Title
Temporally-precise dopamine neuron inhibition disrupts reward seeking behavior

Author
Fischbach-Weiss, Sarah

Publication Date
2013

Peer reviewed|Thesis/dissertation

Temporally Precise Dopamine Inhibition Disrupts Reward Seeking Behavior

by

Sarah Fischbach-Weiss

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Neuroscience

in the

GRADUATE DIVISION

of the

Acknowledgements

This thesis is dedicated to my family. My parents, for always believing I could go wherever I desired (and letting me go too). My sister, for keeping me grounded. My grandparents, for lending me their wisdom. Allllll my aunts, uncles and cousins, for always making me laugh. And of course, to all my marvelous friends and adopted family, thank you for all the distractions and leisure time. And last, but certainly not least, I'd like to thank my wonderful husband, Ivan Weiss, who for some reason decided to fall in love with a first year graduate student and had been paying ever since; your generosity is more than I deserve. If not for their unconditional support, love, and unique flair for simultaneously inspiring both tranquility and borderline lunacy, I would be writing these acknowledgements a much more frazzled, and less whole person.

I am incredibly grateful to my first graduate advisor, Antonello Bonci, for giving me such a great start and always providing sound advice and input, even from the other side of the country. I'd also like to thank my first scientific mentor, Andrea Chiba, who gave me my first research project and first showed me what it was to be a successful woman scientist. Finally, I'd like to thank the members of my committee. My committee chair, Roger Nicoll, who, by always providing the sharpest, and most pertinent criticisms, has helped to improve the work presented here by at least a factor of one thousand. I'd also like to thank Mary Dallman for being a constant champion and inspiring confidence when I needed it most.

I would like to acknowledge members of the Janak lab for their support and contribution to the work presented here. First, T. Michael Gill for your mentorship and

friendship. My fellow graduate students, Elizabeth Steinberg, Alex Loucks, Kate Vitale and Josiah Boivin for insightful discussion and for being the best colleagues I could have ever hoped for. Finally, I'd like to thank Woodward Hopf, Zack Chadick, Ben Saunders and Zayra Millan for their contributions in editing this manuscript.

Last, but certainly not least, I'd like to thank my thesis advisor and mentor, Patricia Janak. I began working with Tricia as a technician and fledgling scientist. Through a series of mishaps and fortunate accidents I am now leaving her lab, 8 years later, with a profound feeling of fulfillment and gratitude; two sentiments that are not always synonymous with graduate school. The work presented here is a direct reflection of her wisdom, mentoring ability, and above all, patience.

Temporally-precise dopamine neuron inhibition disrupts reward-seeking behavior

Sarah C. Fischbach-Weiss

Abstract of the Dissertation

Midbrain dopamine (DA) neurons are phasically activated in response to unexpected reward presentation and are depressed by reward omission; these phasic responses are proposed to contribute to learning. In addition, because DA release in the nucleus accumbens precedes initiation of instrumental actions, and DA receptor blockade in the nucleus accumbens decreases effortful behavioral responding, DA is thought to be necessary to initiate and maintain motivated responding. Hence, midbrain DA signaling may be critically involved in multiple, temporally-specific, behavioral processes. Notably, previous findings were based on long-term perturbations such as dopamine depletion and receptor antagonism, which do not assess the contribution of DA to behavior on short time scales. Therefore, in this study, we used *in vivo* optogenetic techniques to inhibit midbrain DA neurons of mice in real time to investigate how the precise timing of DA neural activity contributes to the decision to exert effort and to maintain cue-reward associations. We found that inhibiting DA neurons while mice were engaged in a bout of operant nose-poke responses decreased the probability of continued nose-poke behavior. Conversely, we found that inhibiting

DA neurons while mice were engaged in off-task behavior decreased the probability of initiating a nose-poke response. Finally, we found that inhibiting DA neurons during reward consumption extinguished learned cue-reward associations, but did not alter effortful responding. We conclude that DA is required at multiple specific phases in behavior to differentially support reward prediction, and both initiation and maintenance of motivation to respond for reward.

Table of contents

Title page i
Copyright page ii
Acknowledgements iii
Abstract v
Table of contents vii
List of figures ix

Introduction
Overview 1
Anatomical and histochemical attributes of VTA DA neurons 1
"Liking" versus "wanting": Historical context and early hypothesis of dopamine function 4
The effort hypothesis of dopamine function 6
Physiological modulation of motivated behavior and dopamine function 7
The reward prediction error hypothesis 8
Reconciling the role of dopamine in learning and motivation 10

Chapter 1
Dopamine neuron activity is required to initiate and maintain reward-seeking behavior
Abstract 15
Introduction 16
Results 19
Discussion 34
Methods 40

Chapter 2
Dopamine neuron inhibition generates an artificial negative reward prediction error signal that drives new learning
Abstract 44
Introduction 45
Results 48
Discussion 64
Methods 68

Chapter 3

Ghrelin signaling in the VTA is required to maintain high-effort responding, but not feeding behavior

Abstract 74
Introduction 75
Results 78
Discussion 90
Methods 92

Chapter 4

Conclusions and remaining questions 96

References 104
Library release statement 123

List of Figures

Figure 1 Major efferent and afferent projections of the VTA 2

Figure 2 Dopamine depletion impairs operant responding for food when response requirements are high 6

Figure 3 Dopamine neurons encode reward prediction errors 9

Figure 4 FR8 task, inhibition of midbrain DA neurons after 100% of nose poke bout initiation trials decreases nose poke bout duration and total nose pokes completed in a 2 hour test session 20

Figure 5 FR8 task, inhibition of midbrain DA neurons after 50% of nose poke bout initiation trials does not decrease the total nose pokes completed in a 2 hour test session 22

Figure 6 FR8, Inhibition of midbrain DA neurons during 100% of off-task behavior trials decreases probability of nose poke initiation 25

Figure 7 PR task, inhibition of midbrain DA neurons in 50% of nose poke bout trials results in a decrease in nose poke bout duration that is dependent on response requirement 27

Figure 8 PR task, inhibition of midbrain DA neurons during 100% of off-task behavior trials decreases probability of nose poke bout initiation 31

Figure 9 PR task, inhibition of midbrain DA neurons in 3, 10 minute blocks does not decrease the total nose pokes completed in a 2 hour test session 33

Figure 10 Experimental design of reward-paired DA neuron inhibition 49

Figure 11 Reward-paired DA neuron inhibition does not decrease motivation 51

Figure 12 Reward-paired DA neuron inhibition decreases reaction to a reward-predictive cue 54

Figure 13 Reward-paired DA neuron inhibition extinguishes continuously reinforced operant responses and cue-reward assocaitions 58

Figure 14 Reward-paired DA neuron inhibition selectively attenuates cue-reward associations 60

Supp Fig 1 Reward-paired DA neuron inhibition does not decrease operant responding or cue-related behavior in YFP mice in a PR task 62

Supp Fig 2 Reward-paired DA neuron inhibition does not decrease operant responding or cue-related behavior in YFP mice in a FR1 task 63

Figure 15 Inhibition of systemic ghrelin signaling is required to maintain feeding behavior, while inhibition of ghrelin signaling in the VTA does not affect feeding behavior 79

Figure 16 Inhibition of systemic ghrelin signaling decreases motivation to obtain high fat and regular food reinforcers 82

Figure 17 Intra-VTA ghrelin increases motivation to obtain high fat and regular food reinforcers 85

Figure 18 Intra-VTA ghrelin receptor antagonists decrease motivation to obtain high fat and regular food reinforcers 87

Figure 19 Intra-VTA ghrelin receptor antagonists do not decrease low-effort responding for high far and regular food reinforcers 89

INTRODUCTION

As we interact with our environment, we constantly learn associations that will help us predict the occurrence of rewarding events such as food, sex and shelter. At the same time, we also express previously learned behaviors, which increase our ability to obtain these rewards. Complex and distributed neural circuits mediate the processes by which the brain learns about cues that predict rewards, forms and stores memories, and utilizes and retrieves these memories to invigorate behavior and increase motivation. However, one neurotransmitter, dopamine (DA), plays a critical role in several of these processes. DA neurons arising from the ventral tegmental area (VTA) have received considerable attention for their role in both motivation and learning of cue-reward associations (Berridge and Robinson, 1998; Berridge, 2012; Fields et al., 2007; Franklin and Paxinos, 2007; Schultz, 2007; Wise, 2006). By taking advantage of new optogenetic technologies that allow us to perturb DA neuron activity at fast time scales previously not possible, the work presented in this thesis contributes to our understanding of how the precise timing of DA neuron activity promotes and maintains motivated behavior and learned associations.

Anatomical and histochemical attributes of VTA DA neurons

A prominent feature of VTA DA neurons are their long and extensive axonal arbors, which results in a broad distribution of DA release within cortical and sub-cortical structures in the forebrain including the nucleus accumbens, prefrontal cortex,

amygdala and hippocampus (Fig.1) (Benjamini and Hoschberg, 1995; Fields et al., 2007; Fuxe et al., 2010; Haber and Fudge, 1997). In addition to broadcasting information to a large area, VTA DA neurons also receive inputs from a diverse set of brain regions carrying a number of different neurotransmitters and modulators. VTA DA neurons receive a reciprocal input from every structure they project to (Bush and Mosteller, 1951; Fields et al., 2007; Hollerman and Schultz, 1998; Salamone et al., 1997; Saunders and Robinson, 2012; Schultz and Dickinson, 2000; Sutton and Barto, 1998). In addition, they receive input from brain structures such as the superior colliculus and the mesopontine parabrachial nucleus that convey sensory information (Coizet et al., 2010; Redgrave and Gurney, 2006), and projections from the hypothalamus that are well

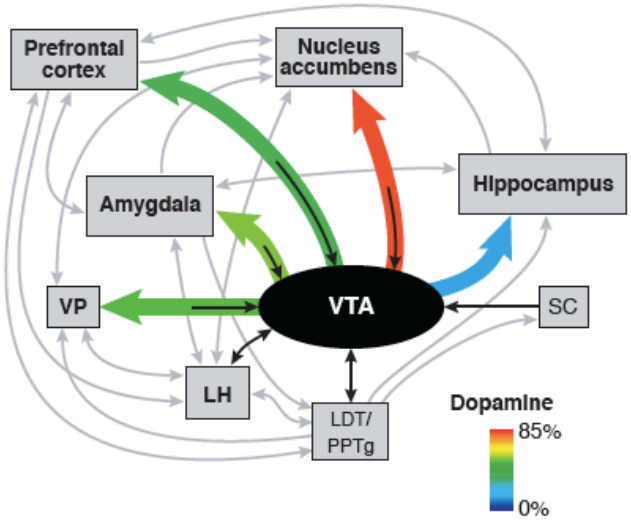


Figure 1 Major efferent and afferent projections of the VTA. Direct connections to and from the VTA are shown in black or color. Other connections are shown in gray. If known, the approximate percentage of projecting neurons that are cytochemically iden tified as dopaminergic is indicated by the color scale shown in the lower right. Abbrevia tions: LDT, laterodorsal tegmental nucleus; PPTg, pedunculopontine tegmental nucleus; LH, lateral hypothalamus; VP, ventral pallidum; SC, superior colliculous. Fields 2007.

suited to informing VTA neurons about the internal state of the animal (Harris et al., 2005; Peyron et al., 1998). Although the majority of inputs to VTA DA neurons are glutamatergic, the VTA also receives GABAergic inputs arising from the ventral pallidum, nucleus accumbens and rostromedial tegmental nucleus (Conrad and Pfaff, 1976; Geisler and Zahm, 2005; Jhou et al., 2009), noradrenergic projections from the locus coeruleus, serotonergic projections from the dorsal raphe nucleus and orexigenic inputs from the lateral hypothalamus (Geisler and Zahm, 2005; Peyron et al., 1998). Studies have also suggested that VTA DA neurons may segregate into distinct circuits based on divergent afferent and efferent connectivity (Fields et al., 2007).

Another striking feature of VTA DA neurons is the diversity that exists within this small group of cells. Although VTA DA neurons are named as such because they all release DA, it is likely that this group of neurons actually represents a diverse set that can be differentiated according to the types of neurotransmitters and peptides they corelease. Recent experiments have demonstrated that a small subset of DA neurons corelease glutamate. Conditional knockout mice that lack the vesicular glutamate transporter 2 selectively in DA neurons are unable to release glutamate from midbrain DA neurons, have decreased tissue dopamine content, and show a greatly reduced locomotor response to cocaine, suggesting that either the ability of glutamate to facilitate DA vesicle filling, or the direct effects of glutamate release from DA neurons are required for normal DA neuron function (Hnasko et al., 2010; Stuber et al., 2010). In addition, subpopulations of DA neurons have been found to corelease the inhibitory neurotransmitter, GABA, as well as several peptides and neurotrophic factors, including

cholecystokinin, neurotensin, neurotrophin3 and brain-derived neurotrophic factor (Hamilton et al., 2000; Seroogy and Gall, 1993; Seroogy et al., 1989; Tritsch et al., 2012). These findings suggest that, by coreleasing a variety of neurotransmitters and peptides, DA neurons may relay distinctive signals at different release sites. Much remains to be identified about the physiological and neuro-chemical properties of each particular subgroup of VTA DA neurons as defined by its projection target. Therefore, in the work presented here, when we refer to manipulations of DA neurons we wish to emphasize that other neurotransmitters and peptides that are coreleased from DA neurons may play a critical role in the behavioral effects we observe.

"Liking" versus "wanting": Historical context and early hypotheses of dopamine function

Early experiments investigating the role of VTA DA neurons in behavior noted that manipulations that either increased or decreased DA signaling in target regions, particularly the nucleus accumbens, had powerful effects on motivation and reinforcement (Berridge and Robinson, 1998; Berridge, 2007). In 1954, Olds & Milner showed that the major rostral efferent pathway of the mesocorticolimbic DA system, the medial forebrain bundle, was a particularly effective site for supporting intracranial electrical self-stimulation (Olds and Milner, 1954). Blockade of DA receptors in DA terminal regions significantly reduced the effectiveness of intracranial self-stimulation (Fouriezos and Wise, 1976; Wise, 2005). Early interpretations of these findings proposed that DA neurons communicated the subjective feeling of reward or "pleasure" to the

rest of the brain, an idea referred to as the hedonia hypothesis (Berridge and Robinson, 1998; Fouriezos and Wise, 1976; Wise et al., 1978). However, subsequent experiments have demonstrated that manipulations of DA function do not influence hedonic reactivity to tastes, nor do they mediate primary food motivation or appetite. 6-OHDA lesions of midbrain DA neurons fail to decrease facial affective expressions that are elicited and modulated by the hedonic impact of natural taste stimuli (Berridge et al., 1989). Conversely, increasing extracellular DA failed to increase hedonic "liking" reactions to sucrose, even as the same manipulation caused mice to work harder to obtain a sucrose reward (Cagniard et al., 2006; Peciña et al., 2003). Additionally, in a T-maze choice paradigm, rats with DA neuron lesions failed to climb over a barrier to obtain a preferred reward but always chose the preferred reward when no barrier was present (Salamone et al., 1994). Finally, the hedonia hypothesis of DA function cannot account for observed increases in DA neuron activity and DA release in terminal regions after exposure to aversive events such as shock or aversive conditioned stimuli. Thus, one current and more refined interpretation of these early experiments is that DA signaling mediates "wanting," or motivation, rather than "liking"/hedonic reactions to rewards (Berridge, 2007; Salamone and Correa, 2012). Currently, two major hypotheses guide the discussion regarding the function of DA neurons in behavior. The effort hypothesis and the reward prediction error hypothesis have each received considerable attention and will be briefly discussed here.

The effort hypothesis of dopamine function

Interference with DA neuron function results in a marked decrease in the amount of effort animals will exert to obtain rewards. Animals with DA neuron lesions will readily consume food rewards and will also perform a low-effort task, such as an operant fixed ratio (FR) 1 task at the same level as sham-lesioned controls (Fig. 2a). However, animals with DA neuron lesions show significant deficits in responding to obtain rewards as the number of lever presses they must perform to earn each reward increases. Rats with DA neuron lesions showed only minor deficits in responding on an FR4 ratio schedule, but were significantly impaired on ratio schedules that required 16

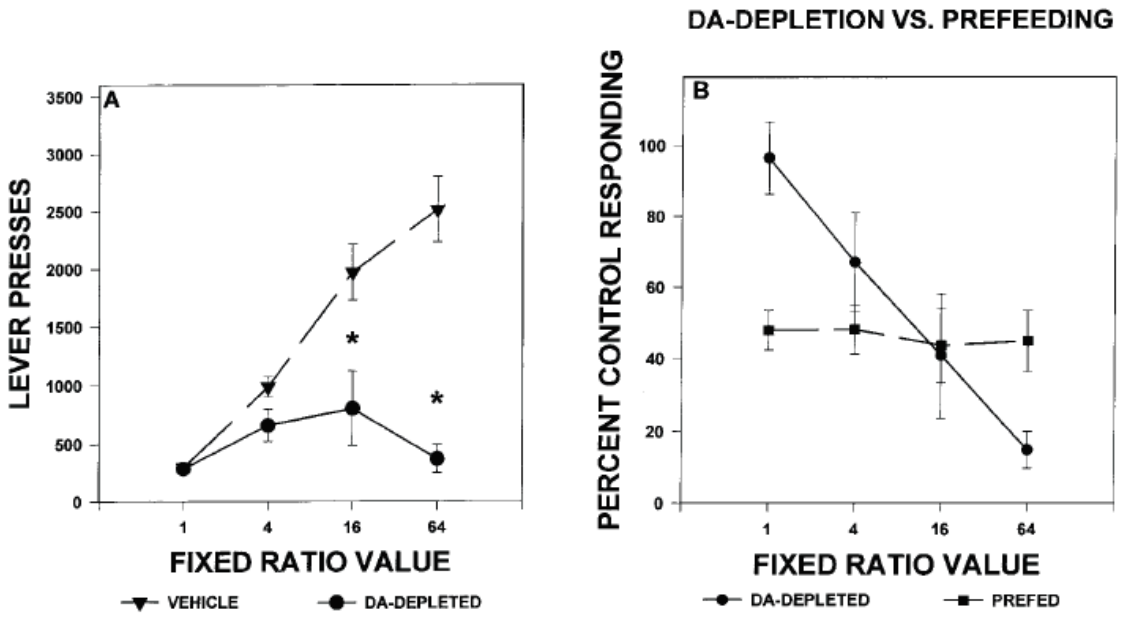


Figure 2 (A) Mean±SEM number of responses per session under each of the four FR schedules, for vehicle controls (dotted line) and DA-depleted reats (solid line). *P<0.05 (B) Comparison of the mean±SEM responses, expressed as a percent of the respective control values, for DA-depleted rats (solid line) and non-lesioned, pre-fed rats (dotted line) .

or 64 levers presses to earn a single reward (Fig. 2a) (Aberman and Salamone, 1999; Salamone et al., 2007). Whereas DA neuron lesions had no effect on FR1 responding, lever pressing for food in this task was significantly decreased after pre-feeding (Fig. 2b), suggesting that manipulations of DA signaling produced specific motivational deficits that were unrelated to appetite or the subjective assessment of reward desirability (Aberman and Salamone, 1999; Salamone et al., 2007). While these results suggest that regulation of effort by midbrain DA neurons is dissociable from the homeostatic state of the animal, there is also substantial evidence that the midbrain DA system can interact with homeostatic signals of hunger and satiety to modulate motivation.

Physiological modulation of motivated behavior and dopamine function

It is well established that food restriction increases motivation to obtain reward. 24-hours of food deprivation increases breakpoint in a progressive ratio task for food reinforcers (Thorpe et al., 2005) and DA release in the nucleus accumbens during sucrose consumption is increased in food-restricted rats compared to unrestricted controls (Rada et al., 2005). Feeding of an unfamiliar food has been shown to stimulate dopamine release in the nucleus accumbens. This release is subject to habituation in non-restricted rats. However, food restriction prevents habituation of dopamine release (Bassareo and Di Chiara, 1999) . In addition, food restriction increases the locomotor response to many drugs of abuse, indicating that food restriction may enhance the sensitivity of the mesolimbic dopamine system (Carr et al., 2001; Carr, 2002).

These experiments demonstrate that dopamine-dependent behaviors can be regulated by the internal state of the animal. However, physiological signals responsible for modulating VTA dopamine neurons and their subsequent behavioral effects are just beginning to be understood. Several physiological modulators of feeding are upregulated during food restriction, including orexins and ghrelin. Orexin receptor antagonists decrease motivation for high-fat reward in food-restricted animals, but had no effect on motivation for regular food. This suggests that, while orexins may modulate responding for high-value rewards, they are not involved in modulating the basic drive to obtain food (Borgland et al., 2009). Ghrelin and orexins have been shown to similarly increase the excitability of VTA dopamine neurons (Abizaid et al., 2006; Borgland et al., 2006). However, it remains unknown whether ghrelin plays a role in modulating motivated behaviors via its actions in the VTA. In the third chapter of this thesis we attempt to define a causal link between ghrelin signaling in the VTA and modulation of motivated behavior that is dissociable from ghrelin's effects in the hypothalamus on appetite and primary food motivation.

The reward prediction error hypothesis

As second prominent hypothesis concerning the role of midbrain DA neuron function proposes that increased DA neuron activity at the time of reward consumption causally contributes to learning about cues that predict rewards. DA neurons are phasically activated at the time of reward consumption when rewards are unpredicted (Fig. 3, top), but show no change in firing to rewards that are fully predicted (Fig. 3,

middle) (Schultz et al., 1997). Importantly, it has also been shown that learning about cues that predict rewards only occurs when the reward is either unexpected or different from predicted, thus demonstrating a correlation between phasic activity of DA neurons and learning (Waelti et al., 2001). The discrepancy between actual reward occurrence and reward prediction is called a "reward prediction error"(Glimcher, 2011; Hollerman and Schultz, 1998). This error is proposed to be generated by DA neuron activity and is thought to be critical for driving learning (Hollerman and Schultz, 1998; Waelti et al., 2001). Recent work from our lab has demonstrated that increasing DA neuron activity at

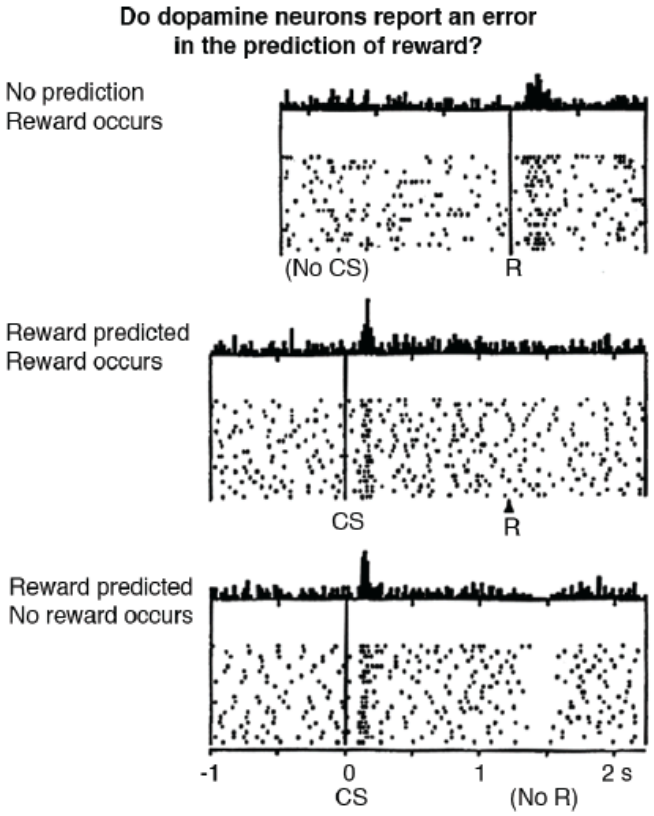


Figure 3 Top, phasic dopamine neuron firing at the time of unexpected reward delivery. Middle, phasic neuron firing at the time of cue presentation of a fully-predicted reward. Bottom, inhibition of dopamine neuron firing at the time of reward when the expected reward is ommited. From Schultz et al., 1997.

the time of reward consumption causally contributes to learning about cues that predict reward (Steinberg et al., 2013). In addition to the phasic activation produced by unexpected reward delivery, DA neurons are also strongly inhibited at the time a predicted reward fails to occur (Fig. 3, bottom) (Schultz et al., 1997). It has been proposed that this pause in DA neuron activity acts a negative reward prediction error that causally contributes to "unlearning" or extinction of learned associations (Glimcher, 2011; Pan et al., 2008). A causal role for inhibition of DA neuron activity at the time of reward expectation in learning has not been demonstrated.

Reconciling the role of dopamine in learning and motivation

Several proposals have been put forth in an attempt to reconcile DA's role in regulating the two different behavioral processes of learning and motivation. We should note that these hypotheses are not necessarily mutually exclusive, and are likely to be complementary in some cases.

One hypothesis that places an overarching role for DA in mediating both effort and learning is the incentive salience hypothesis. This hypothesis proposes that learning about reward-predictive cues can be subdivided into several different processes. In one facet, a cue can convey information about the pleasurable properties of the reward. However, as we have previously discussed, DA neuron activity is not likely to encode this type of information. Alternately, a predictive cue can relate to the specific attributes of the reward, such as the specific taste, shape and feel of a pellet or liquid sucrose reward. A third feature a reward predictive cue can convey is the general "wanting" or

motivational feeling that the anticipation of reward delivery lends to the cue. The incentive salience hypothesis proposes that DA is not required to learn about cue-reward associations, per-se. Rather, DA signaling acts to imbue motivational salience to cues that predict reward. Thus the phasic DA response that occurs during the presentation of a reward-predictive cue signals "wanting" to the animal, but does not carry information about the specific attributes, or the "what," of the reward (Berridge, 2012; Phillips et al., 2007). While this hypothesis is certainly one that merits future study, a direct test of its validity is not within the scope of the work presented in this thesis.

A second hypothesis proposes that the two distinct firing modes of DA neurons, tonic and phasic, may differentially contribute to behavior. In the tonic mode, DA neurons fire in a steady pacemaker-like fashion at a rate of about 3-5Hz (Cagniard et al., 2006; Grace and Bunney, 1984; Grace and Onn, 1989). This firing pattern results in sustained, low-concentration DA release in terminal regions (Grace, 1991; Schultz, 2007). The phasic pattern of DA neuron activity is characterized by short bursts of action potentials that result in much larger, transient increases in DA release in terminal structures (Gonon, 1988; Sombers et al., 2009). It has been proposed that phasic DA release may directly contribute to learning while tonic DA release is required to maintain motivation. For example, inducible knockdown of the dopamine transporter resulted in an increase of the tonic firing rate of DA neurons without affecting phasic firing patterns. These mice showed an increase in motivational drive, measured by the number of nose pokes performed in a progressive ratio task, but acquired a Pavlovian

association at the same rate as control animals (Cagniard et al., 2006). In addition, selectively knocking down NMDA receptors in DA neurons disrupted DA neuron burst firing without affecting tonic firing patterns. These mice took significantly longer to learn in a variety of cue-dependent learning tasks, but they performed similar to controls in a progressive ratio task (Zweifel et al., 2009). Together, these studies suggest that phasic DA neuron activity is not required to maintain motivation, but can promote basic learning of cue-reward associations.

A final proposal, which is examined in part in this thesis, is that the precise timing of DA neuron activity in relation to the current behavioral output of the animal may inform how downstream structures interpret signals received from midbrain DA neurons. For example, phasic DA neuron activity can occur both right before an animal decides to initiate an operate response (Wassum et al., 2012) and at the time an unexpected reward is delivered (Schultz et al., 1997). The pattern of DA neuron firing in these two instances may be very similar and DA may be released downstream to a similar degree. However, in one instance the animal is engaged in preparatory behavior and in the other instance, a consummatory behavior. Therefore, the precise timing of DA neuron activity may interact with other circuits that encode the current behavioral state of the animal to determine how the DA signal might be utilized to guide behavior. In the first and second chapters of this thesis we ask whether DA neuron activity at specific time points in behavior is required to maintain motivation or cue-reward relationships.

In the first chapter of this thesis, we ask whether DA neuron activity is required to maintain motivated behavior. We used optogenetic techniques to selectively inhibit DA neurons at two distinct time points. We inhibited DA neurons either right before an animal decided to initiate a bout of responding, or while a bout of effortful responding was underway. We found that inhibiting DA neurons at either of these time points acutely disrupted task performance. Inhibiting DA neurons while mice were actively engaged in a bout of nose pokes terminated the nose poke bout prematurely while inhibiting DA neurons during off-task behavior decreased the probability that mice would initiate a bout of nose pokes. In addition, inhibiting DA neurons during a bout of nose pokes significantly decreased the total number of nose pokes performed in the behavioral session while DA neuron inhibition during off-task behavior did not have an overall impact on the number of nose pokes completed in the session. These results indicate that DA neuron activity is required to maintain motivation at multiple time points in behavior.

In the second chapter of this thesis, we ask whether DA neuron activity is required to maintain learned associations. We inhibited DA neurons during reward consumption and measured operant responding, motivation and responses to a reward-predictive cue. We found that inhibiting DA neurons specifically during the time of reward consumption generated an artificial negative reward prediction error signal, resulting in selective extinction of continuously reinforced cue- and operant-reward associations but not variably reinforced operant-reward associations. However, reward-paired DA neuron inhibition did not decrease motivation. Based on the studies in this

and previous chapters, we conclude that DA neuron activity plays a critical role in maintaining both motivation and cue-reward relationships and that the precise timing of DA neuron activity influences how these two behaviors are regulated.

Chapter 1

Dopamine neuron activity is required to initiate and maintain reward-seeking behavior.

ABSTRACT

Dopamine's role in maintaining motivated behavior is well established. Disrupting midbrain dopamine (DA) signaling reduces performance on behavioral tasks with high-effort requirements but has no effect on low-effort behavioral tasks. However, previous research is based on long-term perturbations such as DA depletion and receptor antagonism, which cannot assess the contribution of DA in maintaining motivated behavior in real time. In this study, we investigated how the precise timing of DA signaling contributes to the decision to exert effort. We used *in vivo* optogenetic techniques to selectively inhibit midbrain DA neurons of mice in real time. We found that inhibiting DA neurons while mice were engaged in a bout of operant nose poke responses decreased the probability of continued nose poke behavior and also decreased the amount of effort exerted over an entire behavioral session. In addition, we found that inhibiting DA neurons while mice were engaged in off-task behavior decreased the probability of initiating an operant response but had no effect on the amount of effort exerted over the whole behavioral session. We conclude that midbrain DA signaling is required to invigorate behavior and promote initiation of an operant response as well as maintain motivation to continue an ongoing bout of effortful responses.

INTRODUCTION

Midbrain dopamine has long been thought to be critical in maintaining motivated behavior (Ikemoto and Panksepp, 1999; Phillips et al., 2007; Salamone et al., 2007). However, performance of a high-effort behavioral task is not a uniform process. Animals, both in the wild and in controlled laboratory experiments, must often engage in a series of discrete steps to obtain reinforcers such as food or drugs (Bautista et al., 2001; Borgland et al., 2009; Chen et al., 2013; Salamone et al., 1994; Stevens et al., 2005), and experiments attempting to assess the role of DA in maintaining motivated behavior often impose physical barriers or steep response requirements between the experimental subject and the desired reinforcer (Aberman and Salamone, 1999; Salamone et al., 1994; Wanat et al., 2010). How DA contributes to overcoming response costs at specific decision points in behavior such as during approach or while performing an operant response remains the topic of much debate (Berridge and Robinson, 1998; Ikemoto and Panksepp, 1999; Nicola, 2010; Robbins and Everitt, 2007).

Early experiments investigating dopamine's role in reward-related behaviors noted that while control animals would readily perform multiple lever presses to obtain a preferred reinforcer, even while standard lab chow was freely available in the operant chamber, DA-depleted animals would choose to consume the freely available food rather than work to obtain the preferred reinforcer (Cousins and Salamone, 1994; Koch et al., 2000; Nowend et al., 2001). Subsequent experiments demonstrated that DA-depleted animals were able to learn and perform a low-effort operant task, such as a fixed ratio (FR) 1 task, but displayed deficits as response requirements increased,

displaying moderate deficits in FR4 tasks and severe deficits for response requirements above FR16 (Aberman et al., 1998). The effort hypothesis of DA function was developed from these early observations; it states that "the effects of DA antagonism interacts with the kinetic requirements of the instrumental response" such that nucleus accumbens DA is required to enable animals to respond to the work demands imposed by increasing ratio schedules (Salamone et al., 2007).

A similar, but opposing view of dopamine's contribution in maintaining motivated behavior is outlined in the flexible approach hypothesis (Ikemoto and Panksepp, 1999; Nicola, 2010). Proponents of this hypothesis observed that, in addition to high-effort operant tasks, DA depletions decreased performance on some low-effort cue-responding tasks (Di Ciano et al., 2001; Parkinson et al., 2002; Yun et al., 2004a; 2004b). In a series of experiments, the authors demonstrated that DA receptor antagonism decreased the number of rewards earned in an FR1 task with a long, but not short intertrial interval (ITI) (Nicola, 2010). They noted that animals tended to leave the lever and reward port locations when ITIs were long and DA receptor antagonists decreased the likelihood that they would return to the lever operandum after obtaining a reward. They also found that DA receptor antagonists decreased the number of rewards earned in a high-effort FR8 task. However, they observed that once animals began the lever-press sequence, they were able to continue the lever press chain until completion of the response requirement. The authors did observe that while animals with DA receptor antagonism moved to the reward port after the 8^{th} lever press at the same latency as control animals, the latency to reinitiate a lever press after reward

delivery was significantly increased. This increase in the latency to perform a lever press resulted in fewer trials initiated and, subsequently, reduced the number of rewards earned over all. Therefore the authors concluded that nucleus accumbens DA is specifically required during off-task behavior to promote approach and initiate an operant response rather than maintain a bout of responses in progress.

Thus, the major divisive issue between these two hypotheses concerns how the precise timing of DA signaling contributes to the maintenance of effortful behavior. The effort hypothesis posits that DA acting at the time of an operant response is critical in maintaining effortful behavior. On the other hand, the flexible approach hypothesis predicts that DA signaling before an animal engages in operant behavior is critical for initiating an effortful response. However, previous research has relied on long-term perturbations of DA function such as DA depletion and receptor antagonism, which cannot assess the contribution of DA in maintaining motivated behavior in on shot time scales. Advances in *in vivo* optogenetic techniques now allow for selective perturbation of DA neurons as animals engage in motivated behavior on the time scale of milliseconds to minutes (Adamantidis et al., 2011; Tye et al., 2013; Witten et al., 2011).

Here, we inhibited DA neurons at two distinct time points during an operant task. In the first condition, we triggered inhibition just after mice initiated an operant response and were therefore actively engaged in a bout of nose poke responses. In the second condition, we triggered inhibition while mice were engaged in off-task behavior and measured the probability that mice would reinitiate the nose poke response while DA neurons were inhibited. We found that DA neuron inhibition both terminated a bout

of nose poke responses prematurely and decreased the probability that mice would reinitiate a bout of nose pokes. Inhibiting DA neurons after mice had initiated a bout of nose pokes impaired their ability to increase behavioral output in response to increasing response requirements in a progressive ratio (PR) task. Additionally, the total number of nose pokes completed in a 2-hour behavioral session was significantly reduced after inhibiting DA neurons while mice were performing a bout of nose pokes but not when mice were engaged in off-task behavior.

RESULTS

Experiment 1: Inhibiting dopamine neurons after the onset of a nose poke bout decreases bout duration.

To determine whether DA neuron activation is required to maintain an ongoing operant response, we first trained mice on an FR8 nose poke task. We then generated a program that triggered a 532nm laser based on the mouse's behavior in real time. In this condition, the laser pulse was turned on 0.01 seconds after the mouse initiated a nose poke response and remained on for 15 seconds (Fig. 4a). We found that inhibiting DA neurons at the beginning of a bout of nose pokes reduced the number of nose pokes the mouse completed during the laser pulse when compared to a no inhibition control session when the same program was running but the laser was not on ($P=0.0003$, inhibited: 2.48±0.59, no inhibition: 6.59±0.60) (Fig. 4b). No decrease in nose poke behavior during the laser pulse was observed in Th::Cre mice infused with a YFP control virus ($P=0.47$, inhibited: 4.87±0.36, no inhibition: 5.27±0.46, $n=6$). In addition, the

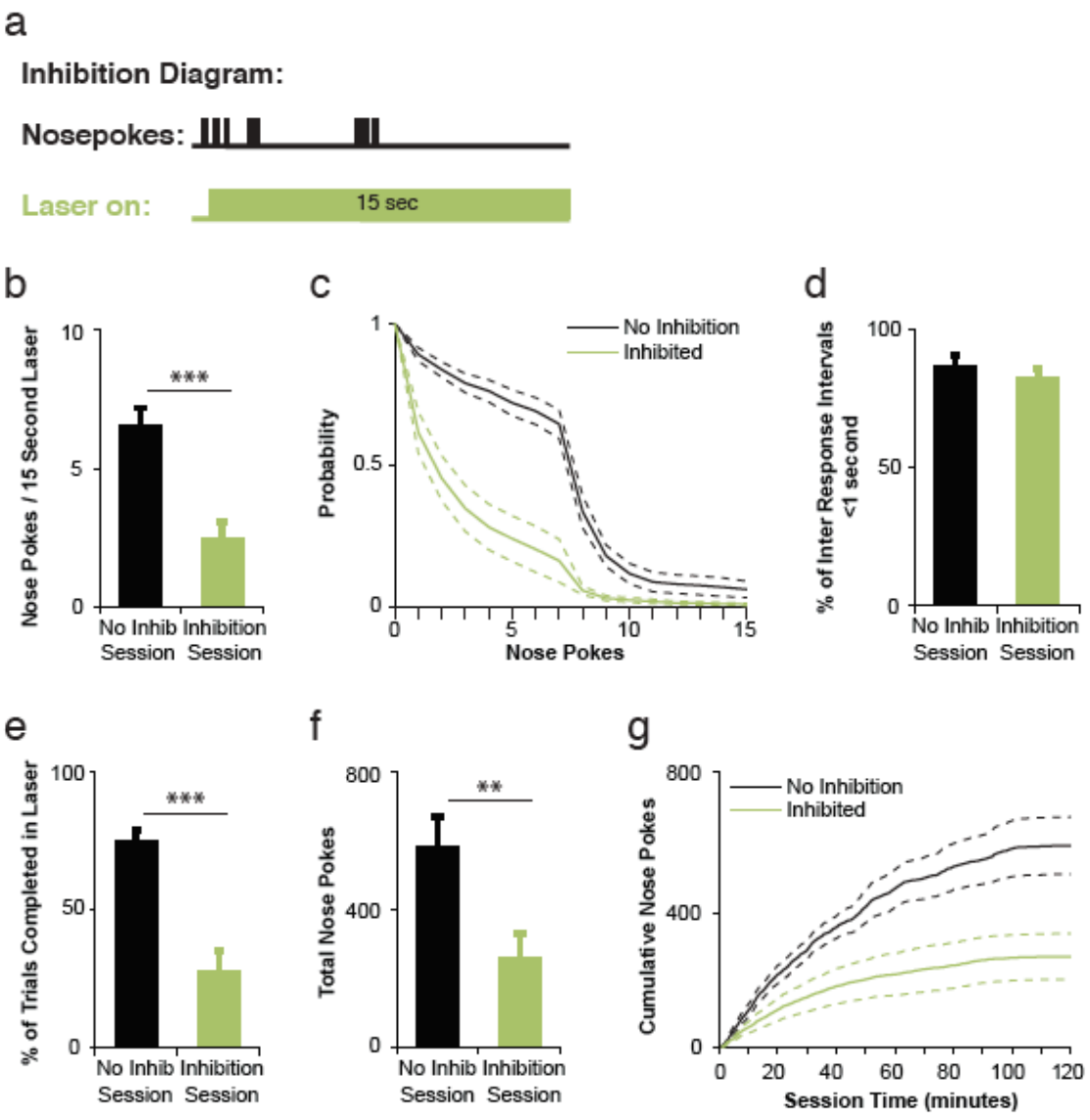


Figure 4 FR8 task, inhibition of midbrain DA neurons after 100% of nose poke bout initiation trials decreases nose poke bout duration and total nose pokes completed in a 2 hour test session. (a) Diagram of parameters for in vivo DA neuron inhibition. A 15 second laser pulse was triggered 0.01 seconds after mice completed a nose poke response in 100% of trials. (b) Mice completed significantly fewer nose pokes in 15 seconds while DA neurons were inhibited compared to a no inhibition session. ***P<0.001. (c) The probability of completing additional nose pokes was decreased in the 15 second laser pulse compared to a no inhibition session. (d) The percentage of short inter response intervals in the 1 second laser pulse was not changed after DA neuron inhibition compared to a no inhibition session. P=0.09. (e) The percen age of trials completed was decreased while DA neurons were inhibited compared to a no inhibition session. ***P<0.001. (f,g) Mice completed significantly fewer nose pokes in the session paired with DA neuron inhibition compared to a no inhibition session. **P<0.01. n=10 mice.

probability of completing additional nosepokes during the laser pulse decreased precipitously when compared to a no inhibition control session (Fig. 4c). To confirm that the observed decrease in nose pokes was not a result of a motor deficit caused by DA neuron inhibition, we measured the interval between nose poke responses while DA neurons were inhibited compared to nose pokes performed in a similar time interval in the no inhibition control session. We found no difference in the percentage of fast inter response intervals (IRI < 1 second) completed during the laser pulse time between the two test conditions (P=0.09, inhibited: 82.42%±3.51%, no inhibition: 86.88%±3.59%) (Fig. 4d). These results agree with a previous report demonstrating that inhibiting DA neurons does not affect locomotor behavior (Tye et al., 2013). Mice also completed significantly fewer FR8 response requirement trials when DA neurons were inhibited than in the no inhibition control session (P<0.0001, inhibited: 27.68%±7.61%, no inhibition: 75.00%±3.96%) (Fig. 4e). Additionally, inhibiting DA neurons after the onset of a nose poke bout resulted in a significant decrease in the number of nose pokes completed in the 2-hour behavioral session (P=0.005, inhibited: 263.4±66.22, no inhibition: 585.10±82.64) (Fig. 4f, g).

Because inhibiting DA neurons in 100% of nose poke bout trials resulted in a marked decrease in overall nose poke behavior, we next aimed to determine if this effect would scale with varying levels of inhibition. In this condition, we inhibited DA neurons after nose poke bout initiation in 50% of trials (Fig. 5a). Once again, we observed a decrease in the number of nose pokes performed while DA neurons were inhibited (P=0.001, inhibited: 4.58±0.63, simulated laser: 6.77±0.51) (Fig. 5b,c) and no

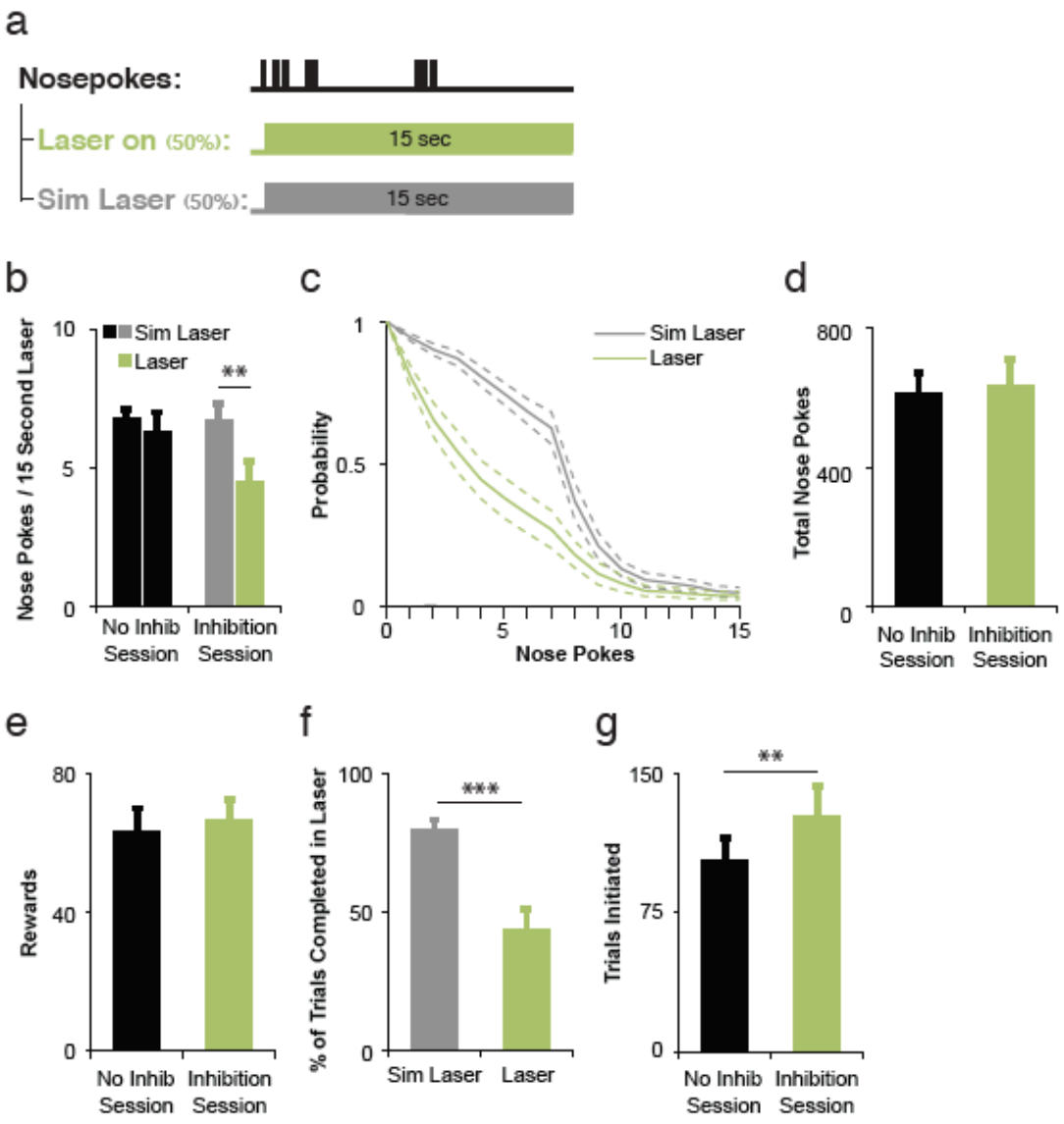


Figure 5 FR8 task, inhibition of midbrain DA neurons after 50% of nose poke bout initiation trials decreases nose poke bout duration but does not decrease the total nose pokes completed in a 2 hour test session. (a) Diagram of parameters for in vivo DA neuron inhibition. A 15 second laser pulse was triggered 0.01 seconds after mice completed a nose poke response in 50% of trials. (b) Mice completed significantly fewer nose pokes in trials paired with DA neuron inhibition compared to no inhibition trials from the same session. **P<0.01. (c) The probability of completing additional nose pokes was decreased in the 15 second laser pulse compared to no inhibition trials from the same session. (d) The total number of nose pokes was not changed after 50% in-bout DA neuron inhibition compared to a no inhibition session. P=0.70. (e) The total number of rewards earned was not changed after 50% in-bout DA neuron inhibition compared to a no inhibition session. P=0.29. (f) The percentage of trials completed was decreased while DA neurons were inhibited compared to no inhibition trials from the same session. ***P<0.001. (g) The number of trials initiated was increased in the 50% in-bout DA neuron inhibition session compared to a no inhibition session. **P<0.01. n=10 mice.

effect in control mice infused with a YFP virus (P=0.15, inhibited: 6.11±0.29, simulated laser: 6.41±0.18, n=6). However, inhibiting DA neurons during 50% of nose poke bouts did not decrease the total number of nose pokes performed (P=0.70, inhibition session: 634.3±78.08, no inhibition session: 615.10±58.18) (Fig. 5d) or the total number of rewards earned (P=0.29, inhibition session: 63.60±6.39, no inhibition session: 66.80±5.57) (Fig. 5e). Because mice completed significantly fewer FR8 trials while DA neurons were inhibited (P=0.0005, inhibited: 43.75%±7.45%, simulated laser: 79.69%±3.63%) (Fig. 5f), we hypothesized that the mice may be compensating by increasing the number of trials initiated throughout the test session. Trial initiation was defined as making a nose poke response after more than 3 seconds had elapsed since the last nose poke response. We chose a 3 second IRI as this was the shortest time period that DA receptor antagonists were shown to impair operant response initiation (Nicola, 2010). We found that mice did indeed initiate more trials in the session in which they received DA neuron inhibition in 50% of nose poke bout trials when compared to a control no inhibition session (P=0.005, inhibition session: 127.4±16.36, no inhibition session: 103.6±12.12) (Fig. 5g). Interestingly, we did not see a similar increase in the number of trials initiated when DA neurons were inhibited after bout initiation in 100% of FR8 trials (P=0.10, inhibition session: 73.20±9.75, no inhibition session: 96.50±12.0).

Experiment 2: Dopamine neuron inhibition during off-task behavior decreases the probability of operant response initiation.

To determine whether DA is required to initiate a nose poke response after a period of off-task behavior, we again tested animals in an FR8 task. Mice were considered to be engaged in off-task behavior if more than 15 seconds had elapsed since collecting a reward pellet and they had not yet reinitiated a nose poke response, at which point a 15-second laser pulse was delivered (Fig. 6a). In this test, DA neurons were inhibited in 100% of eligible trials. We found that mice were significantly less likely to reinitiate nose poking while DA neurons were inhibited when compared to a control session with no inhibition (P=0.003, inhibited: 9.21%±4.49%, no inhibition: 37.37%±6.92%) (Fig. 6b). We did not observe a decrease in the percentage of trials initiated in Th::Cre mice infused with a YFP control virus (P=0.71, inhibition session: 31%±10%, no inhibition session: 32%±10%, n=6). However, we also observed a very fast rebound in the latency to initiate a nose poke response after DA signaling was restored (Fig. 6c). This rebound resulted in no difference in the overall latency to initiate a nose poke response between the inhibited and non-inhibited session (P=0.86, inhibited: 81.59±11.69, no inhibition: 79.63) (Fig. 6d). Furthermore, after inhibiting in 100% of off-task trials, we observed no decrease in the number of nose pokes performed during the 2-hour behavioral session (P=0.70, inhibited: 581.1±108.68, no inhibition 556.10±62.32) (Fig. 6e).

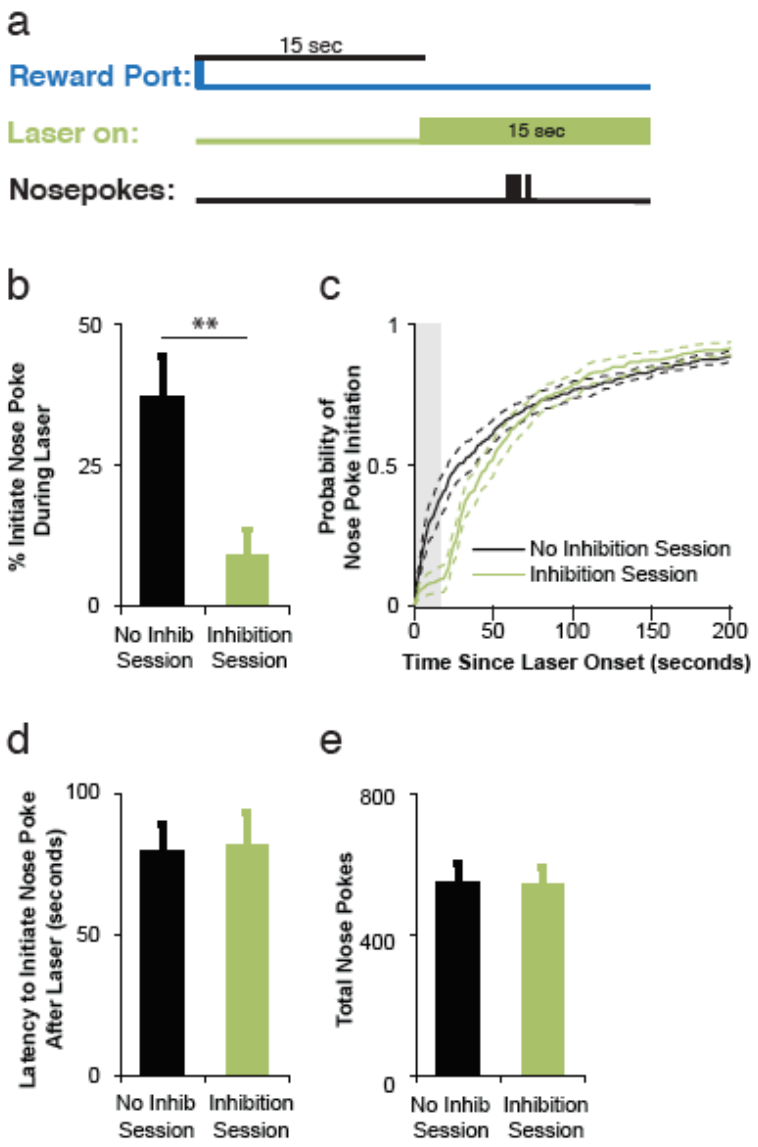


Figure 6 FR8, inhibition of midbrain DA neurons during 100% of off-task behavior trials decreases probability of nose poke initiation but does not decrease the total nose pokes completed in a 2 hour test session. (a) Diagram of parameters for in vivo DA neuron inhibition. A 15 second laser pulse was triggered 15 seconds after mice collected a reward as long as mice had not yet initiated a nose poke response. (b) Mice showed a decrease in the percentage of nose poke bouts initiated when DA neurons were inhibited compared to a no inhibition session. **P<0.01. (c) Mice with DA neuron inhibition were less likely to initiate a nose poke response in the 15 second laser pulse (light grey bar), but initiated a nose poke response quickly after DA neuron activity was restored. (d) The latency to initiate a nose poke response was not changed after DA neuron inhibition compared to a no inhibition session. P=0.86. (e) The total number of nose pokes com pleted was not changed after 100% off-task DA neuron inhibition compared to a no inhibition session. P=0.70. n=10 mice.

Experiment 3: Dopamine neuron inhibition after bout initiation impairs the ability to increase effort to match increasing response requirement demands in a progressive ratio task.

Our results have shown that ongoing DA signaling is required to maintain a bout of operant nose poke responses. However, our previous test assessed the effect of DA neuron inhibition in a behavioral task with a fixed, moderate effort requirement. To expand on this, we tested the effect of DA neuron inhibition after nose poke response initiation in a behavioral task with a varying and increasing effort requirement. Specifically, we aimed to determine whether DA signaling is required to maintain effort in proportion to the current response requirement, or if there is an maximum level of effort an animal will expend, or ceiling effect, when DA neurons are inhibited. We trained mice on a progressive ratio (PR) task, in which the number of nose pokes required to earn a reward is systematically increased after each trial, and which is thought to assess motivation (Hodos, 1961; Richardson and Roberts, 1996). We once again inhibited DA neurons after mice initiated a bout of nose pokes. To be certain that mice were initiating a bout of multiple nose pokes as opposed to performing only one or two nose pokes, in this test we triggered the laser pulse after mice had completed three nose pokes in quick succession (IRI < 0.75 seconds). The laser pulse was triggered in 50% of nose poke bout trials and the duration was 15 seconds (Fig. 7a).

We observed a decrease in the number of nose pokes performed while DA neurons were inhibited when compared to trials with no inhibition (simulated laser) in the same behavioral test session ($P=0.02$, inhibited: 8.93±1.97, simulated laser:

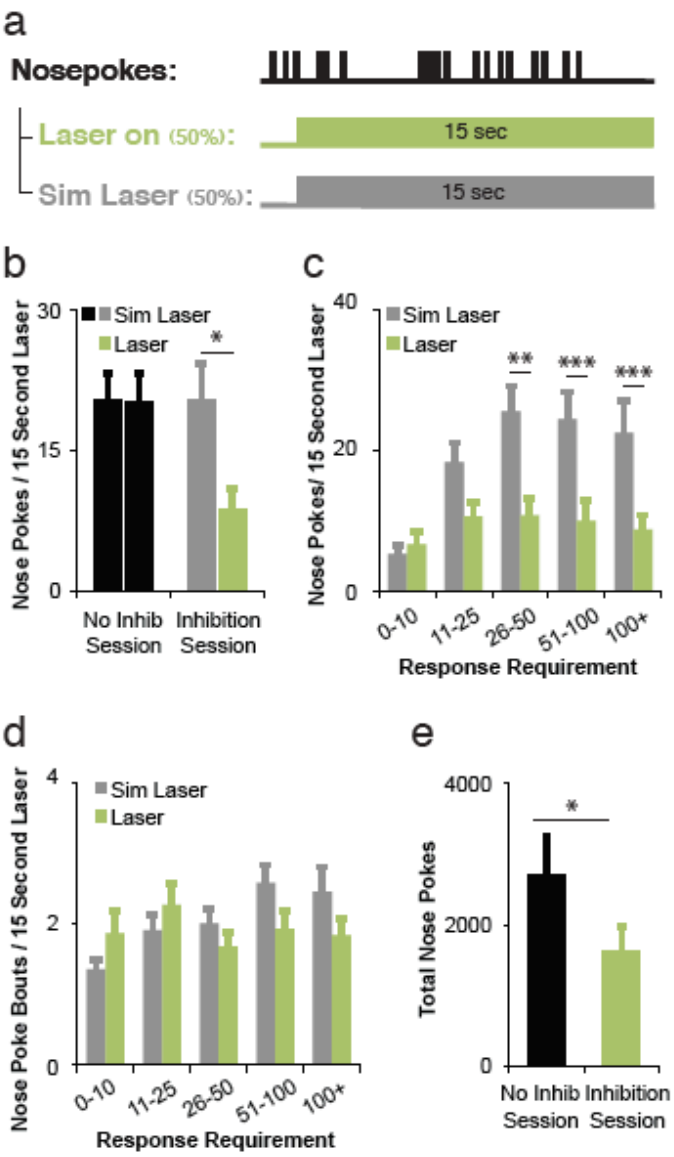


Figure 7 Progressive ratio task, inhibition of midbrain DA neurons in 50% of nose poke bout trials results in a decrease in nose poke bout duration that is dependent on response requirement. (a) Diagram of parameters for in vivo DA neuron inhibition. A 15 second laser pulse was triggered in 50% of trials, 0.01 seconds after mice completed three nose poke responses in succession with an IRI<0.75 seconds. (b) Mice completed significantly fewer nose pokes in trials paired with DA neuron inhibition compared to no inhibition trials from the same session. *P<0.05. (c) Mice with DA neuron inhibition performed fewer nose pokes in the 15 second laser pulse for response requirements (RR) over 26 compared to no inhibition trials from the same session. RR 0-10: P=0.50, RR 11-25: P=0.052, **P<0.01, ***P<0.001. Paired t-test, adjusted for multiple corrections using the Benjamini-Hochberg correction. (d) Mice with DA neuron inhibition did not show a difference in the number of nose poke bouts initiated in the 15 second laser pulse compared to no inhibition trials from the same session. RR 0-10: P=0.16, RR 11-25: P=0.032, RR 26-50: P=0.32, RR 51-100: P=0.13, RR 100+: P=0.16. Paired t-test, adjusted for multiple corrections using the Benjamini-Hochberg correction. (e) Mice completed significantly fewer nose pokes in the session paired with DA neuron inhibition compared to a no inhibition session. *P<0.05. n=10 mice.

20.70±3.51) (Fig. 7b). We did not observe a decrease in nose poke behavior during the laser pulse in control Th::Cre mice infused with a YFP virus (P=0.11, inhibited: 16.62±2.56, simulated laser: 19.25±2.38, n=6). Once again, we saw no difference in the percentage of short IRI's of less than one second between inhibited and non-inhibited conditions, arguing against a motor deficit due to DA neuron inhibition (P=0.16, inhibited: 79.46%±9.07, simulated laser: 91.07%±2.18). We then divided the number of nose pokes performed during the laser pulse according to the current response requirement (Fig. 7c). In trials where DA neurons were not inhibited (simulated laser), mice increased the number of nose pokes performed as the response requirement increased, performing an average of 22.57±4.57 nose pokes in 15 seconds in the most demanding response requirements (greater than 100 nose pokes required to earn each reward). DA neuron inhibition had no effect on the number of nose pokes performed during the laser pulse for response requirements less than 10 (P=0.50, inhibited: 6.78±1.73, simulated laser: 5.39±1.11) (Fig. 7c), and there was a trend toward a decrease for response requirements greater than 11 and less than 50 (P=0.052, inhibited: 10.69±2.03, simulated laser: 18.42±2.72). However, mice performed significantly fewer nose pokes while DA neurons were inhibited when compared to simulated laser trials when response requirements increased above 26 (RR26-50, P=0.01, inhibited: 10.85±2.44, simulated laser: 25.62±3.56; RR51-100, P<0.0001, inhibited: 10.10±2.82, simulated laser: 24.53±2.85). Additionally, for response requirements over 100, mice failed to increase the number of nose pokes performed as response requirements increased, performing an average of 8.82±2.09 nose pokes in

the 15 second laser time compared to 22.57±4.57 in simulated laser trials (P<0.0001) (Fig. 7c). One possibility for the observed failure to increase nose poke behavior as response demands increase is that when DA neurons are inhibited, mice fail to return to the nose poke after a pause in nose poking. To examine this, we measured the number of nose poke bouts performed during each laser pulse across varying response requirements. We defined that a new bout of nose pokes had been initiated if more than one second elapsed since the last nose poke. Mice performed a similar number of nose poke bouts in the 15-second period in both inhibited and non-inhibited trials (RR100+, P=0.16, inhibited: 1.84±0.23, simulated laser: 2.46±0.35) (Fig. 7d). Interestingly, inhibiting DA neurons after 50% of nose poke bout trials resulted in significantly fewer nose pokes performed across the 2-hour test session when compared with a control session with no DA neuron inhibition (P=0.02, inhibition session: 1616.1±360.05, no inhibition session: 2718.5±547.00) (Fig. 7e). To determine if mice attempted to compensate for the deficit of nose pokes caused by DA neuron inhibition, we measured the number of bouts initiated in the inhibition session compared to a no inhibition session. In this test condition, we did not observe an increase in the number of nose poke bouts initiated in the inhibition session (P=0.92, inhibited: 75.5±8.37, no inhibition: 76.2±10.99).

Experiment 4: Dopamine neuron inhibition during off-task behavior decreases the probability of operant response initiation in a progressive ratio task.

We next assessed the requirement for DA neuron activity in response initiation in the PR task. To increase the number of off-task behavior trials, we shifted our parameters for triggering a laser pulse in this test. We categorized off-task behavior as when mice had not made a nose poke response for more than 20 seconds. At that time a 15-second laser pulse was triggered 100% of the time (Fig. 8a). We found that DA neuron inhibition decreased the percentage of nose poke bouts initiated during the laser pulse when compared to a control session where the same program was running but the laser was not turned on (P=0.006, inhibited: 13.47%±4.11, no inhibition: 24.39%±4.36) (Fig. 8b, c). We did not observe a decrease in the percentage of trials initiated in Th::Cre mice infused with a YFP control virus (P=0.42, Inhibition session: 16%±14.6%, No Inhibition: 21%±6.4%, n=6). In addition, we observed a fast rebound in the probability of initiating a nose poke bout as soon as DA signaling was restored (Fig. 8c). As a result, there was no change in the latency to initiate a nose poke after DA neuron inhibition when compared to a control session with no inhibition (P=0.35, inhibited: 131.48±32.19, no inhibition: 129.55±31.83) (Fig. 8d). Interestingly, although the total time of DA neuron inhibition was similar in this test when compared to the PR in-bout inhibition test (p=0.92, PR in-bout: 10.08±1.42 minutes, PR off-task: 10.59±1.08 minutes), we did not observe a decrease in the total number of nose pokes performed when DA neurons were inhibited during off-task behavior (P=0.37, inhibited: 2264.91±557.26, no inhibition: 2479.2±603.89) (Fig. 8e).

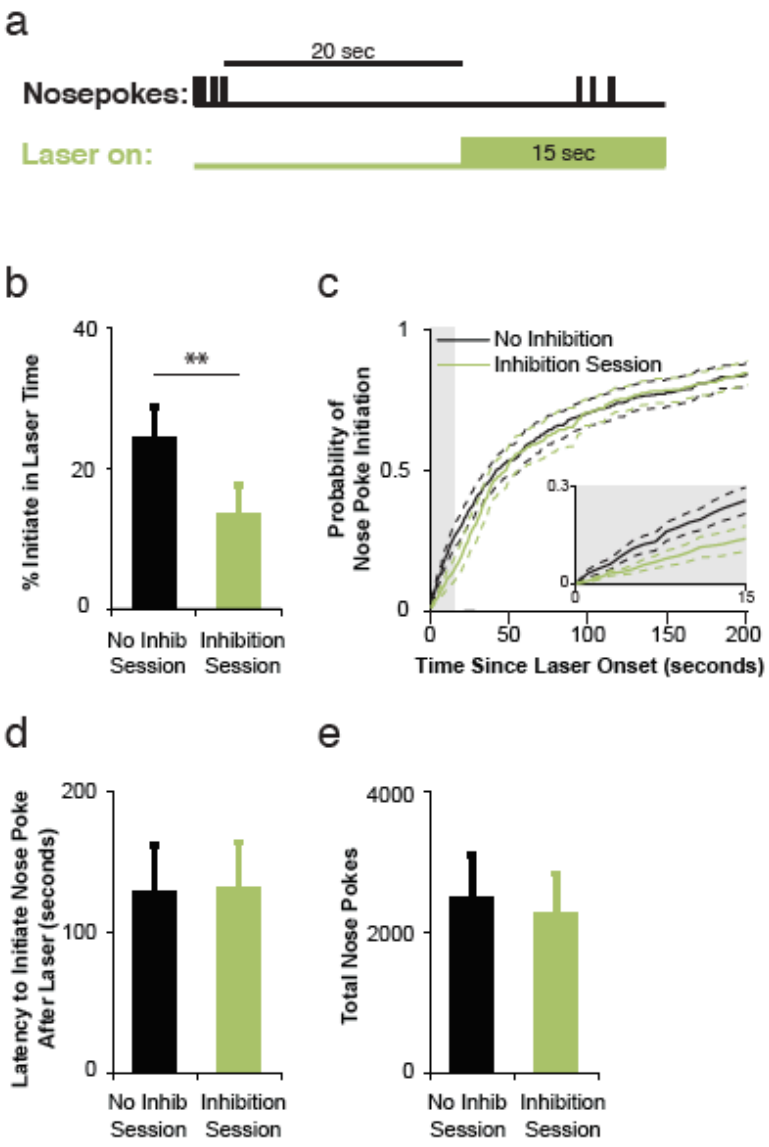


Figure 8 Progressive ratio task, inhibition of midbrain DA neurons during 100% of off-task behavior trials decreases probability of nose poke initiation but does not decrease the total nose pokes completed in a 2 hour test session. (a) Diagram of parameters for in vivo DA neuron inhibition. A 15 second laser pulse was triggered after mice had not performed a nose poke response for 20 seconds. (b) Mice showed a decrease in the percentage of nose poke bouts initiated when DA neurons were inhibited compared to a no inhibition session. **P<0.01. (c) Mice with DA neuron inhibition were less likely to initiate a nose poke response in the 15 second laser pulse (light grey bar and inset), but initiated a nose poke response quickly after DA neuron activity was restored. (d) The latency to initiate a nose poke response was not changed after DA neuron inhibition compared to a no inhibition session. P=0.35. (e) The total number of nose pokes completed was not changed after 100% off-task DA neuron inhibition compared to a no inhibition session. P=0.38. n=11 mice.

Experiment 5: Prolonged DA neuron inhibition does not decrease the number of nose pokes completed in a progressive ratio task.

In several tests presented thus far, we have seen that inhibiting DA neurons during a nose poke bout can significantly impact the total number of nose pokes performed overall in the test session. However, we have not seen a similar effect when inhibiting DA neurons during off-task behavior. One possible reason for this discrepancy could be that delaying response initiation in 15 seconds intervals is not sufficient to have an overall impact on behavior. Therefore, we asked whether inhibiting DA neurons on a much longer time scale would delay response initiation sufficiently to produce an overall effect on nose poke behavior. As mice performed a 2-hour PR task, we triggered a 10-minute laser pulse in 3 blocks starting at 20, 50, and 80 minutes into the test session (Fig. 9a). These inhibition parameters resulted in a total of 30 minutes of DA neuron inhibition, almost double the total inhibition time generated in the FR8 test with inhibition in 100% of nose poke bout trials (FR8 in-bout: 16.56±1.97 minutes). Despite the prolonged inhibition parameters in this test, we did not see a decrease in the total number of nose pokes performed in the 2-hour PR test ($P=0.61$, inhibited: 1528.556±175.17, no inhibition: 1634.11±278.99) (Fig. 9b). The number of nose pokes performed were significantly decreased during each of the 3 10-minute inhibition blocks (blocks 3, 5 & 9) ($P<0.0001$, inhibited: 62.89±35.06, no inhibition: 584.33±77.14) (Fig. 9 c,d). However, we also observed a significant rebound in nose poke behavior during each 10-minute block after DA signaling was restored (blocks 4, 6 &10) ($P=0.01$,

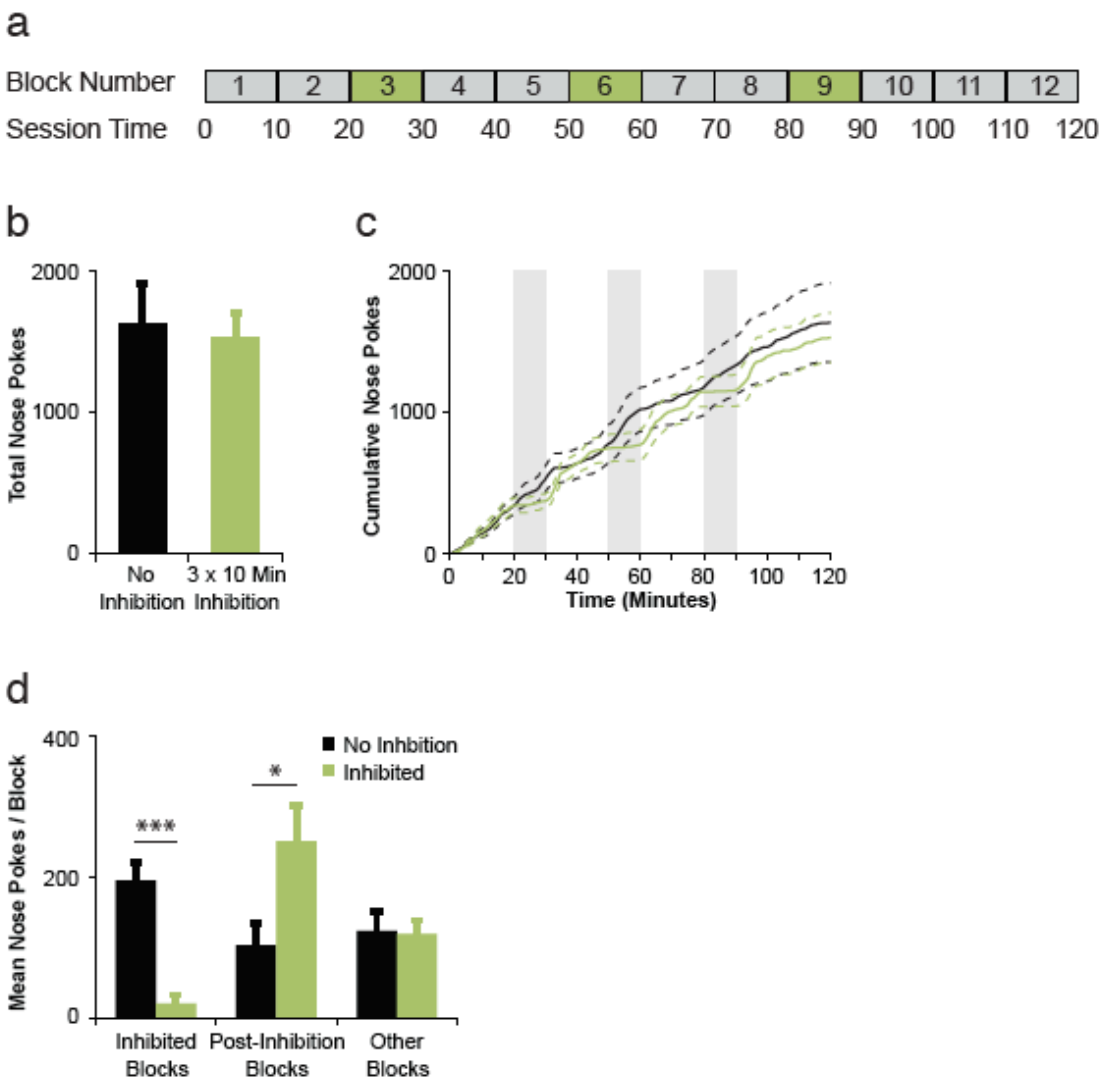


Figure 9 Progressive ratio task, inhibition of midbrain DA neurons in 3, 10 minute blocks does not decrease the total nose pokes completed in a 2 hour test session. (a) Diagram of parameters for in vivo DA neuron inhibition. A 10 minute laser pulse was triggered starting at 20, 50, and 80 minutes into the test session. (b) The total number of nose pokes completed was not changed in mice with DA neuron inhibition compared to a no inhibition session. P=0.61. (c,d) Mice performed significantly fewer nose pokes in the three blocks with DA neuron inhibition (c, light grey bars) and significantly more nose pokes in the three blocks after DA neuron activity was restored when compared to a no inhibition session. *P<0.05, ***P<0.001. n=9 mice.

inhibited: 750.56±151.27, no inhibition: 310.44±92.01) (Fig. 9c,d), resulting in no change in the net number of nose pokes performed between inhibition and control sessions.

DISCUSSION

Here we demonstrate that inhibition of midbrain DA neurons at multiple time points impairs performance of motivated behavior. We first show that DA neuron activity at the time mice are engaged in an operant response promotes the continuation of the operant response through completion of the response requirement. In addition, we show that DA neuron activity at the time mice are engaged in off-task behavior is required to promote reinitiation of an operant response. Finally, because mice readily compensate for nose pokes omitted while DA neurons are inhibited during off-task behavior, but are not able to similarly compensate after DA neurons are inhibited during a bout of nose pokes, we determined that DA neuron activity during an ongoing operant response, but not during off-task behavior is especially critical in maintaining motivation.

In all tests investigating the requirement for DA neuron activity in maintaining a bout of nose poke responses, we found that DA neuron inhibition significantly decreased the number of nose pokes completed during the laser pulse. These results are consistent with many accounts reporting that long-term DA depletions and receptor antagonism progressively impair performance on operant tasks as response costs increase (Aberman et al., 1998; Correa et al., 2002; Mingote et al., 2005; Nicola, 2010; Salamone et al., 2007). However, in the PR test we did not see a deficit in nose poking

behavior as a result of DA neuron inhibition for response requirements less than 10 and only a minor deficit for response requirements less than 25. These results are surprising given the marked decrease in nose poking behavior observed during DA neuron inhibition in the FR8 task. A possible explanation for this discrepancy can be reached by assessing how the response cost is evaluated in each of these tasks. A recent study investigated the effects of DA receptor antagonism on performance of an operant task with varying relative response costs (Ostlund et al., 2012). In this study, rats were trained on an FR1, FR10 or FR20 task. Every training group was then tested on an FR10 task after receiving DA receptor antagonists. Rats that experienced an upshift or no change in relative response cost during the test session (FR1 or FR10 train, FR10 test) showed a substantial deficit in task performance. On the other hand, animals that experienced a downshift in relative response cost (FR20 train, FR10 test) were not impaired in test performance after DA receptor antagonism. Therefore, DA appears to mediate a more subjective and relative evaluation of the current effort requirements of a task rather than the absolute cost of obtaining a reward. With these results in mind, we can infer that mice tested on the FR8 task showed an impairment in nose poke behavior during the laser pulse because the response cost in this task required moderate effort and was held constant. On the other hand, because animals were well trained on the PR task before testing, the subjective response cost for response requirements less than 25 in this test were evaluated relative to the expected much higher response costs later in the session, resulting in an underestimation of moderate response costs and rendering this phase of the task less dependent on DA neuron

activity. In summary, we show that the amount of effort an animal will expend in a bout of nose pokes is dependent on acute DA neuron activity and is relative to the subjective perception of the current response cost.

In contrast with the effort hypothesis of DA function, the flexible approach hypothesis asserts nucleus accumbens DA serves to specifically reinitiate an operant response after a period of off-task behavior (Nicola, 2010). In accordance with these predictions, we found that acute inhibition of midbrain DA neurons while mice were engaged in off-task behavior significantly reduced the probability of reinitiating a nose poke response. These results are consistent with reports that self-initiated action sequences are preceded by DA neuron firing (Jin and Costa, 2010; Romo and Schultz, 1990; Satoh et al., 2003) as well as phasic mesolimbic DA release (Phillips et al., 2003; Stuber et al., 2005a; Wassum et al., 2012). In addition, brief, phasic activation of DA cell body regions have been shown to promote reward-seeking behavior, suggesting a causal role for DA neuron activity in promoting response initiation (Phillips et al., 2003). Our results confirm the importance of DA neuron activity in promoting approach and initiation of an operant response.

In addition to emphasizing a role for NAc DA in promoting approach behavior, the flexible approach hypothesis proposes that DA is not required to maintain effort once an operant response is initiated (Nicola, 2010). This assertion was based on an observation that while performance on an FR8 task was impaired after NAc DA receptor antagonism, this decrease was not due to in an inability to complete the 8 lever press

requirement once a trial had been initiated, but rather due to an increase in time spent away from the lever between trials. In contrast to those observations, we found that inhibiting DA neurons during a bout of nose pokes significantly decreased the ability to complete the FR8 response requirement. A major difference in the methodologies used in these studies could account for the discrepancy in these results. The vast majority of research cited here relies on DA receptor antagonism to eliminate DA signaling in terminal regions such as the nucleus accumbens. While our manipulations selectively inhibit midbrain DA neurons, these neurons have been shown to co-release glutamate as well as DA (Hnasko et al., 2010; 2012; Stuber et al., 2010). Therefore it is possible that some of our behavioral effects may be dependent on decreased glutamate release as well as DA release in terminal regions. To clarify the role of glutamate corelease in modulating behaviors historically considered to be DA-dependent, careful behavioral studies will need to be conducted in mice with selective knockdown of glutamate release from midbrain DA neurons (Hnasko et al., 2010).

In addition to the acute effects on nose poke behavior and response initiation produced by DA neuron inhibition, we also observed an overall effect on motivated behavior in a subset of tests. Inhibiting DA neurons after a bout of nose pokes had been initiated resulted in substantial decreases in overall nose poke behavior despite relatively moderate total inhibition times. In the FR8-100% inhibition task we observed a 54.88%±9.88% reduction in the total number of nose pokes compared to a control session after inhibiting DA neurons for 13.8%±1.64% of the total 120-minute session time (Fig. 4f). In addition, we observed a 32.95%±11.06% reduction in nose poke

behavior in the PR-50% inhibition task after inhibiting DA neurons for 8.40%±1.18% of the total 120-minute session time (Fig. 7e). In contrast to these effects, we found that inhibiting DA neurons in 50% of in-bout trials in an FR8 task, resulting in DA neuron inhibition for 9.79%±1.14% of the 120-minute session, was not sufficient to produce an overall decrease in nose pokes as mice compensated for the nose pokes omitted during DA neuron inhibition by increasing the number of trials initiated throughout the session. This suggests that to a certain extent, mice are able compensate for the acute effects of in-bout DA neuron inhibition.

While we observed pronounced acute effects of inhibiting DA neurons on response initiation, the delays in response initiation did not translate into a reduction in nose poke behavior in any of the tests performed. Our initial test parameters assessed the effect of inhibiting DA neurons in 15-second intervals on response initiation. Because we did not observe an overall effect on nose poke behavior with these inhibition parameters, we hypothesized that delaying response initiation for 15 seconds was not sufficient to produce an overall decrease in nose poke behavior. Therefore, we increased the length of DA neuron inhibition to 3 blocks of 10 minutes each, thus inhibiting DA neurons for a total of 30 minutes, or 25%, of the 120-minute PR task. Although we observed substantial decreases in nose poke behavior while DA neurons were inhibited, even this prolonged inhibition protocol failed to produce an over all decrease in motivated behavior as mice displayed a significant rebound effect in each of the 10 minute blocks after DA neuron activity was restored. These results conflict with a main premise of the flexible approach hypothesis, which asserts that deficits seen in

overall responding in operant tasks with moderate response costs and in low effort tasks with long intertrial intervals are due to an inability to flexibly approach and reinitiate responding and are not due to an inability to perform high rates of operant responses (Nicola, 2010). In contrast to this prediction, we found that inhibiting DA neurons while mice were engaged in a bout of nose poke responses resulted in substantial decreases in overall motivated behavior, while inhibiting DA neurons during off-task behavior was always followed by a quick rebound to compensate for omitted nose pokes and did not result in an overall deficit in motivated behavior in any of the parameters tested. We conclude that, in accordance with the effort hypothesis of DA function, interrupting DA neuron activity at the time of performing a bout of operant responses substantially interferes with the ability of mice to overcome response costs and maintain motivation throughout a behavioral session.

The results presented here confirm that DA neuron activity acts at multiple time points to maintain motivated behavior. DA acting at the time of an ongoing bout of operant responses acts to maintain effortful behavior to complete the current response requirement. Additionally, DA acting at the time of off-task behavior is required to promote flexible behavior and approach to reinitiate a bout of nose poke responses. However, mice display a capacity to compensate for nose poke behavior omitted during off-task DA neuron inhibition but only a limited capacity to compensate for nose poke behavior omitted during in-bout DA neuron inhibition. Therefore, we conclude that DA neuron activity as mice are engaged in an ongoing effortful operant behavior may be the especially critical in expressing the overall motivational state of the animal.

METHODS

Subjects

22 male Th::Cre mice aged 8-12 weeks were housed individually on a reverse 12 h light/dark cycle (lights off at 10:00). All mice began behavioral training with standard rodent chow and water available ad libitum. Mice were tested under food restriction and were brought down to 90% of their free-fed body weight over 5 days before the start of the behavioral test. All experimental procedures were approved by the Institutional Animal Care and Use Committee of the Ernest Gallo Clinic and Research Center at the University of California, San Francisco.

Surgical procedures and virus injection

Mice were anesthetized with 100mg/kg ketamine and 10mg/kg xylazine and placed in a stereotaxic frame (Kopf). Mice received an infusion of either Cre-dependent halo rhodopsin (AAV5/Ef1a-DIO-eNpHR3.0)-eYFP or a control Cre-dependent YFP virus (AAV5/Ef1a-DIO)-eYFP. A custom made 31 gauge infuser was used to deliver 1µl of virus targeting the VTA (target coordinates: ML±0.5mm, AP -3.5mm, DV -4.45mm relative to bregma) at a speed of 0.1µl per minute. The infuser was left in place for an additional 10 minutes following each infusion. Two additional holes were drilled (ML±1.5mm) and two chronic bilateral optical fibers were implanted at an angle of ±13.5° to a targeted depth of DV -4.45mm relative to skull. Optical fibers were made in-house with optical fiber (BFL37-200, Thorlabs) and a ceramic ferrule (MM-FER2007C-2500, Precision Fiber Products) and were secured to the skull surface with two metal screws and dental

cement. As an analgesic, all mice were given ad libitum access to acetaminophen (80mg/100ml water, oral, baby aspirin) for five days following surgery. All mice were given 7 days to recover before starting handling and behavioral procedures. All behavioral tests were conducted at least 5 weeks postsurgery.

Behavioral procedures

Behavioral conditioning: fixed ratio 8 and progressive ratio tests

Mice were trained in standard operant chambers (Med Associates) in ventilated sound-attenuating chambers. They were outfitted with a 2.8W/100mA house light and two nose poke ports flanking a recessed port into which the reward pellet was delivered. The back of one nose poke port was illuminated with a cue light when active and a 500ms, 1kH tone was delivered after each response requirement was completed on the active nose poke, indicating that a reward pellet was available. Mice were first trained on a fixed ratio 1 (FR1) schedule of reinforcement to self-administer a high-fat food (35%, 20mg, Bio-serv) pellet. After acquiring the FR1 task, mice were then progressed to FR3 training for a minimum of 3 days before progressing to an FR8 or PR schedule for testing. Under the PR schedule, the number of nose poke responses required to earn a pellet was increased after each successive reward according to the equation x = (5*e^reward*0.24))-5, which produced the following response requirement schedule: 1,2,4,6,9,12,15,20,25,32,40,50,62,77,95,118… In all schedules, a 500ms, 1kH tone cue signaled the completion of the current ratio requirement and the availability of the reward.

For five days prior to testing and in sessions with optical inhibition of DA neurons, the implanted fiber optic implants were attached to patch cables (MFP_200/240/900-0.22_2.0m_sma, Doric Lenses) that terminated with a ceramic ferrule with a fitted ceramic sleeve (SM-CS125S, Precision Fiber Products). The other end of the patch cable was attached to a bilateral optical commutator (Doric Lenses), which was connected via a second patch cable to a 200mW DPSS 532 nm laser (OEM Laser Systems). The timing of optical inhibition was triggered by a computer running Med PC IV (Med Associates) software, which also recorded nose poke and reward port responses.

Histological verification of NpHR virus expression and fiber optic placement

Mice were deeply anesthetized with pentobarbital and transcardially perfused with phosphate-buffered saline followed by 4% paraformaldehyde. Brains were extracted, post-fixed in paraformaldehyde and cryoprotected in 25% sucrose for 24 hours before slicing. Brains were sectioned coronally on a freezing microtome at a thickness of 50µm. Sections were washed in PBS and incubated with 0.2% bovine serum albumin (BSA) and 0.2% Triton X-100 for 20 minutes. 10% normal donkey serum (NDS) was added for an additional 30 minutes incubation. Sections were then incubated with primary antibodies overnight at 4 degrees Celsius in PBS with BSA and Triton X-100. Concentrations for primary antibodies were as follows: mouse anti-GFP (1:1500, Invitrogen), rabbit anti-TH (1:1500, Fisher Scientific). Sections were then washed with 2% NDS in PBS for 10 minutes before incubating with secondary antibodies for 2 hours at room temperate. Concentrations for secondary antibodies were 1:200 and were conjugated to Alexa Flour

488 or 594 dyes (Invitrogen). Sections were mounted onto glass microscope slides in phosphate buffered water, coverslipped with Vectashield mounting medium (Fisher Scientific), and imaged on a confocal microscope to quantify virus expression and verify fiber placements (Franklin and Paxinos, 2007). Mice with insufficient virus expression or inaccurate fiber placements were discarded from the study.

Statistics

All values are expressed as mean±SEM. Statistical differences within groups were assessed using paired 2-tailed Student's t-tests and statistical differences between groups were assessed using unpaired 2-tailed Student's t tests. A level of confidence of $P < 0.05$ was employed for statistical significance. For data shown in figure 7c,d, p values were adjusted for multiple corrections using the Benjamini–Hochberg correction with a corrected cutoff of P=0.05 (BENJAMINI and HOCHBERG, 1995).

Chapter 2

Dopamine neuron inhibition generates an artificial negative prediction error signal that drives extinction learning

ABSTRACT

Reward prediction errors drive learning by signaling discrepancies between expected and actual reward delivery. When a Pavlovian cue predicts reward delivery but reward is omitted, this generates a negative reward prediction error that drives extinction of conditioned responses. It is hypothesized that inhibitions of dopamine (DA) neuron firing during the omitted reward act as a teaching signal and, over time, causes an animal to cease responding to the previously reinforced cue. However, manipulations using DA receptor antagonists to disrupt and artificially "pause" DA signaling in downstream structures do not produce extinction-like behavior. Thus, a causal role for DA neuron activity in driving extinction learning has not been demonstrated. The development of *in vivo* optogenetic techniques now allows for the selective inhibition of genetically defined DA neurons with unprecedented temporal precision during ongoing behavior. In this study, we hypothesized that inhibiting of DA neurons at the time of reward consumption would generate an artificial negative prediction error signal and drive behavioral extinction of learned associations. We found that inhibiting DA neurons during reward consumption extinguished behavioral responding for continuously reinforced Pavlovian cues and operant responses. In addition, this inhibition did not extinguish variably reinforced operant responses and did not alter effortful responding.

These results suggest that inhibitions of DA neuron activity generate a negative reward prediction error that causally contributes to learning about cue-reward associations.

INTRODUCTION

Associative learning is proposed to be driven by the presence of a signal that informs relevant brain circuits about the discrepancy between the occurrence of unexpected rewards and environmental cues that predict reward delivery (BUSH and MOSTELLER, 1951; Hollerman and Schultz, 1998; Salamone et al., 1997; Saunders and Robinson, 2012; Schultz and Dickinson, 2000; Sutton and Barto, 1998). The phasic activity of midbrain dopamine neurons satisfies many theoretical requirements necessary to act as this reward prediction error signal and guide learning (Berridge and Robinson, 1998; Berridge, 2012; Fields et al., 2007; Franklin and Paxinos, 2007; Glimcher, 2011; Hollerman and Schultz, 1998; Pan et al., 2008; Schultz, 2007; Waelti et al., 2001; Wise, 2006). Midbrain dopamine neurons show strong phasic activation to the presentation of unexpected reward delivery and, in Pavlovian conditioning paradigms, phasic DA neuron activity progressively shifts to the onset of reward-predictive cues (Aberman and Salamone, 1999; Bouton and Moody, 2004; Cohen et al., 2012; Fields et al., 2007; Fuxe et al., 2010; Haber and Fudge, 1997; Myers and Davis, 2002; Pan et al., 2013; Pavlov and Anrep, 1927; Phillips et al., 2007; Salamone and Correa, 2012; Schultz et al., 1997; Waelti et al., 2001). Importantly, recent data from our lab has demonstrated that increasing DA neuron activity at the time of reward delivery causally acts as a positive reward prediction error and drives new learning (Steinberg et al.,

2013). In addition to being phasically activated by reward-predictive cues, DA neurons show time-locked inhibitions to expected reward delivery when the reward is withheld by the experimenter (Fields et al., 2007; Pan et al., 2008; Schultz et al., 1997) or omitted due to behavioral errors (Coizet et al., 2010; Hollerman and Schultz, 1998; Redgrave and Gurney, 2006). This inhibition of DA neuron firing has been proposed to produce a negative prediction error signal that drives extinction learning (Glimcher, 2011; Harris et al., 2005; Hollerman and Schultz, 1998; Pan et al., 2008; Peyron et al., 1998).

Extinction learning is produced by repeatedly presenting conditioned cues in the absence of reward delivery (Bouton and Moody, 2004; Conrad and Pfaff, 1976; Geisler and Zahm, 2005; Jhou et al., 2009; Myers and Davis, 2002; Pavlov and Anrep, 1927). This leads to reduced expression of the conditioned response (i.e. salivation, licking, approach, lever press) and is thought to reflect new, inhibitory learning, rather than forgetting or "unlearning" (Bouton and Moody, 2004; Geisler and Zahm, 2005; Pavlov and Anrep, 1927; Peyron et al., 1998). Consistent with a role for DA neuron inhibitions in mediating extinction learning, both DA neuron activity and phasic DA release in the nucleus accumbens in response to reward-predictive cues diminishes during extinction training (Bush and Mosteller, 1951; Fields et al., 2007; Pan et al., 2008; Schultz and Dickinson, 2000; Stuber et al., 2005b; Sutton and Barto, 1998). This time course closely correlates with the behavioral expression of extinction learning (Glimcher, 2011; Hnasko et al., 2010; Hollerman and Schultz, 1998; Pan et al., 2008; Stuber et al., 2010; Waelti et al., 2001). However, because the baseline firing rate of DA neurons is low (3-5 Hz), it has been argued that brief inhibitions of DA neuron activity cannot sufficiently communicate

a negative prediction error signal to downstream areas (Cohen et al., 2012; Daw et al., 2002; Dayan and Niv, 2008; Glimcher, 2011; Hamilton et al., 2000; Pan et al., 2013; Schultz et al., 1997; Seroogy and Gall, 1993; Seroogy et al., 1989; Tritsch et al., 2012; Waelti et al., 2001). Studies blocking DA signaling in downstream structures such as the nucleus accumbens have reported behavioral effects that mimic extinction learning (Berridge and Robinson, 1998; Berridge, 2007; Chang et al., 1994; Fouriezos and Wise, 1976; Holtzman-Assif et al., 2010; Pan et al., 2008; Schultz et al., 1997). However, others interpret the decrease in behavioral responding after DA receptor antagonism as reflecting a decrease in motivational drive rather than new learning (Hollerman and Schultz, 1998; Olds and Milner, 1954; Salamone et al., 1997; Saunders and Robinson, 2012). Therefore, direct evidence for DA's contribution in communicating negative reward prediction error signals has not been demonstrated.

In this study, we attempt to determine whether midbrain DA neuron inhibition can causally contribute to learning. Therefore, we proposed that inhibiting DA neurons at the time of reward consumption would generate an artificial negative prediction error signal that would drive extinction learning of cue-reward and operant-reward associations. We found that repeatedly pairing DA neuron inhibition with reward consumption across days decreases measures of learned conditioned responses to Pavlovian cues and, in some instances, measures of operant reward associations.

RESULTS

Male transgenic mice expressing Cre recombinase under the control of the tyrosine hydroxylase promoter (Th-cre+ mice) were used to gain control of dopamine neuron activity. Th-cre+ mice were injected with a Cre-dependent virus expressing either halorhodopsin 3.0 (NpHR) or YFP (Fig. 10a). Mice were trained on a progressive ratio (PR) procedure to take advantage of several task features. First, because the number of responses required to receive a reward pellet increases after each successive reward, this task is well suited to gauge motivation (Fig. 10b). Second, because the task has both an instrumental component (a nose poke response) and a Pavlovian component (an auditory cue signaling reward availability), independent measures of each of these learned associations could be assessed. Behavioral testing began once Th-cre+ mice expressing NpHR had reached a stable baseline on the PR task. For mice receiving paired inhibition, a 15 second laser-pulse was triggered 0.01 seconds after entering the reward port to collect a pellet reward. The pellet was dropped into the receptacle 0.05 seconds after mice entered the port (Fig. 10c, paired inhibition). For mice receiving unpaired inhibition, a 15 second laser-pulse was triggered 15 seconds after entering the reward port to collect the pellet reward (Fig. 10c, unpaired inhibition). Thus, mice in the unpaired inhibition group had finished consuming the reward pellet *before* DA neuron inhibition was triggered. All mice were tested over 14 days (Fig. 10d). Baseline behavior was measured with mice attached to the fiber optic cables for two control sessions. Mice were then moved on to 6 days of testing with DA neuron inhibition. DA neurons were inhibited for 15 seconds either with every reward collected

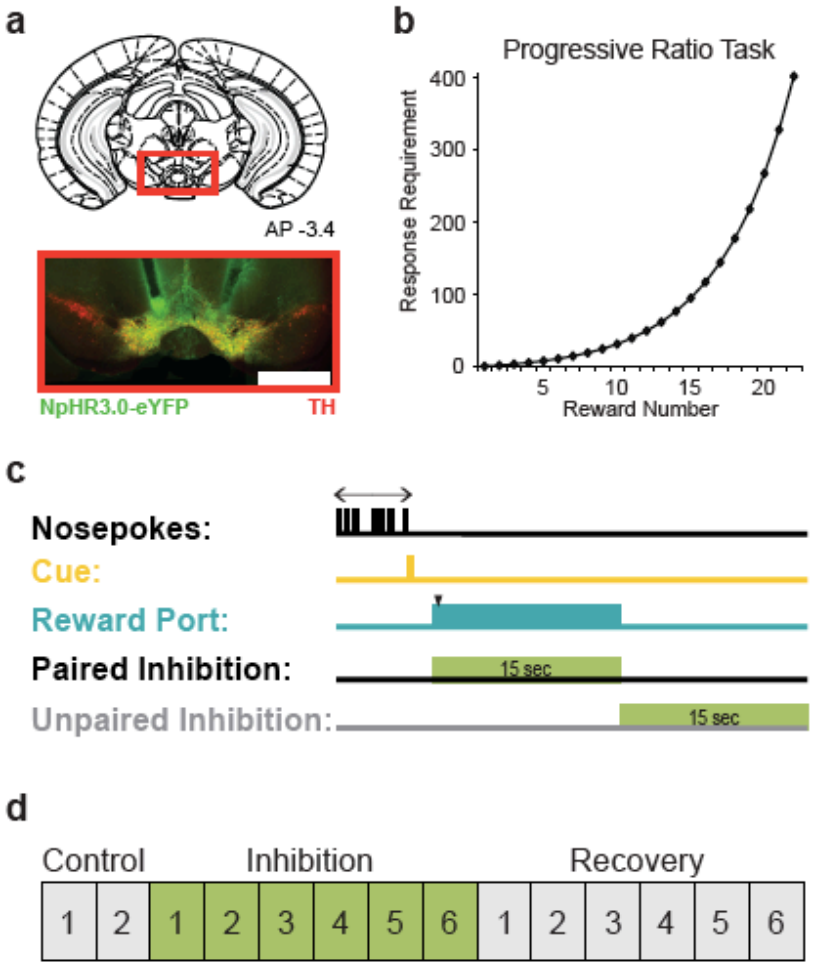


Figure 10 Experimental design. (a) Example histology from a TH-cre+ mouse injected with a Cre-dependent NpHR-containing virus. Vertical tracks indicate fiber optic placements about the VTA. Scale bar represents 1mm. (b) The progressive ratio task. The number of nose pokes to earn each reward increases on each trial. (c) DA neuron inhibition parameters. For the paired group, the laser is turned on 0.01 seconds after entering the reward port. The reward pellet (black triangle) is dropped 0.05 seconds after entering the reward port. For the unpaired group, the laser is turned on 15 seconds after entering the reward port. (d) Baseline behavior is measured for two control sessions, followed by DA neuron inhibition test sessions for 6 days and then 6 days of recovery behavior without DA neuron inhibition.

(paired) or 15 seconds after every reward collected (unpaired). Mice were then tested for an additional 6 recovery days without DA neuron inhibition to assess whether changes in behavior caused by DA neuron inhibition returned to baseline.

Reward-paired DA neuron inhibition does not decrease motivation

Manipulations that disrupt DA signaling in terminal regions have profound effects on motivational drive (Aberman and Salamone, 1999; Bouton and Moody, 2004; Fouriezos and Wise, 1976; Myers and Davis, 2002; Pavlov and Anrep, 1927; Phillips et al., 2007; Salamone and Correa, 2012; Wise, 2005). Therefore, we first assessed whether pairing reward consumption with DA neuron inhibition across 6 days of testing affected motivational drive. We found that DA neuron inhibition did not significantly decrease the number of nose pokes completed across 6 inhibition sessions compared to control sessions in either the paired or unpaired inhibition groups (Kruskal-Wallis test, paired: p=0.245, unpaired: p0.98) (Fig. 11a).

However, we did observe a significant decrease in both the breakpoint (last response requirement completed in the 2-hour session) (Kruskal-Wallis test, control vs. inhibition days, paired: main effect of day p=0.025, unpaired: p=0.98) and the number of rewards earned across the 6 inhibition sessions compared to control sessions (Kruskal-Wallis test, control vs. inhibition days, paired: main effect of day p=0.025, unpaired: p=0.98) (Fig. 11b,c). In addition, mice in the reward-paired group consumed the same percentage of rewards in each session across control and inhibition days (Kruskal-Wallis test, paired: p=0.429) (Fig. 11d). This is important as recent studies have found that

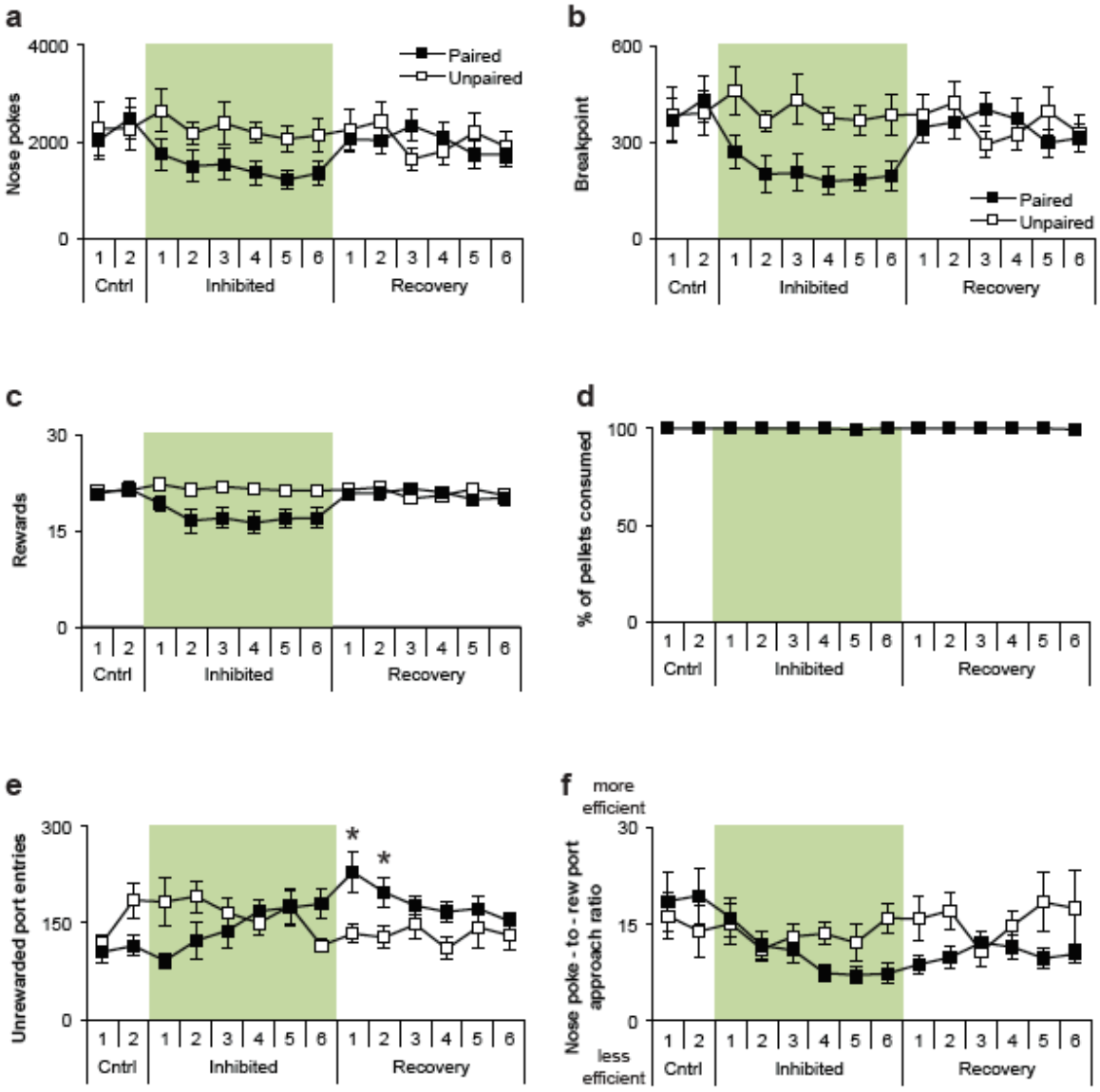


Figure 11 Reward-paired DA neuron inhibition does not decrease motivation. (a) Total nose pokes completed in the 2-hour PR task across 14 days of testing (control vs. inhibition days, paired: p=0.245, unpaired: p0.98). (b) Breakpoint across 14 days of testing (control vs. inhibition days, paired: main effect of day p=0.025, unpaired: p=0.98). (c) Rewards earned across 14 days of testing (control vs. inhibition days, paired: main effect of day p=0.025, unpaired: p=0.98). (d) % of reward pellets consumed (control vs. inhibition days, paired: p=0.429). (e) Number of entries into the reward port when no reward was available (control vs. inhibition days, paired: main effect of day p=0.033, unpaired: p=0.97; inhibition vs. recovery days, paired: main effect of day p=0.007; *p<0.05). (f) Task efficiency measured by the ratio of nose pokes to all reward port entries (control vs. inhibition days, paired: main effect of day p=0.003, unpaired: p=0.82; inhibition vs. recovery days, paired: p=0.08). Paired n=9, unpaired n=8. Kruskal-Wallis test. Data are presented as means and error bars represent s.e.m.

stimulation of VTA GABA neurons inhibits DA neurons and disrupts reward consumption (Berridge and Robinson, 1998; Fouriezos and Wise, 1976; van Zessen et al., 2012b; Wise et al., 1978). A decrease in breakpoint is typically interpreted as being consistent with a decrease in motivation (Berridge et al., 1989; Bouton and Moody, 2004; HODOS, 1961; Pavlov and Anrep, 1927; Richardson and Roberts, 1996). However, we observed that mice receiving reward-paired DA neuron inhibition performed significantly more unrewarded port entries across inhibition days (Kruskal-Wallis test, control vs. inhibition days, paired: main effect of day $p=0.033$, unpaired: $p=0.97$), an effect that persisted across the first two recovery days (Kruskal-Wallis test, inhibition vs. recovery days, paired: main effect of day $p=0.007$) (Fig. 11e). Because mice with reward-paired inhibition checked the reward port more frequently when no reward was available, we hypothesized that the decrease in rewards earned was due to a decrease in task-efficiency rather than due to a decrease in motivation. We assessed task efficiency by measuring the average number of nose pokes performed before checking the reward port (Cagniard et al., 2006; Ostlund et al., 2012; Peciña et al., 2003). We found that mice receiving unpaired-reward DA neuron inhibition maintained consistent efficiency throughout the inhibition test days (Kruskal-Wallis test, control vs. inhibition days, unpaired: $p=0.82$) (Fig. 11f). On the other hand, mice receiving reward-paired DA neuron inhibition showed a significant decrease in task efficiency on days with inhibition when compared to control days (Kruskal-Wallis test, paired: main effect of day $p=0.003$). In addition, task efficiency did not recover to baseline levels across all 6 recovery test days (Kruskal-Wallis test, inhibition vs. recovery days, paired: $p=0.08$). These data

indicate that reward-paired DA neuron inhibition produces lasting effects in task behavior that persist after DA neuron activity has been restored. Additionally, the increase in unrewarded port entries across inhibition days suggests that mice do not find DA neuron inhibition aversive, as they do not avoid the location paired with DA neuron inhibition. We did not observe a change in any behavioral measures in either the NpHR-unpaired (Fig. 11) or YFP-paired inhibition groups (Supplementary Fig. 1).

Reward-paired DA neuron inhibition extinguishes learned cue-reward associations

Inhibitions of DA neuron activity that occur during expected reward omission are proposed to act as a negative reward prediction error signal that drives extinction learning (Glimcher, 2011; Pan et al., 2008; 2013; Salamone et al., 1994). We hypothesized that we could produce the equivalent of an artificial negative reward prediction error signal by holding reward delivery constant while driving DA neuron inhibition. We did not observe a decrease in nose poke behavior (Fig. 11a), indicating that reward-paired DA neuron inhibition does not drive extinction of an operant response in a PR task. Therefore, we next assessed the effects of reward-paired DA neuron inhibition on measures of Pavlovian behavior. In our PR task, an auditory cue signals the completion of the current response requirement and availability of the reward pellet (Fig. 10c). In control sessions, the latency to collect the reward after cue presentation is very fast; it is longer than 5 seconds in just 2%±1% of trials (Fig. 12a, c). However, the percentage of cue-reward latencies greater than 5 seconds is significantly increased across reward-paired inhibition sessions (Kruskal-Wallis test, control vs.

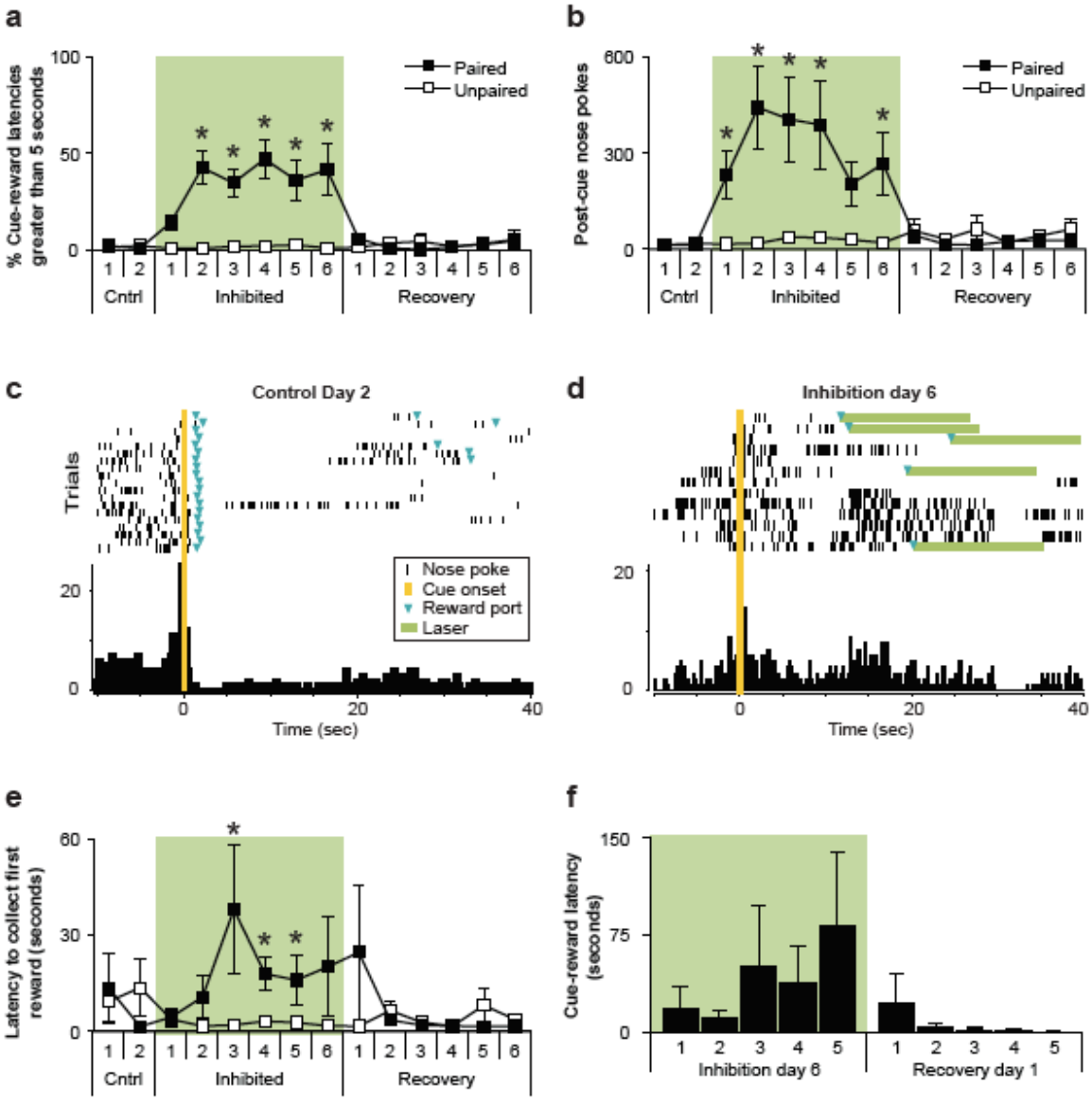


Figure 12 Reward-paired DA neuron inhibition decreases reactions to a reward-predictive cue. (a) Percentage of latencies to collect the reward after cue presentation that are greater than 5 seconds (control vs. inhibition days, paired: main effect of day p=5.26e-07, unpaired: p=0.91; inhibition vs. recovery days, paired: main effect of day p=1.77e-11; *p<0.05). (b) Number of nose pokes performed after cue presentation and before reward collection (control vs. inhibition days, paired: main effect of day p=1.68e-06, unpaired: p=0.80; inhibition vs. recovery days, paired: main effect of day p=3.53e-10; *p<0.05). (c) Behavior during the second control session from a representative reward-paired inhibition Th-cre+ mouse. (d) Behavior during the sixth inhibition session from a representative reward-paired inhibition Th-cre+ mouse. (e) Latency to collect the reward after cue presentation from the first trial of each session (control vs. inhibition days, paired: main effect of day p=1.5e-04, unpaired: p=0.91; inhibition vs. recovery days, main effect of day p=3.06e-07; *p<0.05). (f) Latency to collect the reward after cue presentation from the first 5 trials in the sixth inhibition session (p=0.6752) and the first recovery session (main effect of trial, p=0.018). Paired n=8, unpaired n=8. Kruskal-Wallis test. Data are presented as means and error bars represent s.e.m.

inhibition days, paired: main effect of day p=5.26e-07, unpaired: p=0.91), indicating that mice no longer appropriately respond to the cue indicating reward availability (Fig. 12a,d). To assess behavior in the period between cue presentation and reward port entry we measured the number of nose pokes mice performed after each cue presentation before moving to the reward port to collect the reward. We found that mice in the reward-paired inhibition group performed significantly more post-cue nose pokes when compared to control sessions (Kruskal-Wallis test, control vs. inhibition days, paired: main effect of day p=1.68e-06, unpaired: p=0.80) (Fig. 12b, d). Because mice failed to terminate the nose poke bout when the cue was presented, this suggests that mice no longer perceive that the cue indicates reward availability.

We next sought to determine how reward-paired DA neuron inhibition affected cue-related behavior before any inhibition had occurred during each session. We measured the latency to collect the reward after cue presentation on the first trial of each session and found that the latency was increased compared to control sessions after the third inhibition session (Kruskal-Wallis test, control vs. inhibition days, paired: main effect of day p=1.5e-04, unpaired: p=0.91) (Fig. 12e). Additionally, we found that this measure did not return to baseline levels until the second recovery session (Kruskal-Wallis test, inhibition vs. recovery days, main effect of day p=3.06e-07), suggesting that reward-paired DA neuron inhibition produces deficits in Pavlovian cue-related behavior that persist across test sessions. To further assess the time course of recovery of cue-related behavior we focused on the latencies to collect the first 5 rewards in the sixth inhibition and first control session. In the sixth inhibition session, the cue-to-reward

latency was greater than five seconds on all five trials. However, in the first recovery session mice took longer than five seconds to retrieve the reward pellet in the first two trials, but behavior quickly returned to baseline as animals experienced the reward without DA neuron inhibition (Kruskal-Wallis test, recover day 1, rewards 1-5, paired: main effect of reward, $p=0.018$). The latency to collect the reward was less than five seconds in the third through fifth trials and the cue-to-reward latency on the fifth trial was significantly shorter than on the first trial (Fig. 12f). We did not observe a change in any behavioral measures in either the NpHR-unpaired (Fig. 12) or YFP-paired inhibition groups (Supplementary Fig. 1).

Reward-paired DA neuron inhibition extinguishes continuously reinforced operant responses.

Because reward-paired inhibition in the PR task resulted in an extinction of the cue-reward relationship but not the operant response, we next aimed to elucidate the reason for this discrepancy. We hypothesized two possible scenarios. First, it is possible that cue-reward relationships and operant-reward relationships are maintained by different circuits, with the former being maintained by a VTA-dependent circuit and the latter being maintained by a VTA-independent circuit. The second possibility is that the difference in extinction susceptibility was due to differences in reinforcement schedules. In the PR task, multiple nose pokes are required to earn a reward and the animal does not explicitly know the number of nose pokes required to earn each reward. Therefore, the operant response in the PR paradigm is variably reinforced. On the other hand,

because the Pavlovian cue signals availability of the reward pellet on 100% of trials, it is continuously reinforced. Previous experiments have demonstrated that variably reinforced cues are more resistant to extinction than continuously reinforced cues (Berridge, 2007; Haselgrove et al., 2004; Salamone and Correa, 2012).

Therefore, in our next experiment we trained mice on an FR1 schedule. In this task both the operant response and the Pavlovian cue are reinforced 100% of the time. Over 6 days of reward-paired DA neuron inhibition on this schedule we observed a significant decrease in the number of nose pokes performed (Kruskal-Wallis test, control vs. inhibition days, paired: main effect of day p=0.039, unpaired: 0.997) (Fig. 13a), and the number of rewards earned (Kruskal-Wallis test, control vs. inhibition days, paired: main effect of day p=9.36e-05, unpaired: p=0.89) (Fig. 13b) when compared to control sessions. In addition, we also observed an increase in the percentage of cue-reward latencies greater than 5 seconds in the reward-paired inhibition group (Fig. 13d,e) when compared to control sessions (Kruskal-Wallis test, control vs. inhibition days, paired: main effect of day p=3.12e-05, unpaired: p=0.35) (Fig. 13c,e). However, we did not see an increase in the number of post-cue nose pokes performed across reward-paired inhibition days (Kruskal-Wallis test, control vs. inhibition days, paired: p=0.624, unpaired: 0.99) (Fig. 13f), suggesting that animals disengaged from the nose poke port, but did not immediately approach the reward port after the cue. These results suggest that reward-paired DA neuron inhibition extinguishes both cue-reward associations and operant responses in a continuously reinforced task. We did not observe a change in any

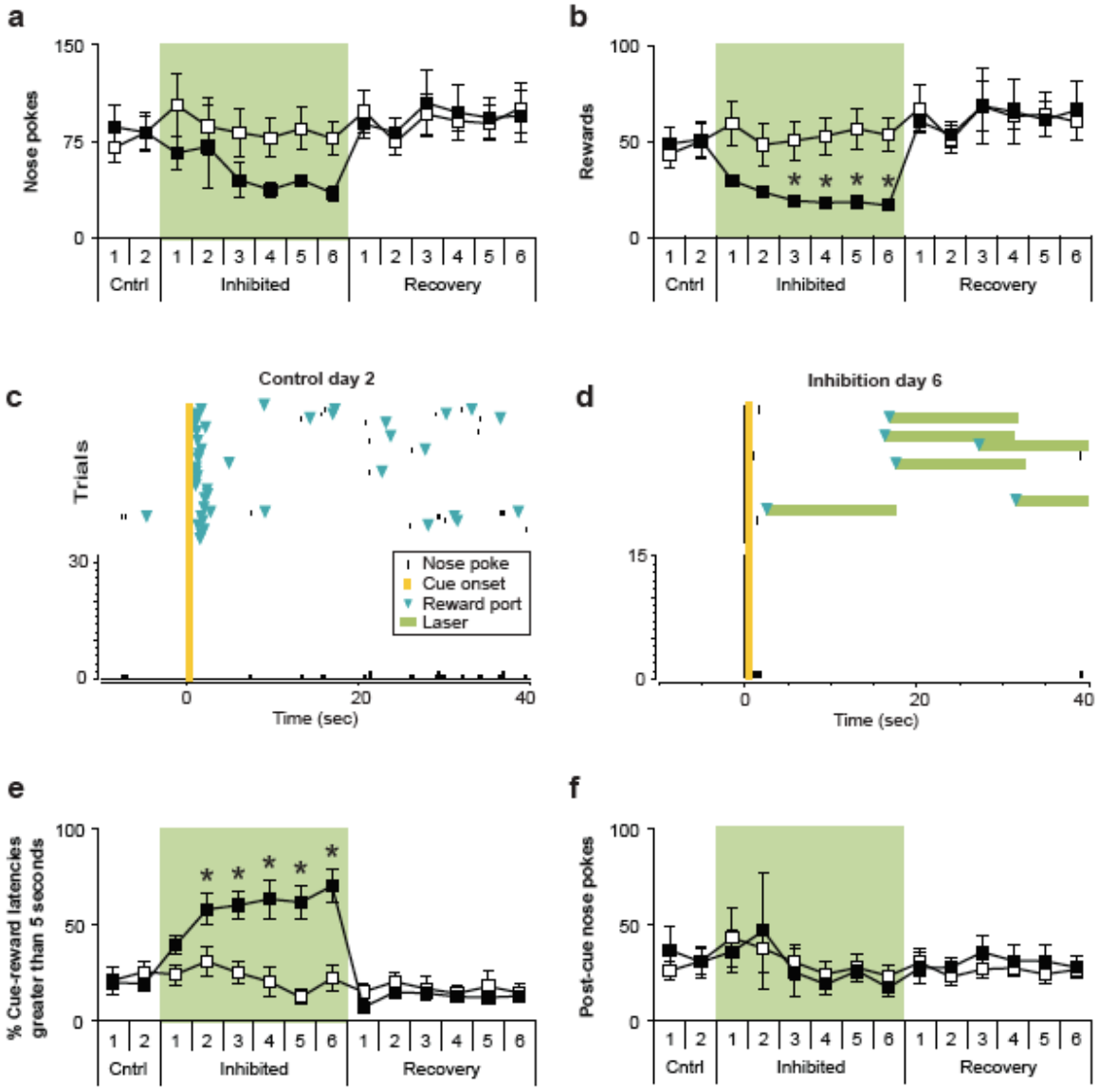


Figure 13 Reward-paired DA neuron inhibition extinguishes continuously reinforced operant responses and cue-reward associations (a) Total nose pokes completed in the 2-hour FR1 task across 14 days of testing (control vs. inhibition days, paired: main effect of day p=0.039, unpaired: 0.997; inhibition vs. recovery days, paired: main effect of day p=2.23e-05). (b) Rewards earned across 14 days of testing (control vs. inhibition days, paired: main effect of day p=9.36e-05, unpaired: p=0.89; inhibition vs. recovery days, paired: main effect of day p=3.23e-11; *p<0.05). (c) Behavior during the second control session from a representative reward-paired inhibition Th-cre+ mouse. (d) Behavior during the sixth inhibition session from a representative reward-paired inhibition Th-cre+ mouse. (e) Percentage of latencies to collect the reward after cue presentation that are greater than 5 seconds (control vs. inhibition days, paired: main effect of day p=3.12e-05, unpaired: p=0.35; inhibition vs. recovery days, paired: main effect of day p=4.82e-11; *p<0.05). (f) Number of nose pokes performed after cue presentation and before reward collection (control vs. inhibition days, paired: p=0.624, unpaired: 0.99). Paired n=8, unpaired n=8. Kruskal-Wallis test. Data are presented as means and error bars represent s.e.m.

behavioral measures in either the NpHR-unpaired (Fig. 13) or YFP-paired inhibition groups (Supplementary Fig. 2).

Extinction produced by reward-paired dopamine neuron inhibition is cue specific

We next sought to determine whether extinction of cue-reward associations produced by pairing DA neuron inhibition with reward consumption would generalize to other cue-reward associations. We trained mice on a two-nose poke FR1 task. In this task, two nose pokes flanked a central reward port. On each trial, the active nose poke was determined on a pseudo-random basis. A light was illuminated in the back of the nose poke port to indicate whether the right or left nose poke was active and each nose poke port produced a distinct tone or white noise cue to indicate reward availability. After performing a correct nose poke, a reward pellet was available in the center reward port. The same reward type was delivered after a nose poke on either side. On reward-paired inhibition days, a 15 second laser-pulse was triggered after mice entered the reward port only after performing either the left or the right nose poke (the nose poke response side that was paired with inhibition was counterbalanced across subjects) (Fig. 14a). With these inhibition parameters we were able to measure both nose poke and cue-reward associations in the same session while only one nose poke response was paired with DA neuron inhibition. We found that inhibiting DA neurons during reward consumption significantly increased the latency to collect the reward after performing a nose poke on the side paired with inhibition (Fig. 14b,d) but had no effect on the cue-to-reward latency on the side not paired with DA neuron inhibition (Fig. 14b,c) (two-way

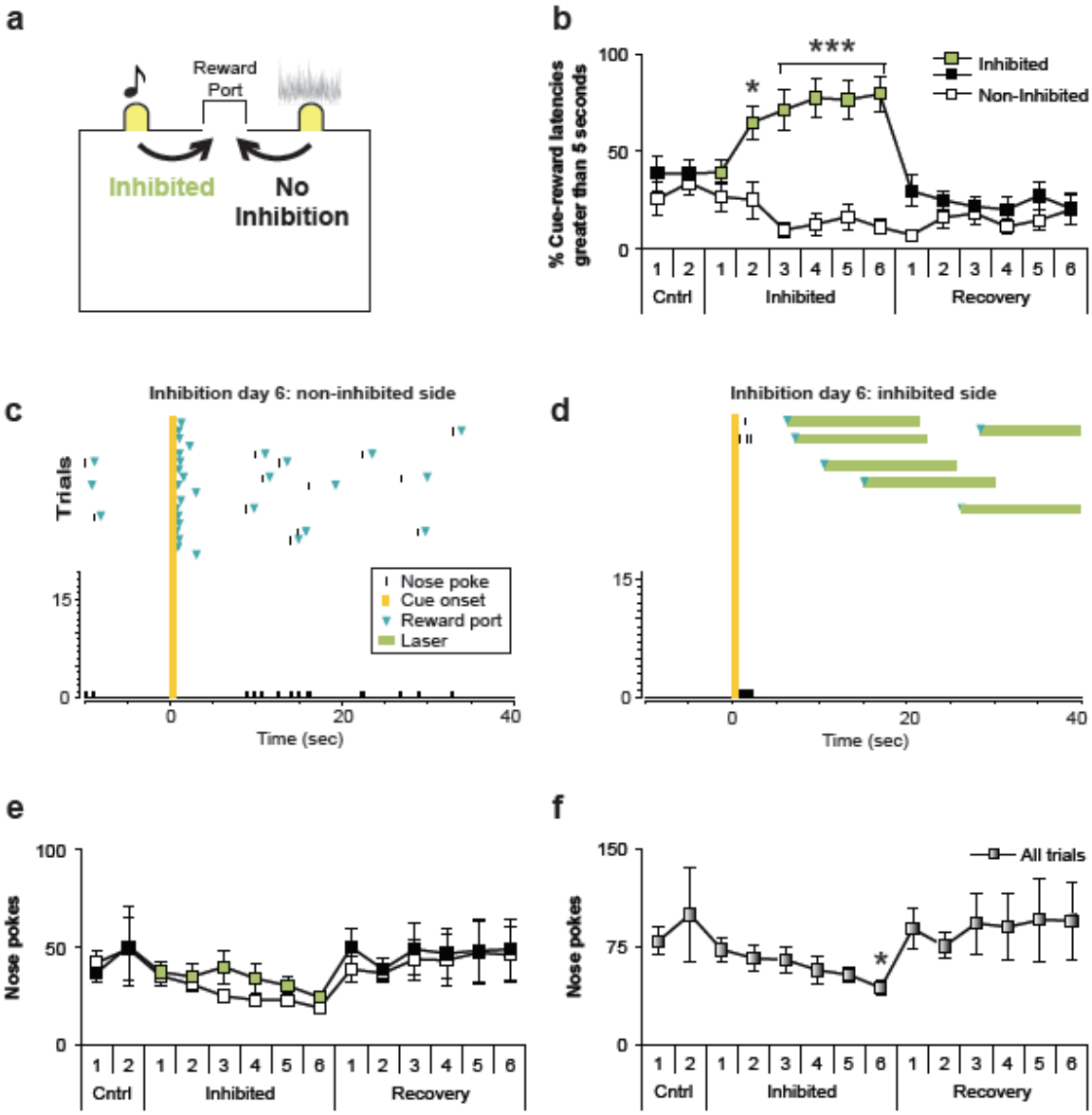
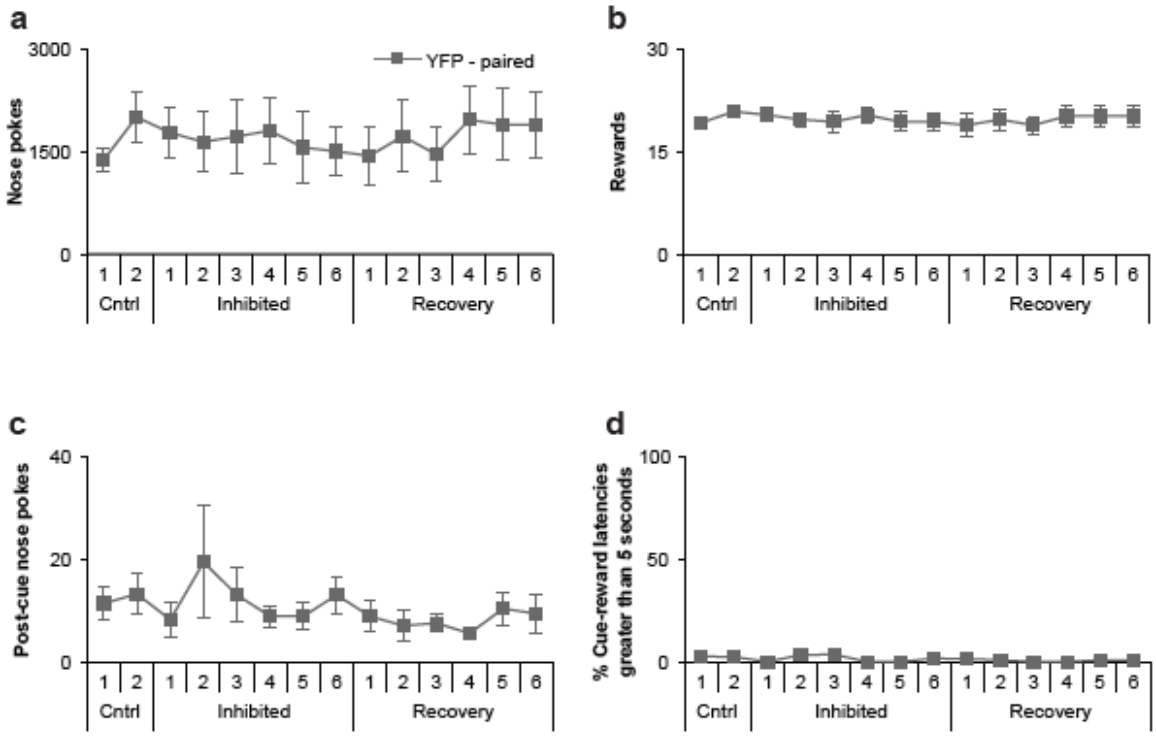
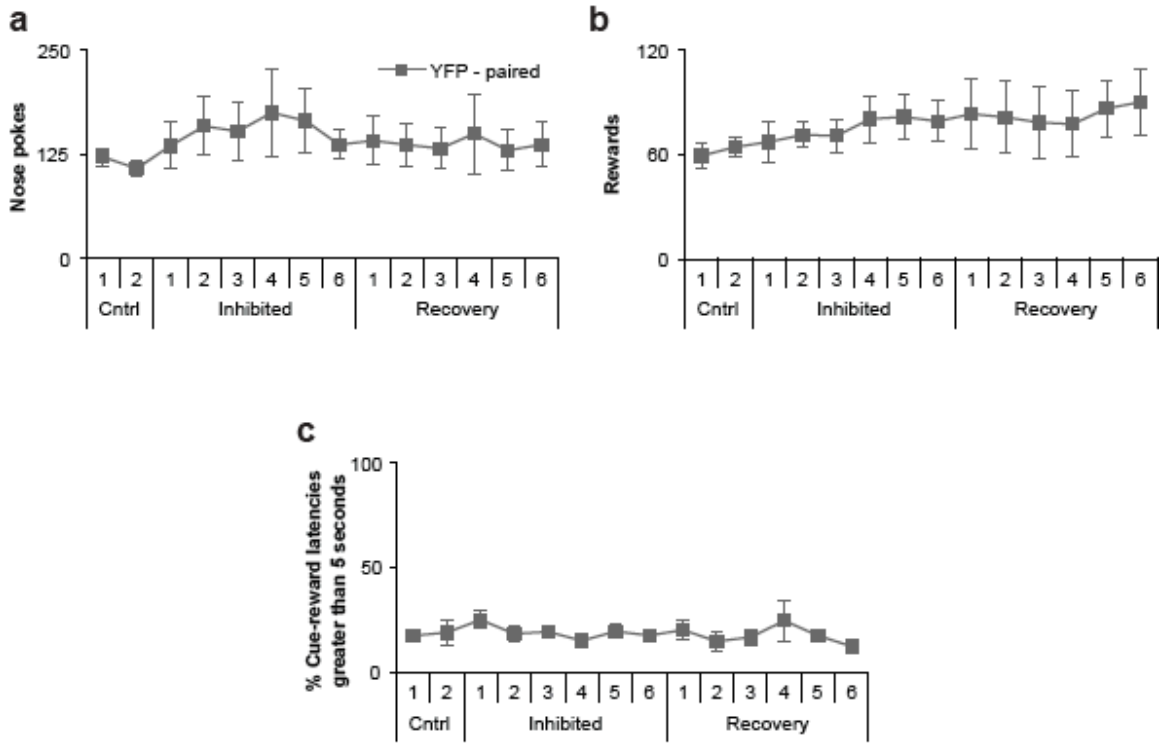


Figure 14 Reward-paired DA neuron inhibition selectively attenuates cue-reward associations. (a) Diagram of test parameters. Two nose pokes flank a center reward port. DA neurons were inhibited during reward consumption after mice performed a nose poke on the left nose poke port. DA neurons were not inhibited after mice performed a nose poke on the right nose poke port. The side that was paired with DA neuron inhibition was counterbalanced across animals. (b) Percentage of latencies to collect the reward after cue presentation that are greater than 5 seconds (2-way ANOVA *p<0.05, ***p<0.0001). (c) Behavior during the second control session from a representative reward-paired inhibition Th-cre+ mouse. (d) Behavior during the sixth inhibition session from a representative reward-paired inhibition Th-cre+ mouse. (e) Total nose pokes completed on each nose poke in the 2-hour 2-nose poke FR1 task task across 14 days of testing (2-way ANOVA, group p=0.18, effect of day p=0.06). (f) Total nose pokes completed on both nose pokes across 14 days of testing. (Kruskal-Wallis test, control vs. inhibition days, main effect of day p=0.003, *p<0.05.) Paired n=8, unpaired n=8. Data are presented as means and error bars represent s.e.m.

repeated-measures ANOVA with Greenhouse-Geisser correction, main effect of group, P=0.001, main effect of day, P=0.03 , group x day, P<0.0001). Because mice were forced to perform a nose poke on the side paired with inhibition before being allowed to earn a reward by nose poking on the side not paired with inhibition, the number of opportunities mice had to interact with the non-inhibited nose poke was constrained by the task parameters. As a result, we also observed a trend toward a decrease in the number of nosepokes performed at both nose poke ports (two-way repeated-measures ANOVA with Greenhouse-Geisser correction, revealed no main effect of group, P=0.18, or day, P=0.06) (Fig. 14e). However, the total number of nose pokes performed on both nose pokes across the six inhibition sessions was significantly decreased compared to control sessions (Kruskal-Wallis test, control vs. inhibition days, main effect of day p=0.003) (Fig. 14f). These results indicate that reward-paired DA neuron inhibition generates an artificial negative reward prediction error that is specific to the cue that preceded inhibition and does not generalize to other cues experienced in the same behavioral session.



Supplementary Figure 1 Reward-paired DA neuron inhibition does not decrease operant responding or cue-realted behavior on a PR task in Th-Cre mice injected with a YFP containing virus. (a) Total nose pokes completed in the 2-hour PR task across 14 days of testing (control vs. inhibition days, p=0.93). (b) Rewards earned across 14 days of testing (control vs. inhibition days, p=93). (c) Number of nose pokes performed after cue presentation and before reward collection (control vs. inhibition days, p=97). (d) Percentage of latencies to collect the reward after cue presentation that are greater than 5 seconds (control vs. inhibition days, p=38). n=6. Kruskal-Wallis test. Data are presented as means and error bars represent s.e.m.



Supplemental Figure 2 Reward-paired DA neuron inhibition does not decrease operant responding or cue-realted behavior on a, FR1 task in Th-Cre mice injected with a YFP containing virus. (a) Total nose pokes completed in the 2-hour FR1 task across 14 days of testing (control vs. inhibition days, p=0.92). (b) Rewards earned across 14 days of testing (control vs. inhibition days, p=0.90). (c) Percentage of latencies to collect the reward after cue presentation that are greater than 5 seconds (control vs. inhibition days, p=0.69). n=6. Kruskal-Wallis test. Data are presented as means and error bars represent s.e.m.

DISCUSSION

We have shown that inhibitions of DA neuron firing directly contribute to learning. By pairing inhibition of DA neurons with reward consumption, we generated an artificial negative reward prediction error signal and examined consequential changes in goal-directed behavior. We found that this pattern of DA neuron inhibition extinguished continuously reinforced Pavlovian cue-reward associations as well as continuously reinforced, but not variably reinforced, operant responses. Finally, extinction produced by reward-paired DA neuron inhibition was specific to the operant response that was antecedent to the inhibition and did not generalize to other learned associations performed in the same behavioral session. Importantly, the precise timing of DA neuron inhibition was critical; DA neuron inhibition triggered just 15 seconds after reward consumption we did not produce a change in any of the behaviors measured. Our results demonstrate that DA neuron activity at the time of reward consumption is required to maintain learned associations and inhibitions of DA neuron activity act as a teaching signal that drives new learning.

Inhibitions of DA neurons that occur when expected rewards are omitted are thought to drive extinction learning (Aberman and Salamone, 1999; Glimcher, 2011; Pan et al., 2008; Salamone et al., 2007; Schultz et al., 1997). Here, we demonstrate that reward-paired DA neuron inhibition resulted in striking deficits in responding to reward-predictive cues. In a PR task, animals failed to disengage from a bout of nose poke behavior when a cue signaled reward availability, resulting in a significant increase in the latency to collect the reward. To compensate for this decrease in cue-directed

behavior, animals switched to a strategy of checking the reward port more frequently, resulting in a decrease in task efficiency. Mice did not decrease the number of nose pokes performed in inhibition sessions or cease nose poking in response to the cue in reward-paired inhibition sessions, suggesting that they were not avoiding the cue or find it aversive.

Importantly, the gradual decrease in operant responding we observed across reward-paired inhibition sessions in the FR1 task bears a striking resemblance to the extinction curves described in extinction studies in which the reward is withheld (Aberman and Salamone, 1999; Myers and Davis, 2002; Pan et al., 2008; Salamone et al., 2007; Tye et al., 2010). However, the work as presented thus far is lacking an important behavioral control. Therefore, current studies are examining how behavior changes during extinction caused by reward omission in both the FR1 and PR tasks. These studies are especially important given the differences we observed in operant behavior after reward-paired DA neuron inhibition in the FR1 and PR task. In the FR1 task, we observed a gradual and consistent decrease in nose pokes that is typical of a classic extinction curve. However, in the PR task we saw marked evidence of extinction of cue-related behavior while we observed little evidence of extinction of the operant response. Previous studies have shown that variably reinforced responses are more resistant to extinction than continuously reinforced responses (Haselgrove et al., 2004; Thorpe et al., 2005). This interpretation can partially account for the discrepancies we observed in changes in operant responding in the FR1 and PR tasks after reward-paired inhibition. However, in the PR task we saw no evidence of decreased operant responding after the

first day of reward-paired DA neuron inhibition. Rather, mice maintained a consistent, if slightly depressed, level of nose poke behavior across all six days of reward-paired inhibition. This suggests that midbrain DA neuron-independent circuits may maintain variably reinforced operant responses, and that these circuits act to support operant behavior while midbrain DA neurons are inhibited. If we withheld the reward in a PR task and observed a decrease in operant behavior beyond that observed here, this would support the hypothesis that continuously reinforced operant responses are maintained by midbrain DA neuron-dependent circuits while variably reinforced operant responses are support by other circuits.

It is of interest to consider whether reward-paired DA neuron inhibitions may drive changes in synaptic plasticity in parallel with the observed changes in behavior. Previous research has shown that learning of a cue-reward association results in an increase in excitatory drive onto DA neurons (Rada et al., 2005; Stuber et al., 2008). Conversely, extinction training has been shown to increase inhibitory drive onto midbrain DA neurons (Bassareo and Di Chiara, 1999; Pan et al., 2013). This study used extracellular recordings during associative learning and found that midbrain GABAergic neurons respond to both rewards and reward-predictive cues at a latency that preceded the onset of phasic DA neuron activity. In addition, GABAergic units that showed the strongest response to the reward developed the strongest responses to the conditioned stimulus after extinction. These results suggest that midbrain DA neurons may be inhibited by phasic GABAergic activity when extinguished reward-predictive cues are presented, resulting in suppression of the conditioned response. In this study, we found

that reward-paired DA neuron inhibition resulted in a decrease in measures of cue-related behavior that mimics extinction learning. However, it remains to be seen if our inhibition parameters result in an increase of inhibitory drive onto DA neurons similar to that observed during extinction caused by reward omission.

The importance of DA signaling in maintaining motivated behavior is well established (Berridge, 2012; Carr et al., 2001; Carr, 2002; Salamone et al., 2007; Wise, 2004). And while reward-paired DA neuron inhibition resulted in a significant decrease in breakpoint in the PR task, we interpreted this decrease as being caused by a decrease in task efficiency rather than a decrease in motivation. It is perhaps surprising that these inhibition parameters, which resulted in substantial deficits in some task-related behaviors, did not produce an overall decrease in motivation. One possible explanation, which we have mentioned above, is that divergent circuits may control variably reinforced and continuously reinforced operant responses. However, manipulations that specifically block DA receptors in the nucleus accumbens have profound deficits on motivated behavior. Because we target DA neuron cell bodies, our inhibition parameters likely decrease DA release in several downstream regions, including the nucleus accumbens. Therefore, it is not likely that alternative circuits are involved in maintaining motivated behavior in these experiments. A second possibility is that the precise timing of DA neuron activity may contribute to whether DA release in terminal regions contributes to maintaining motivation or contributes to maintaining learned associations. The work presented in chapter 1 of this thesis suggests that this may be the case. In chapter 1 we demonstrated that inhibiting DA neuron activity at the time a

mouse is actively engaged in a bout of nose poke responding significantly decreased the total number of nose pokes the mouse performed in a PR task.

The work presented here demonstrates that DA neuron activity specifically at the time of reward consumption is required to maintain learned associations, but not motivation. Future studies comparing the results presented here with behavioral changes resulting from extinction caused by reward omission as well as changes in synaptic plasticity caused by reward-paired DA neuron inhibition will further our understanding of how DA neuron inhibitions drive new learning.

METHODS

Subjects

31 male Th-Cre mice aged 8-12 weeks were housed individually on a reverse 12 h light/dark cycle (lights off at 10:00). All mice began behavioral training with standard rodent chow and water available *ad libitum*. Mice were tested under food restriction and were brought down to 90% of their free-fed body weight over 5 days before the start of the behavioral test. The Institutional Animal Care and Use Committee of the Ernest Gallo Clinic and Research Center at the University of California, San Francisco, approved all experimental procedures.

Surgical procedures and virus injection

Mice were anesthetized with 100mg/kg ketamine and 10mg/kg xylazine and placed in a stereotaxic frame (Kopf). Mice received an infusion of either Cre-dependent halo

rhodopsin (AAV5/Ef1a-DIO-eNpHR3.0)-eYFP or a control Cre-dependent YFP virus (AAV5/Ef1a-DIO)-eYFP. A custom made 31 gauge infuser was used to deliver 1µl of virus targeting the VTA (target coordinates: ML±0.5mm, AP -3.5mm, DV -4.45mm relative to bregma) at a speed of 0.1µl per minute. The infuser was left in place for an additional 10 minutes following each infusion. Two chronic bilateral optical fibers were then implanted (ML±1.5mm) at an angle of ±13.5° to a targeted depth of DV -4.45mm relative to skull. Optical fibers were made in-house with optical fiber (BFL37-200, Thorlabs) and a ceramic ferrule (MM-FER2007C-2500, Precision Fiber Products) and were secured to the skull surface with two metal screws and dental cement. As an analgesic, all mice were given ad libitum access to acetaminophen (80mg/100ml water, oral, baby aspirin) for five days following surgery. All mice were given 7 days to recover before starting handling and behavioral procedures. All behavioral tests were conducted at least 5 weeks postsurgery.

Behavioral procedures

Behavioral conditioning: fixed ratio 1 test

Mice were trained in standard operant chambers (Med Associates) in ventilated sound-attenuating chambers. They were outfitted with a 2.8W/100mA house light and two nose poke ports flanking a recessed port into which a reward pellet was delivered. The back of the active nose poke port was illuminated with a cue light when active and a 500ms, 1kH tone was delivered after each active nose poke was completed on the active nose poke, indicating that a reward pellet was available. Mice were first trained

on a fixed ratio 1 (FR1) schedule of reinforcement to self-administer a high-fat food (35%, 20mg, Bio-serv) pellet.

Behavioral conditioning: Progressive ratio test

After acquiring the FR1 task, mice were then progressed to FR3 training for a minimum of 3 days before progressing to an FR8 or PR schedule for testing. Under the PR schedule, the number of nose poke responses required to earn a pellet was increased after each successive reward according to the equation x = (5*e^reward*0.24)-5, which produced the following response requirement schedule: 1,2,4,6,9,12,15,20,25,32,40,50,62,77,95,118… After completing the current response requirement, a 500ms, 1kH tone cue signaled the completion of the current ratio requirement and the availability of the reward.

Behavioral conditioning: dual-nose poke fixed ratio 1 test

Mice were trained in standard operant chambers (Med Associates) in ventilated sound-attenuating chambers. In this test, each of the two nose poke ports that flanked the center reward port were activated in a pseudo-random fashion. When the left nose poke port was active, the back of the port was illuminated with a cue light and a 500ms, 1kH tone was delivered after each active nose poke was completed on the active nose poke, indicating that a reward pellet was available. When the right nose poke port was active, the back of the port was illuminated with a cue light and a 500ms; white noise sound was delivered after each active nose poke was completed on the active nose

poke, indicating that a reward pellet was available. For both the right and left active nose poke trials, the type of reward pellet and the location of pellet delivery (center reward port) was held constant. On inhibition days (inhibition parameters described below), one nose poke was paired with DA neuron inhibition, such that DA neurons were inhibited at the time of reward consumption after mice had performed an active nose poke on the left side, but not after mice had performed an active nose poke on the right side. The nose poke side that was paired with DA neuron inhibition was counterbalanced across mice.

In vivo optical inhibition parameters

For sessions with optical inhibition of DA neurons, the implanted fiber optic implants were attached to patch cables (MFP_200/240/900-0.22_2.0m_sma, Doric Lenses) that terminated with a ceramic ferrule with a fitted ceramic sleeve (SM-CS125S, Precision Fiber Products). The other end of the patch cable was attached to a bilateral optical commutator (Doric Lenses) and connected – via a second patch cable – to a 200mW DPSS 532 nm laser (OEM Laser Systems). The timing of optical inhibition was triggered by a computer running Med PC IV (Med Associates) software, which also recorded nose poke and reward port responses. Behavior was first recorded across two control test sessions that were not paired with inhibition. This was followed by 6 consecutive days of inhibition training in which a 15 second inhibition laser pulse was triggered each time an animal entered the reward port to collect a reward pellet. The laser was triggered 0.01 seconds after the animals entered the reward port and the pellet was dropped 0.1

seconds after the animal entered the reward port to ensure DA neurons were fully inhibited before the animal received the reward. Following the 6 inhibition test days, animals were then trained for 6 additional "recovery" days in which the laser was not turned on.

Histological verification of NpHR virus expression and fiber optic placement

Mice were deeply anesthetized with pentobarbital and transcardially perfused with phosphate-buffered saline followed by 4% paraformaldehyde. Brains were extracted, post-fixed in paraformaldehyde and cryoprotected in 25% sucrose for 24 hours before slicing. Brains were sectioned coronally on a freezing microtome at a thickness of 50µm. Sections were washed in PBS and incubated with 0.2% bovine serum albumin (BSA) and 0.2% Triton X-100 for 20 minutes. 10% normal donkey serum (NDS) was added for an additional 30 minutes incubation. Sections were then incubated with primary antibodies overnight at 4 degrees Celsius in PBS with BSA and Triton X-100. Concentrations for primary antibodies were as follows: mouse anti-GFP (1:1500, Invitrogen), rabbit anti-TH (1:1500, Fisher Scientific). Sections were then washed with 2% NDS in PBS for 10 minutes before incubating with secondary antibodies for 2 hours at room temperate. Concentrations for secondary antibodies were 1:200 and were conjugated to Alexa Flour 488 or 594 dyes (Invitrogen). Sections were mounted onto glass microscope slides in phosphate buffered water, coverslipped with Vectashield mounting medium (Fisher Scientific), and imaged on a confocal microscope to quantify virus expression and verify

fiber placements (Borgland et al., 2009; Franklin and Paxinos, 2007; Glimcher, 2011; Hollerman and Schultz, 1998; Pan et al., 2008). Mice with insufficient virus expression or inaccurate fiber placements were discarded from the study.

Statistics

All values are expressed as mean±SEM. Statistical differences across testing days were assessed using the non-parametric Kruskal-Wallis one-way ANOVA. Statistical differences within groups in figure 14 were assessed using repeated measures 2-way ANOVA using Greenhouse-Geisser correction. A level of confidence of $P < 0.05$ was employed for statistical significance.

Chapter 3

Ghrelin signaling in the VTA is required to maintain high-effort responding, but not feeding behavior.

ABSTRACT

The orexigenic peptide ghrelin acts on hypothalamic feeding circuits to impact food consummatory behavior. However, ghrelin also interacts with diverse brain systems and is implicated in modulating behaviors such as food palatability and preference, as well as motivation, learning and memory. It has been hypothesized that ghrelin modulates motivation as well as preference for palatable food via its actions in the VTA. We infused ghrelin or ghrelin receptor antagonists directly into the VTA and measured feeding behavior as well as motivation to obtain both a high fat and regular food reinforcer. We found that ghrelin increased, while ghrelin receptor antagonists decreased, motivation to obtain both reinforcer types. After food restricting mice to upregulate endogenous ghrelin signaling, blocking ghrelin signaling systemically reduced feeding behavior. In contrast, ghrelin receptor antagonists infused into the VTA had no effect on feeding behavior. We therefore conclude that ghrelin acts on separate parallel circuits to modulate food preference and feeding via its actions in the hypothalamus and motivation via its actions on the VTA.

INTRODUCTION

The desire to consume food in dissociation from physiological necessity is not a difficult concept to grasp in a society where the availability of rewarding and calorie-dense foods far exceeds any requirement for caloric intake. Indeed, as the global prevalence of obesity continues to increase at unprecedented rates (Feneberg and Malfertheiner, 2012; Popkin et al., 2012), it is imperative that we understand how motivational systems drive us to seek out and consume food, as well as how homeostatic systems evolved to dampen feeding responses signal satiety.

Ghrelin, the only known peripherally produced orexigenic hormone (Kojima and Kangawa, 2005), has been demonstrated to cross the blood-brain barrier (Diano et al., 2006) where its receptor (Growth Hormone Secretagogue Receptor type 1) is expressed in feeding centers in the hypothalamus. Therefore, ghrelin is uniquely situated to convey hunger and feeding signals from the gut to the brain. Several studies have demonstrated the importance of ghrelin signaling on hypothalamic feeding centers in maintaining appetite and feeding behavior (Nakazato et al., 2001; Tamura et al., 2002). However, the ghrelin receptor is also expressed in a variety of brain regions not known for their role in maintaining feeding behavior, including the ventral tegmental area (VTA). Previous work has shown that ghrelin injection into the VTA increases feeding (Naleid et al., 2005; Skibicka et al., 2011). Additionally, it has been shown that ghrelin receptor antagonists administered into the VTA counteract the orexigenic effects of peripherally administered ghrelin (Abizaid et al., 2006), suggesting an important role for VTA ghrelin receptor signaling in feeding. However, it is not known whether ghrelin signaling in the

VTA is required to maintain feeding behavior specifically after increasing endogenous ghrelin and other orexigenic factors via food-restriction.

Ghrelin may play a role in feeding by increasing the hedonic value of palatable reinforcers. In a human fMRI study, a systemic ghrelin injection potentiated VTA activation in response to pictures of highly palatable foods (Malik et al., 2008). Similarly, ghrelin increased the rewarding value of palatable foods in rats (Perello et al., 2009). Several studies have suggested that ghrelin signaling in the VTA may preferentially modulate the hedonic or motivational aspects of palatable foods over more bland food reinforcers. In one study, VTA-lesioned rats treated with i.c.v. ghrelin consumed less high fat food but the same amount of regular food when compared to ghrelin-treated, sham-lesioned rats (Egecioglu et al., 2010). However, while VTA-lesioned rats did consume less high fat food compared to sham-lesioned animals, ghrelin treatment did increase intake of high fat food in VTA-lesioned rats when compared to saline treatment suggesting that ghrelin-induced increases in high fat food are driven only in part by the VTA. However, it is important to note that findings in this study may be confounded by differences in access to food. Food was freely available to the rats when assessing consumption of regular food. On the other hand, when assessing consumption of high fat food, rats were required to extract peanut butter from an eppendorf tube, thus introducing a motivational component. Therefore, it remains unclear whether ghrelin signaling in the VTA is required to increase feeding of high fat foods when freely available. Additionally, it is not known if ghrelin signaling in the VTA is required to maintain motivation to consume regular food reinforcers.

The mesoaccumbens dopamine circuit is critical for promoting motivation in food-seeking (Salamone et al., 2007). The idea that ghrelin may act to increase hunger via its receptors in the hypothalamus and to increase motivation via its receptors in the VTA has been proposed (Kenny, 2011; van Zessen et al., 2012a). Intra-VTA ghrelin receptor antagonists decrease motivation to obtain highly palatable reinforcers such as sucrose (Skibicka et al., 2011) and palatable chocolate-flavored pellets (King et al., 2011). However, experiments directly comparing the effects of VTA ghrelin receptor antagonism on motivation to obtain reinforcers with varying palatability and macronutrient content have not been performed.

In this study we attempt to clarify the role that ghrelin signaling in the VTA plays in modulating feeding and motivation. We also attempt to clarify whether ghrelin acts on the VTA to differentially modulate the hedonic aspects of palatable, high fat reinforcers versus bland, regular food reinforcers. We hypothesize that ghrelin signaling in the VTA is required to maintain motivation to obtain food reinforcers and that it is not required to maintain feeding behavior in response to food restriction. Furthermore, we hypothesize that ghrelin does not modulate the hedonic aspects of food reinforcers via its actions in the VTA and therefore intra-VTA ghrelin receptor antagonists will similarly decrease motivation to obtain both high fat and regular food reinforcers.

RESULTS

Systemic, but not intra-VTA, ghrelin receptor antagonists decrease feeding.

We first determined whether endogenous ghrelin at the systemic level, or at the level of the VTA, mediates feeding behavior. We increased endogenous ghrelin signaling by food restricting mice to 90% of their free-fed body weight. We then confirmed that a systemic ghrelin receptor antagonist could decrease feeding by measuring food intake in the home cage over a 2-hour test period. Mice were given an i.p. injection of either saline or a ghrelin receptor antagonist (D-Lys3-GHRP-6, 1mM, 200µl) 10 minutes prior to being given free access to either standard chow or a high fat pellet. We observed that i.p. ghrelin receptor antagonists decreased the quantity of both high fat and regular food consumed in the first hour (high fat: $P<0.0001$, saline: 1.44±0.10g, D-Lys: 0.29±0.11g; regular food: $P<0.0001$, saline: 0.99±0.06g, D-Lys: 0.21±0.07g) (Fig. 15a,b, left) after drug administration. This demonstrates that endogenous ghrelin is required to maintain a feeding response after prolonged food restriction. In the second hour after drug administration, we observed an increase in the amount of high fat food consumed ($P=0.03$, saline: 0.54±0.07g, D-Lys: 0.73±0.05g) (Fig. 15a, right) and no difference in the amount of regular food consumed ($P=0.33$, saline: 0.60±0.06g, D-Lys: 0.50±0.06g) (Fig. 15b, right) compared to saline treated controls. This is likely an indication that systemically administered ghrelin receptor antagonists have decreased efficacy after the first hour of being administered.

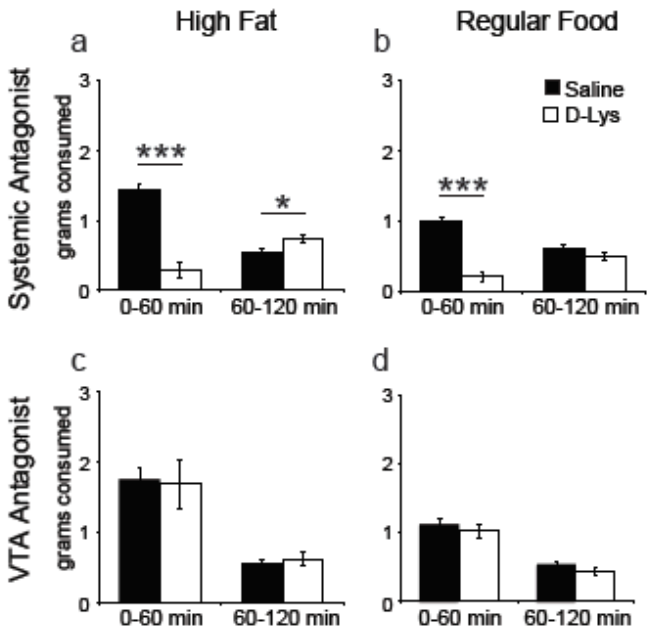


Figure 15 Inhibition of systemic ghrelin signaling is required to maintain feeding behavior, while inhibition of ghrelin signaling in the VTA does not affect feeding behavior. (a) High fat (n=11) consumed during the first (***p<0.001) and second hour (*p<0.05) after i.p. ghrelin receptor antagonists. (b) regular food (n=11) consumed during the first (***p<0.001) and second hour after i.p. ghrelin receptor antagonists. (c) High fat (n=15) and (d) regular food (n=16) consumed during the first and second hour after intra-VTA ghrelin receptor antagonists. All values are represented as mean ±SEM.

While DA arising from the nigrostriatal pathway has been shown to be critical for maintaining feeding behavior (Salamone et al., 1993; Szczypka et al., 2001), DA signaling from the VTA is not thought to be critical for feeding or consumption of rewards (Ikemoto and Panksepp, 1996; Nowend et al., 2001). However, studies have shown that GABAergic signaling from the VTA can contribute to feeding (van Zessen et al., 2012b). Because ghrelin receptors are expressed on both dopaminergic and GABAergic cell types in the VTA (Abizaid et al., 2006), we next determined whether endogenous ghrelin acting at the level of the VTA is necessary to maintain feeding behavior. We administered saline or D-Lys3-GHRP-6 (20mM in saline, 0.5µl/side) into the VTA 10 minutes before performing a feeding test in the home cage. We observed no difference in the amount of high fat or regular food consumed over the two-hour feeding test (high fat, 1hr, P=0.75, saline: 1.75±0.16g, D-Lys: 1.68±0.34g, 2hr, P=0.67, saline: 0.56±0.04g, D-Lys: 0.61±0.10g; regular food, 1hr, P=0.50, saline: 1.12±0.08g, D-Lys: 1.02±0.09g, 2hr, P=0.22, saline: 0.54±0.05, D-Lys: 0.44±0.05) (Fig. 15c,d), indicating that endogenous ghrelin signaling is not required to maintain feeding behavior via its actions in the VTA.

Systemic ghrelin receptor antagonists decrease effort to obtain both high fat and regular food reinforcers.

To determine the requirement for endogenous ghrelin signaling in motivated behavioral responding, we administered i.p. ghrelin receptor antagonists (D-Lys3-GHRP-6, 1mM, 200µl) , thus blocking the actions of ghrelin on all its receptors in addition to

the VTA, before assessing performance in a high-effort operant behavioral task. Mice were tested under a progressive ratio (PR) operant schedule, in which the number of nose pokes to earn a reward is increased after each trial, and which is thought to assess motivation (Hodos, 1961; Richardson and Roberts, 1996). Ghrelin receptor antagonists significantly reduced the total number of nose pokes completed at the end of the two-hour session (high fat, P=0.005, saline: 1886.38±358.20, D-Lys: 549.63±146.48; regular food, P=0.002, saline: 2538.67±366.42, D-Lys: 1023.67±180.48) (Fig. 16a) and also significantly decreased the last response requirement completed (high fat, P=0.005, saline: 313.38±57.31, D-Lys: 103.88±26.25; regular food, P=0.003, saline: 458.44±65.79, D-Lys: 175.00±33.67) (Fig. 16b) compared to saline-treated controls. The observed decrease in nose poking behavior was observed primarily during the first hour after antagonist treatment, with a rebound occurring in the second hour of the test session (Fig. 16c,d), again indicating that the effectiveness of systemically administered D-Lys3-GHRP-6 may be limited to the first hour after administration.

Some findings have suggested that ghrelin signaling in the VTA may mediate preference for "highly palatable" foods such as high fat reinforcers (King et al., 2011; Perello et al., 2009). However, previous work has also shown that VTA dopamine likely mediates motivational aspects of feeding behavior without contributing to hedonic aspects of palatable food consumption (Cannon and Palmiter, 2003; Kenny, 2011). To determine whether i.p. ghrelin receptor antagonists differentially mediate responding for high fat and regular food reinforcers we measured the difference in total nose pokes

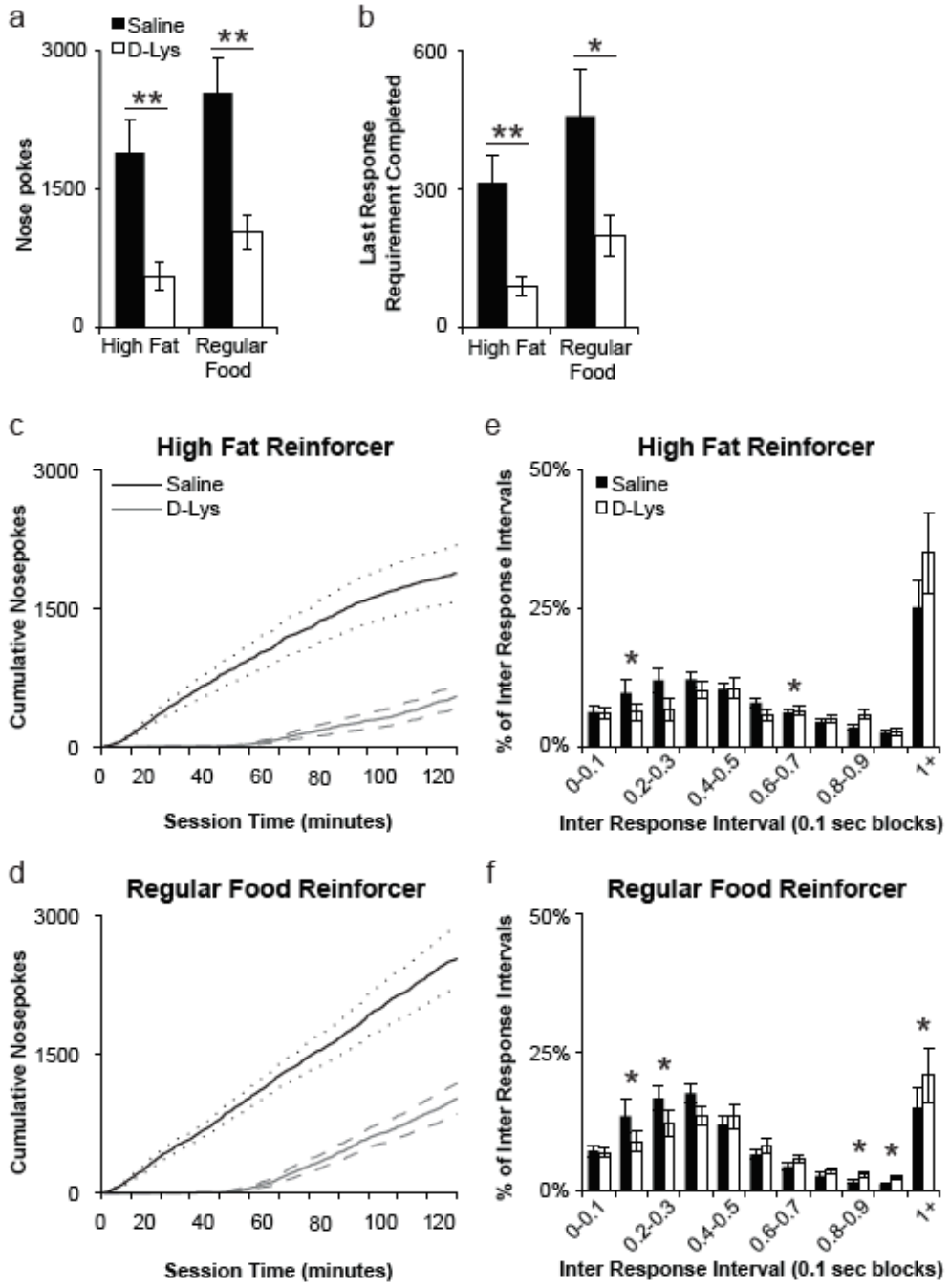


Figure 16 Inhibition of systemic ghrelin signaling decreases motivation to obtain high fat (n=8) and regular food reinforcers (n=6). (a) Total nose pokes completed in the two-hour PR task. **p<0.01. (b) Last response requirement completed at the end of the two-hour PR test. (c,d) Cumulative nose pokes during the PR test for high fat (c) and regular food (d) reinforcers *p<0.05, **p<0.01. Solid lines represent the mean and dashed lines repre sent ±SEM. (e,f) Inter response intervals for high fat (e) and regular food (f) reinforcers expressed in bins as percentage of total responses. *p<0.05. All values are represented as mean ±SEM.

completed after saline versus antagonist treatment. We did not observe a difference in the degree to which i.p. ghrelin receptor antagonists decreased the number of nose pokes completed between animals administering high fat reinforcers and regular food reinforcers (p=0.71, HF: 1336.75±334.41, RF: 1515±330.55).

To confirm that the observed decrease in nose pokes was due to a decrease in motivation, rather than due to a motor deficit as a result of antagonist administration, we measured the inter response interval (IRI) between nose pokes less than one second. We did not see a change in the percentage of IRIs in the fastest time bin measured (0-0.1 seconds) (high fat, P=0.76, saline: 6.19%±1.28%, D-Lys: 6.03%±1.00%; regular food, P=0.65, saline: 7.34%±1.00%, D-Lys: 7.04%±0.87%) (Fig. 16e,f). This indicates that while mice did show a shift toward IRI's longer than 0.5 seconds, this shift was likely due to a change in motivation rather than due to a motor deficit.

Ghrelin administered in the VTA stimulates high-effort food seeking behavior.

We next assessed whether ghrelin regulates motivated behavior via its actions in the VTA. We injected ghrelin (1mM in saline, 0.3µl/side) directly into the VTA, 10 minutes before free-fed mice were tested on a PR task. Ghrelin injection significantly increased both the number of nose pokes (high fat, P=0.03, saline: 624.00±141.67, ghrelin: 1407.67±274.82; regular food, P=0.01, saline: 600.00±12.08, ghrelin: 1540.50±323.35) (Fig. 17a) and the last response requirement completed (high fat, P=0.02, saline: 70.83±10.58, ghrelin: 119.00±14.85; regular food, P=0.006, saline: 63.83±10.72, ghrelin: 118.17±19.02) (Fig. 17b) in mice working for either high fat (Fig.

17c) or regular food reinforcers (Fig. 17d). There was not a significant difference in the difference in the number of nose pokes completed after ghrelin treatment between animals administering high fat or regular food reinforcers (p=0.67, HF: 783.67±263.37, RF: 940.5±247.247.80). These results indicate that ghrelin increases effort to obtain food via its actions in the VTA independently of its orexigenic actions in the hypothalamus and independently of reinforcer type.

Ghrelin signaling in the VTA is required to increase effortful responding for both high fat and regular food reinforcers.

Dopamine signaling from the VTA is recognized to play a crucial role in maintaining motivation when response demands to obtain food reinforcers increase (Aberman and Salamone, 1999; Salamone et al., 2007). Therefore, although we have shown that ghrelin signaling in the VTA is not required to maintain feeding behavior, we hypothesized that endogenous ghrelin signaling in the VTA is required to maintain high-effort behavior. We administered ghrelin receptor antagonists directly into the VTA (20mM in saline, 0.3µl/side) 10 minutes before food-restricted mice performed a PR task. In this test, intra-VTA ghrelin receptor antagonism decreased the total number of nose pokes (high fat, P=0.0002, saline: 1698.17±316.30, D-Lys: 763.58±175.21; regular food, P=0.009, saline: 2411.10±281.28, D-Lys: 1365.09±210.95) (Fig. 18a) and the last response requirement completed (high fat, P=0.0004, saline: 290.33±55.78, D-Lys: 134.33±30.90; regular food, P=0.01, saline: 406.91±48.38, D-Lys: 243.82±34.84) (Fig.

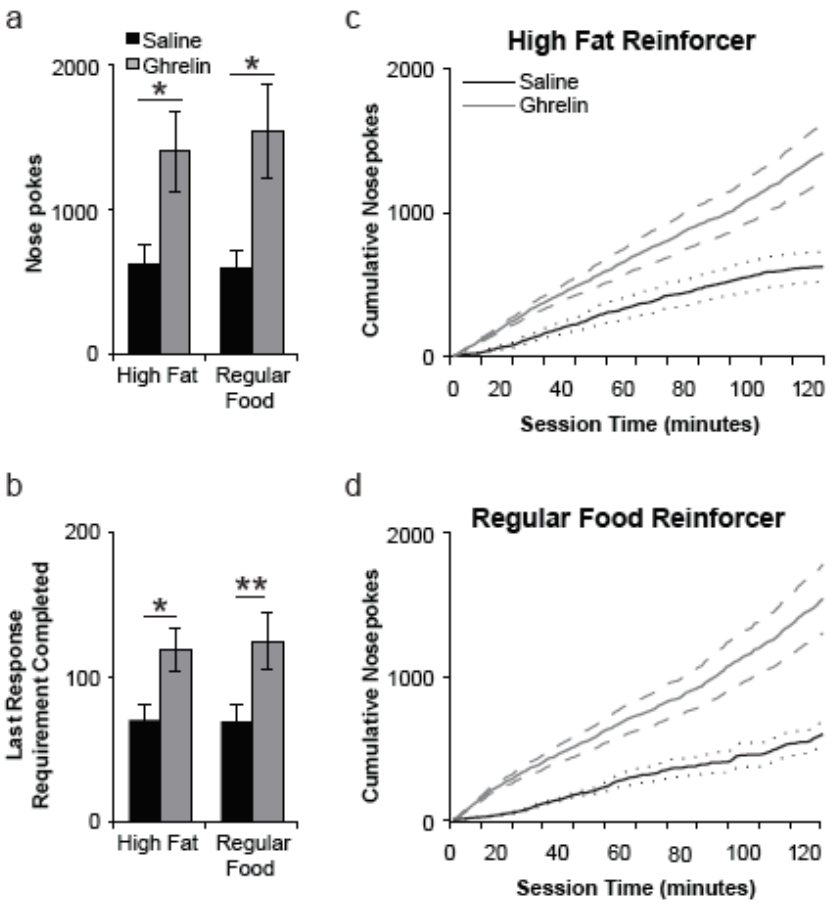


Figure 17 Intra-VTA ghrelin increases motivation to obtain high fat (n=6) and regular food reinforcers (n=6). (a) Total nose pokes completed in the two-hour PR task. *p<0.05.(b) Last response requirement completed at the end of the two-hour PR test. *p<0.05, **p<0.01. All values are represented as mean ±SEM. (c,d) Cumulative nose pokes during the PR test for high fat (c) and regular food (d) reinforcers. Solid lines represent the mean and dashed lines represent ± SEM.

18b) in mice working for either high fat (Fig. 18c) or regular food reinforcers (Fig. 18d). We did not observe a difference in the degree to which intra-VTA ghrelin receptor antagonists decreased the number of nose pokes completed between animals administering high fat reinforcers and regular food reinforcers (p=0.76, HF: 934.58±174.97, RF: 1046±326.32). This indicates that ghrelin signaling in the VTA does not preferentially mediate motivation to obtain highly palatable reinforcers.

To verify that the decrease in nose pokes observed was due to a decrease in motivation rather than a motor deficit, we again measured the IRI between fast nose pokes less than one second. We did not see a change in the percentage of IRIs in the fastest time bin measured (high fat, P=0.69, saline: 5.55%±1.01%, D-Lys: 6.43±1.24%; regular food, P=0.12, saline: 4.98%±0.52%, D-Lys: 6.17%±1.04%) (0-0.1 seconds) (Fig. 18e,f).

Ghrelin signaling in the VTA is not required to maintain low-effort responding.

Previous experiments that have assessed the role of the VTA in maintaining high-effort behaviors have shown that NAc dopamine lesions decrease responding on a high-effort FR64 task but have no effect on a low-effort FR1 task (Aberman and Salamone, 1999). To determine whether the effects of ghrelin receptor antagonism in the VTA on high-effort responding are dissociable from its effect on low-effort responding, we tested food-restricted mice in a two-hour FR1 operant paradigm after administering ghrelin receptor antagonists in the VTA (20mM, 0.3µl/side). We observed no difference in either the number of nose pokes completed (high fat, P=0.22, saline: 97.36±7.63, D-

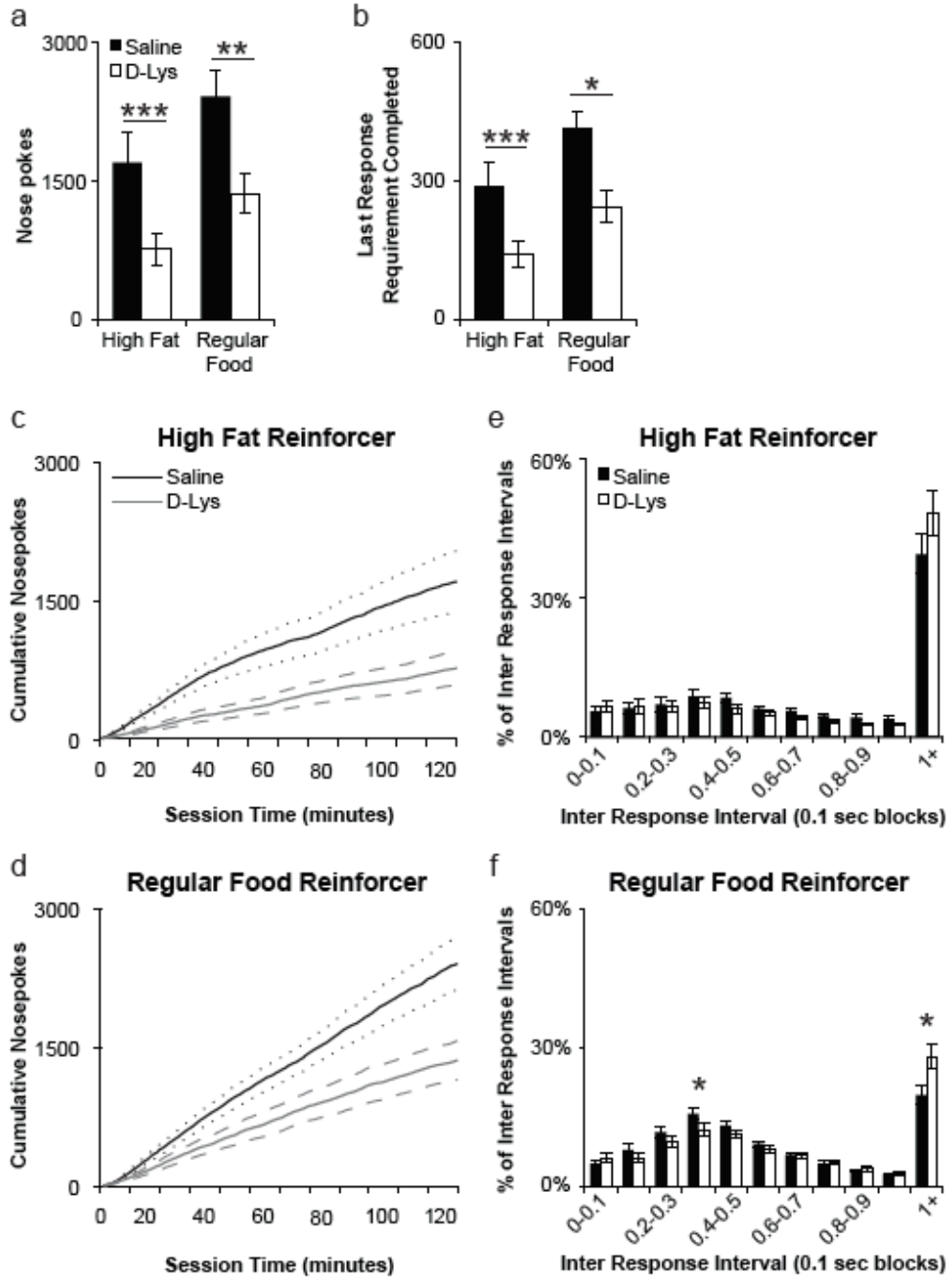


Figure 18 Intra-VTA ghrelin receptor antagonists decrease motivation to obtain high fat (n=13) and regular food reinforcers (n=11). (a) Total nose pokes completed in the two-hour PR task. **p<0.01, ***p<0.001. (b) Last response requirement completed at the end of the two-hour PR test. *p<0.05, ***p<0.001. (c,d) Cumulative nose pokes during the PR test for high fat (c) and regular food (d) reinforcers. Solid lines represent the mean and dashed lines represent ±SEM. (e,f) Inter response intervals for high fat (e) and regular food (f) reinforcers expressed in bins as percentage of total responses. *p<0.05. All values are represented as mean ±SEM.

Lys: 87.64±5.31; regular food, P=0.42, saline: 119.18±17.09, D-Lys: 127.27±17.40) (Fig. 19a), or the number of rewards earned (high fat, P=0.57, saline: 64.45±5.07, D-Lys: 62.36±6.07; regular food, P=0.72, saline: 72.27±7.06, D-Lys: 74.45±6.85) (Fig. 19b) in mice responding for either high fat (Fig. 19c), or regular food (Fig. 19d) reinforcers. This indicates that ghrelin signaling in the VTA is not required to maintain low-effort operant food-seeking behaviors.

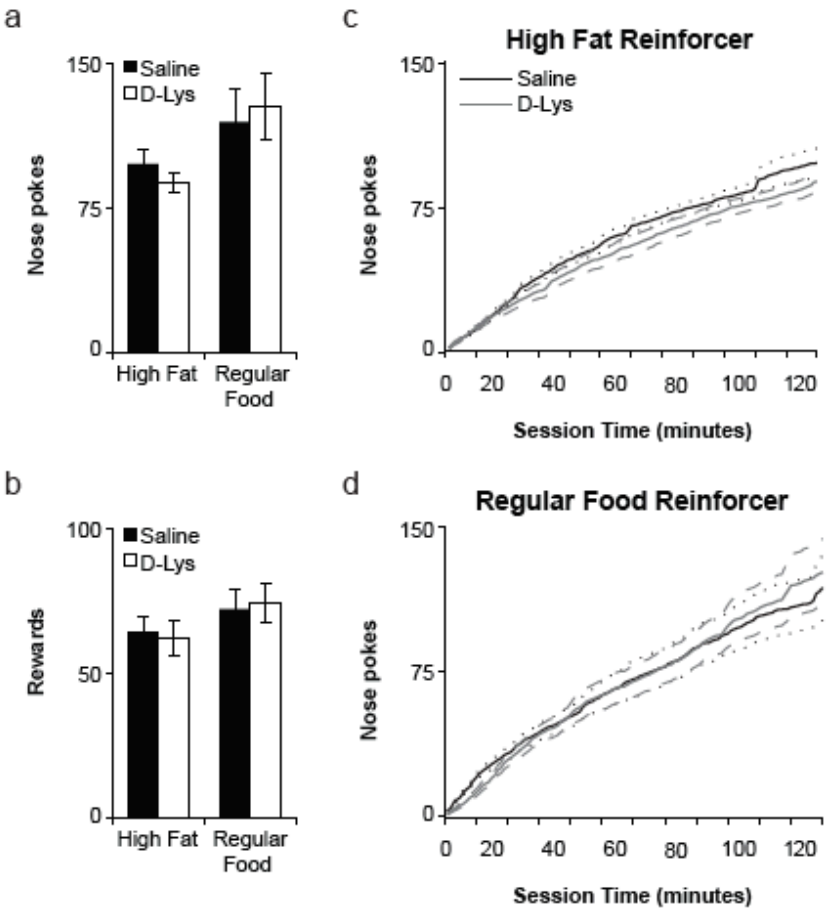


Figure 19 Intra-VTA ghrelin receptor antagonists do not decrease low-effort responding for high fat (n=11) and regular food reinforcers (n=11). (a) Total nose pokes completed in the two-hour PR task. High fat, p=0.22, saline: 97.36±7.63, D-Lys: 87.64±5.31. Regular food, p=0.42, saline: 119.18±17.09, D-Lys: 127.27±17.40. (b) Total number of rewards earned at the end of the two-hour PR test. High fat, p=0.57, saline: 64.45±5.07, D-Lys: 62.36±6.07. Regular food, p=0.72, saline: 72.27±7.06, D-Lys: 74.45±6.85. (c,d) Cumula tive nose pokes during the PR test for high fat (c) and regular food (d) reinforcers. Solid lines represent the mean and dashed lines represent ±SEM. All values are represented as mean ±SEM.

DISCUSSION

This study demonstrates that ghrelin signaling in the VTA is required to increase motivation to obtain food reinforcers in mice, but is not required to maintain feeding behavior. Intra-VTA ghrelin receptor antagonists decreased nose poke responding in a high-effort PR task, whereas intra-VTA antagonists had no effect on a low-effort FR1 task or on feeding. We also showed that intra-VTA ghrelin receptor antagonists have similar effects on nose poke responding for both high-fat and regular food reinforcers, indicating that ghrelin's actions in the VTA are not specific to reinforcer type.

Intra-VTA ghrelin receptor antagonists decreased motivation to obtain food reinforcers. These results are consistent with previous studies that report ghrelin signaling in the VTA is required to increase effort to obtain sucrose or palatable chocolate reinforcers (King et al., 2011; Skibicka et al., 2011). However, these studies emphasized the role of ghrelin in modulating effort directed toward palatable reinforcers. The current results broaden ghrelin's role in modulating motivation for more bland reinforcers. We also showed that administering ghrelin directly into the VTA increased motivation to obtain both high fat and regular food reinforcers, further supporting a role for ghrelin in selectively modulating motivation, and not food palatability, via its actions in the VTA. However, it is possible that ghrelin mediates hedonic aspects of palatable foods through other circuits. Perello et al. found that an i.p. ghrelin injection increased performance on a PR test, an effect that was blocked by orexin antagonists (Perello et al., 2009). Furthermore, work from our lab demonstrated that systemic orexin antagonists decrease motivation for high fat but not regular food

reinforcers in a PR task (Borgland et al., 2009). The ghrelin receptor is expressed on orexin-containing neurons in the lateral hypothalamus and these neurons send dense projections to VTA dopamine neurons (Peyron et al., 1998). Furthermore, orexigenic projections to the VTA are potentiated in rats fed a high fat diet (Borgland et al., 2009). These studies suggest that the orexin system may modulate the incentive value of palatable reinforcers, and also suggests that ghrelin may be well situated to modulate an orexin-VTA motivation circuit.

Antagonizing endogenous ghrelin in food restricted mice at the systemic level, but not at the level of the VTA, decreased the amount of food consumed in a feeding test. While others have shown that ghrelin delivered into the VTA increases feeding behavior (Abizaid et al., 2006), ours is the first study to show that intra-VTA ghrelin receptor antagonists are not sufficient to decrease feeding behavior after increasing endogenous ghrelin signaling via food restriction. Our results demonstrating that intra-VTA ghrelin receptor antagonists do not decrease responding on a low-effort FR1 task provide additional evidence that ghrelin does not modulate feeding via its actions in the VTA.

Therefore, we conclude that ghrelin targets a specific motivational pathway in the VTA that is distinct from circuits mediating ghrelin's effects on feeding, food palatability, or low-effort food seeking behaviors. We suggest the possibilty that ghrelin acts to modulate homeostatic feeding via its actions in the hypothalamus, while it acts in parallel on the VTA to modulate motivation to seek out food reinforcers.

METHODS

Subjects

Male C57Bl/6J mice aged 10 weeks were obtained from Jackson labs and were housed individually on a reverse 12 h light/dark cycle (lights off at 10:00). All mice began behavioral training with standard rodent chow and water available ad libitum. Mice tested under food restriction were brought down to 90% of their free-fed body weight over 5 days before the start of the behavioral test. All experimental procedures were approved by the Institutional Animal Care and Use Committee of the Ernest Gallo Clinic and Research Center at the University of California, San Francisco.

Surgical procedures

Mice were anesthetized with 100mg/kg ketamine and 10mg/kg xylazine and placed in a stereotaxic frame (Kopf). Guide cannulae (26 gauge, Plastics One, Roanoke, VA, USA) were implanted bilaterally into the VTA (target coordinates: ML±0.5, AP -3.5, DV -4.45) using standard stereotaxic procedures. As an analgesic, all mice were given ad libitum access to acetaminophen (80mg/100ml water, oral, baby aspirin) for five days following surgery. All mice were given 7 days to recover before starting handling and behavioral procedures.

Drugs and infusion procedures

For intraperitoneal (i.p.) injections, the ghrelin receptor antagonist (D-Lys3)-GHRP-6 (Tocris) was dissolved in 200µl saline at a concentration of 1mM. On test day, animals

received either saline or antagonist 10 minutes prior to beginning the testing procedure. For intra-VTA injections, ghrelin (Fisher Scientific) was dissolved in saline at a concentration of 1mM, and the ghrelin receptor antagonist, (D-Lys3)-GHRP-6 was dissolved in saline at a concentration of 20mM. Internal cannulae (33 gauge, Plastics One) were connected to 10µl Hamilton syringes with polyethylene tubing filled with distilled water. On test day, animals received infusions of 0.3µl of drug/side, at a rate of 0.2µl/min. Infusers were left in place for 1 min, after which the mice were returned to their home cages for 7-10 minutes before beginning the testing procedure.

Behavioral procedures

Feeding tests

Feeding tests were conducted in the home cage during the dark cycle. Mice were first habituated over 4 days to the feeding procedure. Mice were food-restricted overnight then were given free access to either standard lab chow or a high fat pellet (60%, Bio-serve) for 2 hours. The quantity of food consumed was measured after 1 and 2 hours. At the end of each test mice were given the remaining quantity of food required to maintain 90% body weight. Feeding tests occurred on days 5 and 6. Mice were given saline or ghrelin receptor antagonists, counterbalanced across days. Drugs were administered either i.p. or intra-VTA, depending on the test condition.

Behavioral conditioning: fixed ratio 1 and progressive ratio tests with ghrelin receptor antagonists

Mice were trained in standard operant chambers (Med Associates) in ventilated sound-attenuating chambers. They were outfitted with a 2.8W/100mA house light and two nose poke ports flanking a recessed port into which one of two types of reward pellet was delivered. The back of one nose poke port was illuminated with a cue light when active and a 500ms, 1kH tone was delivered after each response requirement was completed on the active nose poke, indicating that a pellet was available. Mice were first trained on a fixed ratio 1 (FR1) schedule of reinforcement to self-administer either a regular food (grain based, 20mg, Bioserv) or high fat food (35%, 20mg, Bioserv) pellet. After acquiring the FR1 task, a subset of mice were then moved to a FR3 task for 5 days and finally were trained on a progressive ratio (PR) task for testing. Under the PR schedule, the number of nose poke responses required to earn a pellet was increased after each successive reward according to the equation x = (5*e^reward*0.24))-5, which produced the following response requirement schedule: 1,2,4,6,9,12,15,20,25,32,40,50,62,77,95,118… In all schedules, a 500ms, 1kH tone cue signaled the completion of the current ratio requirement and the availability of the reward pellet.

Behavioral conditioning: progressive ratio tests with intra-VTA ghrelin

Because mice receiving intra-VTA infusions of ghrelin were tested free-fed, they were tested under a less rigorous response requirement schedule. Under this PR schedule,

the number of nose poke responses required to earn a pellet was increased after each successive reward according to the following response requirement schedule: 1,2,4,6,9,12,15,19,22,26,30,35,40,45,50,55…

Histological verification of cannula placement

Mice were deeply anesthetized with pentobarbital and transcardially perfused with phosphate-buffered saline followed by 10% formalin (Fisher Scientific). Brains were extracted and sliced on a microtome at a thickness of 50µm. Slices were then stained with thionin and examined under a light microscope to verify accurate cannula placements when compared to an atlas of the mouse brain (Franklin and Paxinos, 2007). Mice with inaccurate cannula placements were discarded from the study.

Statistics

All values are expressed as mean±SEM. Statistical differences within groups were assessed using paired 2-tailed Student's t tests and statistical differences between groups were assessed using unpaired 2-tailed Student's t tests. A level of confidence of $P < 0.05$ was employed for statistical significance.

Chapter 4

Conclusions and remaining questions

The work presented in this thesis represents a significant advancement of our understanding of how the precise timing of DA neuron activity contributes to the maintenance of motivated behavior. We trained Th-cre mice on a variety of appetitive operant paradigms then triggered DA neuron inhibition for a brief period based on their behavior in real time. We hypothesized that the behavioral effects produced by DA neuron inhibition would be dependent on the behavior the mouse was engaged in at the time DA neuron inhibition was triggered. Therefore, we determined that a mouse's behavior while performing an appetitive operant task could be divided into three different behavioral epochs. In the first behavioral epoch, the mouse is actively engaged in a bout of operant responding. We called this epoch the "nose poke bout" epoch. In the second behavioral epoch, the mouse is engaged in consuming the pellet reward. We called this the "reward consumption" epoch. In the third behavioral epoch the mouse is not actively engaged in either the operant response or the reward consumption. We called this the "off-task" epoch. We then designed a program to track the mouse's behavior in real time and used optogenetic techniques to systematically inhibit midbrain DA neurons during each of these behavioral epochs.

In chapter 1 of this thesis, we compared the effect of inhibiting DA neurons in either the nose poke bout epoch or the off-task epoch. We found that inhibiting DA neurons during the nose poke bout significantly decreased the probability that mice

would continue the bout of nose pokes, resulting in early termination of the bout. Additionally, we found that inhibiting DA neurons in the off-task epoch significantly decreased the probability of initiating a bout of nose pokes. These results indicate that DA neuron activity is necessary to both initiate and maintain a bout of effortful operant responses. Interestingly, we only observed a decrease in the total number of nose pokes completed in the behavioral session while inhibiting DA neurons during the nose poke bout epoch. This was true even when we delayed operant response initiation over longer periods by inhibiting DA neurons for 3 blocks of 10 minutes each. We therefore conclude that DA neuron activity at the time of performing a bout of effortful operant responses is especially critical in maintaining motivated behavior.

Chapter 2 of this thesis examined the effect of inhibiting DA neurons during the reward consumption epoch. We found that inhibiting DA neurons during reward consumption across 6 days of testing had profound and long-lasting effects on measures of cue-directed and operant behavior. Reward-paired inhibition of DA neurons during a PR task extinguished responding to a reward-predictive cue but did not extinguish the operant response or reduce motivation. Conversely, reward-paired DA neuron inhibition during an FR1 task extinguished both cue-reward and operant-reward associations. Furthermore, these effects were specific to the operant response and cue antecedent to DA neuron inhibition and did not generalize to other learned associations performed in the same behavioral session. Importantly, inhibiting DA neurons at a time unpaired with reward consumption did not produce an effect in any of the behavioral measures examined, indicating that DA neuron activity specifically at the time of reward

consumption is required to maintain learned cue-reward and operant-reward associations. These results demonstrate that inhibition of DA neurons is sufficient to generate an artificial negative reward inhibition signal and drive new learning.

Finally, in chapter 3 of this thesis we demonstrate that the orexigenic peptide, ghrelin, acts in the VTA to increase motivation. We found that administering a ghrelin receptor antagonist directly into the VTA decreased effort on a PR task. However, intra-VTA ghrelin receptor antagonists did not decrease the amount of food consumed in a feeding test, nor did they decrease performance on a low-effort FR1 task. These results demonstrate that ghrelin signaling in the VTA is required to increase motivation in response to food deprivation and its motivation-enhancing effects in the VTA are dissociable from its orexigenic effects in the hypothalamus.

In the work presented here, we examine the requirement for DA neuron activity in maintaining motivated behavior during three distinct behavioral epochs: nose poke responding, off-task behavior, and reward consumption. However, DA neurons also exhibit robust phasic activations in response to reward predictive cues (Schultz et al., 1997; Waelti et al., 2001). Early in learning, DA neurons respond robustly to reward delivery. As learning progresses, phasic DA neuron activity becomes progressively more pronounced in response to the cue while responses to the now fully predicted reward diminish. It is thought that the increased DA neuron activity at the time of cue presentation serves to promote and invigorate the subsequent conditioned response (Flagel et al., 2011; Glimcher, 2011). Therefore, whether DA neuron activity in response to reward-predictive cues is required to maintain goal-directed motivated behavior is an

important question that remains to be addressed. Because of technical limitations in our task design, a direct test of this question was not practical. In our task, the cue was delivered in response to the animal's behavior and was therefore unpredictable, making it difficult to trigger DA neuron inhibition before cue onset. A better method of assessing this question would be in a Pavlovian paradigm where the timing of both the reward-predictive cue and the reward is under the control of the experimenter. With this protocol, the effects of DA neuron inhibition paired either with cue presentation or reward consumption could be directly compared. Additionally, here we show that reward-paired DA neuron inhibition extinguishes continuously reinforced cue-reward associations in several operant behavioral paradigms. However, we did not directly test whether reward-paired DA neuron inhibition could extinguish variably reinforced cue-reward associations. A pure Pavlovian conditioning paradigm would also be optimal for addressing this question.

Another important issue that deserves further exploration is the role of DA neuron inhibition in aversion and fear learning. Recordings from DA neurons have shown that presentation of aversive events or cues that predict aversive outcomes can both excite or inhibition DA neuron activity (Brischoux et al., 2009; Matsumoto and Hikosaka, 2009; Mileykovskiy and Morales, 2011; Mirenowicz and Schultz, 1996). A recent study found that inhibiting DA neurons while mice explored one side of a conditioning chamber resulted in avoidance of the side paired with inhibition during a test 24 hours later (Tan et al., 2012). The authors concluded that these inhibition parameters produced an aversion to the inhibition-paired chamber. However, an

alternative interpretation of these results could be that mice did not feel any subjective sense of aversion or malaise while exploring the side of the chamber paired with DA neuron inhibition, but rather learned to spend more time in the side paired with intact endogenous DA neuron activity. This interpretation is supported by a large body of research that shows that DA receptor antagonists do not decrease subjective measures of "liking" to rewarding foods, nor do they increase subjective measures of "disliking" to aversive tastes (Berridge and Robinson, 1998; Peciña et al., 2003). In addition, we do not observe that mice avoid the location paired with DA neuron inhibition. In fact, in tests examining reward-paired DA neuron inhibition in a PR task we find that mice approach and enter the reward port (which is the location paired with DA neuron inhibition) significantly more frequently across inhibition sessions. In addition, mice consume the reward pellet while DA neurons are inhibited and reward-paired DA neuron inhibition does not decrease the number of nose pokes performed across all 6 testing days, indicating that mice are motivated to earn the reward.

We propose two studies that may help resolve whether DA neuron inhibition produces subjective feeling of aversion. First, recent work from our lab demonstrated that pairing DA neuron activation with consumption of a flavored sucrose did not increase consumption of that flavor in a home-cage preference test (Steinberg et al., 2013). Therefore, we propose replicating this test with DA neuron inhibition. In this test, mice will be given access to two flavors of Kool-Aid flavored sucrose, cherry or tropical punch, in the home cage. Next, mice will be placed in the operant chamber and each flavor will be delivered in the reward port on alternating days, with one flavor paired

with DA neuron inhibition. On the training days when the inhibition-paired flavor is available, inhibition will be triggered either when the reward is available (paired group) or at a time when the reward is not available (unpaired group). After completing eight sessions of flavor training, flavor preference will be assessed in the home cage. If mice consume similar amounts of the inhibition-paired and non-inhibition-paired flavors then we can conclude that our DA neuron inhibition parameters do not decrease the value or preference for rewards.

Additionally, investigating the contribution of DA neuron activity in fear learning may help elucidate the role of DA neuron inhibitions in aversion. Decreases in DA neuron activity and release have been shown to impair fear learning (Fadok et al., 2009; Zweifel et al., 2011), further suggesting a role for DA in acting as a teaching signal rather than informing subjective measures of pleasure or aversion. To explore the requirement for DA neuron activity in fear learning, we propose pairing DA neuron inhibition with the delivery of shock in a classic Pavlovian fear-learning paradigm. Th-cre+ mice will each receive a single training session in which the delivery of shock will be predicted by an auditory stimulus. The shock-paired group will receive DA neuron inhibition paired with the delivery of each shock while the shock-unpaired group will receive DA neuron inhibition explicitly unpaired with the delivery of each shock. 24 hours later, measures of conditioned responding (freezing) to the shock-predictive cue will be assessed in an extinction test in which the cue is presented without the subsequent shock. If DA neuron inhibition acts to increase subjective feelings of malaise, then we predict that mice in the shock-paired group will exhibit increased freezing behavior to the cue.

However, if DA neuron inhibition acts as a negative reward prediction error signal then we expect that the cue-shock association will be impaired in the shock-paired group and mice will exhibit decreased freezing to the cue.

This work sheds also sheds light on the debate as to whether pauses in DA neuron activity are sufficient to encode a negative reward prediction error signal to downstream structures (Daw et al., 2002; Glimcher, 2011). One caveat of the results presented in this study is that by inhibiting DA neurons for 15 seconds, we generated a pause in DA neuron firing much longer than the 0.5-1 second pause normally observed when an expected reward is withheld (Pan et al., 2008; Schultz et al., 1997). We chose a 15 seconds inhibition time based on the time it takes mice to consume the reward pellet, which was approximately 10-15, and we wanted to ensure that DA neurons were inhibited during the entire reward consummatory epoch. Therefore, it remains to be seen whether shorter periods of DA neuron inhibition would result in behavioral effects similar to those observed here. There are two ways such a study could be designed. First, animals could be trained to earn a pellet reward that takes up to 15 seconds to consume, exactly as we have done in these studies. We could then attempt inhibiting DA neurons for just the first few seconds of reward consumption. One study that examined phasic DA release in the nucleus accumbens during reward consumption in a PR task found that DA transients were increased for just the first 3-5 seconds of reward pellet consumption (Wanat et al., 2010). Therefore, it is possible that DA release just during the first phase of reward consumption is critical in signaling reward prediction error. A potential problem with this experiment design is that DA neuron activity could

potentially rebound after cessation of inhibition if the animal is still engaged in reward consumption. Therefore, a second strategy could involve simply switching to delivery of a liquid sucrose reward. This way, the experimenter could explicitly control the quantity of the reward is delivered so the timing of DA neuron inhibition and reward consumption could be matched. It is worth noting, however, that for the experiments presented in this thesis, the use of the pellet reward offered an important advantage. One recent study found that activation of VTA GABA neurons resulted in both direct inhibition of VTA DA neurons and also disrupted reward consummatory behavior (van Zessen et al., 2012b). Therefore, the use of a discrete reward pellet offered a significant advantage in that we were able to quantify the exact number of pellets consumed in each behavioral session. We found that DA neuron inhibition did not decrease the percentage of rewards consumed across testing days.

References

Aberman, J.E., and Salamone, J.D. (1999). Nucleus accumbens dopamine depletions make rats more sensitive to high ratio requirements but do not impair primary food reinforcement. Neuroscience *92*, 545–552.

Aberman, J.E., Ward, S.J., and Salamone, J.D. (1998). Effects of dopamine antagonists and accumbens dopamine depletions on time-constrained progressive-ratio performance. Pharmacol Biochem Behav *61*, 341–348.

Abizaid, A., Liu, Z.-W., Andrews, Z.B., Shanabrough, M., Borok, E., Elsworth, J.D., Roth, R.H., Sleeman, M.W., Picciotto, M.R., Tschöp, M.H., et al. (2006). Ghrelin modulates the activity and synaptic input organization of midbrain dopamine neurons while promoting appetite. J Clin Invest *116*, 3229–3239.

Adamantidis, A.R., Tsai, H.-C., Boutrel, B., Zhang, F., Stuber, G.D., Budygin, E.A., Touriño, C., Bonci, A., Deisseroth, K., and de Lecea, L. (2011). Optogenetic interrogation of dopaminergic modulation of the multiple phases of reward-seeking behavior. J Neurosci *31*, 10829–10835.

Bassareo, V., and Di Chiara, G. (1999). Modulation of feeding-induced activation of mesolimbic dopamine transmission by appetitive stimuli and its relation to motivational state. Eur J Neurosci *11*, 4389–4397.

Bautista, L.M., Tinbergen, J., and Kacelnik, A. (2001). To walk or to fly? How birds choose

among foraging modes. Proc Natl Acad Sci USA *98*, 1089–1094.

BENJAMINI, Y., and HOCHBERG, Y. (1995). Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B-Methodological *57*, 289–300.

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

Berridge, K.C., Venier, I.L., and Robinson, T.E. (1989). Taste reactivity analysis of 6-hydroxydopamine-induced aphagia: implications for arousal and anhedonia hypotheses of dopamine function. Behavioral Neuroscience *103*, 36–45.

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

Berridge, K.C. (2012). From prediction error to incentive salience: mesolimbic computation of reward motivation. Eur J Neurosci *35*, 1124–1143.

Borgland, S.L., Chang, S.-J., Bowers, M.S., Thompson, J.L., Vittoz, N., Floresco, S.B., Chou, J., Chen, B.T., and Bonci, A. (2009). Orexin A/hypocretin-1 selectively promotes motivation for positive reinforcers. J Neurosci *29*, 11215–11225.

Borgland, S.L., Taha, S.A., Sarti, F., Fields, H.L., and Bonci, A. (2006). Orexin A in the VTA is critical for the induction of synaptic plasticity and behavioral sensitization to

cocaine. Neuron *49*, 589–601.

Bouton, M.E., and Moody, E.W. (2004). Memory processes in classical conditioning. Neuroscience and Biobehavioral Reviews *28*, 663–674.

Brischoux, F., Chakraborty, S., Brierley, D.I., and Ungless, M.A. (2009). Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli. Proc Natl Acad Sci USA *106*, 4894–4899.

BUSH, R.R., and MOSTELLER, F. (1951). A mathematical model for simple learning. Psychol Rev *58*, 313–323.

Cagniard, B., Beeler, J.A., Britt, J.P., McGehee, D.S., Marinelli, M., and Zhuang, X. (2006). Dopamine scales performance in the absence of new learning. Neuron *51*, 541–547.

Cannon, C.M., and Palmiter, R.D. (2003). Reward without dopamine. J Neurosci *23*, 10827–10831.

Carr, K.D., Kim, G.Y., and Cabeza de Vaca, S. (2001). Rewarding and locomotor-activating effects of direct dopamine receptor agonists are augmented by chronic food restriction in rats. Psychopharmacology (Berl) *154*, 420–428.

Carr, K.D. (2002). Augmentation of drug reward by chronic food restriction: behavioral evidence and underlying mechanisms. Physiol Behav *76*, 353–364.

Chang, J.Y., Sawyer, S.F., Lee, R.S., and Woodward, D.J. (1994). Electrophysiological and

pharmacological evidence for the role of the nucleus accumbens in cocaine self-administration in freely moving rats. J Neurosci *14*, 1224–1244.

Chen, B.T., Yau, H.-J., Hatch, C., Kusumoto-Yoshida, I., Cho, S.L., Hopf, F.W., and Bonci, A. (2013). Rescuing cocaine-induced prefrontal cortex hypoactivity prevents compulsive cocaine seeking. Nature *496*, 359–362.

Cohen, J.Y., Haesler, S., Vong, L., Lowell, B.B., and Uchida, N. (2012). Neuron-type-specific signals for reward and punishment in the ventral tegmental area. Nature.

Coizet, V., Dommett, E.J., Klop, E.M., Redgrave, P., and Overton, P.G. (2010). The parabrachial nucleus is a critical link in the transmission of short latency nociceptive information to midbrain dopaminergic neurons. Neuroscience *168*, 263–272.

Conrad, L.C., and Pfaff, D.W. (1976). Autoradiographic tracing of nucleus accumbens efferents in the rat. Brain Res *113*, 589–596.

Correa, M., Carlson, B.B., Wisniecki, A., and Salamone, J.D. (2002). Nucleus accumbens dopamine and work requirements on interval schedules. Behav Brain Res *137*, 179–187.

Cousins, M.S., and Salamone, J.D. (1994). Nucleus accumbens dopamine depletions in rats affect relative response allocation in a novel cost/benefit procedure. Pharmacol Biochem Behav *49*, 85–91.

Daw, N.D., Kakade, S., and Dayan, P. (2002). Opponent interactions between serotonin and dopamine. Neural Networks : the Official Journal of the International Neural Network Society *15*, 603–616.

Dayan, P., and Niv, Y. (2008). Reinforcement learning: the good, the bad and the ugly. Curr Opin Neurobiol *18*, 185–196.

Di Ciano, P., Cardinal, R.N., Cowell, R.A., Little, S.J., and Everitt, B.J. (2001). Differential involvement of NMDA, AMPA/kainate, and dopamine receptors in the nucleus accumbens core in the acquisition and performance of pavlovian approach behavior. J Neurosci *21*, 9471–9477.

Diano, S., Farr, S.A., Benoit, S.C., McNay, E.C., da Silva, I., Horvath, B., Gaskin, F.S., Nonaka, N., Jaeger, L.B., Banks, W.A., et al. (2006). Ghrelin controls hippocampal spine synapse density and memory performance. Nat Neurosci *9*, 381–388.

Egecioglu, E., Jerlhag, E., Salomé, N., Skibicka, K.P., Haage, D., Bohlooly-Y, M., Andersson, D., Bjursell, M., Perrissoud, D., Engel, J.A., et al. (2010). Ghrelin increases intake of rewarding food in rodents. Addict Biol *15*, 304–311.

Fadok, J.P., Dickerson, T.M.K., and Palmiter, R.D. (2009). Dopamine is necessary for cue-dependent fear conditioning. J Neurosci *29*, 11089–11097.

Feneberg, A., and Malfertheiner, P. (2012). Epidemic trends of obesity with impact on metabolism and digestive diseases. Dig Dis *30*, 143–147.

Fields, H.L., Hjelmstad, G.O., Margolis, E.B., and Nicola, S.M. (2007). Ventral tegmental area neurons in learned appetitive behavior and positive reinforcement. Annu Rev Neurosci *30*, 289–316.

Flagel, S.B., Clark, J.J., Robinson, T.E., Mayo, L., Czuj, A., Willuhn, I., Akers, C.A., Clinton, S.M., Phillips, P.E.M., and Akil, H. (2011). A selective role for dopamine in stimulus-reward learning. Nature *469*, 53–57.

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

Franklin, K., and Paxinos, G. (2007). The mouse brain in stereotaxic coordinates (3rd Ed.) with CD-ROM FRANKLIN Keith B., PAXINOS George: Librairie Lavoisier.

Fuxe, K., Dahlström, A.B., Jonsson, G., Marcellino, D., Guescini, M., Dam, M., Manger, P., and Agnati, L. (2010). The discovery of central monoamine neurons gave volume transmission to the wired brain. Prog Neurobiol *90*, 82–100.

Geisler, S., and Zahm, D.S. (2005). Afferents of the ventral tegmental area in the rat-anatomical substratum for integrative functions. J Comp Neurol *490*, 270–294.

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

Gonon, F.G. (1988). Nonlinear relationship between impulse flow and dopamine released by rat midbrain dopaminergic neurons as studied by in vivo electrochemistry. Neuroscience *24*, 19–28.

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*, 1–24.

Grace, A.A., and Bunney, B.S. (1984). The control of firing pattern in nigral dopamine neurons: single spike firing. J Neurosci *4*, 2866–2876.

Grace, A.A., and Onn, S.P. (1989). Morphology and electrophysiological properties of immunocytochemically identified rat dopamine neurons recorded in vitro. J Neurosci *9*, 3463–3481.

Haber, S.N., and Fudge, J.L. (1997). The primate substantia nigra and VTA: integrative circuitry and function. Crit Rev Neurobiol *11*, 323–342.

Hamilton, M.E., Redondo, J.L., and Freeman, A.S. (2000). Overflow of dopamine and cholecystokinin in rat nucleus accumbens in response to acute drug administration. Synapse *38*, 238–242.

Harris, G.C., Wimmer, M., and Aston-Jones, G. (2005). A role for lateral hypothalamic orexin neurons in reward seeking. Nature *437*, 556–559.

Haselgrove, M., Aydin, A., and Pearce, J.M. (2004). A Partial Reinforcement Extinction

Effect Despite Equal Rates of Reinforcement During Pavlovian Conditioning. Journal of Experimental Psychology: Animal Behavior Processes *30*, 240–250.

Hnasko, T.S., Chuhma, N., Zhang, H., Goh, G.Y., Sulzer, D., Palmiter, R.D., Rayport, S., and Edwards, R.H. (2010). Vesicular glutamate transport promotes dopamine storage and glutamate corelease in vivo. Neuron *65*, 643–656.

Hnasko, T.S., Hjelmstad, G.O., Fields, H.L., and Edwards, R.H. (2012). Ventral tegmental area glutamate neurons: electrophysiological properties and projections. J Neurosci *32*, 15076–15085.

HODOS, W. (1961). Progressive ratio as a measure of reward strength. Science *134*, 943–944.

Hollerman, J.R., and Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. Nat Neurosci *1*, 304–309.

Holtzman-Assif, O., Laurent, V., and Westbrook, R.F. (2010). Blockade of dopamine activity in the nucleus accumbens impairs learning extinction of conditioned fear. Learn Mem *17*, 71–75.

Ikemoto, S., and Panksepp, J. (1996). Dissociations between appetitive and consummatory responses by pharmacological manipulations of reward-relevant brain regions. Behavioral Neuroscience *110*, 331–345.

Ikemoto, S., and Panksepp, J. (1999). The role of nucleus accumbens dopamine in

motivated behavior: a unifying interpretation with special reference to reward-seeking. Brain Res Brain Res Rev *31*, 6–41.

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

Jin, X., and Costa, R.M. (2010). Start/stop signals emerge in nigrostriatal circuits during sequence learning. Nature *466*, 457–462.

Kenny, P.J. (2011). Common cellular and molecular mechanisms in obesity and drug addiction. Nat Rev Neurosci *12*, 638–651.

King, S.J., Isaacs, A.M., O'Farrell, E., and Abizaid, A. (2011). Motivation to obtain preferred foods is enhanced by ghrelin in the ventral tegmental area. Horm Behav *60*, 572–580.

Koch, M., Schmid, A., and Schnitzler, H.U. (2000). Role of muscles accumbens dopamine D1 and D2 receptors in instrumental and Pavlovian paradigms of conditioned reward. Psychopharmacology (Berl) *152*, 67–73.

Kojima, M., and Kangawa, K. (2005). Ghrelin: structure and function. Physiol Rev *85*, 495–522.

Malik, S., McGlone, F., Bedrossian, D., and Dagher, A. (2008). Ghrelin modulates brain

activity in areas that control appetitive behavior. Cell Metab *7*, 400–409.

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

Mileykovskiy, B., and Morales, M. (2011). Duration of inhibition of ventral tegmental area dopamine neurons encodes a level of conditioned fear. J Neurosci *31*, 7471–7476.

Mingote, S., Weber, S.M., Ishiwari, K., Correa, M., and Salamone, J.D. (2005). Ratio and time requirements on operant schedules: effort-related effects of nucleus accumbens dopamine depletions. Eur J Neurosci *21*, 1749–1757.

Mirenowicz, J., and Schultz, W. (1996). Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. Nature *379*, 449–451.

Myers, K.M., and Davis, M. (2002). Behavioral and neural analysis of extinction. Neuron *36*, 567–584.

Nakazato, M., Murakami, N., Date, Y., Kojima, M., Matsuo, H., Kangawa, K., and Matsukura, S. (2001). A role for ghrelin in the central regulation of feeding. Nature *409*, 194–198.

Naleid, A.M., Grace, M.K., Cummings, D.E., and Levine, A.S. (2005). Ghrelin induces feeding in the mesolimbic reward pathway between the ventral tegmental area and the nucleus accumbens. Peptides *26*, 2274–2279.

Nicola, S.M. (2010). The flexible approach hypothesis: unification of effort and cue-responding hypotheses for the role of nucleus accumbens dopamine in the activation of reward-seeking behavior. J Neurosci *30*, 16585–16600.

Nowend, K.L., Arizzi, M., Carlson, B.B., and Salamone, J.D. (2001). D1 or D2 antagonism in nucleus accumbens core or dorsomedial shell suppresses lever pressing for food but leads to compensatory increases in chow consumption. Pharmacol Biochem Behav *69*, 373–382.

OLDS, J., and MILNER, P. (1954). Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. J Comp Physiol Psychol *47*, 419–427.

Ostlund, S.B., Kosheleff, A.R., and Maidment, N.T. (2012). Relative Response Cost Determines the Sensitivity of Instrumental Reward Seeking to Dopamine Receptor Blockade. Neuropsychopharmacology.

Pan, W.-X., Brown, J., and Dudman, J.T. (2013). Neural signals of extinction in the inhibitory microcircuit of the ventral midbrain. Nat Neurosci *16*, 71–78.

Pan, W.-X., Schmidt, R., Wickens, J.R., and Hyland, B.I. (2008). Tripartite mechanism of extinction suggested by dopamine neuron activity and temporal difference model. J Neurosci *28*, 9619–9631.

Parkinson, J.A., Dalley, J.W., Cardinal, R.N., Bamford, A., Fehnert, B., Lachenal, G., Rudarakanchana, N., Halkerston, K.M., Robbins, T.W., and Everitt, B.J. (2002).

Nucleus accumbens dopamine depletion impairs both acquisition and performance of appetitive Pavlovian approach behaviour: implications for mesoaccumbens dopamine function. Behav Brain Res *137*, 149–163.

Pavlov, I.P., and Anrep, G.V. (1927). Conditioned reflexes - Ivan Petrovich. Pavlov - Google Books.

Peciña, S., Cagniard, B., Berridge, K.C., Aldridge, J.W., and Zhuang, X. (2003). Hyperdopaminergic mutant mice have higher "wanting" but not “liking” for sweet rewards. J Neurosci *23*, 9395–9402.

Perello, M., Sakata, I., Birnbaum, S., Chuang, J.-C., Osborne-Lawrence, S., Rovinsky, S.A., Woloszyn, J., Yanagisawa, M., Lutter, M., and Zigman, J.M. (2009). Ghrelin Increases the Rewarding Value of High-Fat Diet in an Orexin-Dependent Manner. Biol Psychiatry.

Peyron, C., Tighe, D.K., van den Pol, A.N., de Lecea, L., Heller, H.C., Sutcliffe, J.G., and Kilduff, T.S. (1998). Neurons containing hypocretin (orexin) project to multiple neuronal systems. J Neurosci *18*, 9996–10015.

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

Phillips, P.E.M., Walton, M.E., and Jhou, T.C. (2007). Calculating utility: preclinical evidence for cost-benefit analysis by mesolimbic dopamine. Psychopharmacology (Berl) *191*, 483–495.

Popkin, B.M., Adair, L.S., and Ng, S.W. (2012). Global nutrition transition and the pandemic of obesity in developing countries. Nutr. Rev. *70*, 3–21.

Rada, P., Avena, N.M., and Hoebel, B.G. (2005). Daily bingeing on sugar repeatedly releases dopamine in the accumbens shell. Neuroscience *134*, 737–744.

Redgrave, P., and Gurney, K. (2006). The short-latency dopamine signal: a role in discovering novel actions? Nat Rev Neurosci *7*, 967–975.

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

Robbins, T.W., and Everitt, B.J. (2007). A role for mesencephalic dopamine in activation: commentary on Berridge (2006). Psychopharmacology (Berl) *191*, 433–437.

Romo, R., and Schultz, W. (1990). Dopamine neurons of the monkey midbrain: contingencies of responses to active touch during self-initiated arm movements. J Neurophysiol *63*, 592–606.

Salamone, J.D., Correa, M., Farrar, A., and Mingote, S.M. (2007). Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. Psychopharmacology (Berl) *191*, 461–482.

Salamone, J.D., Cousins, M.S., and Bucher, S. (1994). Anhedonia or anergia? Effects of haloperidol and nucleus accumbens dopamine depletion on instrumental response

selection in a T-maze cost/benefit procedure. Behav Brain Res *65*, 221–229.

Salamone, J.D., Cousins, M.S., and Snyder, B.J. (1997). Behavioral functions of nucleus accumbens dopamine: empirical and conceptual problems with the anhedonia hypothesis. Neuroscience and Biobehavioral Reviews *21*, 341–359.

Salamone, J.D., Mahan, K., and Rogers, S. (1993). Ventrolateral striatal dopamine depletions impair feeding and food handling in rats. Pharmacol Biochem Behav *44*, 605–610.

Salamone, J.D., and Correa, M. (2012). The mysterious motivational functions of mesolimbic dopamine. Neuron *76*, 470–485.

Satoh, T., Nakai, S., Sato, T., and Kimura, M. (2003). Correlated coding of motivation and outcome of decision by dopamine neurons. J Neurosci *23*, 9913–9923.

Saunders, B.T., and Robinson, T.E. (2012). The role of dopamine in the accumbens core in the expression of Pavlovian-conditioned responses. Eur J Neurosci *36*, 2521–2532.

Schultz, W., and Dickinson, A. (2000). Neuronal coding of prediction errors. Annu Rev Neurosci *23*, 473–500.

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

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

Neurosci *30*, 259–288.

Seroogy, K.B., and Gall, C.M. (1993). Expression of neurotrophins by midbrain dopaminergic neurons. Exp Neurol *124*, 119–128.

Seroogy, K., Schalling, M., Brené, S., Dagerlind, A., Chai, S.Y., Hökfelt, T., Persson, H., Brownstein, M., Huan, R., and Dixon, J. (1989). Cholecystokinin and tyrosine hydroxylase messenger RNAs in neurons of rat mesencephalon: peptide/monoamine coexistence studies using in situ hybridization combined with immunocytochemistry. Exp Brain Res *74*, 149–162.

Skibicka, K.P., Hansson, C., Alvarez-Crespo, M., Friberg, P.A., and Dickson, S.L. (2011). Ghrelin directly targets the ventral tegmental area to increase food motivation. Neuroscience.

Sombers, L.A., Beyene, M., Carelli, R.M., and Wightman, R.M. (2009). Synaptic overflow of dopamine in the nucleus accumbens arises from neuronal activity in the ventral tegmental area. J Neurosci *29*, 1735–1742.

Steinberg, E.E., Keiflin, R., Boivin, J.R., Witten, I.B., Deisseroth, K., Janak, P.H. (2013). A causal link between prediction errors, dopamine neurons and learning. Nat Neurosci. In press.

Stevens, J.R., Rosati, A.G., Ross, K.R., and Hauser, M.D. (2005). Will travel for food: spatial discounting in two new world monkeys. Curr. Biol. *15*, 1855–1860.

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

Stuber, G.D., Klanker, M., de Ridder, B., Bowers, M.S., Joosten, R.N., Feenstra, M.G., and Bonci, A. (2008). Reward-predictive cues enhance excitatory synaptic strength onto midbrain dopamine neurons. Science *321*, 1690–1692.

Stuber, G.D., Roitman, M.F., Phillips, P.E.M., Carelli, R.M., and Wightman, R.M. (2005a). Rapid dopamine signaling in the nucleus accumbens during contingent and noncontingent cocaine administration. Neuropsychopharmacology *30*, 853–863.

Stuber, G.D., Wightman, R.M., and Carelli, R.M. (2005b). Extinction of cocaine self-administration reveals functionally and temporally distinct dopaminergic signals in the nucleus accumbens. Neuron *46*, 661–669.

Sutton, R.S., and Barto, A.G. (1998). Reinforcement Learning: An Introduction (Adaptive Computation and Machine Learning). (01 March 1998).

Szczypka, M.S., Kwok, K., Brot, M.D., Marck, B.T., Matsumoto, A.M., Donahue, B.A., and Palmiter, R.D. (2001). Dopamine production in the caudate putamen restores feeding in dopamine-deficient mice. Neuron *30*, 819–828.

Tamura, H., Kamegai, J., Shimizu, T., Ishii, S., Sugihara, H., and Oikawa, S. (2002). Ghrelin stimulates GH but not food intake in arcuate nucleus ablated rats. Endocrinology *143*, 3268–3275.

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

Thorpe, A.J., Cleary, J.P., Levine, A.S., and Kotz, C.M. (2005). Centrally administered orexin A increases motivation for sweet pellets in rats. Psychopharmacology (Berl) *182*, 75–83.

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

Tye, K.M., Cone, J.J., Schairer, W.W., and Janak, P.H. (2010). Amygdala neural encoding of the absence of reward during extinction. J Neurosci *30*, 116–125.

Tye, K.M., Mirzabekov, J.J., Warden, M.R., Ferenczi, E.A., Tsai, H.-C., Finkelstein, J., Kim, S.-Y., Adhikari, A., Thompson, K.R., Andalman, A.S., et al. (2013). Dopamine neurons modulate neural encoding and expression of depression-related behaviour. Nature *493*, 537–541.

van Zessen, R., van der Plasse, G., and Adan, R.A.H. (2012a). Contribution of the mesolimbic dopamine system in mediating the effects of leptin and ghrelin on feeding. Proc Nutr Soc *71*, 435–445.

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

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

Wanat, M.J., Kuhnen, C.M., and Phillips, P.E.M. (2010). Delays conferred by escalating costs modulate dopamine release to rewards but not their predictors. J Neurosci *30*, 12020–12027.

Wassum, K.M., Ostlund, S.B., and Maidment, N.T. (2012). Phasic mesolimbic dopamine signaling precedes and predicts performance of a self-initiated action sequence task. Biol Psychiatry *71*, 846–854.

Wise, R.A., Spindler, J., deWit, H., and Gerberg, G.J. (1978). Neuroleptic-induced "anhedonia" in rats: pimozide blocks reward quality of food. Science *201*, 262–264.

Wise, R.A. (2004). Dopamine, learning and motivation. Nat Rev Neurosci *5*, 483–494.

Wise, R.A. (2005). Forebrain substrates of reward and motivation. J Comp Neurol *493*, 115–121.

Wise, R.A. (2006). Role of brain dopamine in food reward and reinforcement. Philos Trans R Soc Lond, B, Biol Sci *361*, 1149–1158.

Witten, I.B., Steinberg, E.E., Lee, S.Y., Davidson, T.J., Zalocusky, K.A., Brodsky, M., Yizhar, O., Cho, S.L., Gong, S., Ramakrishnan, C., et al. (2011). Recombinase-driver rat lines: tools, techniques, and optogenetic application to dopamine-mediated

reinforcement. Neuron *72*, 721–733.

Yun, I.A., Nicola, S.M., and Fields, H.L. (2004a). Contrasting effects of dopamine and glutamate receptor antagonist injection in the nucleus accumbens suggest a neural mechanism underlying cue-evoked goal-directed behavior. Eur J Neurosci *20*, 249–263.

Yun, I.A., Wakabayashi, K.T., Fields, H.L., and Nicola, S.M. (2004b). The ventral tegmental area is required for the behavioral and nucleus accumbens neuronal firing responses to incentive cues. J Neurosci *24*, 2923–2933.

Zweifel, L.S., Fadok, J.P., Argilli, E., Garelick, M.G., Jones, G.L., Dickerson, T.M.K., Allen, J.M., Mizumori, S.J.Y., Bonci, A., and Palmiter, R.D. (2011). Activation of dopamine neurons is critical for aversive conditioning and prevention of generalized anxiety. Nat Neurosci *14*, 620–626.

Zweifel, L.S., Parker, J.G., Lobb, C.J., Rainwater, A., Wall, V.Z., Fadok, J.P., Darvas, M., Kim, M.J., Mizumori, S.J.Y., Paladini, C.A., et al. (2009). Disruption of NMDAR-dependent burst firing by dopamine neurons provides selective assessment of phasic dopamine-dependent behavior. Proc Natl Acad Sci USA *106*, 7281–7288.

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

Please sign the following statement:

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

Author Signature

6/12/13

Date