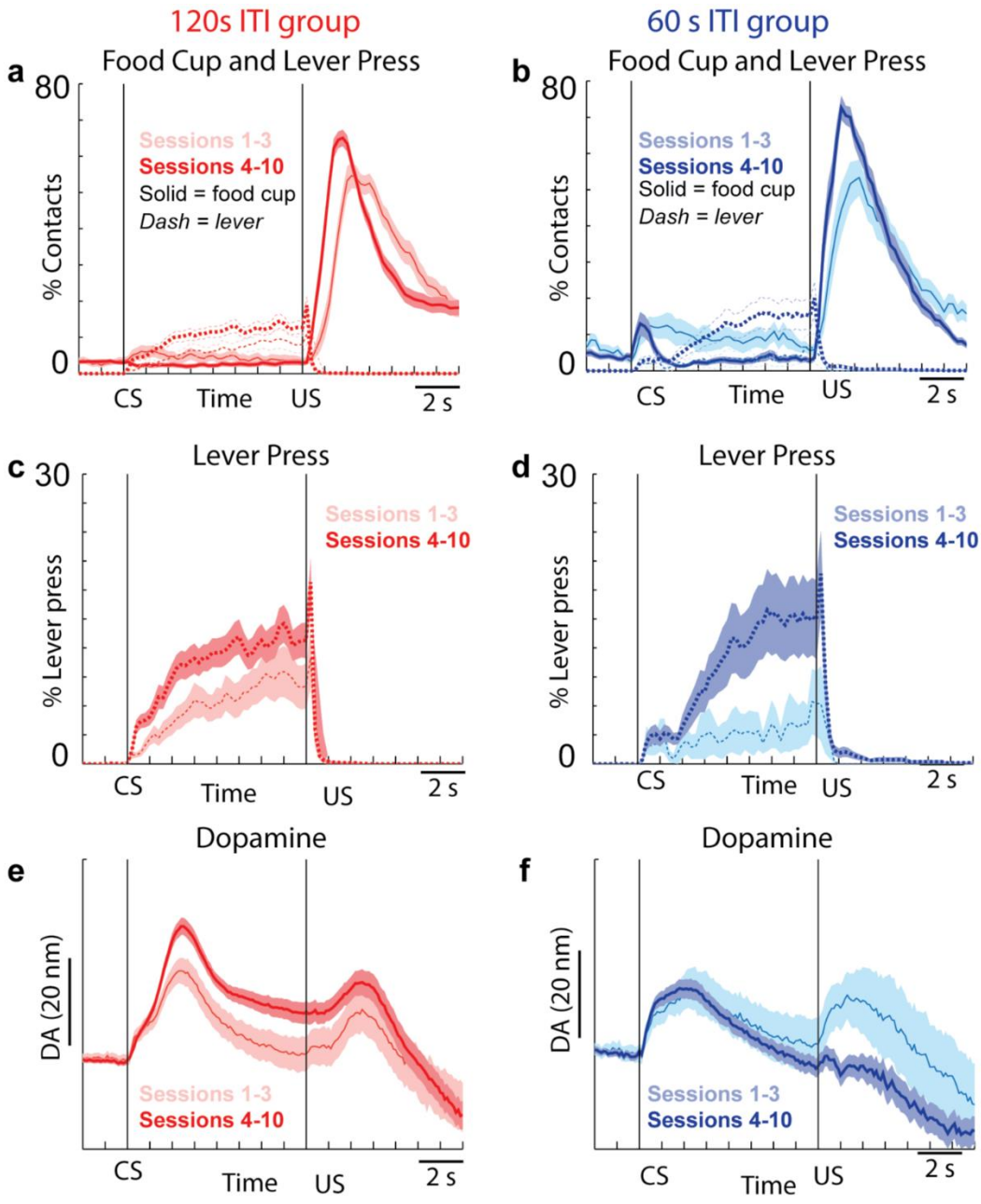


Figure 3. (**A-B**) Average beam break (solid) and lever press (dashed) rate for 60 s (A) and 120 s (B) ITI sessions. (**C-D**) Average lever press rate for 60 s (C) and 120 s (D) ITI sessions. Data is the same as in A and B but with a smaller scale so that differences and timing can be better visualized. (**E-F**) Average DA release over time for 60 s (E) and 120 s (F) ITI sessions. In each of the above (A-F), data is broken down into averages from sessions 1-3 [pale colors; pink (120 s) and turquoise (60s)] and sessions 4-10 [dark colors; red (120 s) and blue (60 s)]. 60 s ITI groups = 7 rats; 120 s ITI group = 11 rats. Error bars represent S.E.M.

last 5 sessions), the 120 s ITI group showed roughly equal food cup entries and lever pressing during the CS period in the first 3 days of training (**Fig. 3A**, thin pink solid vs. thin pink dashed), whereas later in training (days 4-10; red) there was a strong preference for the lever (**Fig. 3A**; thick red dashed versus thick red solid). Indeed, the distribution of PCA indices averaged during days 4-10 were significantly shifted in the positive direction (Wilcoxon; $\mu = 0.27$; $p < 0.05$).

These results suggest that sign-tracking tendencies developed relatively quickly during the first several recording sessions of sessions with a 120 s ITI (**Fig. 3A and C**). This is consistent with the STGT model, which predicts that increasing the ITI duration would increase the global tendency to sign-track within the population and would, thus, speed up the acquisition of lever pressing behavior (Lesaint et al., 2015, 2014). In contrast, the model also predicts that reducing the ITI duration would increase the global tendency to goal-track and would thus slow the acquisition of lever pressing behavior. Interestingly, the behavior of the 60 s ITI group was far more complicated than behavior of the 120s group, with changes in goal- and sign-tracking occurring across training and CS presentation time. Early in training, rats in the 60 s ITI group clearly visited the food cup (**Fig. 3B**; solid turquoise) more than they pressed the lever (**Fig. 3B**; dashed turquoise); food cup entries increased shortly after presentation of the CS and continued throughout the CS period (**Fig. 3B**; solid turquoise). During later sessions (i.e., 4-10), rats in the 60 s ITI group still entered the food cup upon CS presentation – corresponding to goal-tracking behavior predicted by model – but food cup entry only lasted about 2 s, at

which point they transitioned to the lever (**Fig. 3B and D**). In sessions 4-10, none of the behavioral distributions were significantly shifted from zero during the total CS period (Response bias: $\mu = 0.27$, $p = 0.83$; Latency: $\mu = -0.05$, $p = 0.13$; Probability: $\mu = 0.08$, $p = 0.16$; PCA: $\mu = 0.02$, $p = 0.82$) or during the first half of the CS period (Response bias: $\mu = -0.11$, $p = 0.027$; Probability: $\mu = -0.04$, $p = 0.18$; PCA: $\mu = -0.07$, $p = 0.25$); however, when examining the last 4 s of the CS period, distributions were significantly shifted in the positive direction (Response bias: $\mu = 0.32$, $p < 0.05$; Probability: $\mu = 0.28$, $p < 0.05$; PCA: $\mu = 0.24$, $p < 0.05$). Together, this suggests that rats in the 60 s groups were largely goal-trackers early in training, and this goal-tracking behavior could still be seen later in training during the early portion of the CS period, while sign-tracking behavior developed toward the end of the CS period.

Behavioral analyses clearly demonstrate that manipulation of the ITI impacts sign- and goal-tracking behavior and that both groups still learned that the lever predicted reward. Next, we determined how DA patterns changed throughout training. **Figures 3E and F** illustrate DA release averaged across the first 3 days and days 4-10 of sessions with 120 s and 60 s ITIs, respectively. As we have shown previously, both groups started in days 1-3 with modest DA release to both the CS and US. For the 120s ITI group, who would become predominately sign-trackers, DA release was significantly higher to CS presentation later (red) compared to earlier (pink) in learning (**Fig. 3e**; $t = 2.51$; df = 119; $p < 0.05$). This increase of DA to the CS, like that seen in sign-trackers by Flagel and colleagues, is consistent with the reward prediction error hypothesis (Flagel et al., 2011). DA release during US delivery did not significantly differ between early and late phases of training ($t =$

1.27; df = 119; $p = 0.21$). Expanding on the work of Flagel et al., the increase in time available to down-regulate the value associated to the food cup during the ITI may have resulted in a positive surprise at the time of reward delivery, hence preventing the progressive decrease of the DA response to the US across training (Flagel et al., 2011).

In the 60 s ITI group (**Fig. 3F**), who we interpreted as predominately goal-trackers due to their early behavioral response to the CS (**Fig. 2G and 3B**), DA release to the US was initially high during the first 3 days (turquoise) but declined during days 4-10 (blue). Directly comparing DA release during the first 3 days to the remaining days revealed significant differences during the US period ($t = 1.14$; df = 59; $p < 0.05$), but not the CS period ($t = 0.08$, df = 59; $p = 0.93$). This DA pattern, with high DA release to the CS but not the US (**Fig. 3F**, blue), resembles the traditional RPE pattern observed only in sign-trackers by Flagel and colleagues (Flagel et al., 2011). This is a clear demonstration that the dopamine RPE signal can be restored in goal-trackers by manipulating the ITI, as predicted by the STGT model.

Inhibition of VTA to NAc DA neurons during Pavlovian conditioning may increase goal-tracking during learning and potentiate post-learning sign-tracking behavior

The following experiment was carried out as a part of an international collaboration with the lab group of Etienne Coutureau (Decision and Adaptation; DECAD) at the University of Bordeaux within the Institut de Neurosciences Cognitives et Integratives d'Acquitaine (INCIA) of Le Centre National de la Recherche Scientifique (CNRS). This endeavor would have not been possible without

funding and support from the Neuroscience and Cognitive Sciences (NACS) program research training award.

The above results suggest that DA release in the NAc core is critical for both sign- and goal- tracking and that simple manipulations of the ITI can alter DA prediction error encoding and behavior. These results both explain the absence of classic prediction errors signals in goal-trackers described by Flagel and confirm STGT model predictions (Flagel et al., 2011). However, this evidence supporting the computational model is correlational. It still remains unclear if mesolimbic DA (VTA to NAc) connections are critical for the development and expression of sign- and goal-tracking behavior.

To address this issue, I injected a retrograde CAV-2 viral vector containing sites for inhibitory DREADD receptors and a fluorescent mCherry marker in the NAc core of wild type (N = 10) and transgenic (N = 36) rats, with the goal of specifically expressing engineered inhibitory DREADD receptors in VTA to NAc TH+ neurons (**Fig. 4A**). After sufficient viral expression (4 weeks; **Fig. 4B**), rats were then trained on a traditional Pavlovian autoshaping task for 16 consecutive days (**Fig. 4C**). On each of the first 10 days, 45 minutes prior to the start of the behavioral session, rats were injected interperitoneally (i.p.) with either clozapine-n-oxide (CNO) or vehicle, dependent upon genotype and experimental design; this provided us with 1 experimental group (n = 26, TH+ rats injected with CNO) and 2 control groups (n = 10, TH+ rats injected with vehicle; n = 10, WT rats injected with CNO). During each behavioral session, a lever (CS) was introduced into the operant chamber for 10 seconds; after lever retraction, a 45mg grain pellet was delivered into an adjacent

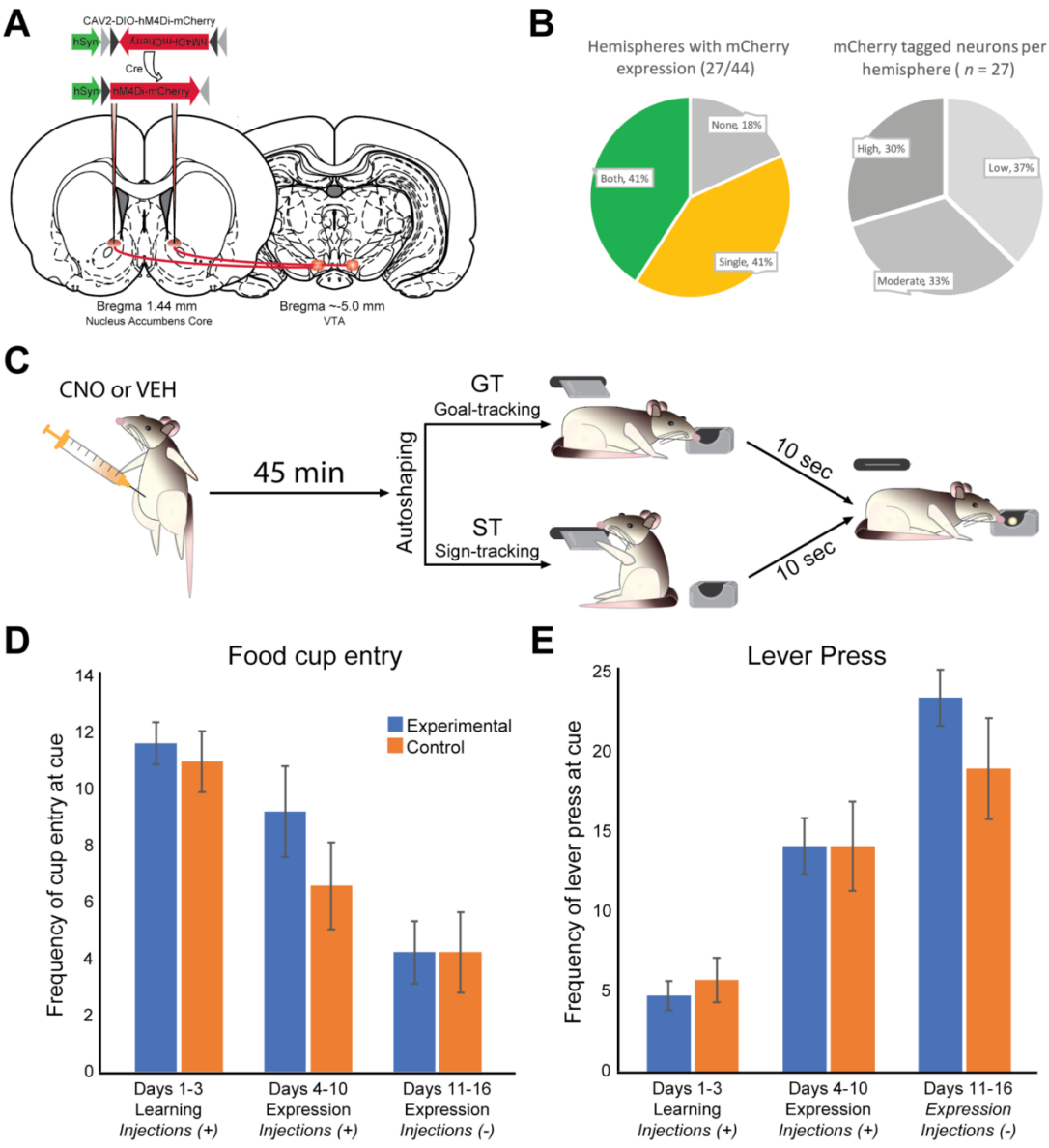


Figure 4. (**A-B**) All animals (N = 46) received bilateral microinjections of the retrograde CAV2-DIO-hM4Di-mCherry viral vector into the NAc core (A). Injection sites within the NAc core and infection specificity of the viral vector to VTA-NAc neurons was verified using a dual immunohistochemical technique (mCherry + TH) (B). Pie charts show group data for mCherry expression levels for half of animals included in the study (N = 22; second group of N = 24 still under analysis in Bordeaux, France). (**C**) After sufficient time for viral expression and recovery (4 weeks), wild type (N = 10) or TH+ animals (N = 36) were injected with either CNO (1mg/kg) or vehicle at least 45 minutes prior to training in a standard Pavlovian conditional-approach behavior task. Rats were trained on this task for 16 days (10 with injections, 6 without injections). Each behavioral session consisted of 25 trials that consisted of the extension of a lever into the chamber for 10 s, followed by pellet

delivery into an adjacent food cup; trials were presented at a variable time-interval of 90 s. Behavior from the animal had no effect on food pellet delivery or ITI timing. (**D-E**) Frequency of food cup entry (D; average food cup entries during the cue presentation) and Frequency of lever press (E; average lever presses during cue presentation) for Experimental (TH+ animals injected with CNO, n = 26; blue) and Control (TH+ animals injected with vehicle or wild-type animals injected with CNO, n = 20; orange) across three different time periods: Learning (Days 1-3; Injections on board), Expression (Days 4-10; Injections on board), and Late Expression (Days 11-16; No injections).

food cup (US). A response on the lever or contact with the food cup had no effect on the outcome (receipt of pellet into the food cup). Each autoshaping session consistedof 25 lever (CS+) and pellet pairings, and each trial was followed by a variable time ITI (VT90); this is the same ITI that was previously used by Flagel (Flagel et al., 2011).

Our overall goal was to determine the effect of specific dopaminergic inhibition on the extent to which rats engaged in sign- vs. goal- tracking in the autoshaping task during three behavioral time points, as defined above: during learning (Days 1-3), after learning during behavioral expression when animals were under inhibition from CNO/VEH (Days 4-10), and later when CNO/VEH was no longer on board (Days 11-16). A two-way analysis of variance yielded no significant differences between our two control groups (n = 10, TH-Cre rats injected with vehicle; n = 10, WT rats injected with CNO) on our two key behavioral metrics: lever pressing frequency during the cue (ANOVA; Group: $F(1,730) = 2.24$, $p = 0.14$; Group x Days: $F(2,730) = 1.54$, $p = 0.22$) and frequency of food cup entry during the cue (ANOVA; Group: $F(1,730) = 1.7$, $p = 0.19$; Group x Days: $F(2,730) = 0.18$, $p = 0.82$). Thus, we chose to combine data from our two control groups in order to increase the power of our experiment. Therefore, all subsequent analyses were done comparing our experimental group (TH-Cre rats injected with CNO; N = 26) to our combined control groups (TH-Cre rats injected with VEH and WT rats injected with CNO; N = 20).

As a first step in quantifying the differences in sign- and goal- tracking behavior exhibited by experimental and control groups, we quantified the overall

frequency of lever pressing and food cup entry during the 10 s presentation of the lever CS across our three previously defined time points: Days 1-3, Days 4-10, and Days 11-16 (**Fig. 4D-E**). When examining food cup entry frequency during the cue, a two-way analysis of variance yielded a significant main effect of Days (ANOVA; $F(2, 730) = 47.67$, $p < 0.001$); this effect was seen across the board for all of our measures (Lever press frequency: $F(2, 730) = 93.16$, $p < 0.001$; Probability index: $F(2, 730) = 122.42$, $p < 0.001$; Response Bias: $F(2, 730) = 115.02$, $p < 0.001$; Latency index: $F(2, 730) = 132.33$, $p < 0.001$; and, PCA index: $F(2, 730) = 126.56$, $p < 0.001$), indicating that behavior of all animals changed across time, from initially displaying more goal-tracking characteristics to developing more sign-tracking tendencies. In addition to a main effect of Days, for food cup entry frequency, we also found a non-significant trend toward a main effect of Group ($F(1, 730) = 3.69$, $p = 0.055$) and a non-significant trend toward an interaction between Days and Group ($F(2, 730) = 2.63$, $p = 0.073$).

In order to break down these results further, we conducted post-hoc t-tests to examine potential group differences within each time point. For food cup entries, we found no differences across days for any of the three timepoints (**Fig. 4D**). During Learning (Days 1-3), there were no differences between groups ($t(44) = 0.50$, $p = 0.62$). During Days 4-10 (when sign- and goal- tracking patterns are thought to stabilize), experimental DA-inhibited animals entered the food cup more often than control animals, but this was not significant ($t(44) = 1.15$, $p = 0.26$). Finally, during Days 11-16, when CNO was no longer on board, food cup entries within experimental animals eventually decreased to match controls ($t(44) = 0.000$, $p =$

1.00). Therefore, in this experiment, inhibiting VTA-DA neurons during learning and expression had no effect on the expression of goal-tracking behaviors; however, by using a different, well-verified technique, it is possible an effect during Days 4-10 would emerge.

When considering lever pressing frequency, as noted above, two-way analysis of variance also yielded a main effect of Days ($F(2, 730) = 92.16$, $p < 0.001$). However, an interaction between Days and Group was also present ($F(2, 730) = 3.88$, $p = 0.02$), indicating that the effect of Day was greater in the experimental group compared with the control group. This suggests that the experimental group may have started exhibiting sign-tracking behaviors more quickly than the control group. However, post-hoc t-test analyses showed no significant differences in the frequency of lever pressing in experimental animals compared to controls (**Fig. 4E**; Days 1-3: $t(44) = 0.62$, $p = 0.54$; Days 4-10: $t(44) = 0.001$, $p = 0.99$, Days 11-16: $t(44) = 1.31$, $p = 0.20$). However, we found that during Days 11-16, when CNO was no longer on board, experimental animals did seem to press the lever more often than control animals, though this was not significant; we think, once again, that a less controversial technique may reveal real differences in lever presses between groups at this time point, once inhibition of VTA-NAc dopamine neurons has been lifted.

To parallel our autoshaping study using FSCV, we next wanted to quantify the degree to which rats engaged in sign- vs goal- tracking behaviors by using the Pavlovian Conditioned Approach (PCA) index (Meyer et al., 2012). Again, the PCA index comprises the average of three ratios: (1) the probability difference, which is ($P_{lever} - P_{receptacle}$), (2) the response bias, which is (Lever Presses – Food Cup Entries) /

(Lever Presses + Food Cup Entries), and (3) the latency index, which is ($\overline{x}$ Cup Entry Latency – $\overline{x}$ Lever Press Latency) / Length of cue (i.e., 10). All of these ratios range from -1.0 to +1.0 (similarly for the resultant PCA index) and are more positive or negative for sign-tracking or goal-tracking behaviors, respectively.

As noted above, a two-way analysis of variance yielded a main effect of Days for PCA score, Probability, Response Bias, and Latency (PCA index: $F(2, 730) = 126.56$, $p < 0.001$, Probability index: $F(2, 730) = 122.42$, $p < 0.001$; Response Bias: $F(2, 730) = 115.02$, $p < 0.001$; and, Latency index: $F(2, 730) = 132.33$, $p < 0.001$). We also found a significant Day x Group interaction for Latency index ($F(2, 730) = 5.04$, $p = 0.007$) and trends toward a Day x Group interaction for both Probability index ($F(2, 730) = 2.55$, $p = 0.08$) and PCA score ($F(2, 730) = 2.87$, $p = 0.06$). There were no effects of Group across measures (PCA: $F(2, 730) = 0.14$, $p = 0.71$; Probability: $F(2, 730) = 0.06$, $p = 0.81$; Response Bias: $F(2, 730) = 0.09$, $p = 0.77$; Latency index: $F(2, 730) = 0.49$, $p = 0.48$), and the Day x Group interaction for Preference was also non-significant ($F(2, 730) = 1.96$, $p = 0.14$).

Though it appears experimental rats may have produced slightly more sign-tracking behaviors compared with controls once DA inhibition was lifted in Days 11-16, post-hoc analyses showed no significant differences between our control group and experimental group across the board on PCA behavioral measures for any time point (**Fig. 5**). There were no differences between groups during Learning (Days 1-3; PCA: $t(44) = 0.46$, $p = 0.64$, Probability: $t(44) = 0.42$, $p = 0.68$, Response Bias: $t(44) = 0.32$, $p = 0.75$, and Latency: $t(44) = 0.83$, $p = 0.41$), during expression under CNO/VEH (Days 4-10; PCA: $t(44) = 0.28$, $p = 0.78$, Probability: $t(44) = 0.33$, $p =$

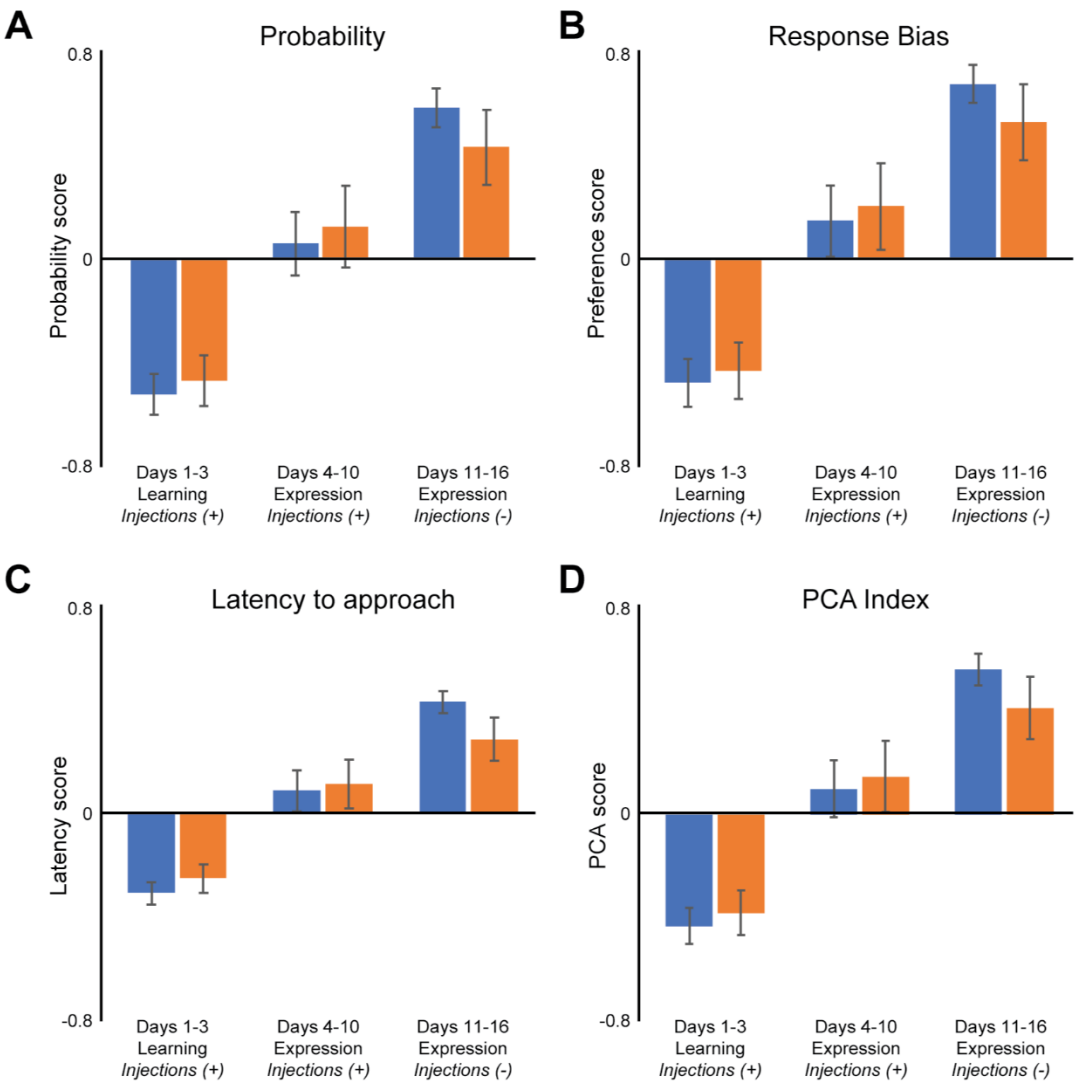


Figure 5. (**A**) The probability difference, which is (Plever – Preceptacle), (**B**) Response bias, which is (Lever Presses – Food Cup Entries) / (Lever Presses + Food Cup Entries), (**C**) Latency index, which is ($\bar{x}$ Cup Entry Latency – $\bar{x}$ Lever Press Latency) / 10, and (**D**) PCA index, which is the average of response bias, probability difference and latency difference indices described in A-C, for Experimental (TH+ animals injected with CNO, n = 26; blue) and Control (TH+ animals injected with vehicle or wild-type animals injected with CNO, n = 20; orange) across three different time periods. All ratios range from -1.0 to +1.0, with more negative values indicating sign-tracking behaviors and more positive values indicating goal-tracking behaviors, respectively. Intermediate behavior is considered to fall within -0.5 and 0.5. All behavioral indices comprise behavior during the entire 8 s CS epoch. Each of the above ratios was computed by session and then averaged according to the time period (Days 1-3, Days 4-10, Days 11-16).

0.74, Response Bias: t(44) = 0.26, p = 0.79, and Latency: t(44) = 0.21, p = 0.83), or during expression once CNO/VEH were offline (Days 11-16; PCA: t(44) = 1.16, p = 0.25, Probability: t(44) = 0.99, p = 0.33, Response Bias: t(44) = 0.97, p = 0.34, and Latency: t(44) = 1.66, p = 0.10). Surprisingly, these behavioral data do not support our hypothesis that inactivation of VTA-NAc neurons during learning and early expression of autoshaping training would promote goal tracking initially, and later promote sign-tracking. However, we suggest these results are verified in the future using another technique.

Chapter Discussion

The voltammetry results reported here support the STGT model proposed by Khamassi et al., predicting that reducing the ITI should lower the amplitude of US DA bursts due to less time for the animal to negatively revise the food cup value; the higher food cup value would consequently lead to an increase in the tendency to GT in the overall population (Lesaint et al., 2015, 2014). Additionally, the model predicts that lengthening the ITI should increase the amplitude of US DA bursts due to more time for the animal to negatively revise the value of the food cup; the lower food cup value would consequently lead to a decreased tendency to GT and an increased tendency to ST in the population, which would be accompanied by a large phasic DA response to the highly salient lever CS. Additionally, we found that specific and temporary inactivation of VTA-NAc neurons during learning and early expression of autoshaping training using DREADDs promoted goal tracking initially; however, once DA was restored, sign-tracking behaviors in experimental animals had been potentiated compared to controls.

Standard Reinforcement Learning (RL) (Richard S Sutton & Barto, 1998) is a widely-used normative framework for modelling learning experiments (Barto, 1995; R.S. Sutton & Barto, 1987). To account for a variety of observations suggesting that multiple valuation processes coexist within the brain, two main valuation systems have been proposed: Model-Based (MB) and Model-Free (MF) frameworks (Clark et al., 2012; Doll, Simon, & Daw, 2012). MB systems employ an explicit, although approximate, internal model of the consequences of actions, which makes it possible to evaluate situations by forward inference. Such systems best explain goal-directed behaviors and rapid adaptation to novel or changing environments (Daw, Gershman, Seymour, Dayan, & Dolan, 2011; Solway & Botvinick, 2012; Yin, Ostlund, Knowlton, & Balleine, 2005; Yin, Knowlton & Balleine, 2006). In contrast, MF systems do not rely on internal models but directly associate stored (cached) values to actions or states based on experience, such that higher valued situations are favored. Such systems best explain habits and persistent behaviors (Daw et al., 2011; Ann M Graybiel, 2008; Yin, Knowlton, & Balleine, 2004). Learning in MF systems relies on a computed reinforcement signal: the reward prediction error (RPE). This signal has been shown to correlate with the phasic response of midbrain dopamine neurons that increase and decrease firing to unexpected appetitive and aversive events, respectively (Bromberg-Martin et al., 2010; W Schultz et al., 1997).

Recent work by Flagel et al. has questioned the validity of classical MF RL methods in Pavlovian conditioning experiments (Flagel et al., 2011). Our goal was to expand on this work using precise FSCV DA measurements and specific pharmacogenetic manipulations of the DA system. To remain as consistent as

possible with the work of Flagel et al., the autoshaping procedures we used here were nearly identical to their procedure, with an 8 s (FSCV) or 10 s (DREADD) retractable-lever CS presentation, immediately followed by food pellet delivery into an adjacent food cup (Flagel et al., 2011). While they used a VT90 s ITI and we replicated this for our DREADD experiment, our FSCV experiment used a short 60 s ITI and a long 120 s ITI as our critical manipulation. Flagel et al. showed that phasic dopamine release in the nucleus accumbens core (NAc) in STs matched RPE signaling and dopamine transmission was necessary for the acquisition of sign-tracking behavior (Flagel et al., 2011). In contrast, though GTs acquired a Pavlovian conditioned approach response, it was not accompanied by the expected RPE-like dopamine signal, nor was the acquisition of a goal-tracking CR blocked by administration of a dopamine antagonist (see also: Danna & Elmer, 2010). Our results are in line with these findings.

To account for these and other results, Khamassi and colleagues proposed a new computational model – the "STGT model" – that explains a large set of behavioral, physiological and pharmacological data obtained from studies on individual variation in Pavlovian conditioned approach (DiFeliceantonio & Berridge, 2012; Flagel et al., 2011, 2009, 2007; Mahler & Berridge, 2009; T. E. Robinson & Flagel, 2009; Saunders & Robinson, 2012). Importantly, the model can reproduce previous experimental data by postulating that both MF and MB learning mechanisms occur within each individual; the main simulated inter-individual variability results from differential weights associated to the contribution of each system. The model accounts for the full spectrum of observed behaviors ranging from one extreme –

sign-tracking associated with a small contribution of the MB system – to the other extreme – goal-tracking associated with a high contribution of the MB system (Meyer et al., 2012). Above all, by allowing the MF system to learn different values associated with different stimuli and taking individual experience into account, the model potentially explains why the lever-CS and the food cup might acquire different motivational values in different individuals, even when they undergo the same training in the same task (Yin et al., 2005).

The STGT model also explains why, here, we observed an RPE-like dopaminergic response in STs but not GTs during our FSCV experiment. It suggests that GTs focus on the reward predictive value of the food cup, which would have been differentially down-regulated during the two ITIs. The model also explains why previous studies inactivating dopamine in the accumbens core or the entire brain results in only partial blockade of behavior: if learning in GTs relies more heavily on the dopamine-independent MB system, dopamine blockade would not impair learning in these individuals (Flagel et al., 2011; Saunders & Robinson, 2012). Importantly, the model has led to a series of new experimentally testable predictions that assess and strengthen the proposed computational theory and allow for a better understanding of the DA-dependent and DA-independent mechanisms underlying inter-individual differences in learning (Lesaint et al., 2015, 2014).

The key computational mechanism of the model is that both the approach and consumption-like behaviors observed towards the lever in sign-trackers (STs) and towards the food cup in goal-trackers (GTs) result from the acquisition of incentive salience by these reward- predicting stimuli. Acquired incentive salience is stimulus-

specific: the stimuli most predictive of reward will be the most "wanted" by the animal. The MF system attributes accumulating salience to the lever or the food cup as a function of the simulated phasic DA signals. In the model simulations, because the food cup is accessible but not rewarding during the ITI, a simulated negative DA error signal occurs each time the animal visits the food cup and does not find a reward; the food cup, therefore, acquires less incentive salience than the lever, which is only presented prior to reward delivery. In simulated STs, behavior is highly subject to incentive salience due to a higher weight attributed to the MF system than to the MB system. As a consequence, STs are more attracted to the lever than the food cup. By contrast, the MB system is weighted more heavily than the MF system for simulated GTs; therefore, GTs prefer the food cup, which is the shortest path to reward in the MB system. Moreover, because the food cup has a lower incentive salience, simulated GTs engage with the food cup less than STs do on the lever, as observed experimentally.

The STGT model also led to specific predictions about what would happen if rats had more exposure to the food cup in the absence of reward. The key prediction of this aspect of the model is that increased access to the food cup during the ITI should decrease the incentive salience associated with it and, conversely, increase the strength of engagement with the lever. This, in turn, would increase the relative proportion of STs compared to GTs in the population. In addition, the model predicts that DA release to the CS would be higher than DA to the US due to the predictive power of the CS, and DA release to the US would remain high after conditioning due to the persistent positive surprise associated to reward delivery in the devalued food

cup. Both of these predictions were confirmed in our current study; rats in the 120 s ITI group were more likely to become sign-trackers, and DA release to the CS was significantly higher than during the US, and US-evoked DA remained after learning. Conversely, the model also predicts that decreased access to the food cup during the ITI would increase the incentive salience associated with the food cup, which would result in more goal-tracking behavior and a DA signal that resembles the RPE pattern observed in sign-trackers. These predictions were partially confirmed in our current study; rats in the 60 s ITI group were more likely to become goal-trackers if they were classified based on their initial approach to the food cup in response to the CS, as in the original study (Flagel et al., 2011). The post-learning DA pattern of the 60 s ITI group showed a high sign-tracking-like RPE response to the CS, not the US. Taken together, these results validate the STGT model.

However, it is worth noting that the observed behavior of the 60 s ITI group goes beyond the predictions of the STGT model. The short ITI group, indeed, showed a more complex behavioral response to lever-CS presentation: an initial food cup approach during the first two seconds after the CS – consistent with a goal-tracking behavior – followed by a more ST-like behavioral engagement with the lever (**Fig. 2G, H** and **Fig. 3B, D**). Late engagement with the lever is not predicted by the computational model, which only attempts to model the first behavioral response of the animals to the CS (Lesaint et al., 2014). This is compatible with the way sign-trackers and goal-trackers were classified based on their initial response to the CS in the original study (Flagel et al., 2011). This simplification in the model still accounts for a full spectrum of inter-individual variability, even animals originally classified in

the “intermediate group”, exhibiting both ST and GT behaviors. Nevertheless, the present results highlight that the STGT model should be extended to account for temporal variability of the animal’s behavior within each trial.

Our DREADD experiment, in which we temporarily and specifically blocked VTA- NAc DA neurons during learning and expression of ST and GT behaviors, complicates the story of dopamine’s role during sign- and goal- tracking. We found that inhibition of DA initially may have increased goal-tracking in experimental animals compared with controls; however, once normal DA function had been restored after prior inhibition, experimental animals exhibited higher sign-tracking tendencies compared to controls. However, we only saw significant differences between our groups during our ANOVA and significance disappeared in further pot-hoc analyses. Taken with a grain of salt, this result contradicts conclusions drawn by a previous study from Flagel et al. that used a systemic dopaminergic antagonist flupenthixol to demonstrate that dopamine is necessary for both learning and performance of sign-tracking CRs but, in contrast, plays no role in the development of goal-tracking CRs (Lesaint et al., 2014). Further, this study suggested that, while dopamine may be necessary for the performance of both sign-tracking and goal-tracking CRs after learning had occurred, it is only necessary for acquisition of a sign-tracking CRs. Surprisingly, our results suggest that specific projections from VTA DA neurons to NAc core do not specifically promote the development or expression of either goal- or sign-tracking behaviors, but instead modulates both.

At the start of training, all animals can be categorized as either intermediate responders or goal-trackers, showing no behavioral preference for lever or food cup

or exhibiting behaviors favoring the food cup, respectively. Thus, at first, rats initially develop an attraction to the food cup since it is the location of food reward (US) delivery and learn to enter the food cup during the cue period before food is delivered. Normally, rats will gradually transition to exhibiting more sign-tracking behaviors like by interacting with the lever (CS) more often during the cue period; sign-tracking behaviors begin to stabilize around day 5 of training; this transition is thought to be mediated by the transfer of dopamine release from the food reward US to the lever CS over time. By disrupting the DA system during learning and early expression, it seems that we slowed this process, in that rats maintained high food cup responding longer when DA neurons were inactivated. This early DA inhibition and the resulting prolonged food cup preference may have led to the formation of a more stable representation of food delivery in the food cup and, hence, diminished the transfer of DA from the food cup to the lever CS. However, it appears that with time and further task exposure, all animals come to display more ST CRs than GT CRs regardless of DA activity. Once DA was restored later in training, it seems enhanced valuation of the food during learning and early expression in DA-inhibited experimental animals may have promoted stronger associations between the CS and US and potentiated ST CRs once CNO DA-inhibition was lifted.

However, it is critical to note that emerging literature has begun to suggest a number of drawbacks to the DREADD technique, specifically surrounding the activation of DREADD receptors using CNO. Perhaps most notably, a recent study has shown that systemic CNO does not seem to cross the blood-brain barrier and its presence is extremely low within the central nervous system after systemic

administration (Gomez et al., 2017b). In addition to this, CNO itself seems to have very low affinity for DREADD receptors (Gomez et al., 2017b). Therefore, it has been postulated that the mechanism of DREADD actuation is not, in fact, CNO but instead converted clozapine, which has high affinity for DREADD receptors and flows freely across the blood-brain barrier; at high concentrations, it may also affect endogenous clozapine binding sites, which disrupts the purported specificity of the technique (Bender, Holschbach, & Stöcklin, 1994; Fang, 2000; Gomez et al., 2017b; B. Ji et al., 2016). Finally, it has been shown that both CNO and clozapine reduce motor activity even at very low, subthreshold doses; however, this generally occurs at a later timepoint (roughly 2-3 hours after administration) (Gomez et al., 2017b).

In our DREADD experiment, we did not see any notable differences in locomotion between CNO and vehicle groups. Taking this information into account, it appears that the effects seen during our DREADD experiment may have been caused by action of clozapine on dopaminergic neurons. Though this potentially disturbs the specificity of our experiment and the exact mechanism of action remains unknown, we still find this hypothesis and these results to be important and worthy of further examination. Future work could employ methods such as optogenetics in order to achieve specific, real-time dopamine receptor inhibition during STGT behavior.

Chapter 5: General Discussion

Summary of Results

Approaching good things, avoiding bad things, and interpreting signs in the environment that predict both are skills that are absolutely critical for survival. Cues and behaviors that help animals acquire something good or escape something bad are reinforcing and will gradually increase the probability that the same behavior will be selected and performed again in the future. Both positive reinforcement (acquiring something good) and negative reinforcement (avoiding something bad) drive behavior, but in seemingly different ways; a major goal of the research presented within this dissertation was to uncover the underlying brain mechanisms encoding positive and negative reinforcement and explore individual differences that arise during learning and performance of both.

Associations between cues and outcomes are encoded across various brain regions within the corticomesolimbic pathway. It is well known that the dopaminergic system and, more broadly, the corticomesolimbic system are involved in the processing of positive and negative reinforcement, but it is still unclear where these signals are colocalized. While this circuit is composed of many regions that work together to produce approach and avoidance behaviors, we chose to explore two regions within this text that we thought would be particularly informative: the ventromedial prefrontal cortex (vmPFC; *Chapter 2*) and DA release in nucleus accumbens (NAc; *Chapters 3 & 4*). The vmPFC is thought to generate contextually appropriate responses, while DA release in NAc reflects prediction error signals

necessary for generating expectancies about future outcomes to motivate behavior. For these reasons, it is easy to see why both regions may be of extreme interest when exploring the neural underpinnings of reinforcement.

A vast library of work compiled over the last three decades has explored the role of DA neurons and accumbal DA release in reinforcement; however, the majority of these studies examine positive reinforcement, exclusively. Negative reinforcement is a more complex behavior, consisting of an initial Pavlovian component that must be overcome before evasive actions can be performed. This multi-component behavioral sequence is more difficult to learn and results in varied performance across animals (i.e., some rats are bad at avoiding negative things in the environment, while others excel) (J. M. Moscarello & LeDoux, 2013). Though negative reinforcement is just beginning to be studied in depth, very few tasks have measured neural correlates of positive and negative reinforcement within the same task. Thus, it has been difficult to know if these reinforcement signals are colocalized within regions such as vmPFC or NAc, or if they are encoded separately. The first objective of my research was to determine if positive and negative reinforcement signals are colocalized or are encoded separately by the aforementioned vmPFC (*Chapter 2*) and the NAc core (*Chapter 3*).

To explore this question, we measured single-unit activity and phasic dopamine release in the vmPFC and NAc, respectively, while rats performed a combined approach and avoidance task. Critically, our combined approach and avoidance task presented opportunities for reward and punishment avoidance within the same session and included a neutral trial type as a control. Additionally, each trial

type required the same action sequence to be performed and a time delay separated the cue onset from lever access, both of which allowed us to control for and separate out changes related to the action alone.

Behaviorally, in both studies, we found that rats pressed more often and were faster to press for reward cues compared with neutral or shock cues; rats were also slower to press for shock cues compared with neutral or reward cues, but still pressed more for shock cues compared with neutral cues. When we separated our rats based on their performance during shock trials, we found that animals that were bad at avoiding shock also exhibited increased freezing to shock cues during both press and no press trials; importantly, animals that were good at avoiding shock still froze during shock press trials but less often, suggesting that these animals were better at overcoming any initial freezing response to the shock cue. It seems that the overall behavior of poor avoiders mirrored the previously reported group behavior, with faster and more frequent lever pressing after reward cues compared with shock cues, while good avoiders pressed fast and frequently for all trial types. This suggests that perhaps good avoiders are behaving in a more habitual manner, while poor avoiders seem to be focusing on the specific outcomes predicted by the cues, suggesting they are more goal-directed.

In *Chapter 2*, we showed that neurons within the vmPFC preferentially encoded reward-related signals. While cues predicting reward significantly modulated the baseline firing rates of our vmPFC cells, cues predicting shock or neutral trials did not. When we compared firing rate during reward and shock trials to firing rate during neutral trials, we found a large sub-population of vmPFC cells whose firing

rate was modulated by reward (N = 70); these cells were not modulated by shock or neutral in the same way. Though modulation by reward was more common, we did find a small sub-population of neurons that was modulated more by shock cues than by neutral cues (N = 17); however, modulation by shock for these cells was not as robust as that seen in cells modulated by reward. Importantly, we wanted to know if these changes in firing rate to reward and shock cues were happening within the same session and, thus, within the same cells; we found no correlation between shock and reward cue firing, indicating that cells encoding positive and negative reinforcement are likely occurring in separate neural populations within vmPFC.

Since there is an abundance of literature suggesting that vmPFC is also involved in extinction learning and expression, we also measured changes in firing for reward selective neurons during extinction press and no press trials. We found sub-populations of cells that were selective for outcome during extinction or conditioning, showing modulation only during reinforced press trials or only during extinction press trials. Additionally, we found sub-populations of cells that were response-selective during extinction, showing modulation for either press or no press, but not both. We were unable to perform these analyses in shock-modulated cells with sufficient power due to the low number of cells that were modulated by shock cues in our task.

Alternatively, in *Chapter 3*, we showed that positive and negative reinforcement signals seem to be colocalized within the same microdomain of the NAc core. In the overall population, we found increased dopamine release to cues that predict both the acquisition of reward and the avoidance of punishment compared

to baseline and to neutral cues. We also showed that DA release to cues remained elevated for lever extension and reward delivery on reward trials. This is different from what was shown in the study by Oleson et al., since we did not see increased DA release to the safety period when animals successfully avoided shock (E. B. Oleson et al., 2012); we think this is due to the fact that the Oleson et al. study used another discriminatory cue to signal the safety period, whereas our safety period was not signaled. Importantly, we also found that changes in DA release to both reward and shock cues were correlated, indicating that these changes were occurring within the same session and, hence, within the same microdomain; this was not true for changes in DA release to reward and shock during the lever epoch.

Additionally, distinct differences in dopaminergic patterns emerged based on performance during shock trials. Behaviorally, animals who did better on shock trials than neutral trials (i.e., good avoidance), pressed quickly and at a high rate for all trial types. Critically, though animals showing good avoidance behaviors seem to be responding in a habitual manner, dopamine patterns clearly delineated trial type. During good avoidance, we saw increased DA release to cues predicting shock and reward, while DA was slightly inhibited to cues predicting a neutral trial; we also saw increased DA during lever extension and reward delivery. In contrast, during poor avoidance, DA release to the cue was significantly increased above baseline for all trial types; in these animals, indiscriminately high DA release to all cues may confuse downstream action areas and promote poor avoidance.

It is important to note that we can confirm that animals exhibiting poor avoidance behavior still seemed to clearly understand the meanings and consequences

of each cue type, even though this was not reflected in the dopamine signal; they froze only to shock cues, were slow to press for shock, and were fast to press for reward. It is possible that these animals were too focused on the possibility of negative outcome to behave appropriately. In fact, behavior was only correlated to DA release to the cue for neutral and shock cues, but not for reward. It is possible that in rats exhibiting good avoidance, control of these behaviors had already been transferred to downstream areas involved with habitual responding, such as DLS, though DA was still keeping track of cue values. Thus, in this scenario, when possibility for reward and punishment are both present within the same environment, it may be beneficial to be more habit-driven and less goal-oriented.

Following this, we were very interested in continuing to explore the development and expression of individual differences that arise within the DA system in response to cues. While our previous studies examined individual differences in shock avoidance during a complex combined approach and avoidance task, we next wanted to parse apart the Pavlovian aspect of this behavior. To do this, we used a simple autoshaping task. In this type of task, each trial presents a cue followed by reward delivery to an adjacent food cup; importantly, the animal's behavior towards the cue has no effect on the outcome. Each trial is followed by some inter-trial interval (ITI), usually 90 s in a standard autoshaping task (Flagel et al., 2011). This task has been shown to reliably produce two distinct behavioral phenotypes during the cue presentation: sign-tracking (ST), where animals reliably approach and interact with the cue, and goal-tracking (GT), where animals reliably approach and interact with the reward location (i.e., the food cup).

Results from previous studies that systemically antagonized dopamine receptors during autoshaping suggest that the development of sign-tracking behaviors, but not goal-tracking behaviors, is dependent upon dopamine signals within the brain (Flagel et al., 2011; Saunders & Robinson, 2012, 2013a). These results showed that sign-trackers develop the traditional PE signal (i.e., transfer of DA from the reward to the cue), while goal-trackers never fully develop a PE signal and maintain DA release to the reward delivery even after the task is well-trained and behavioral patterns become solidified (Flagel et al., 2011). However, a recently proposed computational model – the Lesaint-Khamassi model – suggests that dopamine is indeed involved in both sign- and goal- tracking, just to varying degrees (Lesaint et al., 2015, 2014).

In short, this model suggests that all animals revise the value of the food cup multiple times during the ITI, when the cue is not available; during this time, no food is available in the food cup, the animal experiences a negative PE, and the food cup loses value. Thus, in animals that only visit the food cup during the ITI, this would lead to the progressive downward revision of the food cup; this would lead to fewer food cup visits in subsequent trials and more ST behavior. Since GTs visit the food cup both during the ITI (when it is being downwardly revised) and during the trial (when the food cup maintains an expectation of reward due to the simultaneous presentation of the cue), this downward revision is offset. The authors suggest this mechanism may also prevent the transfer of reward value (and, consequently, DA release) from the time of reward to the time of the cue in GTs, preventing the traditional PE pattern from forming.

Importantly, this model has testable predictions. The authors suggest that by changing the length of the ITI, it should be possible to shift behavioral responses and DA responses. For example, using a short ITI (e.g., 60 s) would provide less time for the animal to negatively revise the food cup value during the ITI, which would lead to smaller PEs to the reward delivery and higher food cup value during the trial; this would promote more food cup visits in the future (i.e., GT behavior). A long ITI (e.g., 120 s) would give the animal more time to negatively revise food cup value during ITI, which would lead to larger PEs to food delivery and lower food cup value during the trial; this would promote fewer food cup visits in the future (i.e., ST behavior). The next objective of my research was to test these predictions by manipulating the ITI during a standard autoshaping task to determine if we could influence behavioral and dopaminergic patterns that were seen by Flagel et al. (Flagel et al., 2011); we also attempted to shift the development and expression of sign- and goal- tracking behavioral patterns by specifically and temporarily inhibiting VTA-NAc DA neurons during this task.

In *Chapter 4*, we found evidence to support the Lesaint-Khamassi model. Behaviorally, rats in the 120 s ITI group showed sustained lever pressing behavior and no increases in food cup entry during cue presentation; in addition to increased frequency of lever press during the cue, the rats in the long ITI group were also shifted along the PCA index, showing significantly more sign-tracking tendencies (i.e., preference for the lever, higher probability to press the lever, and quicker to press the lever than enter the food cup). Rats in the 60 s group showed more complex behavioral patterns, initially entering the food cup during the first 2 s of the cue, then

switching to lever pressing for the remainder of the cue period; these animals were also shifted on the PCA index towards goal-tracking, but this seems to have been driven solely by quicker entry into the food cup. Interestingly, when we measured DA release in these animals using FSCV, we found that animals in the short ITI group showed normal PE signals (with low DA release to the reward delivery and high release to the cue), while animals in the long ITI group showed peak DA release to both the cue and the reward. In this sense, dopamine patterns of animals in the 120 s ITI group resembled patterns of GTs in the study conducted by Flagel et al. using the traditional 90 s ITI, and dopamine patterns of animals in the 60 s ITI group resembled patterns seen in STs.

To investigate the development of these behavioral and dopaminergic effects across time, we next broke down this average effect into early learning (Sessions 1-3) and expression (Sessions 4-10). Behaviorally, we found that ST-like approach to the lever develops early and is maintained in the long ITI group; in the short ITI group, however, food cup entries are sustained early in learning, and this pattern shifts during expression to mimic that in the average: food cup entries during the first 2s of the cue with a transition to lever pressing for the remainder. When examining DA patterns between these two groups across days, we see that the long ITI group maintains two peaks (to cue and reward delivery) across both learning and expression, while the short ITI group indeed develops a traditional RPE signal over time, with dopamine release disappearing to the reward and increasing to the cue across learning. Together, this suggests that normal DA PE patterns could likely be restored in Flagel's aforementioned goal-trackers by manipulating the ITI and,

consequently, the animal's ability to downwardly revise the food cup value between trials (Flagel et al., 2011).

Finally, in *Chapter 4*, we also specifically and temporarily inhibited VTA-NAc neurons during learning and expression of a standard (90 s ITI) autoshaping task. Here, we found subtle effects suggesting that the inhibition of these cells initially increases the propensity towards goal-tracking behavior and slows the onset of sign-tracking behaviors. Specifically, we found that inhibited animals entered the food cup slightly more often than controls during expression (Days 4-10). However, when inhibition of these cells was lifted (Days 11-16), we saw a slight potentiation in sign-tracking behaviors (i.e., increased lever pressing frequency and ST-shifted scores for all PCA index effects). Unfortunately, none of the specific comparisons we examined during our post-hoc analyses were found to be significantly different from chance alone, and we will discuss possible explanations for this in the following section, "*Potential Limitations*."

Potential Limitations

While the experiments carried out in this dissertation will critically help advance the knowledge of the field, most of our studies are examining the neural correlates of behavior, and it remains difficult to infer direction or causation. Detailed in *Chapter 4*, an experiment carried out in collaboration with our French colleagues at the University of Bordeaux, aimed to use the DREADD pharmacogenetic technique to address this issue by temporarily inhibiting specific VTA-NAc DA neurons during learning and expression of sign- and goal- tracking behaviors. However, for many reasons, this experiment has limitations of its own.

Due to the international nature of this collaboration, analysis of the data remains gradual and ongoing; specifically, the immuno-histochemical experiments for our second subgroup (WT animals receiving CNO and TH+ animals receiving CNO) remain incomplete and viral expression has not been quantified in the processed tissue of our first subgroup (TH+ animals receiving VEH or CNO). To try to address this, we have included an initial qualitative analysis of viral expression for our first subgroup (*Chapter 4*, Fig. 4B; TH+ animals receiving VEH or CNO). Based on these forthcoming results, quantifying viral expression in the processed tissue for all of the animals used in our DREADD experiment, it is possible that some animals included in our DREADD experiment analyses may need to be excluded from the study. If so, not only could this significantly alter the results we have presented here, but it also may reduce the power of our experiments and require replication in more animals. Next, I will discuss a few critical limitations associated with the DREADD technique itself, which only came to light after the completion of the experiment we discuss in *Chapter 4* and my return from France.

Use of DREADDs

Broadly, pharmacogenetics is a set of tools that use a previously biologically unrecognizable small molecule to activate proteins, such as receptors, that have been specifically engineered to only respond to this inert small molecule as its ligand (Roth, 2016). DREADDs (Designer Receptors Exclusively Activated by Designer Drugs) are a specific class of pharmacogenetic tool that were recently developed to allow for the temporary, cell-type specific manipulation of the signal transduction of g-protein coupled receptors in freely-moving animals (Zhu & Roth, 2014). The

DREADD technique uses a viral vector to insert either an excitatory or inhibitory g-protein coupled receptor (based on different human muscarinic receptors), which can then only be activated by a synthetic, biologically-inert pharmaceutical small molecule, clozapine-n-oxide (CNO). CNO is a metabolite of the antipsychotic drug clozapine. It can be delivered systemically or intracranially and has been shown to have rapid CNS penetration and distribution in mice, with a minimum of 60 minute residency (Bender et al., 1994; Roth, 2016). These viral vectors can also drive either retrograde or anterograde expression (e.g., CAV-2 and AAV, respectively), include fluorescent tracers to mark expression patterns, and can be linked to certain neural cell types using a floxed system in animals expressing Cre-recombinase. This lock and key style of neural control and the biologically relevant time course of CNO have caused the DREADD technique to quickly become one of the most sought-after and widely-used pharmacogenetic techniques in the last decade (Armbruster et al., 2007).

However, a recent study revealed an array of potentially fatal flaws with the technique and our interpretation of what DREADDs are actually influencing within the brain (Gomez et al., 2017a). Notably, using saturation binding experiments, Gomez et al. found that CNO does not enter the brain after systemic injection and shows very little affinity for DREADD receptors (Gomez et al., 2017a). Instead, it seems that upon injection, CNO is rapidly back-converted to clozapine, which rapidly enters the brain, shows high affinity for DREADD receptors, and is highly effective in activating them (Gomez et al., 2017a). Unfortunately, clozapine can also bind to many endogenous receptor types within the brain (histamine H1, serotonergic 5-$HT2_a$, muscarinic M1, M3 and M4, adrenergic A1, and dopaminergic D1 and D2 type

with particularly high affinity for D4 receptors), which negates the specificity of the DREADD technique (National Center for Biotechnology Information, 2017). Thus, it seems likely that the effects presented in the hundreds of DREADD experiments conducted over the last decade are likely due to the effect of clozapine, not CNO.

Here, in *Chapter 4*, we utilized DREADD technology by inserting a retrograde viral vector (CAV-2) containing a modified inhibitory human muscarinic receptor (hM4Di) and the mCherry fluorescent tag into the nucleus accumbens core of wild-type or transgenic Th-Cre rats. Our goal was to specifically and temporarily inhibit DA neurons projecting from VTA to NAc core during the acquisition and expression of sign- and goal- tracking behaviors in a traditional autoshaping task. We used a moderate dose of CNO (1mg/kg) as our actuator. However, we saw mixed success with viral expression in VTA cell bodies (and also found some expression in NAc fibers) and saw very few significant effects of VTA-NAc DA inhibition on the development or expression of sign- and goal- tracking behaviors. This was very surprising to us, as DA antagonism studies have shown DA to be a critical factor in the development and expression of sign-tracking behaviors, and it is thought that DA is likely involved in both sign- tracking and goal-tracking responses (Flagel et al., 2011, 2009; Lesaint et al., 2014).

Based on this recent work by Gomez et al., we believe that converted clozapine, instead of CNO, may be driving our effects in this experiment and fouling the purported specificity of the DREADD technique in inhibiting only VTA-NAc neurons (Gomez et al., 2017a). Since clozapine not only has high affinity for engineered DREADD receptors but to a variety of endogenous receptors (serotonin,

dopamine, histamine, muscarinic acetylcholine, adrenergic) within the brain, this means that if clozapine is indeed our actuator instead of CNO, we cannot say that we are manipulating the system as we intended. However, it is important to note that we did utilize the proper controls in our experimental design (wild-type animals receiving CNO). For this reason, though our results are promising, we think it is extremely important to replicate our DREADD experiment using another, well-tested method that employs a similar level of specificity, such as optogenetic TH+ cell-specific inhibition during the same task.

Exploring parallels between avoidance and sign- and goal- tracking behaviors and DA patterns

Individual differences naturally arise in animals' behavioral and dopaminergic patterns during both avoidance and autoshaping training. In both tasks, one behavioral pattern is beneficial, while the other is maladaptive or illogical. Here, I aim to ask if there might be important common denominators among those who do well and those who do poorly in each task. It could be that there are important similarities and differences in the processing of dopamine and the execution of behavioral responses that emerge during avoidance (good vs. poor) and autoshaping (sign- vs. goal- tracking) procedures. If so, these similarities in dopaminergic and behavioral patterns across tasks could lend insight into the underlying mechanisms behind these individual differences.

At first glance, the behaviors produced within these tasks may not seem directly comparable, but abnormal DA patterns develop in certain subsets of animals during each behavioral task and these brain differences are reflected in differing

behavioral responses. Maybe there is a connection between differences in DA and differences in behavior among the two tasks. My goal here is to compare good and poor avoiders to sign- and goal- trackers to explore any such similarities. By exploring these individual differences, we could gain critical insight into how the dopamine system can be altered and manipulated during behavioral dysfunction and what this means in varying contexts.

When we examine dopaminergic patterns of each sub-group, we first notice that both poor avoiders and goal-trackers seem to encode cues somewhat abnormally, with poor avoiders showing indiscriminately high DA release to cues of different values and goal-trackers failing to develop the traditional RPE pattern over time. Therefore, it seems that both of these groups, poor avoiders and goal trackers, may experience complications when encoding and updating the value of environmental cues within the DA system. While muddled, uninformative DA signals to cues and outcomes in poor avoiders may confuse downstream areas, alternatively, increased DA to both cues and outcomes in goal-trackers continues to drive goal-directed behavior toward reward retrieval. It seems, here, context matters. In comparison, good avoiders and sign-trackers seem to have "properly functioning" dopamine systems, with good avoiders accurately discriminating between and encoding the value of cues and sign-trackers developing the traditional RPE, where DA release transfers from the reward delivery to the cue with learning. In this case, too, it seems that these dopamine signals can be either beneficial or distracting, depending on the context.

When we examine the behavior of each sub-group, we find two groups whose behavior benefits them in the task and two groups whose behavior does not. In our approach and avoidance task, we find that good avoiders press the lever quickly and often while poor avoiders freeze more to shock cues and, thus, are slow or fail to press for this trial type. During autoshaping, we find that some animals, goal-trackers, focus attention on the location of reward delivery, while other animals, sign-trackers, focus attention on an inconsequential environmental stimulus. Thus, it seems that both poor avoiders and goal-trackers are more focused on the outcome, which can be a negative or a positive thing depending on the environment and circumstances; for poor avoiders, when the outcome is potentially aversive, this behavior is not beneficial, but it is for goal-trackers. In contrast, it seems that both good avoiders and sign-trackers become cue-focused, potentially having formed strong S-R associations and seem to interact habitually with the predictive cues. Perhaps, when there is potential for negative consequences, like shock, behaviors that appear to be more habit-driven are beneficial, while when there is only the potential to gain a positive outcome, goal-directed behaviors are best.

Changing behavioral expression through shifts in task parameters

It has been proposed that changing the length of the ITI during an autoshaping task and, thus, an animal's exposure to the food cup in the absence of food (when its value is low), may shift dopamine release and behavioral patterns during subsequent trials (Lesaint et al., 2015, 2014). Traditionally, autoshaping tasks use a 90 s inter-trial- interval (ITI). It is thought that using a short ITI (60 s, compared with 90 s) would provide less time for the animal to negatively revise the food cup value outside

of the trial, which would lead to smaller PEs to the reward delivery and higher food cup value during the trial; this would, in turn, promote more food cup visits in the future (i.e., GT behavior). Alternatively, a longer ITI (120 s, compared with 90 s) would give the animal more time to negatively revise food cup value outside of the trial, which would lead to larger PEs to food delivery and lower food cup value during the trial; this would, consequently, promote fewer food cup visits in the future (i.e., ST behavior). By changing the ITI in our autoshaping task (*Chapter 4*) and providing more (120 s) or less (60 s) time to interact with a devalued food cup, we showed that we were, indeed, able to influence the behavior of our animals and dopamine patterns within NAc, when compared with previous results (Flagel et al., 2011).

These results beg the question: would it be possible to turn poor avoiders into good avoiders by changing certain task parameters? Since we think these animals may be similar to goal-trackers in the sense that they are focused on the outcome, could we make them more like sign-trackers by using the same method of ITI manipulation and, hence, improve their behavior in our combined approach and avoidance task? Perhaps it would not be as simple as lengthening the ITI, in this case, due to the complexity of the task.

However, it has been shown that training animals under different reinforcement schedules can influence the development of goal-directed versus habitual behaviors. Specifically, it has been hypothesized that animals use correlations between the rate of responding and the rate of reward during training, and degrading this correlation can promote habitual responding (A. Dickinson, 1985). For

example, animals are more likely to develop habits if trained to press a lever to gain reward on a random interval schedule as opposed to a random ratio schedule (A. Dickinson, 1985; Anthony Dickinson, Nicholas, & Adams, 1983; Gremel & Costa, 2013; O'Hare et al., 2016). Taking this research into account, it seems that the temporal uncertainty of outcomes (at least, rewards) may increase the likelihood of habit development, and switching animals to a less temporally-uncertain schedule may result in more goal-directed behaviors. Our autoshaping study utilizes and further supports this hypothesis. By increasing the ITI, we are increasing the temporal uncertainty associated with the task; in turn, we saw increased habit-like behaviors, similar to those seen in sign-trackers. By decreasing the ITI, we are decreasing the temporal uncertainty associated with task and reward delivery; in turn, we saw increased goal-directed behaviors, similar to those that would be seen in goal-trackers.

In the context of good and poor avoiders, we have previously suggested that poor avoiders may be acting in a more goal-directed manner, focusing on the outcomes, which could be detrimental when these outcomes are potentially harmful. In contrast, we found that good avoiders seemed to be acting in a more habitual manner, pressing fast and often for all trial types. In order to transform a poor avoider into a good avoider, perhaps we need to make the behavior of those demonstrating poor avoidance more habitual. According to the aforementioned scheduling hypotheses regarding the formation of habit and our studies manipulating ITI in sign- and goal-trackers, it may be possible to switch poor avoiders to good avoiders by increasing the ITI and, thus, increasing the temporal uncertainty of the task.

One could also imagine that it might be possible to shift behavioral and dopaminergic patterns by manipulating shock cue parameters and the shock itself. For instance, in earlier tasks where the safety period after shock avoidance was cued, a different dopaminergic pattern emerged (i.e., increased dopamine release to the cue and to the cued safety period) and learning was easier and faster (E. B. Oleson et al., 2012). This suggests that task parameters are indeed critical to consider and that by manipulating certain aspects of our combined approach and avoidance task, we may be able to convert poor avoiders into good avoiders.

Implications for the corticomesolimbic circuit

Though the studies we present here have focused on neural correlates we measured within the NAc core and vmPFC, our results should also be considered within the context of the corticomesolimbic circuit, as other regions within this pathway are also likely to be affected. To start, differences in DA release in the NAc for good and poor avoiders, as well as sign- and goal- trackers, would likely be reflected in similar differences in dopamine neuron firing within the VTA. For instance, DA cells in poor avoiders would likely also encode expectancies incorrectly, firing at a high rate to all cue types and failing to encode accurate PEs; a similar scenario would also be true for goal-trackers in our autoshaping task. So how could these poorly encoded prediction error signals within DA neurons and the NAc be reciprocally affecting cortical and downstream areas within this circuit?

The orbitofrontal cortex (OFC) would be interpreting and using these signals, so it follows that signaling in this region would also be affected. The OFC is thought to encode, compare, and integrate outcome values, both aversive and appetitive. For

instance, it has been shown that neurons in the OFC fire strongly when an animal anticipates a desirable outcome (Schoenbaum et al., 1998; Schoenbaum, Chiba, & Gallagher, 1999), but firing rates will decrease if that outcome is presented alongside a chance for another, better outcome (Wallis & Miller, 2003) or if that outcome is devalued through satiation (Rolls et al., 1989). It seems, then, that firing of OFC neurons reflects the combined value conveyed by environmental cues, both appetitive and aversive. In a study by Roesch and Olson, they found that OFC neurons fired most for cues that predicted a large reward with a small penalty and the least for cues that predicted a small reward with a small penalty, compared with neutral (Roesch & Olson, 2004). Importantly, it was also found that OFC neurons do not respond differently based on mismatched expectations, like DA neurons do (Roesch & Olson, 2004; Schoenbaum, Roesch, Stalnaker, & Takahashi, 2009), though VTA DA prediction error signals do depend on information from OFC (Y. K. Takahashi et al., 2017). Thus, if the OFC is receiving inaccurate prediction errors from VTA/NAc, it follows that signals within the OFC would likely fail to properly encode, compare, and integrate outcome values for shock, reward, and neutral cues in our task; based on their dopaminergic signal, poor avoiders may produce a signal in OFC stating that each cue is important and valued, even though the outcome for shock cues could be negative and neutral cues predict nothing.

While our results from NAc during combined approach avoidance and autoshaping were in line with previous studies and current theories, the signals we recorded in ventromedial prefrontal cortex (vmPFC) were somewhat surprising to us. Traditionally, the vmPFC has been most often associated with aversive contexts and

fear extinction paradigms (Maren & Quirk, 2004; Milad & Quirk, 2002). Specifically, increases in vmPFC activity have been shown to occur in response to the recall of fear extinction and to cues signaling shifts from shock to no-shock (Milad & Quirk, 2002; Schiller & Delgado, 2010). However, a recent study has also implicated the vmPFC in the extinction of appetitive reinforcers (Eddy et al., 2016). Here, we found additional evidence that certain populations of vmPFC neurons can encode appetitive cues and reinforcers. We found that, in our combined approach and avoidance task, vmPFC cells preferentially encoded reward-related cues and, additionally, became response- and outcome- selective during extinction.

The vmPFC is also thought to help suppress Pavlovian freezing responses originating in the amygdala. Because of this, we expected to see increased vmPFC activity during shock avoidance trials, when rats successfully overcome freezing to press to avoid shock; however, we saw very little encoding of shock-related cues. This naturally led us to wonder what the amygdala might be doing in this experiment. It seems that, while animals are very good at this task overall at the time of recording, the CeA probably remains quite active. Importantly, both poor and good avoiders continued to freeze to shock cues. Perhaps there is a mechanism within the BMA or ITC that allows for a somewhat leaky inhibition of CeA, just in case the threat becomes real again or changes slightly in the future. It is also possible that vmPFC to BMA/ITC connections were no longer being used to inhibit CeA at the time of our recordings; since we saw very few cells in vmPFC modulated by shock cues at this stage in training, it might be the case that neural control of shock-related responses had already been transferred to more habitual brain regions. It is important to recall

that, while CeA is most often associated with the formation of Pavlovian threat responses, it has also been implicated in reward and appetitive behaviors, as well as the overall saliency of rewards. For this reason, as well, we might assume that CeA still remains quite active in this task.

Finally, the basolateral amygdala (BLA) is modulated by the predictability of both appetitive and aversive events, specifically when expectancies are violated (Belova, Paton, Morrison, & Salzman, 2007; Calu, Roesch, Haney, Holland, & Schoenbaum, 2010; Roesch, Calu, Esber, & Schoenbaum, 2010). It is also thought that the BLA integrates information about appetitive and aversive events, including their saliency. As we recall from earlier, the BLA and the OFC have strong reciprocal connections within the circuit. Since it has been shown that BLA lesions or DREADD inhibition interferes with outcome expectancy encoding in the OFC, which we proposed earlier may be faulty, it might be the case that BLA activity is reduced in this task (Lichtenberg et al., 2017; F. Lucantonio et al., 2015). Taken together, it is clear that our results must be considered and interpreted within the context of the larger corticomesolimbic circuit.

Future Directions

In the work we presented here, we found that vmPFC neurons preferentially encoded reward-related cues during our combined approach and avoidance task and were also response- and outcome- specific during extinction; but, this story remains incomplete. It is thought that the vmFPC, specifically the IL, is responsible for suppressing amygdala-driven freezing responses during fear conditioning and shock avoidance tasks. Though we did not see much modulation by shock cues in our

recordings of vmPFC neurons, we hypothesize that vmPFC may still be modulated by shock avoidance cues during earlier time points. To address this, future studies should record from IL during early learning of this task; it would be interesting to also record from PL during early learning and performance, since this region has been often associated with reward cues (Burgos-Robles et al., 2013; Sangha et al., 2014b). These studies would provide even more insight into the role of vmPFC during combined approach and avoidance.

Additional work must also be carried out to determine the specific role of dopamine in the development and expression of sign- and goal- tracking behaviors and, furthermore, to explore possible parallels in DA processing that arise during avoidance (good vs. poor) and autoshaping (sign- vs. goal- tracking) procedures. As mentioned above, in our section "*Exploring parallels in behavioral and dopaminergic patterns during avoidance and autoshaping*", it is possible that there could be important similarities between the dopamine patterns underlying individual differences that emerge. For instance, both poor avoiders and goal-trackers seem to encode cues somewhat abnormally, with poor avoiders showing indiscriminately high DA release to cues of different value and sign-trackers failing to develop the traditional RPE pattern over time. Therefore, problems encoding and updating the value of cues within the DA system may be a universal problem among these groups. Considering the behavior of each group, it is possible that both poor avoiders and goal-trackers are more focused on the outcome, which can be a good or a bad thing based on the environment and circumstances. Perhaps, when there is potential for

negative consequences, habit-driven behaviors are more beneficial, while when there is only potential to gain a positive outcome, goal-directed behaviors are best.

In the future, a specific set of experiments could be performed to test these ideas. First, it will be critical to replicate the findings of our DREADD experiment using a technique that has been thoroughly verified, such as optogenetics. In the proposed study, we could infuse a floxed, retrograde viral vector encoding for a light-gated chloride (inhibitory) channel and an eYFP fluorescent marker into the NAc core of TH-Cre animals; an optic ferrule would also be placed into the VTA, which would be used to activate these inhibitory channels at the time of the cue during learning and performance of our autoshaping task. As hypothesized earlier, we believe that specifically inhibiting VTA to NAc neurons would slow down the acquisition of sign-tracking behaviors, or may confirm our results which suggested the promotion of goal-tracking behavior early on and potentiation of sign-tracking once inhibition was lifted.

As a follow-up, it might also be interesting to perform simultaneous FSCV recordings to measure DA release in the NAc during VTA-NAc inhibition in both learning, expression, and later when inhibition has been lifted. If these experiments were successful, a next step could be stimulating VTA-NAc neurons (using a retrograde ChR2 virus) in goal-trackers during the cue in early learning to see if we could speed up the development of sign-tracking behaviors. The sum results of these experiments would provide us with a more complete picture of the role of DA during the development and expression of sign- and goal- tracking behaviors and should provide more evidence supporting its necessity for both behavioral phenotypes.

Next, it would be interesting to see if behavior could be shifted in animals designated as poor avoiders by altering dopamine signals within the NAc. First, using a similar optogenetic technique as described above, we could stimulate or inhibit VTA-NAc DA in poor avoiders during the cue in our combined approach avoidance task. Since dopamine release to cues in these animals was shown to be uninformative and muddled across trial types, perhaps we could artificially train this teaching signal. For instance, if we stimulate VTA-NAc DA neurons during cues predicting reward and shock, while inhibiting these same DA neurons during cues predicting neutral events, this could potentially switch poor avoiders to good avoiders across time. By repairing these signals, I would predict that we would alter behavior in these animals to begin demonstrating good avoidance behavior. Additionally, it would be interesting to put good and poor avoiders in an autoshaping task, to see if they would uphold our predictions and behave as a sign- or goal- tracker, respectively; after we artificially retrained the DA PE signal of poor avoiders, we could then retest them in the autoshaping procedure to see if their behavior in this task also shifted.

The ultimate goal of these proposed studies is to further our understanding of how dopaminergic neural signals can influence cue-driven learning and performance, specifically relating to how differences in dopamine transmission can lead to individual behavioral differences across a number of tasks and contexts. Individual differences that arise in our tasks, specifically relating to poor avoidance and the formation of sign- vs. goal-tracking, are critical for furthering our understanding of certain psychopathologies. For instance, it is easy to imagine that if good avoiders/sign-trackers are driven more by cues than by outcomes, they may be more

susceptible to developing addictive and compulsive behaviors. These groups may also show deficits in delay-discounting, stop-signal, and go/no-go tasks, which have been historically employed to investigate choice and action impulsivity, as seen in disorders such as ADHD, Tourette's syndrome, and OCD. Measuring stress hormones and activity in other brain regions that are thought to be involved in the development of anxiety disorders (e.g., insula, amygdala) in poor avoiders would also be a valuable next step. Thus, in future work, it may be useful to examine individual differences in good/poor avoiders and sign- and goal- trackers during other tasks used to model deficits related to these psychiatric disorders. A more complete understanding of these mechanisms will be critical for the advancement of the field. Hopefully the work presented here, along with results from future studies based on this research, will provide insight that will help in the production of more specific and effective treatments for individuals suffering from disorders involving reinforcement and value-encoding dysfunction.

References

Adhikari, A., Lerner, T. N., Finkelstein, J., Pak, S., Jennings, J. H., Davidson, T. J., … Deisseroth, K. (2015). Basomedial amygdala mediates top-down control of anxiety and fear. *Nature*, *527*, 179–185. https://doi.org/10.1038/nature15698

Anderson, A. K., Christoff, K., Stappen, I., Panitz, D., Ghahremani, D. G., Glover, G., … Sobel, N. (2003). Dissociated neural representations of intensity and valence in human olfaction. *Nature Neuroscience*, *6*(2), 196–202. https://doi.org/10.1038/nn1001

Anstrom, K. K., Miczek, K. A., & Budygin, E. A. (2009). Increased phasic dopamine signaling in the mesolimbic pathway during social defeat in rats. *Neuroscience*, *161*(1), 3–12. https://doi.org/10.1016/j.neuroscience.2009.03.023

Armbruster, B. N., Li, X., Pausch, M. H., Herlitze, S., & Roth, B. L. (2007). Evolving the lock to fit the key to create a family of G protein-coupled receptors potently activated by an inert ligand. *Proceedings of the National Academy of Sciences*, *104*(12), 5163–5168. https://doi.org/10.1073/pnas.0700293104

Asede, D., Bosch, D., Lüthi, A., Ferraguti, F., & Ehrlich, I. (2015). Sensory inputs to intercalated cells provide fear-learning modulated inhibition to the basolateral amygdala. *Neuron*, *86*(2), 541–54. https://doi.org/10.1016/j.neuron.2015.03.008

Badrinarayan, a., Wescott, S. a., Vander Weele, C. M., Saunders, B. T., Couturier, B. E., Maren, S., & Aragona, B. J. (2012). Aversive Stimuli Differentially Modulate Real-Time Dopamine Transmission Dynamics within the Nucleus Accumbens Core and Shell. *Journal of Neuroscience*, *32*(45), 15779–15790.

https://doi.org/10.1523/JNEUROSCI.3557-12.2012

Bailey, K., & Crawley, J. (2009). Anxiety-Related Behaviors in Mice. In J. Buccafusco (Ed.), *Methods of Behavior Analysis in Neuroscience* (2nd ed.). Boca Raton: CRC Press. Retrieved from http://www.ncbi.nlm.nih.gov/books/NBK5221/

Balleine, B. W., & O'Doherty, J. P. (2010). Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology*, *35*(1), 48–69. https://doi.org/10.1038/npp.2009.131

Barker, J. M., Torregrossa, M. M., & Taylor, J. R. (2013). Bidirectional modulation of infralimbic dopamine D1 and D2 receptor activity regulates flexible reward seeking. *Frontiers in Neuroscience*, *7*, 126. https://doi.org/10.3389/fnins.2013.00126

Barto, A. G. (1995). Adaptive Critics and the Basal Ganglia, 215–232. Retrieved from http://scholarworks.umass.edu/cs_faculty_pubs

Belova, M. a, Paton, J. J., Morrison, S. E., & Salzman, C. D. (2007). Expectation modulates neural responses to pleasant and aversive stimuli in primate amygdala. *Neuron*, *55*(6), 970–84. https://doi.org/10.1016/j.neuron.2007.08.004

Bender, D., Holschbach, M., & Stöcklin, G. (1994). Synthesis of n.c.a. carbon-11 labelled clozapine and its major metabolite clozapine-N-oxide and comparison of their biodistribution in mice. *Nuclear Medicine and Biology*, *21*(7), 921–5. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9234345

Beninger, R. J., Mason, S. T., Phillips, A. G., & Fibiger, H. C. (1980). The Use of

Conditioned Suppression to Evaluate the Nature of Neuroleptic-Induced Avoidance Deficits. *Journal of Pharmacology and Experimental Therapeutics*, *213*(3), 623–627.

Berretta, S., Pantazopoulos, H., Caldera, M., Pantazopoulos, P., & Paré, D. (2005). Infralimbic cortex activation increases c-Fos expression in intercalated neurons of the amygdala. *Neuroscience*, *132*(4), 943–53. https://doi.org/10.1016/j.neuroscience.2005.01.020

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

Bissonette, G. B., Burton, A. C., Gentry, R. N., Goldstein, B. L., Hearn, T. N., Barnett, B. R., … Roesch, M. R. (2013). Separate Populations of Neurons in Ventral Striatum Encode Value and Motivation. *PLoS ONE*, *8*(5), e64673. https://doi.org/10.1371/journal.pone.0064673

Bissonette, G. B., Gentry, R. N., Padmala, S., Pessoa, L., & Roesch, M. R. (2014). Impact of appetitive and aversive outcomes on brain responses: linking the animal and human literatures. *Frontiers in Systems Neuroscience*, *8*(March), 24. https://doi.org/10.3389/fnsys.2014.00024

Bissonette, G. B., Powell, E. M., & Roesch, M. R. (2013). Neural structures underlying set-shifting: Roles of medial prefrontal cortex and anterior cingulate cortex. *Behavioural Brain Research*, *250*, 91–101. https://doi.org/10.1016/j.bbr.2013.04.037

Björklund, A., & Dunnett, S. B. (2007). Dopamine neuron systems in the brain: an

update. *Trends in Neurosciences*, *30*(5), 194–202. https://doi.org/10.1016/j.tins.2007.03.006

Blair, K., Marsh, A. A., Morton, J., Vythilingam, M., Jones, M., Mondillo, K., … Blair, J. R. (2006). Choosing the Lesser of Two Evils, the Better of Two Goods: Specifying the Roles of Ventromedial Prefrontal Cortex and Dorsal Anterior Cingulate in Object Choice. *Journal of Neuroscience*, *26*(44), 11379–11386. https://doi.org/10.1523/JNEUROSCI.1640-06.2006

Boehme, R., Deserno, L., Gleich, T., Katthagen, T., Pankow, a., Behr, J., … Schlagenhauf, F. (2015). Aberrant Salience Is Related to Reduced Reinforcement Learning Signals and Elevated Dopamine Synthesis Capacity in Healthy Adults. *Journal of Neuroscience*, *35*(28), 10103–10111. https://doi.org/10.1523/JNEUROSCI.0805-15.2015

Bolles, R. C. (1969). Avoidance and escape learning: simultaneous acquisition of different responses. *Journal of Comparative and Physiological Psychology*, *68*(3), 355–358. https://doi.org/10.1037/h0027536

Bolles, R. C. (1970). Species-specific defence reactions and avoidance learning. *Psychological Review*, *77*(1), 32–48. https://doi.org/10.1037/h0028589

Bolles, R. C., & Popp, R. J. (1964). Parameters affecting the acquisition of sidman avoidance. *Journal of the Experimental Analysis of Behavior*, *7*(4), 315–21.

Bossert, J. M., Stern, A. L., Theberge, F. R. M., Cifani, C., Koya, E., Hope, B. T., & Shaham, Y. (2011). Ventral medial prefrontal cortex neuronal ensembles mediate context-induced relapse to heroin. *Nature Neuroscience*, *14*(4), 420–2. https://doi.org/10.1038/nn.2758

Bravo-Rivera, C., Roman-Ortiz, C., Brignoni-Perez, E., Sotres-Bayon, F., & Quirk, G. J. (2014). Neural Structures Mediating Expression and Extinction of Platform-Mediated Avoidance. *Journal of Neuroscience*, *34*(29), 9736–9742. https://doi.org/10.1523/JNEUROSCI.0191-14.2014

Bravo-Rivera, C., Roman-Ortiz, C., Montesinos-Cartagena, M., & Quirk, G. J. (2015). Persistent active avoidance correlates with activity in prelimbic cortex and ventral striatum. *Frontiers in Behavioral Neuroscience*, *9*, 184. https://doi.org/10.3389/fnbeh.2015.00184

Breiter, H. C., Aharon, I., Kahneman, D., Dale, A., & Shizgal, P. (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron*, *30*(2), 619–39. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11395019

Brischoux, F., Chakraborty, S., Brierley, D. I., & Ungless, M. A. (2009). Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(12), 4894–9. https://doi.org/10.1073/pnas.0811507106

Britt, J. P., Benaliouad, F., McDevitt, R. A., Stuber, G. D., Wise, R. A., & Bonci, A. (2012). Synaptic and Behavioral Profile of Multiple Glutamatergic Inputs to the Nucleus Accumbens. *Neuron*, *76*(4), 790–803. https://doi.org/10.1016/j.neuron.2012.09.040

Bromberg-Martin, E. S., Matsumoto, M., & Hikosaka, O. (2010). Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron*, *68*(5), 815–34. https://doi.org/10.1016/j.neuron.2010.11.022

Brush, F. R. (2003). Selection for Differences in Avoidance Learning: The Syracuse Strains Differ in Anxiety, Not Learning Ability. *Behavior Genetics*, *33*(6), 677–696. https://doi.org/10.1023/A:1026135231594

Bryden, D. W., Johnson, E. E., Tobia, S. C., Kashtelyan, V., & Roesch, M. R. (2011). Attention for Learning Signals in Anterior Cingulate Cortex. *Journal of Neuroscience*, *31*(50), 18266–18274. https://doi.org/10.1523/JNEUROSCI.4715-11.2011

Budygin, E. a., Park, J., Bass, C. E., Grinevich, V. P., Bonin, K. D., & Wightman, R. M. (2012). Aversive stimulus differentially triggers subsecond dopamine release in reward regions. *Neuroscience*, *201*, 331–337. https://doi.org/10.1016/j.neuroscience.2011.10.056

Burgos-Robles, A., Bravo-Rivera, H., & Quirk, G. J. (2013). Prelimbic and Infralimbic Neurons Signal Distinct Aspects of Appetitive Instrumental Behavior. *PLoS ONE*, *8*(2), 1–7. https://doi.org/10.1371/journal.pone.0057575

Burgos-Robles, A., Vidal-Gonzalez, I., Santini, E., & Quirk, G. J. (2007). Consolidation of Fear Extinction Requires NMDA Receptor-Dependent Bursting in the Ventromedial Prefrontal Cortex. *Neuron*, *53*(6), 871–880. https://doi.org/10.1016/j.neuron.2007.02.021

Burton, A. C., Bissonette, G. B., Lichtenberg, N. T., Kashtelyan, V., & Roesch, M. R. (2014). Ventral Striatum Lesions Enhance Stimulus and Response Encoding in Dorsal Striatum. *Biological Psychiatry*, *75*(2), 132–139. https://doi.org/10.1016/j.biopsych.2013.05.023

Cai, H., Haubensak, W., Anthony, T. E., & Anderson, D. J. (2014). Central amygdala

PKC-δ(+) neurons mediate the influence of multiple anorexigenic signals. *Nature Neuroscience*, *17*(9), 1240–8. https://doi.org/10.1038/nn.3767

Calu, D. J., Roesch, M. R., Haney, R. Z., Holland, P. C., & Schoenbaum, G. (2010). Neural correlates of variations in event processing during learning in central nucleus of amygdala. *Neuron*, *68*(5), 991–1001. https://doi.org/10.1016/j.neuron.2010.11.019

Calvert, A. L., Green, L., & Myerson, J. (2011). Discounting in Pigeons When the Choice is between Two Delayed Rewards: Implications for Species Comparisons. *Frontiers in Neuroscience*, *5*, 96. https://doi.org/10.3389/fnins.2011.00096

Camp, M., Norcross, M., Whittle, N., Feyder, M., D'Hanis, W., Yilmazer-Hanke, D., … Holmes, A. (2009). Impaired Pavlovian fear extinction is a common phenotype across genetic lineages of the 129 inbred mouse strain. *Genes, Brain and Behavior*, *8*(8), 744–752. https://doi.org/10.1111/j.1601-183X.2009.00519.x

Cardinal, R. N., & Cheung, T. H. (2005). Nucleus accumbens core lesions retard instrumental learning and performance with delayed reinforcement in the rat. *BMC Neuroscience*, *6*(1), 9. https://doi.org/10.1186/1471-2202-6-9

Cardinal, R. N., Parkinson, J. a, Lachenal, G., Halkerston, K. M., Rudarakanchana, N., Hall, J., … Everitt, B. J. (2002). Effects of selective excitotoxic lesions of the nucleus accumbens core, anterior cingulate cortex, and central nucleus of the amygdala on autoshaping performance in rats. *Behavioral Neuroscience*, *116*(4), 553–567. https://doi.org/10.1037/0735-7044.116.4.553

Carelli, R. M., & Deadwyler, S. A. (1994). A comparison of nucleus accumbens

neuronal firing patterns during cocaine self-administration and water reinforcement in rats. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *14*(12), 7735–46. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7996208

Carter, R. M., MacInnes, J. J., Huettel, S. A., & Adcock, R. A. (2009). Activation in the VTA and nucleus accumbens increases in anticipation of both gains and losses. *Frontiers in Behavioral Neuroscience*, *3*, 21. https://doi.org/10.3389/neuro.08.021.2009

Carvalho, P. F., & Goldstone, R. L. (2015). The benefits of interleaved and blocked study: Different tasks benefit from different schedules of study. *Psychonomic Bulletin & Review*, *22*(1), 281–288. https://doi.org/10.3758/s13423-014-0676-4

Chang, C. H., Berke, J. D., & Maren, S. (2010). Single-unit activity in the medial prefrontal cortex during immediate and delayed extinction of fear in rats. *PLoS ONE*, *5*(8). https://doi.org/10.1371/journal.pone.0011971

Cheer, J. F., Aragona, B. J., Heien, M. L. a V, Seipel, A. T., Carelli, R. M., & Wightman, R. M. (2007). Coordinated Accumbal Dopamine Release and Neural Activity Drive Goal-Directed Behavior. *Neuron*, *54*(2), 237–244. https://doi.org/10.1016/j.neuron.2007.03.021

Cheer, J. F., Heien, M. L. A. V., Garris, P. A., Carelli, R. M., & Wightman, R. M. (2005). Simultaneous dopamine and single-unit recordings reveal accumbens GABAergic responses: Implications for intracranial self-stimulation. *Proceedings of the National Academy of Sciences*, *102*(52), 19150–19155. https://doi.org/10.1073/pnas.0509607102

Choi, J.-S., Cain, C. K., & LeDoux, J. E. (2010). The role of amygdala nuclei in the expression of auditory signaled two-way active avoidance in rats. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *17*(3), 139–147. https://doi.org/10.1101/lm.1676610

Choi, J. M., Padmala, S., & Pessoa, L. (2012). Impact of state anxiety on the interaction between threat monitoring and cognition. *NeuroImage*, *59*(2), 1912–1923. https://doi.org/10.1016/j.neuroimage.2011.08.102

Choi, J. M., Padmala, S., Spechler, P., & Pessoa, L. (2014). Pervasive competition between threat and reward in the brain. *Social Cognitive and Affective Neuroscience*, *9*(6), 737–750. https://doi.org/10.1093/scan/nst053

Clark, J. J., Hollon, N. G., & Phillips, P. E. M. (2012). Pavlovian valuation systems in learning and decision making. *Current Opinion in Neurobiology*, *22*(6), 1054–1061. https://doi.org/10.1016/j.conb.2012.06.004

Clark, J. J., Sandberg, S. G., Wanat, M. J., Gan, J. O., Horne, E. a, Hart, A. S., … Phillips, P. E. M. (2010). Chronic microsensors for longitudinal, subsecond dopamine detection in behaving animals. *Nature Methods*, *7*(2), 126–9. https://doi.org/10.1038/nmeth.1412

Cooper, B. R., Howard, J. L., Grant, L. D., Smith, R. D., & Breese, G. R. (1974). Alteration of avoidance and ingestive behavior after destruction of central catecholamine pathways with 6-hydroxydopamine. *Pharmacology, Biochemistry, and Behavior*, *2*(5), 639–649. https://doi.org/10.1016/0091-3057(74)90033-1

Cooper, J. C., & Knutson, B. (2008). Valence and salience contribute to nucleus

accumbens activation. *NeuroImage*, *39*(1), 538–47. https://doi.org/10.1016/j.neuroimage.2007.08.009

Corbit, L. H., Muir, J. L., & Balleine, B. W. (2001). The role of the nucleus accumbens in instrumental conditioning: Evidence of a functional dissociation between accumbens core and shell. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *21*(9), 3251–60. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11312310

Coutureau, E., & Killcross, S. (2003). Inactivation of the infralimbic prefrontal cortex reinstates goal-directed responding in overtrained rats. *Behavioural Brain Research*, *146*(1–2), 167–74. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14643469

Cragg, S. J. (2003). Variable dopamine release probability and short-term plasticity between functional domains of the primate striatum. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *23*(10), 4378–85. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12764127

Cragg, S. J., Baufreton, J., Xue, Y., Bolam, J. P., & Bevan, M. D. (2004). Synaptic release of dopamine in the subthalamic nucleus. *The European Journal of Neuroscience*, *20*(7), 1788–802. https://doi.org/10.1111/j.1460-9568.2004.03629.x

Cruz, E., López, A. V., & Porter, J. T. (2014). Spontaneous Recovery of Fear Reverses Extinction-Induced Excitability of Infralimbic Neurons. *PLoS ONE*, *9*(8), e103596. https://doi.org/10.1371/journal.pone.0103596

Danna, C. L., & Elmer, G. I. (2010). Disruption of conditioned reward association by

typical and atypical antipsychotics. *Pharmacology, Biochemistry, and Behavior*, *96*(1), 40–7. https://doi.org/10.1016/j.pbb.2010.04.004

Danna, C. L., & Elmer, G. I. (2011). Disruption of conditioned reward association by typical and atypical antipsychotics. *Pharmacology Biochemistry and Behavior*, *72*(2), 181–204. https://doi.org/10.1038/nature13314.A

Darvas, M., Fadok, J. P., & Palmiter, R. D. (2011). Requirement of dopamine signaling in the amygdala and striatum for learning and maintenance of a conditioned avoidance response. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *18*(3), 136–143. https://doi.org/10.1101/lm.2041211

Davis, M., Walker, D. L., Miles, L., & Grillon, C. (2010). Phasic vs sustained fear in rats and humans: role of the extended amygdala in fear vs anxiety. *Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology*, *35*(1), 105–35. https://doi.org/10.1038/npp.2009.109

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-Based Influences on Humans' Choices and Striatal Prediction Errors. *Neuron*, *69*(6), 1204–1215. https://doi.org/10.1016/j.neuron.2011.02.027

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. *Nature Neuroscience*, *10*(8), 1020–1028. https://doi.org/10.1038/nn1923

Dayan, P. (2012). Instrumental vigour in punishment and reward. *European Journal of Neuroscience*, *35*(7), 1152–1168. https://doi.org/10.1111/j.1460-

9568.2012.08026.x

Dayan, P., & Berridge, K. C. (2014). Model-based and model-free Pavlovian reward learning: Revaluation, revision, and revelation. *Cognitive, Affective & Behavioral Neuroscience*, *14*(2), 473–92. https://doi.org/10.3758/s13415-014-0277-8

Dayan, P., & Niv, Y. (2008). Reinforcement learning: The Good, The Bad and The Ugly. *Current Opinion in Neurobiology*, *18*(2), 185–196. https://doi.org/10.1016/j.conb.2008.08.003

Delgado, M. R., Beer, J. S., Fellows, L. K., Huettel, S. A., Platt, M. L., Quirk, G. J., & Schiller, D. (2016). Viewpoints: Dialogues on the functional role of the ventromedial prefrontal cortex. *Nature Neuroscience*, *19*(12), 1545–1552. https://doi.org/10.1038/nn.4438

Delgado, M. R., Li, J., Schiller, D., & Phelps, E. a. (2008). The role of the striatum in aversive learning and aversive prediction errors. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *363*(1511), 3787–800. https://doi.org/10.1098/rstb.2008.0161

Delgado, M. R., Nystrom, L. E., Fissell, C., Noll, D. C., & Fiez, J. A. (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. *Journal of Neurophysiology*, *84*(6), 3072–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11110834

Dias-Ferreira, E., Sousa, J. C., Melo, I., Morgado, P., Mesquita, A. R., Cerqueira, J. J., … Sousa, N. (2009). Chronic Stress Causes, *325*(July), 621–625.

Dickinson, A. (1985). Actions and Habits: The Development of Behavioural

Autonomy. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *308*(1135), 67–78. https://doi.org/10.1098/rstb.1985.0010

Dickinson, A., Nicholas, D. J., & Adams, C. D. (1983). The effect of the instrumental training contingency on susceptibility to reinforcer devaluation. *The Quarterly Journal of Experimental Psychology Section B*, *35*(1), 35–51. https://doi.org/10.1080/14640748308400912

DiFeliceantonio, A. G., & Berridge, K. C. (2012). Which cue to "want"? Opioid stimulation of central amygdala makes goal-trackers show stronger goal-tracking, just as sign-trackers show stronger sign-tracking. *Behavioural Brain Research*, *230*(2), 399–408. https://doi.org/10.1016/j.bbr.2012.02.032

Doll, B. B., Simon, D. A., & Daw, N. D. (2011). The ubiquity of model-based reinforcement learning. *Current Opinion in Neurobiology*, *72*(2), 181–204. https://doi.org/10.1038/nature13314.A

Doll, B. B., Simon, D. A., & Daw, N. D. (2012). The ubiquity of model-based reinforcement learning. *Current Opinion in Neurobiology*, *22*(6), 1075–81. https://doi.org/10.1016/j.conb.2012.08.003

Douglass, A. M., Kucukdereli, H., Ponserre, M., Markovic, M., Gründemann, J., Strobel, C., … Klein, R. (2017). Central amygdala circuits modulate food consumption through a positive-valence mechanism. *Nature Neuroscience*, *20*(10), 1384–1394. https://doi.org/10.1038/nn.4623

Eddy, M. C., Todd, T. P., Bouton, M. E., & Green, J. T. (2016). Medial prefrontal cortex involvement in the expression of extinction and ABA renewal of instrumental behavior for a food reinforcer. *Neurobiology of Learning and*

Memory, *128*, 33–9. https://doi.org/10.1016/j.nlm.2015.12.003

Ethridge, L. E., Brahmbhatt, S., Gao, Y., McDowell, J. E., & Clementz, B. A. (2009). Consider the context: blocked versus interleaved presentation of antisaccade trials. *Psychophysiology*, *46*(5), 1100–7. https://doi.org/10.1111/j.1469-8986.2009.00834.x

Everitt, B. J., Parkinson, J. a., Olmstead, M. C., Arroyo, M., Robledo, P., & Robbins, T. W. (1999). Associative processes in addiction and reward. The role of amygdala-ventral striatal subsystems. *Annals of the New York Academy of Sciences*, *877*, 412–438. https://doi.org/10.1111/j.1749-6632.1999.tb09280.x

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

Fang, J. (2000). Metabolism of clozapine by rat brain: the role of flavin-containing monooxygenase (FMO) and cytochrome P450 enzymes. *European Journal of Drug Metabolism and Pharmacokinetics*, *25*(2), 109–14. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11112091

Fernando, a, Urcelay, G., Mar, a, Dickinson, a, & Robbins, T. W. (2014). Free-Operant Avoidance Behavior by Rats after Reinforcer Revaluation Using Opioid Agonists and D-Amphetamine. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *34*(18), 6286–6293. https://doi.org/10.1523/JNEUROSCI.4146-13.2014

Flagel, S. B., Akil, H., & Robinson, T. E. (2009). Individual differences in the attribution of incentive salience to reward-related cues: Implications for

addiction. *Neuropharmacology*, *56 Suppl 1*, 139–48. https://doi.org/10.1016/j.neuropharm.2008.06.027

Flagel, S. B., Clark, J. J., Robinson, T. E., Mayo, L., Czuj, A., Akers, C. a, … Akil, H. (2011). A selective role for dopamine in reward learning. *Nature*, *469*(7328), 53–57. https://doi.org/10.1038/nature09588.A

Flagel, S. B., Watson, S. J., Robinson, T. E., & Akil, H. (2007). Individual differences in the propensity to approach signals vs goals promote different adaptations in the dopamine system of rats. *Psychopharmacology*, *191*(3), 599–607. https://doi.org/10.1007/s00213-006-0535-8

Flupentixol. (2005). Retrieved from https://www.drugbank.ca/drugs/DB00875

Francois, J., Huxter, J., Conway, M. W., Lowry, J. P., Tricklebank, M. D., & Gilmour, G. (2014). Differential Contributions of Infralimbic Prefrontal Cortex and Nucleus Accumbens during Reward-Based Learning and Extinction. *Journal of Neuroscience*, *34*(2), 596–607. https://doi.org/10.1523/JNEUROSCI.2346-13.2014

Freund, T. F., Powell, J. F., & Smith, A. D. (1984). Tyrosine hydroxylase-immunoreactive boutons in synaptic contact with identified striatonigral neurons, with particular reference to dendritic spines. *Neuroscience*, *13*(4), 1189–215. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6152036

Gallagher, M., Graham, P. W., & Holland, P. C. (1990). The amygdala central nucleus and appetitive Pavlovian conditioning: lesions impair one class of conditioned behavior. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *10*(6), 1906–11. Retrieved from

http://www.ncbi.nlm.nih.gov/pubmed/2355257

Gentry, R. N., Lee, B., & Roesch, M. R. (2016a). Phasic dopamine release in the rat nucleus accumbens predicts approach and avoidance performance. *Nature Communications*, *7*, 13154. https://doi.org/10.1038/ncomms13154

Gentry, R. N., Lee, B., & Roesch, M. R. (2016b). Phasic dopamine release in the rat nucleus accumbens predicts approach and avoidance performance. *Nature Communications*, *7*, 13154. https://doi.org/10.1038/ncomms13154

Giustino, T. F., Fitzgerald, P. J., & Maren, S. (2016). Fear Expression Suppresses Medial Prefrontal Cortical Firing in Rats. *PLOS ONE*, *11*(10), e0165256. https://doi.org/10.1371/journal.pone.0165256

Giustino, T. F., & Maren, S. (2015). The Role of the Medial Prefrontal Cortex in the Conditioning and Extinction of Fear. *Frontiers in Behavioral Neuroscience*, *9*, 298. https://doi.org/10.3389/fnbeh.2015.00298

Gomez, J. L., Bonaventura, J., Lesniak, W., Mathews, W. B., Sysa-Shah, P., Rodriguez, L. A., … Michaelides, M. (2017a). Chemogenetics revealed: DREADD occupancy and activation via converted clozapine. *Science*, *357*(6350), 503–507. https://doi.org/10.1126/science.aan2475

Gomez, J. L., Bonaventura, J., Lesniak, W., Mathews, W. B., Sysa-Shah, P., Rodriguez, L. A., … Michaelides, M. (2017b). Chemogenetics revealed: DREADD occupancy and activation via converted clozapine. *Science (New York, N.Y.)*, *357*(6350), 503–507. https://doi.org/10.1126/science.aan2475

Gourley, S. L., & Taylor, J. R. (2016). Going and stopping: dichotomies in behavioral control by the prefrontal cortex. *Nature Neuroscience*, *19*(6), 656–664.

https://doi.org/10.1038/nn.4275

Grace, A. A. (1991). Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: a hypothesis for the etiology of schizophrenia. *Neuroscience*, *41*(1), 1–24. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1676137

Grace, A. A. (1995). The tonic/phasic model of dopamine system regulation: its relevance for understanding how stimulant abuse can alter basal ganglia function. *Drug and Alcohol Dependence*, *37*(2), 111–29. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7758401

Graybiel, A. M. (2008). Habits, rituals, and the evaluative brain. *Annual Review of Neuroscience*, *31*(January), 359–387. https://doi.org/10.1146/annurev.neuro.29.051605.112851

Graybiel, A. M., Aosaki, T., Flaherty, A. W., & Kimura, M. (1994). The basal ganglia and adaptive motor control. *Science (New York, N.Y.)*, *265*(5180), 1826–31. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8091209

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

Haber, S. N. (2003). The primate basal ganglia: parallel and integrative networks. *Journal of Chemical Neuroanatomy*, *26*(4), 317–30. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14729134

Haber, S. N., Fudge, J. L., & McFarland, N. R. (2000). Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum.

The Journal of Neuroscience : The Official Journal of the Society for Neuroscience, *20*(6), 2369–82. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10704511

Haber, S. N., & Knutson, B. (2010). The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology*, *35*(1), 4–26. https://doi.org/10.1038/npp.2009.129

Hatfield, T., Han, J. S., Conley, M., Gallagher, M., & Holland, P. (1996). Neurotoxic lesions of basolateral, but not central, amygdala interfere with Pavlovian second-order conditioning and reinforcer devaluation effects. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *16*(16), 5256–65. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8756453

Hayden, B. Y., Heilbronner, S. R., Pearson, J. M., & Platt, M. L. (2011). Surprise signals in anterior cingulate cortex: neuronal encoding of unsigned reward prediction errors driving adjustment in behavior. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *31*(11), 4178–87. https://doi.org/10.1523/JNEUROSCI.4652-10.2011

Hefner, K., Whittle, N., Juhasz, J., Norcross, M., Karlsson, R.-M., Saksida, L. M., … Holmes, A. (2008). Impaired fear extinction learning and cortico-amygdala circuit abnormalities in a common genetic mouse strain. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *28*(32), 8074–85. https://doi.org/10.1523/JNEUROSCI.4904-07.2008

Heien, M. L. a V, Johnson, M. a., & Wightman, R. M. (2004). Resolving

neurotransmitters detected by fast-scan cyclic voltammetry. *Analytical Chemistry*, *76*(19), 5697–5704. https://doi.org/10.1021/ac0491509

Heien, M. L. a V, Khan, A. S., Ariansen, J. L., Cheer, J. F., Phillips, P. E. M., Wassum, K. M., & Wightman, R. M. (2005). Real-time measurement of dopamine fluctuations after cocaine in the brain of behaving rats. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(29), 10023–10028. https://doi.org/10.1073/pnas.0504657102

Heien, M. L. a V, Phillips, P. E. M., Stuber, G. D., Seipel, A. T., & Wightman, R. M. (2003). Overoxidation of carbon-fiber microelectrodes enhances dopamine adsorption and increases sensitivity. *The Analyst*, *128*(12), 1413–1419. https://doi.org/10.1039/b307024g

Heimer, L., Alheid, G. F., de Olmos, J. S., Groenewegen, H. J., Haber, S. N., Harlan, R. E., & Zahm, D. S. (1997). The accumbens: beyond the core-shell dichotomy. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *9*(3), 354–381. https://doi.org/10.1176/jnp.9.3.354

Heinz, a. (2002). Dopaminergic dysfunction in alcoholism and schizophrenia--psychopathological and behavioral correlates. *European Psychiatry : The Journal of the Association of European Psychiatrists*, *17*(1), 9–16. https://doi.org/S0924933802006284 [pii]

Hikida, T., Kimura, K., Wada, N., Funabiki, K., & Nakanishi, S. (2010). 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

Hikosaka, O. (2007). Basal Ganglia Mechanisms of Reward-Oriented Eye Movement. *Annals of the New York Academy of Sciences*, *1104*(1), 229–249. https://doi.org/10.1196/annals.1390.012

Holmes, A., Fitzgerald, P. J., MacPherson, K. P., DeBrouse, L., Colacicco, G., Flynn, S. M., … Camp, M. (2012). Chronic alcohol remodels prefrontal neurons and disrupts NMDAR-mediated fear extinction encoding. *Nature Neuroscience*, *15*(10), 1359–1361. https://doi.org/10.1038/nn.3204

Hosokawa, T., Kato, K., Inoue, M., & Mikami, A. (2007). Neurons in the macaque orbitofrontal cortex code relative preference of both rewarding and aversive outcomes. *Neuroscience Research*, *57*(3), 434–45. https://doi.org/10.1016/j.neures.2006.12.003

Hull, C. L. (1943). *Principles of Behavior: An Introduction to Behavior Theory. The Journal of Abnormal and Social Psychology* (Vol. 39). https://doi.org/10.1037/h0051597

Ikegami, M., Uemura, T., Kishioka, A., Sakimura, K., & Mishina, M. (2014). Striatal dopamine D1 receptor is essential for contextual fear conditioning. *Scientific Reports*, *4*, 3976. https://doi.org/10.1038/srep03976

Ikemoto, S., & Panksepp, J. (1999). The role of nucleus accumbens dopamine in motivated behavior: A unifying interpretation with special reference to reward-seeking. *Brain Research Reviews*, *31*(1), 6–41. https://doi.org/10.1016/S0165-0173(99)00023-5

Janak, P. H., Chen, M.-T., & Caulder, T. (2004). Dynamics of neural coding in the accumbens during extinction and reinstatement of rewarded behavior.

Behavioural Brain Research, *154*(1), 125–135. https://doi.org/10.1016/j.bbr.2004.02.003

Ji, B., Kaneko, H., Minamimoto, T., Inoue, H., Takeuchi, H., Kumata, K., … Higuchi, M. (2016). Multimodal Imaging for DREADD-Expressing Neurons in Living Brain and Their Application to Implantation of iPSC-Derived Neural Progenitors. *The Journal of Neuroscience*, *36*(45), 11544–11558. https://doi.org/10.1523/JNEUROSCI.1279-16.2016

Ji, G., & Neugebauer, V. (2012). Modulation of medial prefrontal cortical activity using in vivo recordings and optogenetics. *Molecular Brain*, *5*(1), 36. https://doi.org/10.1186/1756-6606-5-36

Kalenscher, T., & van Wingerden, M. (2011). Why We Should Use Animals to Study Economic Decision Making – A Perspective. *Frontiers in Neuroscience*, *5*, 82. https://doi.org/10.3389/fnins.2011.00082

Kapur, S. (2003). Psychosis as a state of aberrant salience: A framework linking biology, phenomenology, and pharmacology in schizophrenia. *American Journal of Psychiatry*, *160*(1), 13–23. https://doi.org/10.1176/appi.ajp.160.1.13

Katoh, A., Nabeshima, T., Kuno, A., Wada, M., Ukai, R., & Kameyama, T. (1996). Changes in striatal dopamine release in stress-induced conditioned suppression of motility in rats. *Behavioural Brain Research*, *77*(1–2), 219–221. https://doi.org/10.1016/0166-4328(95)00228-6

Kaveri, S., & Nakahara, H. (2014). Dual reward prediction components yield Pavlovian sign- and goal-tracking. *PLoS ONE*, *9*(10). https://doi.org/10.1371/journal.pone.0108142

Keistler, C., Barker, J. M., & Taylor, J. R. (2015). Infralimbic prefrontal cortex interacts with nucleus accumbens shell to unmask expression of outcome-selective Pavlovian-to-instrumental transfer. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *22*(10), 509–13. https://doi.org/10.1101/lm.038810.115

Keithley, R. B., Heien, M. L., & Wightman, R. M. (2009). Multivariate concentration determination using principal component regression With residual analysis. *Trends in Analytical Chemistry*, *28*(9), 1127–1136. https://doi.org/10.1016/j.trac.2009.07.002.Multivariate

Kemp, J. M., & Powell, T. P. (1971). The connexions of the striatum and globus pallidus: synthesis and speculation. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *262*(845), 441–57. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/4399123

Kennerley, S. W., Behrens, T. E. J., & Wallis, J. D. (2011). Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. *Nature Neuroscience*, *14*(12), 1581–1589. https://doi.org/10.1038/nn.2961

Killcross, S., & Coutureau, E. (2003). Coordination of actions and habits in the medial prefrontal cortex of rats. *Cerebral Cortex (New York, N.Y. : 1991)*, *13*(4), 400–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12631569

Kim, H., Lee, D., & Jung, M. W. (2013). Signals for previous goal choice persist in the dorsomedial, but not dorsolateral striatum of rats. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *33*(1), 52–63. https://doi.org/10.1523/JNEUROSCI.2422-12.2013

Kim, J., Zhang, X., Muralidhar, S., LeBlanc, S. A., & Tonegawa, S. (2017).

Basolateral to Central Amygdala Neural Circuits for Appetitive Behaviors. *Neuron*, *93*(6), 1464–1479.e5. https://doi.org/10.1016/j.neuron.2017.02.034

Kishioka, A., Fukushima, F., Ito, T., Kataoka, H., Mori, H., Ikeda, T., … Mishina, M. (2009). A novel form of memory for auditory fear conditioning at a low-intensity unconditioned stimulus. *PLoS ONE*, *4*(1). https://doi.org/10.1371/journal.pone.0004157

Kramer, R. (1998). *Chemometric techniques for quantitative analysis*. Marcel Dekker, Inc.

Kreitzer, A. C., & Malenka, R. C. (2008). Striatal plasticity and basal ganglia circuit function. *Neuron*, *60*(4), 543–54. https://doi.org/10.1016/j.neuron.2008.11.005

Kumar, V., Bhat, Z. A., & Kumar, D. (2013). Animal models of anxiety: A comprehensive review. *Journal of Pharmacological and Toxicological Methods*, *68*(2), 175–183. https://doi.org/10.1016/j.vascn.2013.05.003

LaLumiere, R. T., Smith, K. C., & Kalivas, P. W. (2012). Neural circuit competition in cocaine-seeking: roles of the infralimbic cortex and nucleus accumbens shell. *European Journal of Neuroscience*, *35*(4), 614–622. https://doi.org/10.1111/j.1460-9568.2012.07991.x

Lammel, S., Ion, D. I., Roeper, J., & Malenka, R. C. (2011). Projection-specific modulation of dopamine neuron synapses by aversive and rewarding stimuli. *Neuron*, *70*(5), 855–62. https://doi.org/10.1016/j.neuron.2011.03.025

Lavoie, B., Smith, Y., & Parent, A. (1989). Dopaminergic innervation of the basal ganglia in the squirrel monkey as revealed by tyrosine hydroxylase immunohistochemistry. *The Journal of Comparative Neurology*, *289*(1), 36–52.

https://doi.org/10.1002/cne.902890104

Lazaro-Munoz, G., LeDoux, J. E., & Cain, C. K. (2010). Sidman Instrumental Avoidance Initially Depends on Lateral and Basal Amygdala and is Constrained by Central Amygdala-mediated Pavlovian Processes. *Biological Psychiatry*, *67*(12), 1120–1127. https://doi.org/10.1016/j.biopsych.2009.12.002.Sidman

LeDoux, J. E. (2000). Emotion Circuits in the Brain. *Annual Review of Neuroscience*, *23*(1), 155–184. https://doi.org/10.1146/annurev.neuro.23.1.155

Lesaint, F., Sigaud, O., Clark, J. J., Flagel, S. B., & Khamassi, M. (2015). Experimental predictions drawn from a computational model of sign-trackers and goal-trackers. *J Physiol Paris*, *8*(12), 1699–1712. https://doi.org/10.1016/j.rasd.2014.08.015.Social

Lesaint, F., Sigaud, O., Flagel, S. B., Robinson, T. E., & Khamassi, M. (2014). Modelling individual differences in the form of pavlovian conditioned approach responses: a dual learning systems approach with factored representations. *PLoS Computational Biology*, *10*(2), e1003466. https://doi.org/10.1371/journal.pcbi.1003466

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

Lichtenberg, N. T., Pennington, Z. T., Holley, S. M., Greenfield, V. Y., Cepeda, C., Levine, M. S., & Wassum, K. M. (2017). Basolateral amygdala to orbitofrontal cortex projections enable cue-triggered reward expectations. *The Journal of Neuroscience*, *37*(35), 486–17. https://doi.org/10.1523/JNEUROSCI.0486-

17.2017

Likhtik, E., Popa, D., Apergis-Schoute, J., Fidacaro, G. A., & Paré, D. (2008). Amygdala intercalated neurons are required for expression of fear extinction. *Nature*, *454*(7204), 642–5. https://doi.org/10.1038/nature07167

Lissek, S., & van Meurs, B. (2014). Learning models of PTSD: Theoretical accounts and psychobiological evidence. *International Journal of Psychophysiology*. https://doi.org/10.1016/j.ijpsycho.2014.11.006

Litt, A., Plassmann, H., Shiv, B., & Rangel, A. (2011). Dissociating Valuation and Saliency Signals during Decision-Making. *Cerebral Cortex*, *21*(1), 95–102. https://doi.org/10.1093/cercor/bhq065

Lovibond, P. F., Saunders, J. C., Weidemann, G., & Mitchell, C. J. (2008). Evidence for expectancy as a mediator of avoidance and anxiety in a laboratory model of human avoidance learning. *Quarterly Journal of Experimental Psychology (2006)*, *61*(8), 1199–216. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18938780

Lucantonio, F., Gardner, M. P. H., Mirenzi, A., Newman, L. E., Takahashi, Y. K., & Schoenbaum, G. (2015). Neural Estimates of Imagined Outcomes in Basolateral Amygdala Depend on Orbitofrontal Cortex. *Journal of Neuroscience*, *35*(50), 16521–16530. https://doi.org/10.1523/JNEUROSCI.3126-15.2015

Lucantonio, F., Stalnaker, T. a, Shaham, Y., Niv, Y., & Schoenbaum, G. (2012). The impact of orbitofrontal dysfunction on cocaine addiction. *Nature Neuroscience*, *15*(3), 358–366. https://doi.org/10.1038/nn.3014

Mahler, S. V, & Berridge, K. C. (2009). Which cue to “want?” Central amygdala

opioid activation enhances and focuses incentive salience on a prepotent reward cue. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *29*(20), 6500–13. https://doi.org/10.1523/JNEUROSCI.3875-08.2009

Maia, T. V. (2010). Two-factor theory, the actor--critic model, and conditioned avoidance. *Learning & Behavior*, *38*(1), 50–67. https://doi.org/10.3758/LB.38.1.50

Maia, T. V, & Frank, M. J. (2011). From reinforcement learning models to psychiatric and neurological disorders. *Nature Neuroscience*, *14*(2), 154–162. https://doi.org/10.1038/nn.2723

Maren, S., & Quirk, G. J. (2004). Neuronal signalling of fear memory. *Nature Reviews. Neuroscience*, *5*(11), 844–852. https://doi.org/10.1038/nrn1535

Maroun, M., Kavushansky, A., Holmes, A., Wellman, C., & Motanis, H. (2012). Enhanced Extinction of Aversive Memories by High-Frequency Stimulation of the Rat Infralimbic Cortex. *PLoS ONE*, *7*(5), e35853. https://doi.org/10.1371/journal.pone.0035853

Martinez, R. C. R., Gupta, N., Lázaro-Muñoz, G., Sears, R. M., Kim, S., Moscarello, J. M., … Cain, C. K. (2013). Active vs. reactive threat responding is associated with differential c-Fos expression in specific regions of amygdala and prefrontal cortex. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *20*(8), 446–52. https://doi.org/10.1101/lm.031047.113

Matsumoto, M., & Hikosaka, O. (2009). Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature*, *459*(7248), 837–841.

https://doi.org/10.1038/nature08028

McCullough, L. D., Sokolowski, J. D., & Salamone, J. D. (1993). A neurochemical and behavioral investigation of the involvement of nucleus accumbens dopamine in instrumental avoidance. *Neuroscience*, *52*(4), 919–925.

McDonald, A. J. (1998). Cortical pathways to the mammalian amygdala. *Progress in Neurobiology*, *55*(3), 257–332. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9643556

McDonald, A. J., Mascagni, F., & Guo, L. (1996). Projections of the medial and lateral prefrontal cortices to the amygdala: a Phaseolus vulgaris leucoagglutinin study in the rat. *Neuroscience*, *71*(1), 55–75. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8834392

Meyer, P. J., Lovic, V., Saunders, B. T., Yager, L. M., Flagel, S. B., Morrow, J. D., & Robinson, T. E. (2012). Quantifying individual variation in the propensity to attribute incentive salience to reward cues. *PLoS ONE*, *7*(6). https://doi.org/10.1371/journal.pone.0038987

Milad, M. R., & Quirk, G. J. (2002). Neurons in medial prefrontal cortex signal memory for fear extinction. *Nature*, *420*(6911), 70–74. https://doi.org/10.1038/nature01138

Milad, M. R., Wright, C. I., Orr, S. P., Pitman, R. K., Quirk, G. J., & Rauch, S. L. (2007). Recall of Fear Extinction in Humans Activates the Ventromedial Prefrontal Cortex and Hippocampus in Concert. *Biological Psychiatry*, *62*(5), 446–454. https://doi.org/10.1016/j.biopsych.2006.10.011

Mishara, A. L., & Fusar-Poli, P. (2013). The phenomenology and neurobiology of

delusion formation during psychosis onset: Jaspers, truman symptoms, and aberrant salience. *Schizophrenia Bulletin*, *39*(2), 278–286. https://doi.org/10.1093/schbul/sbs155

Mogenson, G. J., Jones, D. L., & Yim, C. Y. (1980). From motivation to action: functional interface between the limbic system and the motor system. *Progress in Neurobiology*, *14*(2–3), 69–97. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6999537

Monosov, I. E. (2017). Anterior cingulate is a source of valence-specific information about value and uncertainty. *Nature Communications*, *8*(1), 134. https://doi.org/10.1038/s41467-017-00072-y

Monosov, I. E., & Hikosaka, O. (2012). Regionally Distinct Processing of Rewards and Punishments by the Primate Ventromedial Prefrontal Cortex. *Journal of Neuroscience*, *32*(30), 10318–10330. https://doi.org/10.1523/JNEUROSCI.1801-12.2012

Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *Journal of Neuroscience*, *16*(5), 1936–1947. https://doi.org/10.1.1.156.635

Moorman, D. E., & Aston-Jones, G. (2015a). Prefrontal neurons encode context-based response execution and inhibition in reward seeking and extinction. *Proceedings of the National Academy of Sciences*, *112*(30), 9472–9477. https://doi.org/10.1073/pnas.1507611112

Moorman, D. E., & Aston-Jones, G. (2015b). Prefrontal neurons encode context-based response execution and inhibition in reward seeking and extinction.

Proceedings of the National Academy of Sciences, *112*(30), 201507611. https://doi.org/10.1073/pnas.1507611112

Morgan, M. A., & LeDoux, J. E. (1995). Differential contribution of dorsal and ventral medial prefrontal cortex to the acquisition and extinction of conditioned fear in rats. *Behavioral Neuroscience*, *109*(4), 681–688.

Morris, R. G. . (1974). Pavlovian conditioned inhibition of fear during shuttlebox avoidance behavior. *Learning and Motivation*, *5*(4), 424–447. https://doi.org/10.1016/0023-9690(74)90002-2

Morrison, S. E., Saez, A., Lau, B., & Salzman, C. D. (2011). Different Time Courses for Learning-Related Changes in Amygdala and Orbitofrontal Cortex. *Neuron*, *71*(6), 1127–1140. https://doi.org/10.1016/j.neuron.2011.07.016

Morrison, S. E., & Salzman, C. D. (2011). Representations of appetitive and aversive information in the primate orbitofrontal cortex. *Annals of the New York Academy of Sciences*, *1239*(1), 59–70. https://doi.org/10.1111/j.1749-6632.2011.06255.x

Moscarello, J. M., & LeDoux, J. E. (2013). Active Avoidance Learning Requires Prefrontal Suppression of Amygdala-Mediated Defensive Reactions. *Journal of Neuroscience*, *33*(9), 3815–3823. https://doi.org/10.1523/JNEUROSCI.2596-12.2013

Moscarello, J. M., & LeDoux, J. E. (2013). Active avoidance learning requires prefrontal suppression of amygdala-mediated defensive reactions. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *33*(9), 3815–23. https://doi.org/10.1523/JNEUROSCI.2596-12.2013

Mowrer, O. H., & Lamoreaux, R. R. (1946). Fear as an intervening variable in

avoidance conditioning. *Journal of Comparative Psychology*, *39*(1), 29–50. https://doi.org/10.1037/h0060150

National Center for Biotechnology Information. (2017). Clozapine. https://doi.org/2818

Nestler, E. J. (2005). Is there a common molecular pathway for addiction? *Nature Neuroscience*, *8*(11), 1445–9. https://doi.org/10.1038/nn1578

Nishi, A., Kuroiwa, M., & Shuto, T. (2011). Mechanisms for the Modulation of Dopamine D1 Receptor Signaling in Striatal Neurons. *Frontiers in Neuroanatomy*, *5*, 43. https://doi.org/10.3389/fnana.2011.00043

Niv, Y., & Schoenbaum, G. (2008). Dialogues on prediction errors. *Trends in Cognitive Sciences*, *12*(7), 265–72. https://doi.org/10.1016/j.tics.2008.03.006

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable Roles of Ventral and Dorsal Striatum in Instrumental Conditioning. *Science*, *304*(5669), 452–454. https://doi.org/10.1126/science.1094285

O'Doherty, J. P. (2011). Contributions of the ventromedial prefrontal cortex to goal-directed action selection. *Annals of the New York Academy of Sciences*, *1239*(1), 118–129. https://doi.org/10.1111/j.1749-6632.2011.06290.x

O'Hare, J. K., Ade, K. K., Sukharnikova, T., Van Hooser, S. D., Palmeri, M. L., Yin, H. H., & Calakos, N. (2016). Pathway-Specific Striatal Substrates for Habitual Behavior. *Neuron*, *89*(3), 472–9. https://doi.org/10.1016/j.neuron.2015.12.032

Oleson, E. B., & Cheer, J. F. (2013). On the role of subsecond dopamine release in conditioned avoidance. *Frontiers in Neuroscience*, *7*(7 JUN), 1–9.

https://doi.org/10.3389/fnins.2013.00096

Oleson, E. B., Gentry, R. N., Chioma, V. C., & Cheer, J. F. (2012). Subsecond Dopamine Release in the Nucleus Accumbens Predicts Conditioned Punishment and Its Successful Avoidance. *Journal of Neuroscience*, *32*(42), 14804–14808. https://doi.org/10.1523/JNEUROSCI.3087-12.2012

Öngür, D., Ferry, A. T., & Price, J. L. (2003). Architectonic subdivision of the human orbital and medial prefrontal cortex. *The Journal of Comparative Neurology*, *460*(3), 425–449. https://doi.org/10.1002/cne.10609

Ostlund, S. B., & Balleine, B. W. (2005). Lesions of Medial Prefrontal Cortex Disrupt the Acquisition But Not the Expression of Goal-Directed Learning. *Journal of Neuroscience*, *25*(34), 7763–7770. https://doi.org/10.1523/JNEUROSCI.1921-05.2005

Padoa-Schioppa, C., & Assad, J. a. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature*, *441*(7090), 223–6. https://doi.org/10.1038/nature04676

Padoa-Schioppa, C., & Assad, J. A. (2008). The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. *Nature Neuroscience*, *11*(1), 95–102. https://doi.org/10.1038/nn2020

Parker, N. F., Cameron, C. M., Taliaferro, J. P., Lee, J., Choi, J. Y., Davidson, T. J., … Witten, I. B. (2016). Reward and choice encoding in terminals of midbrain dopamine neurons depends on striatal target. *Nature Neuroscience*, *19*(6), 845–854. https://doi.org/10.1038/nn.4287

Parkinson, J. A., Robbins, T. W., & Everitt, B. J. (2000). Dissociable roles of the central and basolateral amygdala in appetitive emotional learning. *The European*

Journal of Neuroscience, *12*(1), 405–13. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10651899

Paxinos, G., & Watson, C. (2007). *The Rat Brain in Stereotaxic Coordinates* (6th ed.). Burlington, MA: Elsevier Inc.

Peters, J., LaLumiere, R. T., & Kalivas, P. W. (2008). Infralimbic Prefrontal Cortex Is Responsible for Inhibiting Cocaine Seeking in Extinguished Rats. *Journal of Neuroscience*, *28*(23), 6046–6053. https://doi.org/10.1523/JNEUROSCI.1045-08.2008

Peters, J., Pattij, T., & De Vries, T. J. (2013). Targeting cocaine versus heroin memories: divergent roles within ventromedial prefrontal cortex. *Trends in Pharmacological Sciences*, *34*(12), 689–695. https://doi.org/10.1016/j.tips.2013.10.004

Phillips, P. E. M., Robinson, D. L., Stuber, G. D., Carelli, R. M., & Wightman, R. M. (2003). Real-time measurements of phasic changes in extracellular dopamine concentration in freely moving rats by fast-scan cyclic voltammetry. *Methods Mol Med*, *79*, 443–464. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12506716

Pickens, C. L., Saddoris, M. P., Gallagher, M., & Holland, P. C. (2005). Orbitofrontal Lesions Impair Use of Cue-Outcome Associations in a Devaluation Task. *Behavioral Neuroscience*, *119*(1), 317–322. https://doi.org/10.1037/0735-7044.119.1.317

Pickens, C. L., Saddoris, M. P., Setlow, B., Gallagher, M., Holland, P. C., &

Schoenbaum, G. (2003). Different roles for orbitofrontal cortex and basolateral amygdala in a reinforcer devaluation task. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *23*(35), 11078–84. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14657165

Pinard, C. R., Mascagni, F., & McDonald, A. J. (2012). Medial prefrontal cortical innervation of the intercalated nuclear region of the amygdala. *Neuroscience*, *205*, 112–24. https://doi.org/10.1016/j.neuroscience.2011.12.036

Plassmann, H., O'Doherty, J. P., & Rangel, A. (2010). Appetitive and Aversive Goal Values Are Encoded in the Medial Orbitofrontal Cortex at the Time of Decision Making. *Journal of Neuroscience*, *30*(32), 10799–10808. https://doi.org/10.1523/JNEUROSCI.0788-10.2010

Purgert, R. J., Wheeler, D. S., McDannald, M. a., & Holland, P. C. (2012). Role of Amygdala Central Nucleus in Aversive Learning Produced by Shock or by Unexpected Omission of Food. *Journal of Neuroscience*, *32*(7), 2461–2472. https://doi.org/10.1523/JNEUROSCI.5090-11.2012

Quirk, G. J., & Mueller, D. (2008). Neural Mechanisms of Extinction Learning and Retrieval. *Neuropsychopharmacology*, *33*(1), 56–72. https://doi.org/10.1038/sj.npp.1301555

Redish, A. D. (2004). Addiction as a computational process gone awry. *Science (New York, N.Y.)*, *306*(5703), 1944–7. https://doi.org/10.1126/science.1102384

Rescorla, R. A., & Solomon, R. L. (1967). Two-process learning theory: relationships between Pavlovian conditioning and instrumental training. *Psychological Review*, 151–83. https://doi.org/10.1016/j.ajog.2010.07.025

Rice, M. E., Patel, J. C., & Cragg, S. J. (2011). Dopamine release in the basal ganglia. *Neuroscience*, *198*, 112–37. https://doi.org/10.1016/j.neuroscience.2011.08.066

Robbins, T. W., & Sahakian, B. J. (1983). Behavioral effects of psychomotor stimulant drugs: clinical and neuropsychological implications. In *Stimulants: Neurochemical, Behavioral, and Clinical Perspective* (pp. 301–338).

Robinson, M. J. F., Warlow, S. M., & Berridge, K. C. (2014). Optogenetic Excitation of Central Amygdala Amplifies and Narrows Incentive Motivation to Pursue One Reward Above Another. *Journal of Neuroscience*, *34*(50), 16567–16580. https://doi.org/10.1523/JNEUROSCI.2013-14.2014

Robinson, T. E., & Flagel, S. B. (2009). Dissociating the Predictive and Incentive Motivational Properties of Reward-Related Cues Through the Study of Individual Differences. *Biological Psychiatry*, *65*(10), 869–873. https://doi.org/10.1016/j.biopsych.2008.09.006

Robinson, T. E., & Flagel, S. B. (2010). Dissociating the predictive and incentive motivational properties of reward-related cues through the study of individual differences. *Biological Psychiatry*, *65*(10), 869–873. https://doi.org/10.1016/j.biopsych.2008.09.006.Dissociating

Roesch, M. R., Calu, D. J., Esber, G. R., & Schoenbaum, G. (2010). Neural Correlates of Variations in Event Processing during Learning in Basolateral Amygdala. *Journal of Neuroscience*, *30*(7), 2464–2471. https://doi.org/10.1523/JNEUROSCI.5781-09.2010

Roesch, M. R., & Olson, C. R. (2004). Neuronal activity related to reward value and motivation in primate frontal cortex. *Science (New York, N.Y.)*, *304*(5668), 307–

10. https://doi.org/10.1126/science.1093223

Roesch, M. R., Singh, T., Brown, P. L., Mullins, S. E., & Schoenbaum, G. (2009). Ventral Striatal Neurons Encode the Value of the Chosen Action in Rats Deciding between Differently Delayed or Sized Rewards. *Journal of Neuroscience*, *29*(42), 13365–13376. https://doi.org/10.1523/JNEUROSCI.2572-09.2009

Rogan, S., & Roth, B. (2011). Remote control of neuronal signaling. *Pharmacological Reviews*, *63*(2), 291–315. https://doi.org/10.1124/pr.110.003020.291

Roitman, M. F., Wheeler, R. a, Wightman, R. M., & Carelli, R. M. (2008). Real-time chemical responses in the nucleus accumbens differentiate rewarding and aversive stimuli. *Nature Neuroscience*, *11*(12), 1376–1377. https://doi.org/10.1038/nn.2219

Rolls, E. T., Sienkiewicz, Z. J., & Yaxley, S. (1989). Hunger Modulates the Responses to Gustatory Stimuli of Single Neurons in the Caudolateral Orbitofrontal Cortex of the Macaque Monkey. *European Journal of Neuroscience*, *1*(1), 53–60. https://doi.org/10.1111/j.1460-9568.1989.tb00774.x

Roth, B. L. (2016). DREADDs for Neuroscientists. *Neuron*, *89*(4), 683–94. https://doi.org/10.1016/j.neuron.2016.01.040

Roy, M., Shohamy, D., & Wager, T. D. (2012). Ventromedial prefrontal-subcortical systems and the generation of affective meaning. *Trends in Cognitive Sciences*, *16*(3), 147–56. https://doi.org/10.1016/j.tics.2012.01.005

Saddoris, M. P., Gallagher, M., & Schoenbaum, G. (2005). Rapid Associative

Encoding in Basolateral Amygdala Depends on Connections with Orbitofrontal Cortex. *Neuron*, *46*(2), 321–331. https://doi.org/10.1016/j.neuron.2005.02.018

Salamone, J. D., & Correa, M. (2002). Motivational views of reinforcement: Implications for understanding the behavioral functions of nucleus accumbens dopamine. *Behavioural Brain Research*, *137*(1–2), 3–25. https://doi.org/10.1016/S0166-4328(02)00282-6

Salamone, J. D., & Correa, M. (2012). The Mysterious Motivational Functions of Mesolimbic Dopamine. *Neuron*, *76*(3), 470–485. https://doi.org/10.1016/j.neuron.2012.10.021

Salgado, S., & Kaplitt, M. G. (2015). The Nucleus Accumbens: A Comprehensive Review. *Stereotactic and Functional Neurosurgery*, *93*(2), 75–93. https://doi.org/10.1159/000368279

Sangha, S., Robinson, P. D., Greba, Q., Davies, D. A., & Howland, J. G. (2014a). Alterations in Reward, Fear and Safety Cue Discrimination after Inactivation of the Rat Prelimbic and Infralimbic Cortices. *Neuropsychopharmacology*, *39*(10), 2405–2413. https://doi.org/10.1038/npp.2014.89

Sangha, S., Robinson, P. D., Greba, Q., Davies, D. a, & Howland, J. G. (2014b). Alterations in reward, fear and safety cue discrimination after inactivation of the rat prelimbic and infralimbic cortices. *Neuropsychopharmacology*, *39*(10), 2405–2413. https://doi.org/10.1038/npp.2014.89

Saunders, B. T., & Robinson, T. E. (2012). The role of dopamine in the accumbens core in the expression of Pavlovian-conditioned responses. *The European Journal of Neuroscience*, *36*(4), 2521–32. https://doi.org/10.1111/j.1460-

9568.2012.08217.x

Saunders, B. T., & Robinson, T. E. (2013a). Individual variation in resisting temptation: Implications for addiction. *Neuroscience and Biobehavioral Reviews*, *37*(9), 1955–1975. https://doi.org/10.1016/j.neubiorev.2013.02.008

Saunders, B. T., & Robinson, T. E. (2013b). The role of dopamine in the accumbens core in the expression of Pavlovian conditioned responses. *European Journal of Neuroscience*, *36*(4), 2521–2532. https://doi.org/10.1111/j.1460-9568.2012.08217.x.The

Schiller, D., & Delgado, M. R. (2010). Overlapping neural systems mediating extinction, reversal and regulation of fear. *Trends in Cognitive Sciences*, *14*(6), 268–276. https://doi.org/10.1016/j.tics.2010.04.002

Schoenbaum, G., Chiba, A. A., & Gallagher, M. (1998). Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. *Nature Neuroscience*, *1*(2), 155–9. https://doi.org/10.1038/407

Schoenbaum, G., Chiba, A. A., & Gallagher, M. (1999). Neural encoding in orbitofrontal cortex and basolateral amygdala during olfactory discrimination learning. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *19*(5), 1876–84. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10024371

Schoenbaum, G., & Roesch, M. (2005). Orbitofrontal Cortex, Associative Learning, and Expectancies. *Neuron*, *47*(5), 633–636. https://doi.org/10.1016/j.neuron.2005.07.018

Schoenbaum, G., Roesch, M. R., Stalnaker, T. a, & Takahashi, Y. K. (2009). A new

perspective on the role of the orbitofrontal cortex in adaptive behaviour. *Nature Reviews. Neuroscience*, *10*(12), 885–92. https://doi.org/10.1038/nrn2753

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology*, *80*(1), 1–27. https://doi.org/10.1007/s00429-010-0262-0

Schultz, W. (2000). Multiple reward signals in the brain. *Nature Reviews. Neuroscience*, *1*(December), 199–207. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11257908

Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. *Annual Review of Psychology*, *57*, 87–115. https://doi.org/10.1146/annurev.psych.56.091103.070229

Schultz, W., Apicella, P., Scarnati, E., & Ljungberg, T. (1992). Neuronal activity in monkey ventral striatum related to the expectation of reward. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *12*(12), 4595–610. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1464759

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science (New York, N.Y.)*, *275*(5306), 1593–1599. https://doi.org/10.1126/science.275.5306.1593

Schwabe, L., & Wolf, O. T. (2009). Stress prompts habit behavior in humans. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *29*(22), 7191–7198. https://doi.org/10.1523/JNEUROSCI.0979-09.2009

Schwabe, L., & Wolf, O. T. (2011). Stress-induced modulation of instrumental behavior: From goal-directed to habitual control of action. *Behavioural Brain Research*, *219*(2), 321–328. https://doi.org/10.1016/j.bbr.2010.12.038

Schwartz, N., Miller, C., & Fields, H. L. (2017). Cortico-Accumbens Regulation of Approach-Avoidance Behavior Is Modified by Experience and Chronic Pain. *Cell Reports*, *19*(8), 1522–1531. https://doi.org/10.1016/j.celrep.2017.04.073

Senn, V., Wolff, S. B. E., Herry, C., Grenier, F., Ehrlich, I., Gründemann, J., … Lüthi, A. (2014). Long-range connectivity defines behavioral specificity of amygdala neurons. *Neuron*, *81*(2), 428–37. https://doi.org/10.1016/j.neuron.2013.11.006

Seo, D., Funderburk, S. C., Bhatti, D. L., Motard, L. E., Newbold, D., Girven, K. S., … Bruchas, M. R. (2016). A GABAergic Projection from the Centromedial Nuclei of the Amygdala to Ventromedial Prefrontal Cortex Modulates Reward Behavior. *The Journal of Neuroscience*, *36*(42), 10831–10842. https://doi.org/10.1523/JNEUROSCI.1164-16.2016

Setlow, B., Schoenbaum, G., & Gallagher, M. (2003). Neural encoding in ventral striatum during olfactory discrimination learning. *Neuron*, *38*(4), 625–36. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12765613

Seymour, B., Singer, T., & Dolan, R. (2007). The neurobiology of punishment. *Nature Reviews. Neuroscience*, *8*(4), 300–311. https://doi.org/10.1038/nrn2119

Shackman, A. J., & Fox, A. S. (2016). Contributions of the Central Extended Amygdala to Fear and Anxiety. *Journal of Neuroscience*, *36*(31), 8050–8063. https://doi.org/10.1523/JNEUROSCI.0982-16.2016

Singh, T., McDannald, M. A., Haney, R. Z., Cerri, D. H., & Schoenbaum, G. (2010). Nucleus Accumbens Core and Shell are Necessary for Reinforcer Devaluation Effects on Pavlovian Conditioned Responding. *Frontiers in Integrative*

Neuroscience, *4*, 126. https://doi.org/10.3389/fnint.2010.00126

Small, D. M., Gregory, M. D., Mak, Y. E., Gitelman, D., Mesulam, M. M., & Parrish, T. (2003). Dissociation of neural representation of intensity and affective valuation in human gustation. *Neuron*, *39*(4), 701–11. https://doi.org/10.1016/S0896-6273(03)00467-7

Soares, J. M., Sampaio, a, Ferreira, L. M., Santos, N. C., Marques, F., Palha, J. a, … Sousa, N. (2012). Stress-induced changes in human decision-making are reversible. *Translational Psychiatry*, *2*(7), e131. https://doi.org/10.1038/tp.2012.59

Soler-Cedeño, O., Cruz, E., Criado-Marrero, M., & Porter, J. T. (2016). Contextual fear conditioning depresses infralimbic excitability. *Neurobiology of Learning and Memory*, *130*, 77–82. https://doi.org/10.1016/j.nlm.2016.01.015

Solway, A., & Botvinick, M. (2011). Goal-directed decision making as probabilistic inference: A computational framework and potential neural correlates. *Psychological Review*, *72*(2), 181–204. https://doi.org/10.1038/nature13314.A

Solway, A., & Botvinick, M. M. (2012). Goal-directed decision making as probabilistic inference: a computational framework and potential neural correlates. *Psychological Review*, *119*(1), 120–54. https://doi.org/10.1037/a0026435

Song, C., Ehlers, V. L., & Moyer, J. R. (2015). Trace Fear Conditioning Differentially Modulates Intrinsic Excitability of Medial Prefrontal Cortex-Basolateral Complex of Amygdala Projection Neurons in Infralimbic and Prelimbic Cortices. *Journal of Neuroscience*, *35*(39), 13511–13524.

https://doi.org/10.1523/JNEUROSCI.2329-15.2015

Sotres-Bayon, F., & Quirk, G. J. (2010). Prefrontal control of fear: more than just extinction. *Current Opinion in Neurobiology*, *20*(2), 231–5. https://doi.org/10.1016/j.conb.2010.02.005

Stalnaker, T. A., España, R. A., & Berridge, C. W. (2009). Coping behavior causes asymmetric changes in neuronal activation in the prefrontal cortex and amygdala. *Synapse*, *63*(1), 82–85. https://doi.org/10.1002/syn.20583

Stalnaker, T. A., Roesch, M. R., Calu, D. J., Burke, K. A., Singh, T., & Schoenbaum, G. (2007). Neural Correlates of Inflexible Behavior in the Orbitofrontal Amygdalar Circuit after Cocaine Exposure. *Annals of the New York Academy of Sciences*, *1121*(1), 598–609. https://doi.org/10.1196/annals.1401.014

Surmeier, D. J., Song, W. J., & Yan, Z. (1996). Coordinated expression of dopamine receptors in neostriatal medium spiny neurons. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *16*(20), 6579–91. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8815934

Suto, N., Laque, A., De Ness, G. L., Wagner, G. E., Watry, D., Kerr, T., … Weiss, F. (2016). Distinct memory engrams in the infralimbic cortex of rats control opposing environmental actions on a learned behavior. *eLife*, *5*. https://doi.org/10.7554/eLife.21920

Sutton, R. S., & Barto, A. G. (1987). A temporal-difference model of classical conditioning. In *Proceedings of the ninth annual conference of the cognitive science society* (pp. 355–378).

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement Learning: An Introduction*.

Retrieved from https://pdfs.semanticscholar.org/aa32/c33e7c832e76040edc85e8922423b1a1db77.pdf

Takahashi, Y. K., Stalnaker, T. A., Roesch, M. R., & Schoenbaum, G. (2017). Effects of inference on dopaminergic prediction errors depend on orbitofrontal processing. *Behavioral Neuroscience*, *131*(2), 127–134. https://doi.org/10.1037/bne0000192

Takahashi, Y., Roesch, M. R., Stalnaker, T. A., & Schoenbaum, G. (2007). Cocaine exposure shifts the balance of associative encoding from ventral to dorsolateral striatum. *Frontiers in Integrative Neuroscience*, *1*, 11. https://doi.org/10.3389/neuro.07.011.2007

Taylor, S. B., Anglin, J. M., Paode, P. R., Riggert, a. G., Olive, M. F., & Conrad, C. D. (2014). Chronic stress may facilitate the recruitment of habit- and addiction-related neurocircuitries through neuronal restructuring of the striatum. *Neuroscience*, *280*, 231–242. https://doi.org/10.1016/j.neuroscience.2014.09.029

Tye, K. M., Cone, J. J., Schairer, W. W., & Janak, P. H. (2010). Amygdala Neural Encoding of the Absence of Reward during Extinction. *Journal of Neuroscience*, *30*(1), 116–125. https://doi.org/10.1523/JNEUROSCI.4240-09.2010

Tye, K. M., & Janak, P. H. (2007). Amygdala neurons differentially encode motivation and reinforcement. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *27*(15), 3937–45. https://doi.org/10.1523/JNEUROSCI.5281-06.2007

van Meurs, B., Wiggert, N., Wicker, I., & Lissek, S. (2014). Maladaptive behavioral

consequences of conditioned fear-generalization: A pronounced, yet sparsely studied, feature of anxiety pathology. *Behaviour Research and Therapy*, *57*(1), 29–37. https://doi.org/10.1016/j.brat.2014.03.009

Vertes, R. P. (2004). Differential Projections of the Infralimbic and Prelimbic Cortex in the Rat. *Synapse*, *51*(1), 32–58. https://doi.org/10.1002/syn.10279

Volman, S. F., Lammel, S., Margolis, E. B., Kim, Y., Richard, J. M., Roitman, M. F., & Lobo, M. K. (2013). New insights into the specificity and plasticity of reward and aversion encoding in the mesolimbic system. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *33*(45), 17569–76. https://doi.org/10.1523/JNEUROSCI.3250-13.2013

Wallis, J. D., & Miller, E. K. (2003). Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *The European Journal of Neuroscience*, *18*(7), 2069–81. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14622240

Warren, B. L., Mendoza, M. P., Cruz, F. C., Leao, R. M., Caprioli, D., Rubio, F. J., … Hope, B. T. (2016). Distinct Fos-Expressing Neuronal Ensembles in the Ventromedial Prefrontal Cortex Mediate Food Reward and Extinction Memories. *The Journal of Neuroscience*, *36*(25), 6691–6703. https://doi.org/10.1523/JNEUROSCI.0140-16.2016

Wendler, E., Gaspar, J. C. C., Ferreira, T. L., Barbiero, J. K., Andreatini, R., Vital, M. a B. F., … Da Cunha, C. (2013). The roles of the nucleus accumbens core, dorsomedial striatum, and dorsolateral striatum in learning: Performance and extinction of Pavlovian fear-conditioned responses and instrumental avoidance

responses. *Neurobiology of Learning and Memory*, *109*, 27–36. https://doi.org/10.1016/j.nlm.2013.11.009

Wenzel, J. M., Rauscher, N. a., Cheer, J. F., & Oleson, E. B. (2015a). A Role for Phasic Dopamine Release within the Nucleus Accumbens in Encoding Aversion: A Review of the Neurochemical Literature. *ACS Chemical Neuroscience*, *6*(1), 16–26. https://doi.org/10.1021/cn500255p

Wenzel, J. M., Rauscher, N. A., Cheer, J. F., & Oleson, E. B. (2015b). A role for phasic dopamine release within the nucleus accumbens in encoding aversion: A review of the neurochemical literature. *ACS Chemical Neuroscience*, *6*(1), 16–26. https://doi.org/10.1021/cn500255p

Wightman, R. M., Heien, M. L. a V, Wassum, K. M., Sombers, L. a., Aragona, B. J., Khan, A. S., … Carelli, R. M. (2007). Dopamine release is heterogeneous within microenvironments of the rat nucleus accumbens. *European Journal of Neuroscience*, *26*(7), 2046–2054. https://doi.org/10.1111/j.1460-9568.2007.05772.x

Wilber, A. A., Walker, A. G., Southwood, C. J., Farrell, M. R., Lin, G. L., Rebec, G. V, & Wellman, C. L. (2011). Chronic stress alters neural activity in medial prefrontal cortex during retrieval of extinction. *Neuroscience*, *174*, 115–31. https://doi.org/10.1016/j.neuroscience.2010.10.070

Winton-Brown, T. T., Fusar-Poli, P., Ungless, M. a., & Howes, O. D. (2014). Dopaminergic basis of salience dysregulation in psychosis. *Trends in Neurosciences*, *37*(2), 85–94. https://doi.org/10.1016/j.tins.2013.11.003

Wise, R. a. (2004). Dopamine, learning and motivation. *Nature Reviews.*

Neuroscience, *5*(6), 483–94. https://doi.org/10.1038/nrn1406

Yager, L. M., Garcia, A. F., Wunsch, A. M., & Ferguson, S. M. (2015). The ins and outs of the striatum: Role in drug addiction. *Neuroscience*, *301*, 529–541. https://doi.org/10.1016/j.neuroscience.2015.06.033

Yin, H. H., Knowlton, B. J., & Balleine, B. W. (2004). Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *The European Journal of Neuroscience*, *19*(1), 181–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14750976

Yin, H. H., Ostlund, S. B., & Balleine, B. W. (2008). Reward-guided learning beyond dopamine in the nucleus accumbens: the integrative functions of cortico-basal ganglia networks. *European Journal of Neuroscience*, *28*(8), 1437–1448. https://doi.org/10.1111/j.1460-9568.2008.06422.x

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

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

Zeeb, F. D., Baarendse, P. J. J., Vanderschuren, L. J. M. J., & Winstanley, C. A. (2015). Inactivation of the prelimbic or infralimbic cortex impairs decision-making in the rat gambling task. *Psychopharmacology*, *232*(24), 4481–4491.

https://doi.org/10.1007/s00213-015-4075-y

Zeeb, F. D., & Winstanley, C. A. (2013). Functional Disconnection of the Orbitofrontal Cortex and Basolateral Amygdala Impairs Acquisition of a Rat Gambling Task and Disrupts Animals' Ability to Alter Decision-Making Behavior after Reinforcer Devaluation. *Journal of Neuroscience*, *33*(15), 6434–6443. https://doi.org/10.1523/JNEUROSCI.3971-12.2013

Zhu, H., & Roth, B. L. (2014). DREADD: a chemogenetic GPCR signaling platform. *The International Journal of Neuropsychopharmacology*, *18*(1). https://doi.org/10.1093/ijnp/pyu007