Ventral Tegmental Area Neuronal Ensembles Accurately Encode Action Number

by

Jesse Wood

B.S. Psychology, Virginia Commonwealth University, 2007

Submitted to the Graduate Faculty of the

Kenneth P. Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2015

UNIVERSITY OF PITTSBURGH

Dietrich School of Arts and Sciences

This dissertation was presented

by

Jesse Wood

It was defended on

November 06, 2014

and approved by

Dr. Hagai Bergman, Professor, Hebrew University, Department of Physiology

Dr. Marlene Cohen, Assistant Professor, University of Pittsburgh, Department of Neuroscience

Dr. Susan Sesack, Professor, University of Pittsburgh, Department of Neuroscience

Dr. Robert E. Kass, Professor, Carnegie Mellon University, Department of Statistics

Dr. Robert Turner, Associate Professor, University of Pittsburgh, Department of Neurobiology

Dissertation Advisor: Dr. Bita Moghaddam, Professor, University of Pittsburgh, Department of Neuroscience

Ventral Tegmental Area Neuronal Ensembles Accurately Encode Action Number

Jesse Wood, Ph.D.

The University of Pittsburgh, 2015

Ventral Tegmental Area Neuronal Ensembles Accurately Encode Action Number

Jesse Wood, PhD

University of Pittsburgh, 2015

Real-time updates to behavioral strategy require animals to understand how many actions have been executed toward completion of a goal. These operations are essential for optimizing behavior and have been linked to dopaminergic innervation of prefrontal cortex networks (Gallistel & Gibbon, 2000; Allman *et al.*, 2011; Lustig, 2011). It is an open question how networks of dopaminergic and non-dopaminergic neurons in the ventral tegmental area (VTA) encode information when multiple or complex behaviors are required to earn rewards (Niv *et al.*, 2006; Dayan & Niv, 2008; Niv & Schoenbaum, 2008). Most electrophysiological studies have focused on the averaged activity of dopamine neurons during reward prediction error signaling in simple behavioral paradigms. Thus, VTA neuronal correlates of executive processes and complex behavior remain elusive.

In the current experiment, rats learned to repetitively execute actions (nose pokes) to receive rewarding outcomes (sugar pellets). These actions were randomly rewarded, and all actions were identically valued because each was equally likely to be reinforced. Actions

differed only by their number within a trial. While animals executed serial actions, many VTA neurons were activated or suppressed by unique subsets of actions within a trial. Some neurons fired preferentially during low numbered actions while others preferred high numbered actions. A population averaging approach, which is conventionally used for analysis of dopaminergic neuronal activity, offered poor decoding of action number. In contrast, action number within a trial was accurately decoded from the entire pool of unique activity patterns, considering each neuron independently. These results suggest that the collective activity of VTA neuronal ensembles signals real-time information about ongoing action number—a critical component of behavioral organization.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

It is truly humbling to reflect upon all of the help that I have received while completing my doctoral studies. I would like to thank my mother, Suzanne, father, John, and brother, Michael. My parents both struggled to provide a home for me when I was growing up. I never wanted for a single thing that I needed. As an adult, I realize that we had so few resources that this could not have happened without the tireless work, planning, and resourcefulness of parents that sacrificed a great deal to raise two children. I thank them from the bottom of my heart. As I write this, my brother is in his first year of study in a chemistry graduate program. Few people in my life have been more interested to hear my thoughts on the process of being a scientist. I thank him for the interest in my work. I hope it helped. I would also like to thank my grandfather Thomas, who lent unyielding support to me until his last days. I am the first person in my family to graduate from college or pursue an advanced degree. None of that would have been possible without the sacrifices my family made or the support they lent me.

My friends back home in Richmond, Virginia supported me a great deal throughout this process as well. I cannot mention them all here, but I would like to thank two of my closest friends, Paul McLeod and Julie McLeod, in particular. They have never failed to remind me of how proud they are that I have done this and their support means a great deal to me.

I would like to thank my friends and loved ones in Pittsburgh, both in the graduate program and outside of it. I was lucky to have befriended a number of students, which were a great source of advice over the years. Dr. Kate Cosgrove, Warren Anderson, Dr. Layla Banihashemi, Dr. Shawn Kotermanski, Mary Patton, Dr. Ericka Holmstrand, Dr. Deanne Buffalari, Dr. Sam Ewing, and Dr. Kristin Stedenfeld, were all great friends throughout this process, as were my former labmates Dr. Adria Martig, Dr. Nelson Totah, and Dr. David Sturman. My classmates Dr. Chris Divito and Dr. Erin Kirschman also deserve special thanks for their help and support. I would also like to thank Dr. Joseph Porter for helping me get started in research, while I attended Virginia Commonwealth University. I'm very appreciative of Dr. Charles Bradberry for his interest in my work and the interesting discussions about his work. I would also like to thank Dr. Hank Jedema for advice, the many discussions, and interest in my work. I owe a special thanks to the extremely talented departmental electronics specialist Jim Burhman. Jim's excellent technical support over the years was instrumental in allowing me to perform these experiments. He also taught me a tremendous amount about electronics and measurement hardware, for which I am very grateful. I truly appreciate the friendship of Thom and Katherine Gulish, Brett and Sophie Zoric, and Nichole Merritt. I especially appreciate the love and support of Maryellen Baverso, who was a constant source of encouragement while finishing this dissertation.

I would like to thank the talented group of people in the laboratory of Dr. Bita Moghaddam for their help, support, and encouragement: Meredyth Wegener, Sima Lotfi, Junchol Park, Susie Sonnenschein, and Sweyta Lohani, as well as former lab members Dan Mandell and Alicia Lisowitz, and visiting scholars Alberto Del Arco and Lezio Beuno. I would also like to thank Dr. Alan Pehrson who I worked with in the lab of Dr. Joseph Porter at Virginia

Commonwealth University and in Dr. Bita Moghaddam's lab. Alan was instrumental in fostering my interest in science and pursuit of a doctorate. Dr. Yunbok Kim and I worked together closely for six years. He taught me a great deal about electrophysiology, and I am especially grateful for his patience and the time that he sacrificed to do this. I would also like to thank Dr. Nicholas W. Simon for all of his help throughout my graduate studies. He has been a constant source of advice and help in all aspects of this profession. He has also been a great friend for many years. Finally, I'd like to thank Dr. Marguerite Matthews. Marguerite was my first friend in Pittsburgh and we worked together from August of 2007, until she completed her doctoral studies in July of 2012. Marguerite was a constant source of entertainment, encouragement, reason, and advice. I am forever grateful that we had the chance to work together for so many years.

I would like to thank my committee for all of their hard work, patience, and advice throughout this process. I thank Dr. Robert Turner for serving as an excellent committee chair, and interest in my work dating back to the beginning of my graduate studies. I met Rob in 2007, when was leading my first graduate school journal club. He has since served as a member of several committees, and met with me numerous times to provide technical advice. I also thank Dr. Marlene Cohen for her excellent advice on some of the analyses and interpretations in this dissertation. I am certain that her advice and comments improved the quality of this work. I first interacted with Dr. Robert Kass early in my graduate studies, when he agreed to instruct me in an independent studies class I had designed. Our interactions continued over the years and Rob Kass lent his technical support to analyses that are very important to this work. I would like to thank Rob for teaching me a great deal over the years, as well as helping with this project. Dr. Susan Sesack has been an important part of my graduate studies dating back to my first year in the program. She has served on numerous committees and written letters of recommendation that

were greatly appreciated. In addition to double checking my assessment of many electrode placements in this dataset, I thank her for the numerous discussions and valuable insights to neuroanatomy over the years. Many things we talked about are scattered throughout the introduction and discussion sections of this document. Finally, I would like to Dr. Hagai Bergman for graciously agreeing to serve as my outside examiner. Dr. Bergman's work made a strong impression upon me very early in my graduate studies, and I was thrilled to learn that he would lend his expertise to this process.

Finally, I would like to thank my advisor Dr. Bita Moghaddam. She has taught me a great deal, and I am sure that I will learn a great deal from her for many years to come. It goes without saying that I could not have done this work without her financial, intellectual, or scientific support. She allowed me to attend many conferences, granted me a great deal of intellectual freedom, and helped me to fulfill a number of unique roles in her research group over the years. I am very grateful for these opportunities. However, if there is one thing that I am most grateful it is this: Bita always encouraged me to tackle the biggest problems for which the tools were ready, and I was the guy to address them. Bita always emphasized to me that creativity is the most important part of this process, and pushed me to consider more important ideas and problems. As much as I appreciate the support she gave me, I appreciate the times she challenged me more.

1.0 INTRODUCTION

There are several aspects of ventral tegmental area (VTA) function that future research can address. Electrophysiological recordings of the firing patterns of dopaminergic VTA neurons have taught us a great deal about how these neurons encode reward-related information. In contrast, neurochemical, lesion, and pharmacological studies have focused on the role of dopamine in stress, locomotion, motivation, and executive functions (Salamone & Correa, 2002; Seamans & Yang, 2004; Wise, 2004; Berridge, 2007; Robbins & Arnsten, 2009). It is an open question how these fields can be unified into a comprehensive view of dopamine function. Further, while a number of studies have identified and recorded from non-dopaminergic VTA neurons during behavior (Steffensen *et al.*, 1998; Lee *et al.*, 2001; Steffensen *et al.*, 2001; Kim *et al.*, 2010; Cohen *et al.*, 2012; Kim *et al.*, 2012), the non-dopaminergic population of the VTA is less frequently studied than the dopaminergic system. There is not yet a cohesive framework for understanding what information these neurons encode (Sesack & Grace, 2010; Creed *et al.*, 2014).

This thesis addresses a topic with major relevance to our knowledge of the VTA: how VTA neurons encode real-time information about an ongoing series of behaviors performed to receive rewarding outcomes. Special attention has been paid to considering these data within both reward prediction error and behavioral organization frameworks. Both dopaminergic and non-dopaminergic neurons were recorded in these experiments. Neurons were not prescreened

for any particular firing correlates, so that a more complete representation of the activity of VTA neurons could be derived (Fiorillo *et al.*, 2003; Tobler *et al.*, 2005). Network-wide patterns of information processing are not often considered in VTA recordings, but the collective activity of neuronal ensembles is a point of focus in this dissertation. These experimental and interpretational approaches may render a more unified approach to understanding the current findings.

NEUROANATOMY OF THE VENTRAL TEGMENTAL AREA

The VTA is a complex neuronal network with unique anatomical features. The region integrates input from diverse sources, and broadcasts a multi-neurotransmitter signal to the telencephalon, diencephalon, and hindbrain (Thierry *et al.*, 1973; Swanson, 1982; Goldman-Rakic, 1998; Carr & Sesack, 2000b). The VTA contains mostly dopamine neurons (approximately 60%) and γ-amminobutyric acid (GABA, approximately 30%) neurons, but glutamate neurons and dopamine neurons which co-release GABA and glutamate are also present (Swanson, 1982; Carr & Sesack, 2000b; Yamaguchi *et al.*, 2007; Nair-Roberts *et al.*, 2008; Sesack & Grace, 2010; Stuber *et al.*, 2010; Tritsch *et al.*, 2012; Tritsch *et al.*, 2014). The region receives sparse, intermingled inputs from a continuous band of cells that stretches from the prefrontal cortex to the medulla oblongata (Ramon-Moliner & Nauta, 1966; Geisler & Zahm, 2005; Sesack & Grace, 2010; Watabe-Uchida *et al.*, 2012). Thus, VTA receives projections from sensory, motor, and associative regions, and possesses an enormous capacity for integrating information from throughout the brain (Geisler & Zahm, 2005; Omelchenko & Sesack, 2007). Glutamatergic interneurons, GABAergic interneurons, and electrical connections are also present and may contribute to local information

processing (Vandecasteele *et al.*, 2005; Allison *et al.*, 2006; Lassen *et al.*, 2007; Vandecasteele *et al.*, 2008; Omelchenko & Sesack, 2009; Sesack & Grace, 2010; Steffensen *et al.*, 2011). These patterns of connectivity are critical to understanding the function of the region, and may emphasize the need to approach the VTA from a new perspective, one based upon ensembles or network based information processing (Geisler & Zahm, 2005).

FUNCTION OF THE VTA DOPAMINE SYSTEM

Dopamine is often conceptualized in energetic terms, as dopaminergic manipulations lead to locomotor, effort, and motivational disruptions (Wise, 2004; Floresco & Magyar, 2006; Berridge, 2007; Salamone *et al.*, 2007; Gruber *et al.*, 2009). The mesocortical dopamine projection is also directly implicated in behavioral organization via working memory, attention, cognitive flexibility, and decision-making (Sawaguchi & Goldman-Rakic, 1991; Watanabe *et al.*, 1997; Phillips *et al.*, 2004; Seamans & Yang, 2004; Robbins & Roberts, 2007; Vijayraghavan *et al.*, 2007; Robbins & Arnsten, 2009). There is little understanding of how dopaminergic neurons encode information for these functions. Dopaminergic neurons do not directly encode a neuronal correlate of working memory (Schultz *et al.*, 1993). Similarly, recordings from dopaminergic neurons during decision-making tasks suggest that the phasic responses of these neurons occur after behavioral decisions have been made (Morris *et al.*, 2006; Niv *et al.*, 2006; Roesch *et al.*, 2007). Our knowledge of the function of the VTA dopamine system is far more detailed than that of the GABAergic or glutamatergic systems, with no consensus on the role of the latter in cognition (Creed *et al.*, 2014).

EARLY INVESTIGATIONS INTO THE FUNCTION OF DOPAMINE NEURONS

Despite the aforementioned role of the dopamine system in behavioral organization, almost all of our knowledge of how dopamine neurons encode information through their firing patterns is singularly focused on value prediction. Early attempts to understand dopamine neuronal activity focused on motor movements, novel stimuli, or high intensity stimuli in the environment (Steinfels *et al.*, 1983a; b; Strecker & Jacobs, 1985; Freeman & Bunney, 1987; Nishino *et al.*, 1987; Schultz & Romo, 1987). Several years later, Wolfram Schultz and colleagues found that unpredictable rewards evoked short latency, short duration, phasic increases in dopamine neuron firing rates (Romo & Schultz, 1990; Schultz & Romo, 1990). An early study noted that when reinforcing outcomes could be predicted, the outcomes themselves did not modulate neuronal activity (Schultz & Romo, 1990). It was later observed that this property emerged with learning; initially, outcomes evoked dopaminergic activity, but after learning this was no longer the case (Ljungberg *et al.*, 1992; Schultz *et al.*, 1993). This finding was of particular importance to the field, as it demonstrated that these neuronal responses were plastic, and did not occur because of changes in the animal's actions, outcomes, or the environment. Thus, the phasic dopamine response cannot be related solely to sensory attributes, movement, reward, or economic value.

THEORETICAL MODELS OF DOPAMINERGIC NEURONAL ACTIVITY

A particularly compelling dataset from Wolfram Schultz's group demonstrated that 78% of dopamine neurons responded to unpredicted juice delivery, but 0% of these neurons responded to predictable delivery of the same outcome (Mirenowicz & Schultz, 1994). Similarly themed

findings were also reported from the same group around this time (Ljungberg *et al.*, 1992; Schultz *et al.*, 1993; Hollerman & Schultz, 1998). The work by Mirenowicz and Schultz (1994) is particularly noteworthy because it marked one of the earliest attempts to interpret phasic dopamine responses in the context of a formal model of associative learning. Portending a critical development in the field, the authors noted the importance of prediction errors in mediating learning, and a possible role for dopamine in this context (Mirenowicz & Schultz, 1994). Learning is inextricably linked to understanding how stimuli or behavior predict subsequent events, and psychologists have long linked errors in predicting outcomes to driving learning (Bush & Mosteller, 1951; Rescorla & Wagner, 1972). For an excellent review, see (Glimcher, 2011). Though Mirenowicz and Schultz (1994) did not explore the details of this relationship in detail, the notion that dopaminergic activity could be related to predictions about rewards, and ultimately learning, was an important turning point in our understanding of the dopamine system.

A detailed theoretical framework for understanding the role of dopaminergic neuronal activity in learning would emerge several years later. Montague, Dayan, and Sejnowski first conceptualized dopaminergic responses as a component of the temporal difference (TD) algorithm, the TD error signal (Montague *et al.*, 1996). This algorithm iteratively updates predictions about future outcome values at each time step within a trial via the TD error (Dayan & Abbott, 2001; Glimcher, 2011). The TD error quantifies the difference between the value of future rewards (bracketed terms below) and the predicted current value (the final term). Following Dayan and Abbott (2001), the algorithm for calculating the TD error is as follows in Equation 1:

$$\delta(t) = [r(t) + v(t+1)] - v(t)$$

Here, $\delta(t)$ is the TD error at time (t). The current reward value is $r(t)$. Estimated reward at the next time step (an estimator of future value) is $v(t+1)$. The predicted current reward value is $v(t)$. Far more detailed considerations of this model exist (Sutton, 1988; Barto & Sutton, 1990). The reader may also see Appendix A for additional discussion on how the TD error is calculated.

THE TEMPORAL DIFFERENCE ERROR MIMICS THE ACTIVITY OF DOPAMINE NEURONS

The TD error is a reward prediction error. A wealth of evidence suggests that the TD prediction error models the activity of dopaminergic neurons. As stated above, dopaminergic responses are not reflective of reward value, per se. Instead, dopaminergic responses reflect errors in predicting value. Errors in value prediction indicate that the value of rewards has not been fully learned. Because these errors are bidirectional and proportional to the magnitude of the error in prediction, they provide an account of how much, and in what direction, value estimates should be adjusted (Dayan & Abbott, 2001). For these reasons, dopaminergic prediction errors are thought to play an important role in value learning. Dopamine bidirectionally controls corticostriatal synaptic plasticity, and reward prediction errors are thought to strengthen or weaken synaptic weights (Reynolds & Wickens, 2002). Taken together, these theoretical and biological considerations detail how the dopaminergic prediction error signal could modulate associative learning.

The TD error signal follows several axiomatic patterns. Unexpected positively or negatively valued outcomes cause a positive or negative prediction error, respectively. These prediction errors manifest as phasic increases in neuronal activity (positive prediction error), or pauses in the firing of dopamine neurons (negative prediction error), just after the outcome is delivered. Unexpected outcomes generate prediction errors because the actual value at the time of outcome delivery is greater or less than predicted (Schultz, 1998; Dayan & Abbott, 2001; Glimcher, 2011). Fully predicted outcomes do not cause a prediction error because the predicted and actual values are roughly equivalent (Schultz, 1998; Dayan & Abbott, 2001; Glimcher, 2011).

The second term of the TD error calculation, $v(t+1)$, is an estimate of future reward value. New information about impeding outcome value is incorporated into the prediction error through this term. For this reason, stimuli that predict the value of an outcome also cause prediction errors. Prediction errors occur at the earliest moment in a trial that an organism is informed of future value. If multiple predictive stimuli are presented successively, only the first causes a prediction error, because the latter stimuli add no new information about outcome value (Schultz, 1998; Dayan & Abbott, 2001; Glimcher, 2011). The animal must learn that a stimulus is predictive of outcome value. It is expected that in early trials outcomes produce a prediction error. As the predictive relationship is learned the error occurs when the stimulus is presented. Thus, the prediction error moves to earlier predictors of the outcome. This phenomenon is termed ‘back propagation’ (Schultz, 1998; Dayan & Abbott, 2001; Glimcher, 2011).

A great deal of work suggests phasic increases in dopaminergic firing rates approximate the above mentioned attributes of the TD error signal, and play a causal role in associative learning (Schultz *et al.*, 1997; Schultz, 1998; Dayan & Abbott, 2001; Waelti *et al.*, 2001;

Schultz, 2002; Dayan & Niv, 2008; Niv & Schoenbaum, 2008; Schultz, 2010; Glimcher, 2011; Steinberg *et al.*, 2013). There is no consensus if negatively valued outcomes evoke activity reflecting TD errors, or if sustained changes in firing rates reflect TD errors, as conflicting results and theories have emerged (Mirenowicz & Schultz, 1996; Fiorillo *et al.*, 2003; Ungless *et al.*, 2004; Bayer & Glimcher, 2005; Joshua *et al.*, 2008; Brischoux *et al.*, 2009; Matsumoto & Hikosaka, 2009; Bromberg-Martin *et al.*, 2010a; Glimcher, 2011; Mileykovskiy & Morales, 2011; Fiorillo, 2013; Fiorillo *et al.*, 2013a; Fiorillo *et al.*, 2013b; Totah *et al.*, 2013). This dissertation is mainly concerned with phasic changes in VTA neuronal activity in response to rewarding outcomes and instrumental action.

THEORETICAL ACCOUNTS OF DOPAMINERGIC NEURONAL ACTIVITY IN INSTRUMENTAL BEHAVIOR

Pavlovian conditioning refers to situations in which environmental stimuli predict that an outcome will be delivered to an animal, and no explicit behavior is required of the animal. In contrast, instrumental behaviors require animals to execute actions in order to receive an outcome. The TD framework is very well suited to describe Pavlovian conditioning (Dayan & Abbott, 2001; Glimcher, 2011), though it can be applied to instrumental behaviors as well. In fact, the data initially suggesting that dopaminergic neurons encode TD error signals were derived from operant behavioral experiments (Romo & Schultz, 1990; Ljungberg *et al.*, 1992; Schultz *et al.*, 1993; Montague *et al.*, 1996).

There are 2 classes of models describing how actions are selected and executed: the actor-critic model and Q-value based models. The most widely used, the actor-critic model, suggests the dopaminergic TD error directly influences action selection (Houk *et al.*, 1995; Dayan & Abbott, 2001; Joel *et al.*, 2002; Morris *et al.*, 2006; Niv *et al.*, 2006). The few attempts to evaluate this idea, however, implicate Q-value based models (Morris *et al.*, 2006; Niv *et al.*, 2006; Roesch *et al.*, 2007). These models use the TD error to estimate the future value of actions (termed Q-values), but not for directly selecting actions (Morris *et al.*, 2006; Niv *et al.*, 2006; Roesch *et al.*, 2007). They suggest the TD error represents the value of the action that an animal selects or the value of the best available option. Each of these response patterns have been detected in dopamine neurons (Morris *et al.*, 2006; Roesch *et al.*, 2007). These data suggest that dopamine neurons are not directly implicated in deciding which action to select, but rather encode reward prediction errors, which are critical to learning the value of response strategies (Morris *et al.*, 2006; Niv *et al.*, 2006; Roesch *et al.*, 2007). In the aforementioned studies, this information was encoded just before action execution.

Many studies have not focused on the execution of the instrumental action itself. The studies that have depicted these data, suggest actions weakly modulate neuronal activity (Schultz, 1986; Romo & Schultz, 1990; Schultz & Romo, 1990; Ljungberg *et al.*, 1992). This may seem puzzling at first, as the dopaminergic signal theoretically conveys prediction errors about action values. The experimental designs used in these studies are of great importance to understanding these data. Many of these experiments utilize a cue at trial start, which often generates a phasic response, and there are often only seconds between cue, action, and outcome delivery (Miller *et al.*, 1981; Schultz, 1986; Romo & Schultz, 1990; Mirenowicz & Schultz, 1996; Nakahara *et al.*, 2004; Matsumoto & Hikosaka, 2007; Roesch *et al.*, 2007; Bromberg-

Martin & Hikosaka, 2009; Takahashi *et al.*, 2011; Totah *et al.*, 2013). In this situation, the cue accurately predicts future outcome value, and actions do not provide new information about outcome value. Thus, actions should not evoke phasic responses resembling TD errors.

PERFORMANCE OF SERIAL ACTIONS

It is not well understood how actions modulate neuronal activity when they are more relevant predictors of outcome delivery, and thus, future value. This is exactly the case when an animal is required to perform multiple actions within a trial that is progressing in an uncertain fashion. This distinction is critical because in such settings, information necessary for real-time behavioral organization must be encoded as behaviors are being executed, and environmental stimuli at the beginning of a behavioral series cannot be used for this purpose.

The current work investigates how VTA neurons encode information when many actions are performed in a series. In the only previous examination of VTA neuronal activity during serial actions, dopaminergic and non-dopaminergic VTA neurons had elevated and "irregularly" varying firing rates while actions were executed (Nishino *et al.*, 1987). Cues at the beginning of each trial did not modulate neuronal activity in most neurons (Nishino *et al.*, 1987). Those data suggest that networks of dopaminergic and non-dopaminergic VTA neurons encode information critical for serial action execution. Shared and complementary roles in encoding information by both types of VTA neuron have subsequently been suggested (Seamans & Yang, 2004; Kim *et al.*, 2010; Kim *et al.*, 2012), but are largely unexplored. The "irregular" variations in firing rate suggest that there is structure in the neuronal activity that has not been previously explored or

theorized, and are suggestive of a much more complex role for VTA in behavioral organization than has previously been described.

GOALS OF THE CURRENT WORK

This dissertation explores how VTA encodes information during the performance of serial instrumental actions. This topic is vital to our theoretical understanding of behavioral organization, but is poorly understood. This dissertation examines the collective activity of networks of dopaminergic and non-dopaminergic neurons. Previous electrophysiological experiments have seldom addressed non-dopaminergic neurons or network function. The anatomical structure of the VTA, however, suggests that distributed network function could be critical to how the region processes information. Thus, this work will fill several voids in our theories of VTA function.

2.0 METHODS

SUBJECTS AND APPARATUS

All procedures were conducted in accordance with the National Institute of Health's *Guide to the Care and Use of Laboratory Animals*, and approved by the University of Pittsburgh Institutional Animal Care and Use Committee. Ten adult male Sprague-Dawley rats (Harlan, Frederick, MD) were utilized in the current study, and behavioral and neuronal data were collected from all animals. Each animal was housed on a 12-hour light cycle (lights on at 7pm). At the time of surgery, rats weighed approximately 350 grams. Under isoflurane anesthesia, each rat was implanted with 8 or 16 channel 50 μm stainless steel Teflon insulated microelectrode arrays, in left VTA (relative to Bregma: -5.30 mm posterior, 0.8 mm lateral, and 8.3 mm ventral) for chronic recording. A subset of rats was also implanted in left prelimbic cortex (data not presented). Implants were made through small craniotomies and head caps were sculpted from dental cement attached to the skull surface and bonded to skull screws for stability. All experiments were run in a standard operant chamber (Coulbourn Instruments, Allentown, PA). The operant chamber had a wire grate floor, with a custom-made adjustable food trough on one wall and a single nose poke port on the opposite wall. Both the food trough and nose poke port could be illuminated and were equipped with an infrared beam, which detected the animal’s entry. A house light was located at the top of the operant chamber. The operant chamber system

controller was configured to send the time of behavioral and environmental events to the recording interface via standard TTL pulses and a digital interface.

BEHAVIOR

Each rat was given 7 days to recover from surgery and food restricted to approximately 90% of their free feeding body weights. Rats were habituated to handling for 5 minutes per day for 3 consecutive days, before being habituated to being handled and connected to a headstage cable in the procedure room. Each day of habituation, a rat was given 10 sugar pellets in his home cage. Following habituation, rats were given a single 30 minute magazine training session in the operant chamber, in which sugar pellets were delivered on a variable time 75 second reinforcement schedule and the only environmental stimulus present was a small light outside the operant chamber which provided low level ambient illumination. When each pellet was delivered, the pellet trough was illuminated for 4 seconds. The animal's behavior had no programmed consequences in the magazine training session.

Following the magazine training session, each animal began instrumental conditioning. During all instrumental conditioning sessions, each trial began with illumination of the nose poke port (cue light onset). This served as a discriminative stimulus that reinforcing outcomes (sugar pellets) were available (response period), contingent upon the animal executing actions (nose pokes into the lit port). In each trial, actions were reinforced randomly, according to a predetermined probability. When an action was executed, the behavioral system controller randomly drew an outcome state (either reinforcement or no programmed consequence) with

replacement, according to the probability of reinforcement. Each action was reinforced randomly and independently of the animal's action history within that trial or session. When an action was reinforced, the cue light was immediately extinguished (cue light offset) and nose pokes had no additional programmed consequences. A 0.500 sec delay between the final action and outcome delivery was instituted to temporally separate these events, as done in previous work (Schultz *et al.*, 1993). Following the delay, the outcome was delivered to the animal and the food trough was illuminated. Outcomes were delivered into the food trough from a standard pellet magazine, via the operation of a smaller stepper motor and dispenser. The operation of the motor created a sound that was audible throughout the operant chamber. The food trough remained illuminated and the task did not progress until the animal retrieved the outcome. Once the animal retrieved the outcome, a variable length intertrial interval (ITI) of 10-12 seconds was initiated. Following this, the next trial began with the onset of the discriminative stimulus cue light. In each session 180 trials were administered.

In the first instrumental conditioning session, actions were reinforced with a probability of 1 (each action was reinforced) equivalent to a fixed ratio 1 (FR01) reinforcement schedule. In the second session, the probability that an action was reinforced was decreased across three blocks of trials. In the first block of 60 trials, actions were reinforced with a probability of 1 (FR01). In the second block of 60 trials, each action had a 1 in 3 chance of being reinforced (random ration 3, RR03). In the third block of 60 trials, the probability was further decreased to 0.2 (random ratio 5, RR05). In sessions 3 and 4, actions were reinforced with a 0.2 probability for all trials (RR05). In sessions 5, 6 and 7, actions were reinforced with a probability of 0.1 for all trials (random ratio 10, RR10). In all trials but the FR01 trials, each animal was required to execute an unknown, varying, and randomly determined number of actions per trial. Except in

trials in which the animal was reinforced on the first action, the animal was required to execute a series of actions to earn the outcome. Because actions were reinforced unpredictably, the RR schedules limited the ability of the animal to anticipate reward delivery. Actions differed from each other only in terms of their location within the action series in each trial (the action number within a trial, e.g. 1st action, 2nd action, 3rd action, etc.).

BEHAVIORAL ANALYSIS

In each trial, each animal's action rate was calculated as the number of actions (nose pokes executed while the nose poke port was illuminated) divided by the duration of the response period (time the nose poke port was illuminated). This served as a measure of behavioral conditioning and performance. Response latency was measured as the delay between the onset of the cue light at the start of each trial, and the first response in the trial, in seconds. This served as a measure of attention to the task, and learning about the action-outcome relationship. Outcome retrieval latency was measured as the delay between outcome delivery and retrieval in seconds. This served as a general measure of motivation. For all three measures, between-sessions changes in these metrics were assessed with repeated measure analysis of variance (ANOVA), and repeated measures contrasts were applied as appropriate.

HISTOLOGY

Following the completion of experiments, animals were perfused with saline and brains were extracted. Each brain was stored in a mixture of sucrose and formalin. The brains were then frozen and sliced in 60 µm coronal sections on a cryostat, before being stained with cresyl-violet. The location of each implant was histologically verified under light microscope according to Swanson's brain atlas (Swanson, 2004).

ELECTROPHYSIOLOGY

During experiments, animals were attached to a flexible headstage cable and motorized commutator that allowed the animal to move freely about the operant chamber, with limited disruption of behavior (Plexon, Dallas, TX). Neural data were recorded via the PlexControl software package, operating a 64-channel OmniPlex recording system (Plexon, Dallas, TX). Briefly, neural data were buffered by a unity gain headstage and then a preamplifier. The digitized broadband signal was then band-pass filtered (100 Hz – 7 KHz). High-pass filtering can affect spike waveform shapes and neuronal identification, but with freely moving animals it is necessary to apply these filters to remove artifacts from the neuronal signal (Ungless & Grace, 2012). The filter pass bands that were utilized in the current manuscript are consistent with those that have previously been used to record from dopamine containing brain regions (Schultz *et al.*, 1993; Fiorillo *et al.*, 2003; Tobler *et al.*, 2005), and were chosen to be most consistent with previous work. Data were digitized at 40 KHz and continuously recorded to hard disk. Voltage thresholds were applied to the digitized spike data offline and spikes were sorted into well-

isolated units (Offline Sorter, Plexon, Dallas, TX). Single units were sorted using standard techniques, and were utilized only if they had a signal to noise ratio in excess of 2/1, and were clearly separated from noise clusters and other single unit clusters.

VTA neurons were classified as putative dopaminergic and non-dopaminergic neurons. A VTA neuron was classified as dopaminergic if it had broad action potentials, greater than 1.4 ms in duration, and a mean intertrial interval firing rate less than 10 Hz. These criteria are similar to those used in previous studies (Hyland *et al.*, 2002; Fiorillo *et al.*, 2003; Anstrom & Woodward, 2005; Pan *et al.*, 2005; Tobler *et al.*, 2005; Anstrom *et al.*, 2007; Totah *et al.*, 2013). All remaining neurons were classified as non-dopaminergic. These criteria may be subject to both false positive and false negative classification errors (Margolis *et al.*, 2006). These are standard criteria for identifying dopamine neurons and represent the most widely used solution with extracellular recordings, where direct identification (e.g. juxtacellular labeling) is not feasible. While some work suggests that dopamine agonists can be used to verify that a VTA neuron is dopaminergic (Bunney *et al.*, 1973; Grace & Bunney, 1983a; Johnson & North, 1992), this approach was not suitable for the current experiments. Manipulations of the dopamine system can modulate behavioral performance, motor activity, and memory consolidation (Krivanek & McGaugh, 1969; Robbins *et al.*, 1983; Oades *et al.*, 1986; McGaugh, 2000; Setlow & McGaugh, 2000; Wise, 2004; Simon & Setlow, 2006). Thus, drug administration would likely disrupt behavioral performance or learning in the task. All analyses were also conducted on the entire population of neurons that were recorded to provide a classification-free examination of the data. Should neurons have been misclassified, it is unlikely that this strongly affects the conclusions of the current work. In general, qualitatively similar responses were observed in both groups of neurons. Units recorded in different sessions were considered separate units, as methods to

estimate neuronal identity between sessions are not widely used with VTA recordings. Since fixed electrode arrays were utilized in the current experiments, it is likely that some of the same neurons were serially recorded. Though this is the case in most chronic recording experiments, it introduces some problems for data analysis. It is unclear which neurons would have been serially recorded, and thus, it is impossible to treat those neurons as repeated measures. While all units were assumed to be independent in the current work, a better solution would be to treat serially recorded neurons as a repeated measure and unique neurons as independent. This would necessitate a mixed-model design that cannot currently be implemented, because it is difficult, if not impossible to identify VTA neurons between recording sessions. Thus, care should be taken when interpreting these data, as some statistical assumptions may not be justified. The current analyses represent a compromise between opposing factors such as technical feasibility and some statistical assumption.

NEURONAL DATA ANALYSIS

Neuronal activity evoked by environmental stimuli:

Each single unit's spike times were binned into spike counts (0.025 sec bins) within a trial. Binned spike counts were aligned to all relevant environmental events (e.g. cue light onset, delay between cue light offset and outcome delivery, and outcome delivery). A four second portion of the ITI (5 seconds to 1 second prior to cue light onset) served as the neuronal activity baseline. Single unit firing rates were Z-score normalized relative to baseline and zero-centered before each unit's smoothed activity (3 bin rectangular window) was averaged together. In addition to

analyzing the entire population of neurons, data were split into putative dopamine or non-dopamine neurons, so that functional dissociations between neuronal subtypes could be assessed. Each unit's normalized activity was examined in 0.250 sec windows around experimental events (cue onset: +0.050 – 0.300 sec, relative to cue onset; delay between last action and outcome delivery: +0.150 – 0.400, relative to execution of the last action; outcome delivery: +0.050 – 0.300 sec, relative to delivery). To assess between-session changes in population-level evoked activity, windowed activity was compared with a between groups two way ANOVA, with session number and neuron type (dopamine or non-dopamine) as grouping variables. In all cases, protected Fisher's least significant difference tests were applied as appropriate.

A unit was classified as being activated or suppressed by an event if it met 2 criteria: I) a significant paired samples t-test comparing raw (non-normalized) baseline firing rates with raw evoked firing rates, and II) 3 or more consecutive bins of mean activity within the event-window, that were in excess of a 95% confidence interval around the baseline mean. All classifications were inspected visually to verify the validity of these criteria. The overwhelming majority of cells that were recorded either increased their firing rates or did not respond to the task events. Decreased firing rates were seldom, if ever, observed in any of the sessions. It should be noted that less stringent criteria may have yielded more frequent classification of cells as significantly suppressed. With respect to a given task event, the proportions of units classified as being activated or suppressed were calculated. Differences in the proportions of activated units between sessions were compared with a Chi-squared test of independence. Insufficient numbers of suppressed neurons (in some sessions zero) were obtained to permit reliable statistical analyses of this class of responses.

Neuronal activity evoked by the execution of actions:

The previously described analyses involved environmental stimuli, which are by definition, solely external to the animal. In contrast, actions involve the animal's motoric output in the context of consistent environmental stimuli. Neuronal activity around the time of the action may be evoked by the action, stimuli in the environment (e.g. the nose poke port, the cue light inside this port, etc.), or some interaction between these factors. This experiment was not designed to separate the contributions of these factors to neuronal activity. The terminology "action-evoked" neuronal responses refers to all of these factors collectively, without assuming that the action is solely responsible for evoking this neuronal response.

The time of action execution was defined as the moment that an animal broke the infrared photodetector beam located inside the nose poke port. Neuronal activity was aligned to action execution and data were windowed (-0.125 - +0.125 sec, relative to the time of action execution). Data were normalized as described above. Changes in action evoked neuronal activity were assessed with a two way ANOVA, with session and neuron type as grouping variables. For these analyses, all actions were grouped together. Similar to above, neurons were classified as activated or suppressed by action execution if action evoked neuronal activity was significantly different from baseline activity levels and 3 consecutive bins of activity were in excess of a 95% confidence interval around baseline activity levels.

Each unit's activity was examined as a function of action number (a unit's mean response to each n^{th} numbered action within a trial, across all trials). These analyses were restricted to the RR10 sessions (sessions 5-7). RR10 sessions required larger numbers of actions per trial, on average, and would ensure that there were a sufficient number of higher numbered actions (e.g. actions 18, 19, 20, etc.) for analysis. All action number analyses utilized actions 1 through 20.

While even higher numbered actions occurred in some trials, these actions occurred less frequently and were excluded from action number analyses, as there was insufficient sample size for reliable statistical analysis. To remove any effects of impending reward delivery on the action evoked neuronal responses, only unrewarded actions were used in this analysis. Preliminary analyses suggested that including rewarded actions had little effect on the results. Neuronal responses were collected in the same windows around action execution as described above. Mean normalized population activity as a function of action number is presented in the Results section. Action evoked neuronal responses were binned into 4 bins of 5 consecutive action numbers (actions 1-5, 6-10, 11-15, and 16-20). Statistical significance of differences between action number bins or neuron type were assessed with a two way repeated measures ANOVA, with action number as a repeated measure and neuron type as a between groups variable. Linear correlations between dopaminergic and non-dopaminergic neuronal responses and action numbers 1-20 were examined via a Pearson's product-moment correlation. Similar to other analyses, a neuron's response was defined as activated or suppressed by a bin of consecutive actions if the raw baseline firing rates were significantly different from action evoked firing rates, and 3 or more consecutive bins of mean activity that were in excess of a 95% confidence interval around the baseline mean.

Individual neurons preferred different subsets of action numbers. To visualize the various tunings of VTA neurons to action number, each neuron's mean activity as a function of action number is displayed. Because evoked firing rates could span a large range of values, each tuning curve was scaled so that the maximum evoked activity was equal to 1 and the minimum evoked activity was equal to 0. Scaled tuning curves allow effective visualization of all neuronal responses simultaneously. These data are used only for visualization, and unscaled tuning curves

are displayed for clarity. If a neuron's action evoked firing rates significantly differed across action number bins, as assed with a between groups ANOVA, then it was defined as significantly modulated by action number. Tuning curve maximum – minimum depth was calculated as the difference between the absolute maximum and minimum value of each tuning curve. Each tuning curve was fit with a cubic smoothing spline (smoothing parameter 0.001) and maxima were detected as points in the spline with derivatives equal to 0.

Functional principal components analysis of tuning curves:

Tuning curves across action number are inherently high dimensional and visualizing similarities between tuning curves requires transforming these data to a lower dimensional space. This necessitated the use of functional principal components analysis of the tuning curves of each neuron. Each neuron's activity evoked by the A^{th} action can be treated as a varying function of action number (tuning curve). Tuning curves were projected onto smooth functional principal components, for visualizing and characterizing the neuronal variability in tuning. The linear combination of a small number of orthogonal principal components tends to explain a large proportion of variability between tuning curves and produce relatively accurate reconstruction of the original data. The principal components must all be orthogonal, and their squares must integrate to 1. The principal components scores, f_k^i, are defined in Equation 2:

$$f_k^i = \int X_k(a) R^i(a) da$$

Here, $X_k(a)$ is the principal component k, at action number (a). $R^i(a)$ is the mean response of neuron i at action number (a). The principal component projection is derived from the scores f_k^i, which maximize Equation 3:

$$\lambda_k = \frac{1}{N-1}\sum_{i=1}^{N}(f_k^i)^2$$

Here, N is the number of neurons, and λ_k is proportional to the variability in tuning curves, between units, that principal component k explains. Thus, the functional principal components represent functions, which the tuning curves can be projected onto, to maximize variability between the tuning curves.

Decoding action number from neuronal activity:

A Bayesian decoder classified binned action number using either the population average spike count (population average decoder) or the ensemble of individual-neuron spike counts, where individual neurons were assumed to spike independently (naive Bayesian decoder). Decoding accuracy was evaluated using cross-validation (Kass *et al.*, 2014). Binned action number, A, was defined as one of 4 bins of 5 consecutive actions (1-5, 6-10, 11-15, or 16 - 20). Bayesian decoding finds the action number bin (A) that maximizes $p(data|A)\ p(A)$, where $p(data|A)$ is the probability density function for the data under action A. The prior probability, $p(A)$, was 0.25 for all 4 action number bins. In cross-validation, the whole data set is decomposed, repeatedly, into test data and training data. The multiple sets of test data are used to evaluate the performance of each classifier while, for each set of test data, the corresponding training data are used to estimate $p(data|A)$.

The data set consisted of $N = 156$ units combined from 3 consecutive RR10 sessions. We created pseudo-data for each action number: to create one set of spike counts from N units, we resampled 1 spike count from each of the units recorded in each session. Each set of N spike counts, which reflects the experimental structure of independence across sessions with

simultaneous recording of units within sessions, was treated as a vector of test data. The remaining, non-sampled data were used to train each classifier. We repeated the process 500 times to create 500 test data vectors for each action number. We let R^i denote the random variable representing the test data spike count for unit i, and R^{i*} its observed value, for $i = 1$ to N, then also let R^* denote the population average spike count as in Equation 4:

$$N^{-1}R^* = \sum_{i=1}^{N} R^{i*}$$

and R denote the corresponding random variable. For the population average decoder, $p(data|A) = p(R = R^*|A)$, the probability density $p(R|A)$ was estimated from the training data by resampling 300 sets of pseudo-data and then applying a Gaussian kernel density estimate (Gaussian filter, standard deviation = 0.04 spikes).

For the naive Bayesian decoder (Equation 5):,

$$p(data|A) = \prod_{i=1}^{N} p(R^i = R^{i*}|A)$$

and we estimated $p(R^i = R^{i*}|A)$ as the empirical proportion of counts for which $R^i = R^{i*}$ within the training data. For both classifiers, the cross-validated classification probability was the percent correct out of the 500 test data vectors, for each action number.

Statistical Testing of Decoders:

Correct classification may be considered as success in a Bernoulli trial, and the sum of Bernoulli trials is binomially distributed. We evaluated decoder performance, first, by comparing with

chance levels, taking the null hypothesis to be that the binomial probability of success was .25. To test for a significant difference between decoders, we fit a binomial generalized linear model and controlled for action number. For each trial of action bin, tested on decoder, the log odds of the probability of correct classification was given by the regression function as in Equation 6:

$$\text{logit}\left(\mathbb{E}\left[I_A\left(\hat{A}\right) \mid (I_d, A)\right]\right) = \log\left(\frac{p_d(A)}{1-p_d(A)}\right) = \beta_0 + \beta_d \cdot I_d + \beta_A \cdot A.$$

Here, $I_A\left(\hat{A}\right)$ is an indicator signifying correct or incorrect classification, and I_d is an indicator for the two decoders. The data provides evidence for improved classification using the naive Bayes decoder when β_d is found to be significantly greater than zero. Correct classification can be conceptualized as success in a Bernoulli trial. The sum of Bernoulli trials is binomially distributed. Thus, the proportion of correct classifications at each action number was compared to a binomial distribution in which correct classifications (successes) occurred at chance levels (0.25). The resulting probability of observing a classification rate as extreme, or more extreme, than the empirical classification rate was significant if less than 0.05.

3.0 RESULTS

This dissertation examines how VTA neurons encoded information during execution of serial behaviors, a topic that is poorly understood. Animals were trained to execute actions (nose pokes into a lit cue port) in order to receive rewarding outcomes (sugar pellets). Because the most obvious difference between actions was their corresponding action number within a trial (action number), special attention was paid to analyzing neuronal activity as a function of action number. Stimuli predicting trial start (cue light onset) and outcome delivery are also examined to understand how VTA neurons process information about these events during trials requiring serial action execution.

TASK DESCRIPTION

In the task, animals learned to execute actions to earn rewards. In the first session, each action was reinforced. In session 2, the probability that an action was rewarded was decreased from 1 to 0.2, in a block-wise fashion. In sessions 3 and 4, actions were reinforced at a probability of 0.2. In sessions 5 – 7, actions were reinforced at a probability of 0.1 (Figure 3.1). Random reinforcement ensured that each action was equally likely to lead to reinforcement, and thus, was equally valued. Further, reward anticipatory effects would likely be minimized in such a design. In all trials that were randomly reinforced, unpredictable and varying numbers of actions were

required. For example, a given trial may require only 1 action or dozens of actions. The probability that an action would be reinforced did not change based upon how many actions had been completed up to that point in the trial. In each trial, onset of a cue light signaled rewards were available to be earned. Rats then began nose poking for rewards. If an action was not reinforced, there was no change in the environment. When an action was reinforced the cue light turned off and reward delivery occurred after a half second delay (Figure 3.1).

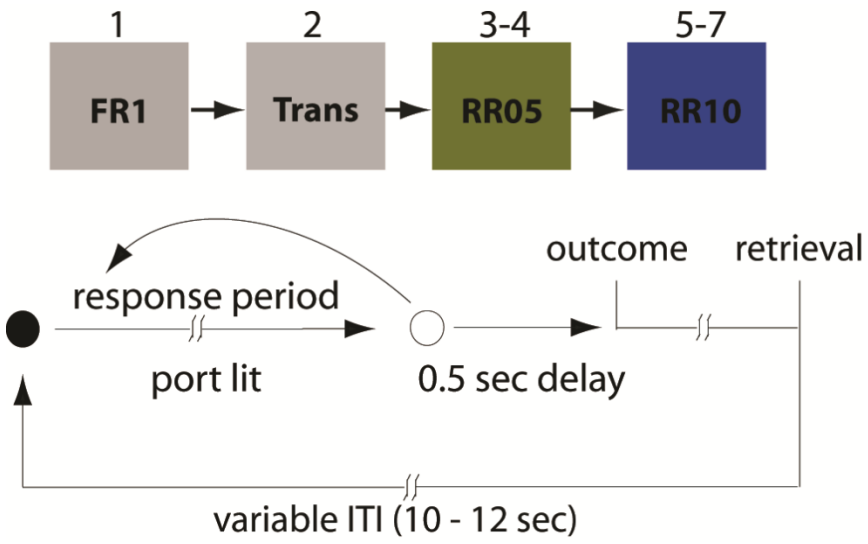

Figure 3.1. Experimental Design.

Instrumental actions (nose pokes into a lit port) were reinforced probabilistically with delivery of sucrose pellets (outcomes). In session 1, each action was reinforced (fixed ratio 1). During session 2 the probability that an action was reinforced was decreased in 3 blocks of trials (block 1, p = 1; block 2, p = 0.3; block 3, p = 0.2; transition session). In sessions 3 and 4, the probability of reinforcement was 0.2 (random ratio 5 reinforcement schedule). In the final 3 sessions (5-7), actions were reinforced at a probability of 0.1 (random ratio 10). These sessions are referred to as FR01, TRANS, RR05 and RR10. There were 180 trials per session. A cue light was illuminated at the start of each trial, signaling that reinforcement could be earned through action execution. If a rat executed an unreinforced action, there were no changes in the environment. When an action was reinforced (white circle), the cue light was immediately extinguished, and following a 0.5 sec delay, the outcome was delivered. At outcome delivery, the reward trough was illuminated and a small stepper motor turned to dispense sucrose pellets. The sugar pellet then rolled into the reward trough. The motor created a clearly audible mechanical sound. Thus, outcome delivery was associated with visual and auditory stimuli that immediately preceded the actual delivery of the outcome. The task proceeded to the ITI once the animal retrieved the outcome. Only one action was reinforced per trial.

ANIMALS LEARNED TO EXECUTE SERIAL ACTIONS

Animals learned to perform and sustain serial actions until outcomes were earned in each trial. In nearly every session, all 180 trials were routinely completed and all outcomes consumed. Action rate, defined as the number of actions executed per second when outcomes were available, was a behavioral index of learning and performance in each reinforcement schedule. As expected, action rates were greater in all random ratio sessions than in the fixed ratio 1 session (Figure 3.2). Action rates in the last two random ratio 10 sessions were significantly higher than all other sessions, and action rates in the random ratio 5 sessions were significantly higher than the preceding sessions (Figure 3.2; $F_{(6,24)} = 4.726$, $p = 0.003$; repeated measures contrasts $p < 0.05$). Thus, consistent with behavioral theory, increasing the average action requirement increased action rate (Reynolds, 1975). The latency to begin executing actions once the cue light was illuminated was significantly faster in all sessions than during the initial fixed ratio 1 session ($F_{(6,24)} = 8.996$, $p = 0.003$; all repeated measures contrasts $p < 0.05$). This effect is consistent with animals learning that cue onset predicted outcome delivery. The latency to retrieve outcomes, measured from the time of delivery to retrieval, did not differ between sessions ($F_{(6,24)} = 1.928$, $p = 0.226$). Thus, rewarding outcomes equally motivated animals in all sessions, and differences in behavioral performance across sessions are most likely unrelated to affective processes. Taken together, these data suggest that animals readily learned to perform serial actions and that behavioral performance was sensitive to the action-outcome contingency. Contingency is reflective of the average number of actions performed to receive an outcome, and these data

suggest that the animals were also sensitive to this aspect of the FR01, RR05, and RR10 reinforcement schedules.

There was no correlation between the latency to retrieve outcomes and the number of actions in the current trial ($r_{(3468)} = -0.017$, $p = 0.309$). Similarly, the latency to begin responding in the next trial (measured from onset of the cue light) was not correlated with the number of actions performed in the previous trial ($r_{(3445)} = 0.015$, $p = 0.389$). The lack of correlation between action number and latency to retrieve outcomes or initiate the next trial suggests that performing more or less actions did not alter fatigue, attention, or motivation. Inter-action interval increased significantly at higher action number bins (Figure 3.3; $F_{(3,66)} = 8.665$, $p < 0.001$). This outcome suggests that animals were sensitive to the number of actions performed within each trial. Inter-action intervals in the first half of each session were not statistically different from those in the second half of each session ($t_{(21)} = 1.458$, $p = 0.160$). These data, coupled with the lack of correlation between retrieval latency and the number of actions in a trial, further suggest that increasing inter-action intervals were not a byproduct of fatigue or decreased motivation.

Taken together, the change in action rate between reinforcement schedules, and the increased inter-action interval at higher action numbers, may suggest that animals perceived the accumulation of successive actions. It should also be noted that the increased inter-action interval for higher numbered actions was not associated with overt changes in the way that actions were performed (e.g. according to visual observation, actions were still performed with similar patterns of motor output, etc.) This is important because it suggests that actions did not systematically differ from each other in a way that would confound the analyses of neuronal activity as a function of action number.

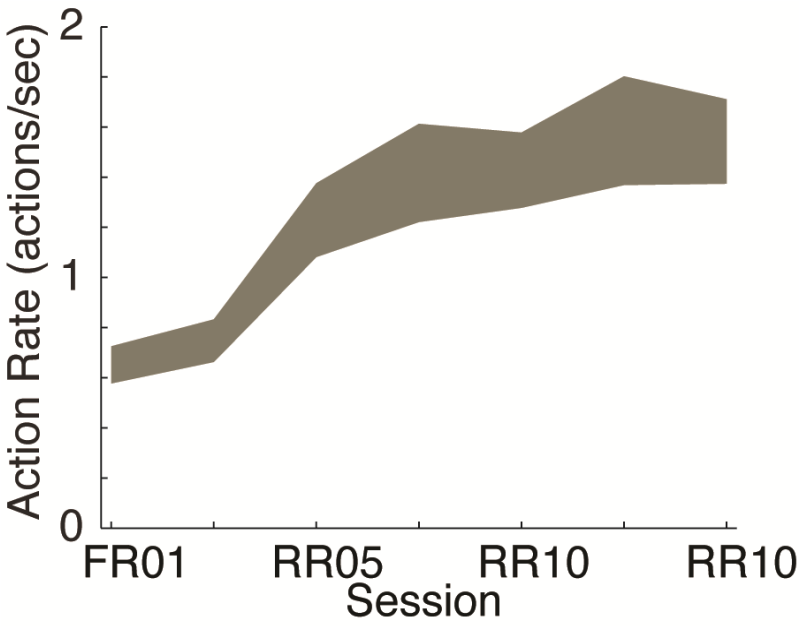

Figure 3.2. Action Rate.

Data depict the mean ± SEM action rate of all animals in each recording session. Animals learned to perform serial actions. Action rates were greater in RR sessions than the FR1 session. RR10 action rates were greater than RR05 action rates. There were no differences in action rate during the final RR sessions ($F_{(6,24)} = 4.726$, $p = .003$; repeated measures contrasts $p < 0.05$).

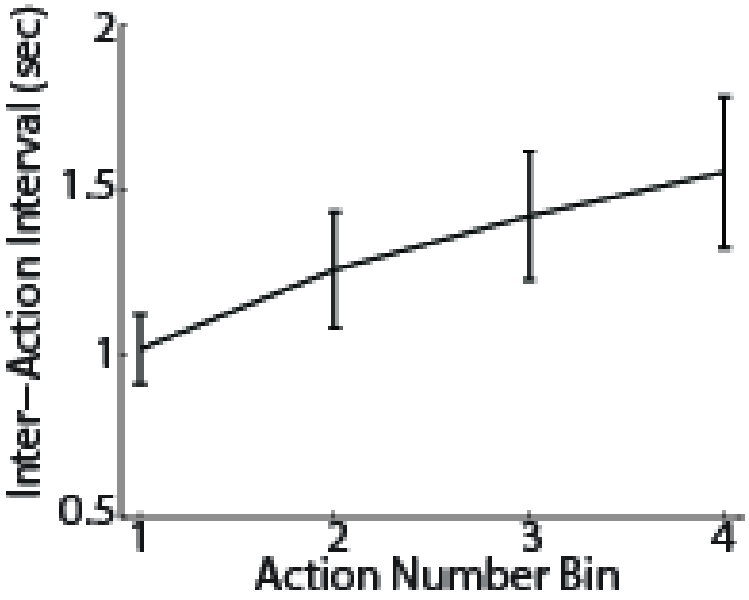

Figure 3.3. Inter-Action Interval as a Function of Action Number.

Data depict the mean ± SEM inter-action interval all animals in all RR10 recording sessions as a function of binned action number. Action number has been binned into 4 bins of 5 consecutive action numbers. Note that there were significantly greater inter-action intervals at higher action numbers ($F_{(3,66)} = 8.665$, $p < 0.001$).

NEUROPHYSIOLOGICAL CLASSIFICATION

The current dataset consists of 375 units recorded from 10 rats in 7 sessions. All recording arrays were verified under light microscope to be located in the VTA via standard histological approaches (Figure 3.4). Cells were identified as putative VTA dopamine units ('dopamine neurons', n = 155) if they had spike waveforms wider than 1.4 ms and baseline firing rates below 10 Hz (Fiorillo *et al.*, 2003; Tobler *et al.*, 2005) (Figure 3.4). The remaining 220 units were classified as putative non-dopamine units ('non-dopamine neurons') (Figure 3.5). It should be noted that there were not natural distinctions, or 'clusters', formed between these two subpopulations using this approach. This ultimately means that the classification is somewhat arbitrary. For this reason, all neuronal analyses are also performed on all neurons, without respect to neuronal classification. While some previous work has suggested that interspike intervals can be used to classify neurons as dopaminergic or non-dopaminergic (Hyland *et al.*, 2002), there was no evidence of this in the current dataset. Both classes of neurons (according to the above described criteria) had similar coefficients of variation of baseline interspike intervals (dopamine neurons: 1.13 ± 0.04, non-dopamine neurons: 1.19 ± 0.04; $t_{(373)} = -1.039$, $p = 0.300$). Examples of cell sorting and raw voltage traces demonstrate typical unit separation in principal component space, and typical signal to noise ratios (Figure 3.6). Accompanying those raw data are example rasters demonstrating task-evoked neuronal responses for each event in the experiment (Figure 3.6).

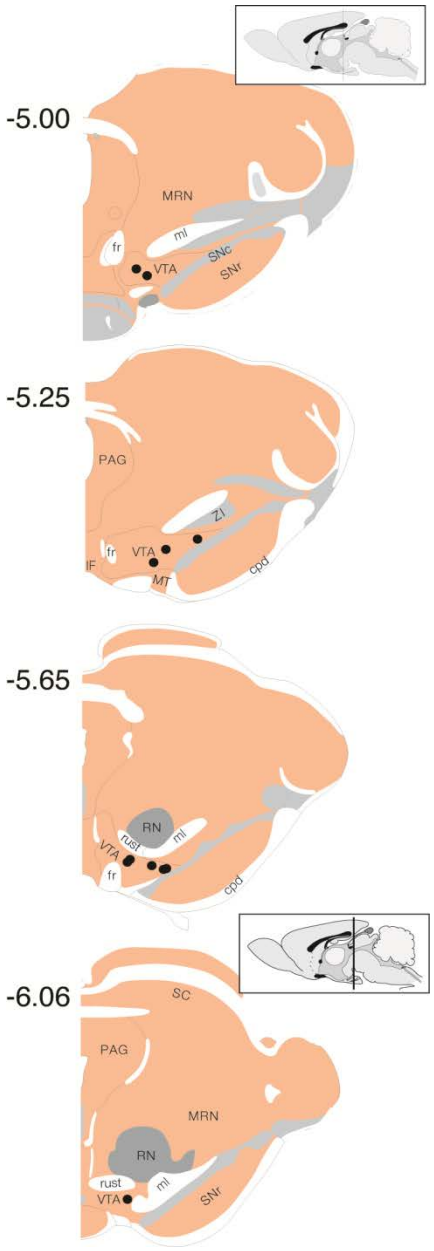

Figure 3.4. Locations of Recording Electrodes within VTA.

Each section shows estimated locations of each recording array plotted as a black circle. All placements were verified under light microscope. Insets show midsagittal diagram of rodent brain with vertical lines representing the approximate location of the most anterior and posterior coronal sections shown.

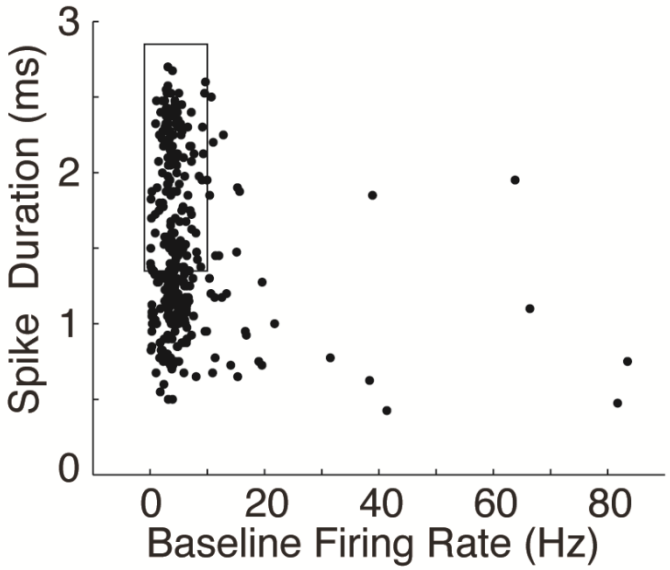

Figure 3.5. Electrophysiological Classification of VTA Units.

Each point represents a single unit and data from all sessions are depicted. Units with wide spike widths (>1.4 ms) and low firing rates (< 10 Hz) are enclosed in the box. These units were considered putative dopamine neurons. The remaining units were considered putative non-dopamine neurons.

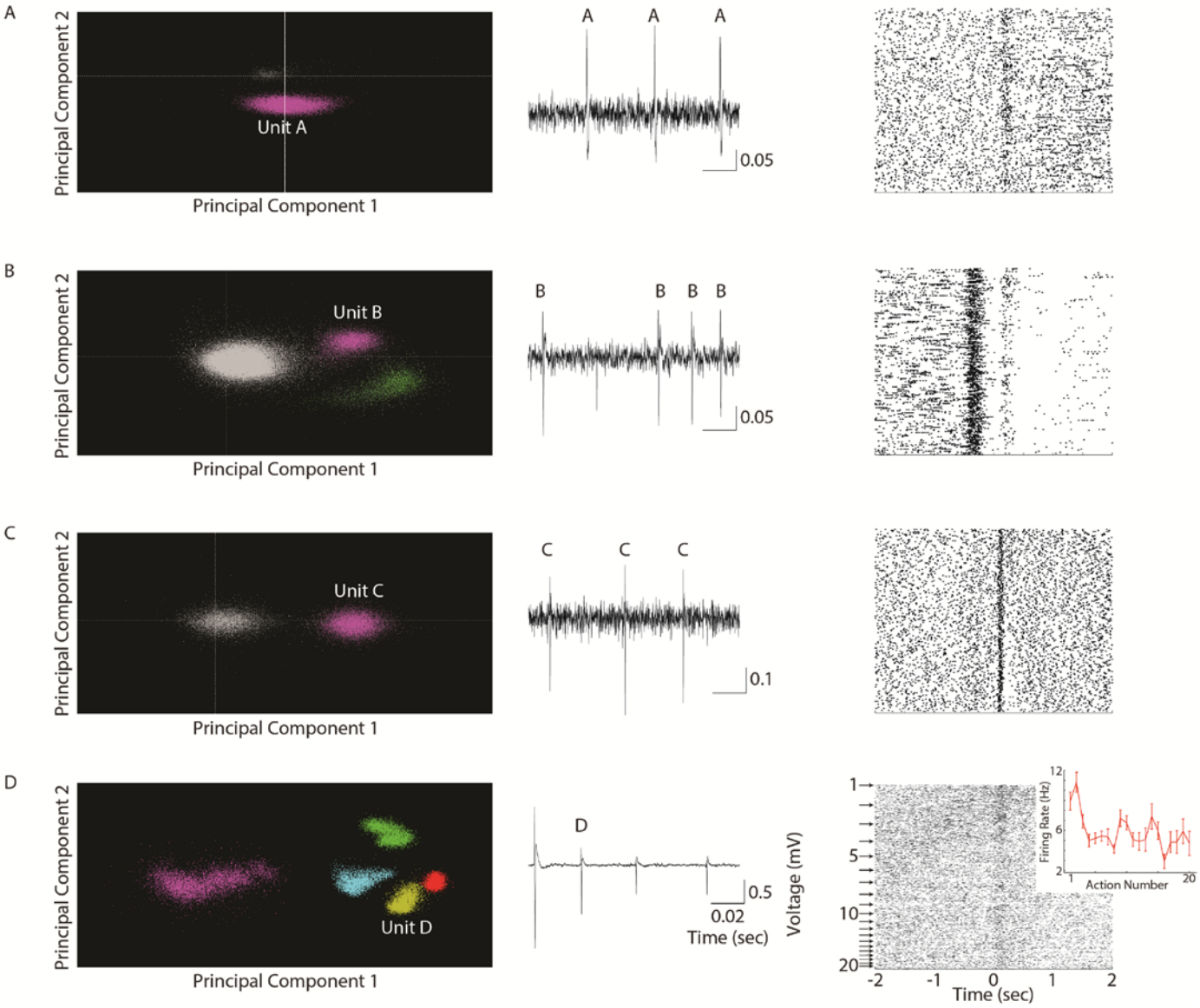

Figure 3.6. Representative Examples of Sorting, Recordings, and Responses.

(A) Spikes from unit A, a dopamine neuron, sorted according to the first 2 principal components of all threshold crossing waveforms (left). Purple points represent spikes that were assigned to unit A, and gray points represent noise that was not sorted into single unit spikes. A raw voltage trace (band pass filtered between 100 Hz and 7 KHz) corresponding to the same unit is depicted in the middle column. Examples of spikes belonging to unit A are notated in the trace. Unit A represents a typical cue-responsive unit. Raster plot depicts the unit's response aligned to cue onset (right). Each dash represents a single spike, and each row represents a single trial (first trial in the top row).

Note the increased spike density just after cue onset (time 0) across all trials. (B) Representative delay period responsive, non-dopaminergic, neuron. Data plotted with the same conventions as (A). In this example recording, two units were simultaneously recorded (left). Raster (right) depicts delay period (-0.5 – 0s) spikes aligned to the time of reward delivery. Note the consistent increase in spike density after cue offset and preceding outcome delivery. (C) Data from a representative non-dopaminergic, outcome delivery responsive neuron is plotted with similar conventions as (A). Raster (right) depicts neuronal activity aligned to the time of outcome delivery. Note the consistent delivery evoked response. (D) Data from a typical dopaminergic neuron that preferred low action numbers. Spike sorting (left) plotted with similar conventions as (A). Several units were simultaneously recorded and the example voltage trace contains spikes from multiple units (middle). Only a spike corresponding to unit D (yellow) is notated. The raster (right) shows spikes aligned to the time of action execution (time 0). Each row of the raster represents one action evoked response and rows are arranged by action number. Each arrow on the right represents action numbers 1-20. For each action number, the earlier occurrences of an Nth numbered action are arranged toward the top. Thus, the first row of the raster represents the first occurrence of an action number 1, and the second row represents the second occurrence of action number 1 (i.e. trials one and two, respectively). The inset depicts the averaged response across action numbers 1-20 ± SEM. Note that the neuron most strongly prefers actions 1 and 2, which is reflected in the tuning curve (inset) and the spike density in the raster.

ACTIONS EVOKED MODEST INCREASESES IN POPULATION ACTIVITY

Neuronal responses were divided into those from dopamine and non-dopamine neurons and population averaged activity was aligned to a 0.250 sec window centered on the time of action execution. In this analysis, all actions were considered together, without respect to action number. Action execution evoked a modest increase in neuronal activity. The magnitude of the evoked population responses was not statistically different between sessions or dopaminergic and non-dopaminergic neurons (Figure 3.7 A, B; main effect of session, $F_{(6,361)} = 0.919$, $p = 0.481$; main effect of neuron type, $F_{(1,361)} = 2.204$, $p = 0.139$; interaction, $F_{(6,361)} = 0.660$, $p = 0.682$). It should be noted a phasic increase in the activity of dopamine neurons was present in the first session, though this was not significantly greater than other sessions (Figure 3.7 A). When single neurons were examined, suppressed firing rates were observed in only a minority of single neurons (Figure 3.7 C). The most common response evoked by action execution was activation (Figure 3.7 C). The proportion of dopaminergic or non-dopaminergic neurons that were activated in the peri-action window did not differ across sessions (Figure 3.7 C; dopamine neurons, $X^2_{(6)} = 4.323$, $p = 0.633$; non-dopamine neurons, $X^2_{(6)} = 2.318$, $p = 0.888$). Too few neurons were suppressed in the peri-action window to permit reliable statistical analysis. Taken together, these data suggest that action execution, on the whole, weakly increased population averaged activity via consistent proportions of dopaminergic and non-dopaminergic neurons being activated during action execution.

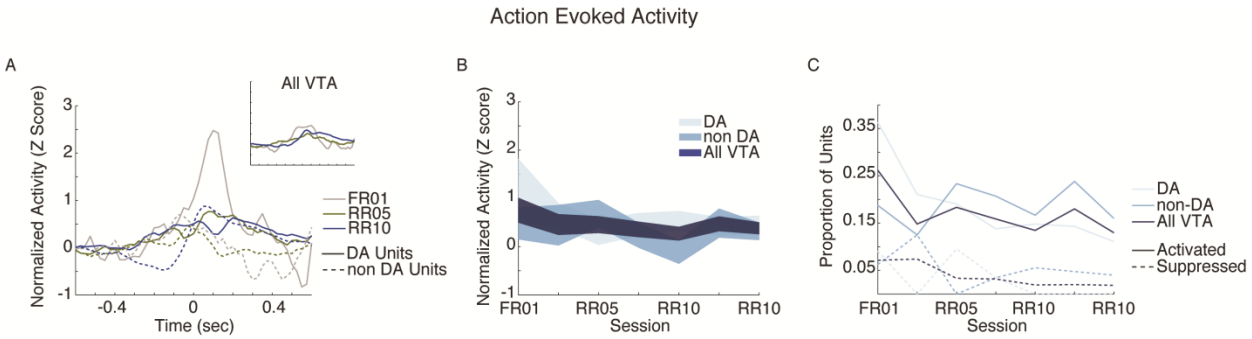

Figure 3.7. Action-Evoked Neuronal Activity.

(A) The mean population response evoked by action execution in select sessions. Data are depicted for the FR01 session (session 1) the final RR05 session (session 4) and the final RR10 session (session 7). All actions in all trials were utilized for these analyses. The main figure depicts the normalized population response for all dopamine units (solid line) and non-dopamine units (dashed line), aligned to the time of action execution (0.025 sec bin). Inset depicts data plotted identically, for all neurons grouped together (both dopamine and non-dopamine neurons). The same axes are used in the inset. The legend for (A) appears to the lower right of that panel. (B) Mean ± SEM neuronal response evoked by action execution. Each unit's data were averaged inside a 0.250 sec time window centered on the action. Data are depicted separately for all putative dopamine and non-dopamine neurons, as well as all VTA units pooled together. Note that in each grouping of units, the evoked population response was stable across sessions, with no difference between groups in action evoked neuronal response (main effect of session, $F_{(6,361)}$ =.919, p = .481; main effect of neuron type, $F_{(1,361)}$ = 2.204, p = .139; interaction, $F_{(6,361)}$ =.660, p = .682). (C) The proportion of units classified as either significantly activated (solid lines) or suppressed (dashed lines) by action execution. Data are depicted across all sessions, for putative dopamine, non-dopamine and all VTA neurons. Note that for all groupings of neurons, suppression was rarely evoked and the proportion of activated neurons did not change across sessions (All VTA units, $X^2_{(6)}$ =3.977, p = .680; dopamine units, $X^2_{(6)}$ =4.323, p = .633; non-dopamine units, $X^2_{(6)}$ =2.318, p = .888). Note that the legend for (C) appears to the lower right.

SERIAL ACTIONS EVOKE UNIQUE ACTIVITY PATTERNS IN DIFFERENT VTA NEURONS

The previous analysis examined how actions modulated neuronal activity, but did not elucidate how neuronal activity was modulated throughout a series of actions. To understand this, neuronal activity was aligned to the time of action execution and examined as a function of action number within a trial (e.g. all first, second, or third actions that occurred across all trials). The random ratio 10 sessions were used for this analysis, as this reinforcement schedule resulted in greater average numbers of actions per trial, and larger samples of each action number. The random nature of reinforcement resulted in a geometric distribution of the number of times that each action number would be required before outcomes were delivered. Thus, higher numbered actions occurred less frequently than lower numbered actions. In a small number of trials, very high numbered actions were required (e.g. 50 or 60 actions). To ensure that a sufficient number of observations of each action number were collected, neuronal activity was only analyzed as a function of action number for actions 1 – 20.

Examination of individual neurons revealed that few had firing rates that were uniformly modulated by the execution of all actions (Figure 3.8). Instead, unique subsets of actions activated or suppressed individual neurons (Figure 3.8). For instance, the pair of simultaneously recorded neurons in the top row of Figure 3.8 preferred only the lowest numbered actions and the highest numbered actions, respectively (Figure 3.8). Some neurons, such as the pair in the middle row of Figure 3.8, were activated by a larger number of actions. Other neurons preferred more complex combinations of actions, such as the pair of simultaneously recorded neurons in the

bottom row of Figure 3.8. Note that in Figure 3.8, the most obvious patterns of modulation are activation, and not suppression, around individual actions. These examples were chosen to illustrate the diversity of neuronal responses, and the non-uniformity of firing rate modulation across action numbers. In some cases, neurons had firing rates that increased at some actions and decreased during others (see below). The fact that heterogeneous patterns of activation and suppression were observed amongst neurons recorded simultaneously, suggests that these diverse patterns of neuronal activity were not attributable to behavioral idiosyncrasies, attention, motivation, or response vigor. Because heterogeneous activity patterns were present in multiple animals, no single animal could account for the amount of diversity observed in the entire population.

One alternative possibility is that neuronal activity could also reflect how much time had elapsed since the start of each trial. As elapsed time and action number are correlated, it may be impossible to discount this interpretation entirely. In order to gain insight into this possibility, action evoked spike counts were regressed on these two predictor variables in a generalized linear model (Poisson regression). In this model, action number was the sole significant predictor of spike count in 29% (45/156) of the neurons (Figure 3.9). In contrast, time elapsed since trial start was a significant predictor of spike count in only 7% (11/156) of the neurons (Figure 3.9). Owing to the fact that action number and time elapsed are correlated, both variables were significant predictors of spike in an additional 22% (35/156) of the neurons (Figure 3.9). Thus, time elapsed since trial start rarely modulated spike count by itself, while action number modulated the activity of a much larger proportion of the population. The results of this regression analysis strongly suggest that the aforementioned action evoked neuronal activity was representing something related to action number, and not time elapsed in a trial, in the majority

of neurons firing selectively to subsets of action numbers. The fact time elapsed since trial start was seldom a strong predictor of neuronal activity, and most frequently predicted spike counts only when action number also was a strong predictor of neuronal activity, further buttresses the notion that action number, but not time, modulated neuronal activity. This likely owes to the fact that actions, and not time, were causally related to outcome delivery. However, these data should be interpreted carefully, as action number and time elapsed are somewhat confounded. Taken together, these data suggest that differently numbered actions modulated VTA neuronal activity in a heterogeneous fashion.

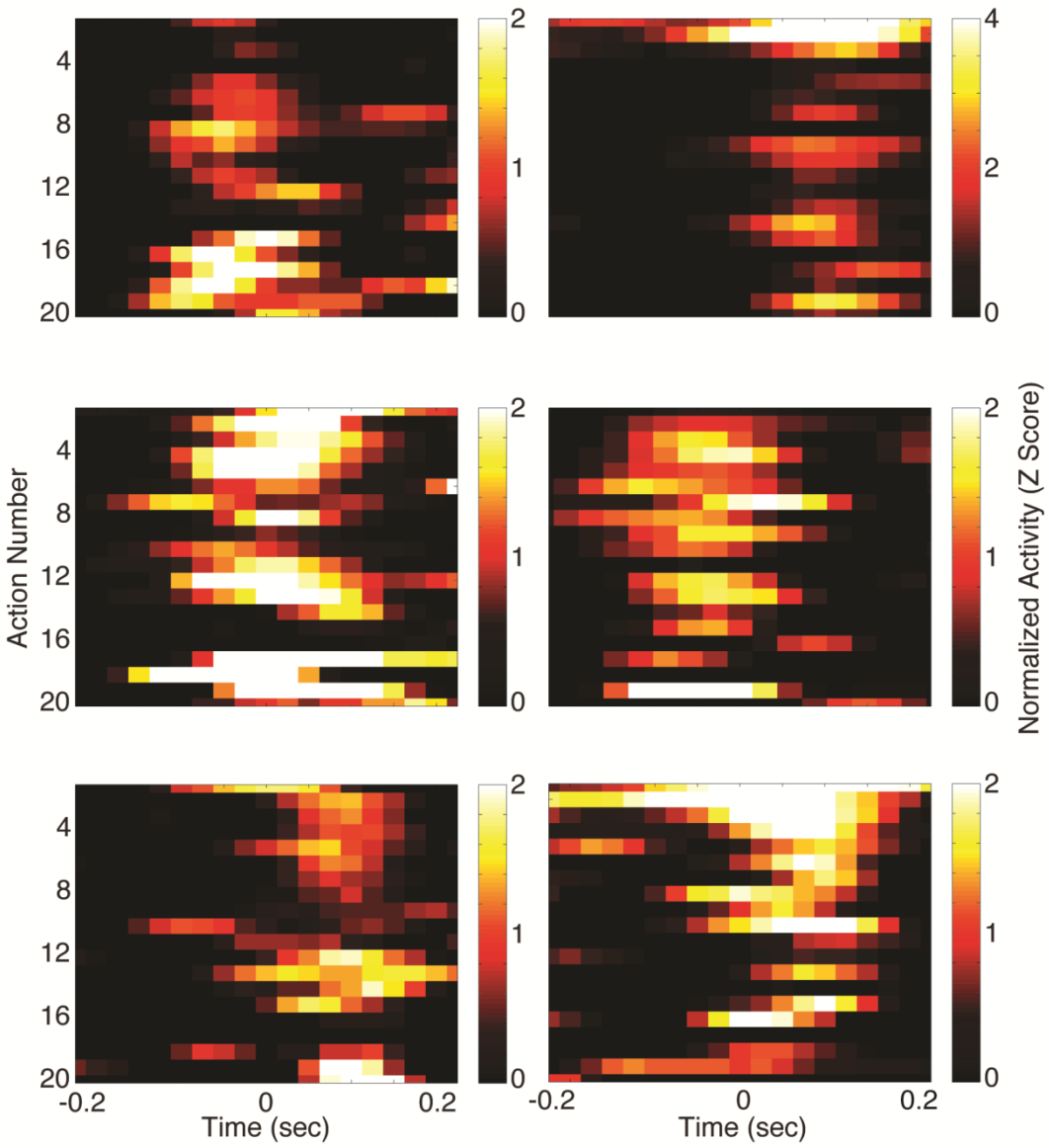

Figure 3.8. Example Neuronal Responses Around Actions 1 - 20.

Normalized firing rate is plotted as a function of color, and aligned to the time of action execution (0.025 sec bins). Data are depicted as the average response evoked by each nth action, across all occurrences of that action. Neurons responded to unique subsets of action numbers. The neurons depicted in each row were recorded simultaneously, and each row represents a pair from a different sessions. These 3 pairs of neurons were chosen to highlight the diversity of responses that were found in the data, and are highly representative of the patterns of activation found throughout the dataset. For simplicity, bi-directionally modulated or suppressed neurons are not depicted but, were present in the dataset.

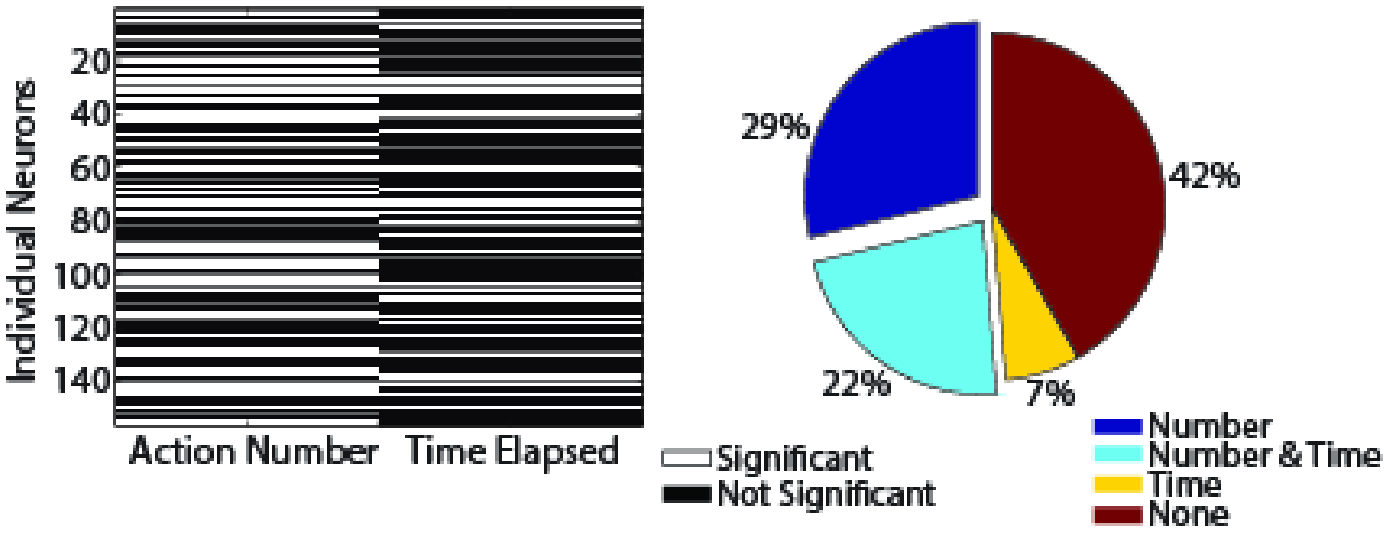

Figure 3.9. Action Number and Elapsed Time as Predictors of Neuronal Activity.

Action evoked spike counts from each neuron were regressed on action number and time elapsed since trial start (Poisson regression). This model was utilized to determine what proportion of neurons activity had significantly predicted by action number or time. Color plot indicates the significance (white versus black) of each predictor variable for each neuron in dataset. Each row represents one neuron. Note that more neurons have activity that is predicted by action number (left column), and that very few neurons have activity predicted solely by elapsed time (right column). A summary of the data is presented in a pie chart, which depicts the percentage of units with activity predicted only by action number (dark blue), only elapsed time (yellow), both variables (cyan), or neither (brown). Note that over half the units have activity predicted by action number, and that time alone only predicts the activity of a small percentage of neurons.

To visualize each neuron's tuning curve as a function of action number, neuronal activity for every neuron that was recorded during the random ratio 10 sessions are plotted (Figure 3.10 A). Examination of scaled tuning curves (smallest evoked response equal to zero and the largest evoked response equal to one) indicated individual neurons had firing rates that were maximized and minimized around different subsets of actions (Figure 3.10 A). Further, each action was associated with unique patterns of activation and suppression of activity amongst different neurons (Figure 3.10 A). For clarity, unscaled tuning curves are plotted similarly to Figure 3.10 A. Note that while transforming the neuronal data to scaled tuning curves improves visualization of these effects, they are also apparent without this transformation (Figure 3.10 B). Diverse activity patterns occurred in both dopaminergic and non-dopaminergic neurons (Figures 3.11 A, B). Tuning curve peak to valley depths did not differ by neuron type (Figure 3.12 A). Approximately half of the neurons that were recorded were significantly modulated by action number. Of these neurons, most had a single maximum in their tuning curve peaks (Figure 3.12 B). This did not differ by neuron type. Different groups of neurons had tuning curve global maxima in each bin of actions numbers (actions 1-5, 6-10, 11-15, 16-20) with substantial proportions of neurons preferring each action number bin (Figure 3.13). Thus, unexpected levels of heterogeneity and a mosaic of activity patterns were evoked by execution of serial actions. The fact that both populations of neurons carry comparable signals suggests that they both process similar information. These data suggest that together, both populations of VTA neurons transmit a complex, multi-neurotransmitter signal to target brain regions.

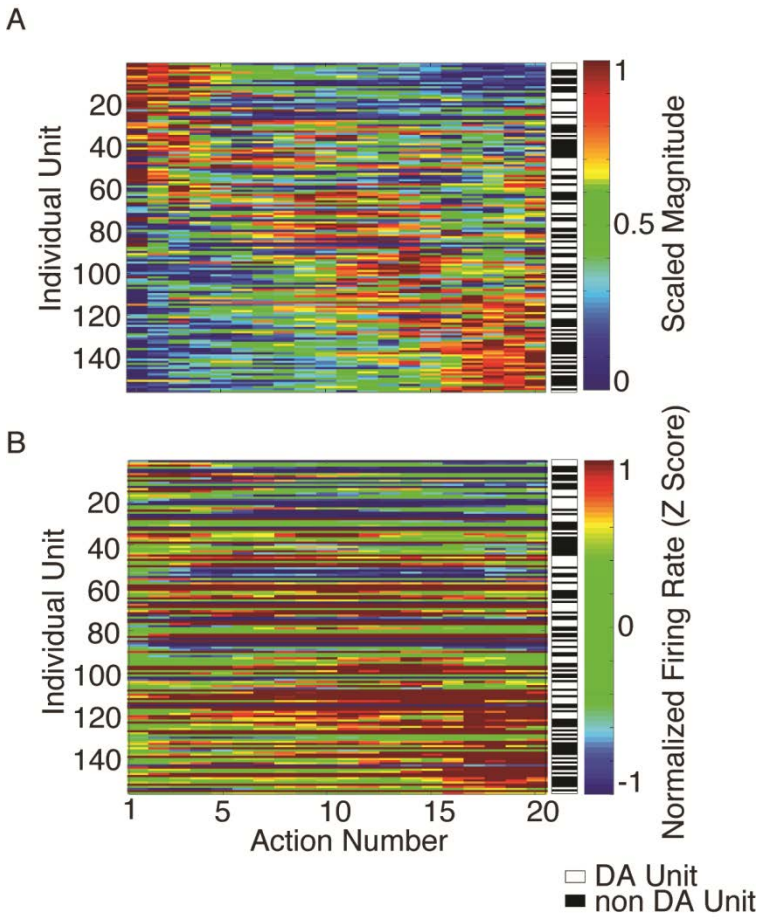

Figure 3.10. Tuning Curves of Neuronal Responses Across Action Number.

(A) Scaled tuning curves of neuronal activity across action number (0.25 sec window centered on action execution). Because neurons had different magnitude activations and suppressions of activity around individual actions, data are scaled so each neuron's full range of evoked firing rates span 0 – 1 (plotted by color). Figures depict neurons recorded from each of the three RR10 sessions. Each neuron is depicted in a separate row. This transformation is done solely for data visualization, and subsequent analyses of neuronal activity as a function of action number utilize raw spike counts. Data are sorted by the location of the peak of the tuning function, with higher action number peaks towards the bottom. Dopaminergic (white) or non-dopaminergic neurons (black) are indicated by inner color bar at the right. Note that different neurons are tuned to prefer different subsets of action numbers. Each action maximizes or minimizes the firing rates of a subset of neurons. (B) Tuning curves of neuronal activity across action number. Data plotted similarly to (A), except that data are displayed un-scaled. Note that while scaling the data improved visualization of the effect, the same trend is present in (B) as in (A).

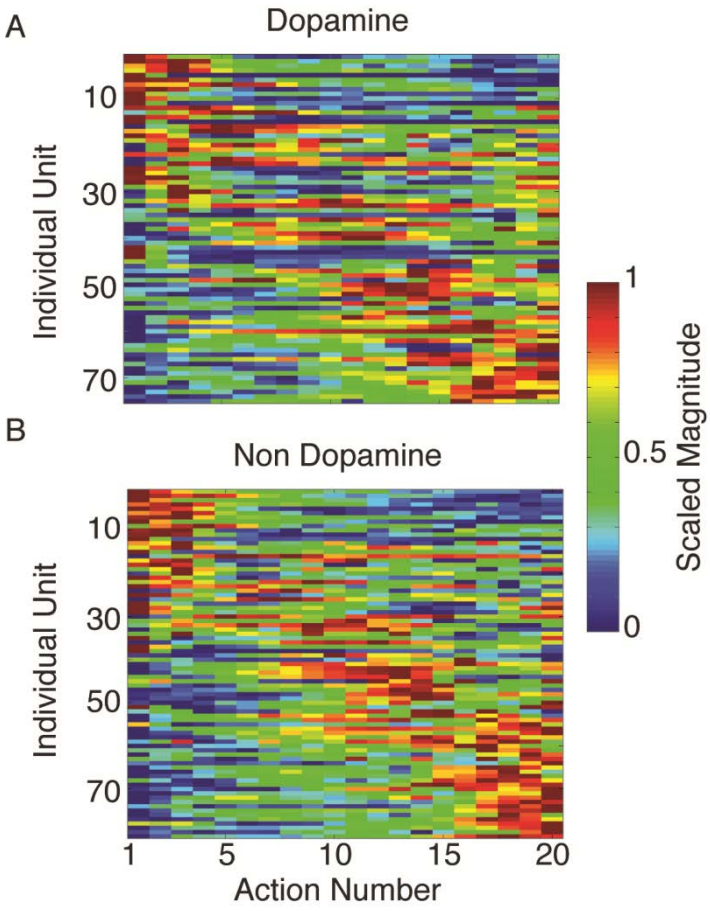

Figure 3.11. Tuning Curves of Dopaminergic and Non-Dopaminergic Neurons.

(A) Scaled tuning curves of neuronal activity evoked by actions (0.25 sec window centered on action execution) across action numbers 1-20, for each dopamine neuron recorded in the three RR10 sessions. Each neuron is depicted in a separate row. Data are sorted by the location of the peak of the tuning function, with higher action number peaks towards the bottom. Note that even amongst only dopamine neurons there is an unexpected degree of tuning function heterogeneity. (B) Data depicted identically for non-dopamine neurons. Note the similarity between (A) and (B), which suggests that both pools of neurons encode similar information about serial actions.

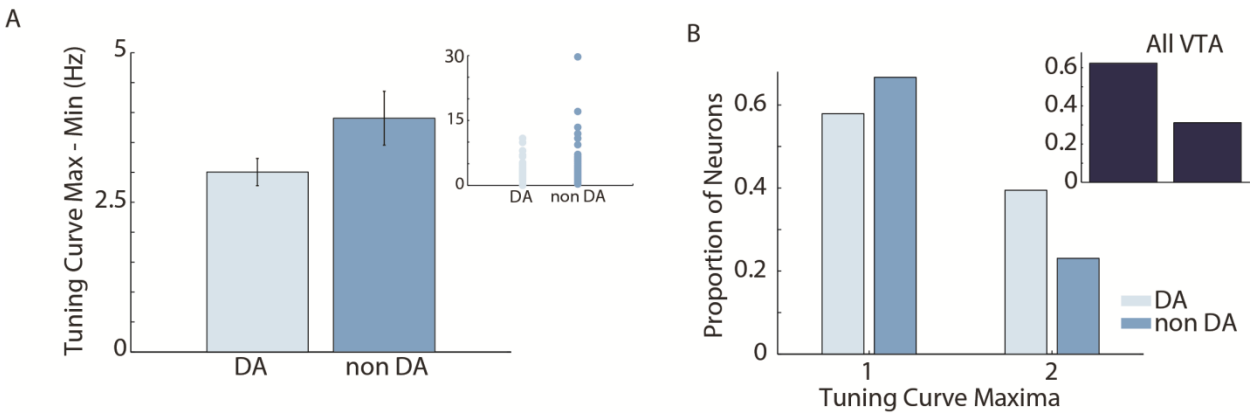

Figure 3.12. Quantification of Tuning Curve Attributes.

(A) Mean ± SEM difference between each tuning curve's maximum and minimum. Inset depicts the entire distribution of values, which approached, but did not reach statistical significance ($t_{(154)} = 3.224$, $p = 0.084$). (B) The proportion of neurons that were significantly modulated by binned action number with either 1 or 2 local maxima in their tuning curve. 49.4% of neurons were significantly modulated by action number, with no effect of neuron type ($\chi^2_{(1)} = 0.099$, $p = 0.753$). Most neurons had a single maxima in their tuning curve ($\chi^2_{(1)} = 8.00$, $p = 0.005$), with no effect of neuron type ($\chi^2_{(1)} = 1.779$, $p = 0.182$). Also note that amongst neurons that were significantly modulated by action number, a smaller but substantial proportion contained two maxima in their tuning curves. The inset depicts all VTA neurons (dopamine and non-dopamine) plotted identically to main figure. Note the same trend is present.

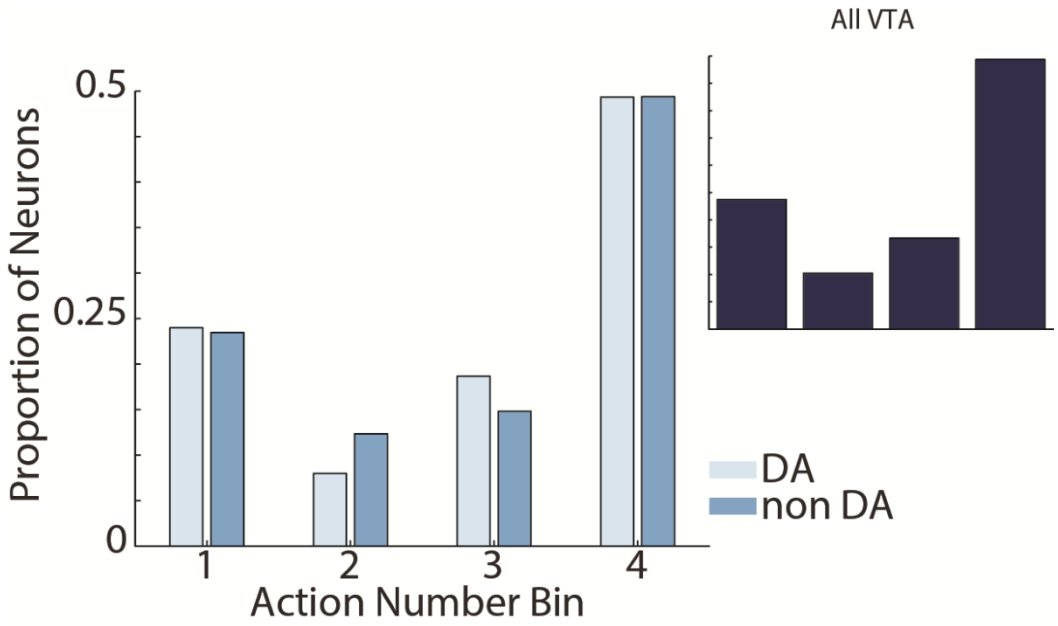

Figure 3.13. Location of Tuning Curve Global Maxima.

Data depict the proportion of neurons with a tuning curve global maximum in each action number bin. The main figure presents this analysis for dopamine and non-dopamine neurons, separately. Note the similarity of each population ($\chi^2_{(3)} = 1.069$, $p = 0.785$) and note that each action number bin contains a sizeable proportion of neurons that fire most preferentially for actions in that bin. This distribution was non-uniform ($\chi^2_{(3)} = 55.026$, $p < 0.001$), with more neurons preferring the highest numbered actions. The inset to the upper right depicts the same data for each VTA neuron. Note the similarity in the data.

Many patterns of action-evoked activity were present in the dataset. Functional principal components analysis was utilized to inspect the tuning curves in low dimensional space. Each tuning curve is plotted as a single point against functional principal components 1 and 2 (Figure 3.14). The first principal component is the linear vector that accounts for the most variation between tuning curves. The second principal component accounts for the second most variation between tuning curves and is orthogonal to the first. These two components accounted for a total of 68% of the variation between tuning curves. Principal components transformation did not produce strong clustering of the data (Figure 3.14). Taken together with the variety of tuning curves that were observed, these data suggest there is not a preponderance of any particular shape of tuning curve in the dataset. There is a diverse array of tunings to action number in the current data. It bears mentioning that the lack of clear clustering in these data must be interpreted cautiously. These data could be visualized along another set of dimensions that may reveal a pattern that is not readily apparent in function principal component space. Alternatively, if additional recordings were added to the current dataset, a more robust clustering that is not currently apparent could emerge.

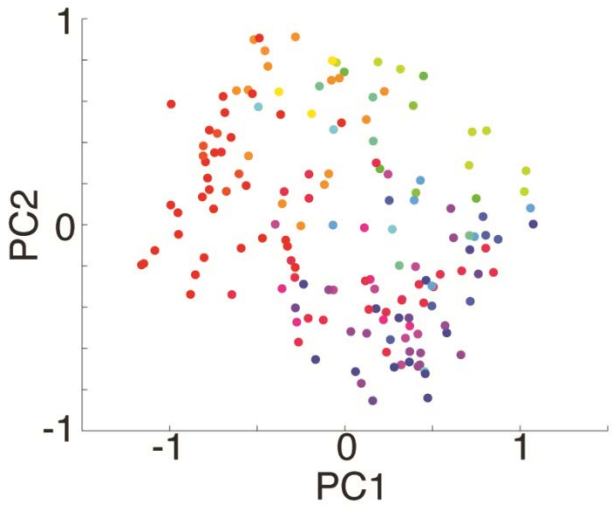

Figure 3.14. Functional Principal Components Transformation of the Tuning Curves.

Each tuning curve plotted in lower dimensional functional principal component space to determine if some aspect of the tuning functions of these neurons appeared regularly. Because tuning curves across action number were inherently high dimensional, functional principal components analysis was utilized to plot each scaled tuning curve in a two dimensional space that would capture a large proportion of the variance between tuning curves. Each point represents one neuron's tuning curve. The weightings of the first two principal components are plotted for all neurons from all RR10 sessions. Color represents the location of the peak of the tuning function (hot colors represent peaks at higher action numbers). Note that there is not a clear clustering of points, but rather points are scattered, indicating that there is no preponderance of similar tuning functions. The first principal component explains the most variability between tuning curves. This dimension correlates highly with whether or not the tuning functions are maximal at low numbered action numbers.

POPULATION AVERAGE ACTIVITY IS INVARIANT WITH ACTION NUMBER

Most analyses of VTA or dopaminergic neuronal activity focus on population averaging. As a first step, this traditional approach was utilized to examine how information may be processed during the execution of serial actions. As expected from the previously described exploratory data analyses, averaging the action evoked neuronal activity in either group of neurons concealed the unique firing patterns of these neurons, and produced modestly elevated population activity (Figures 3.15 A, B). There was not a significant difference in population activity between action numbers or neuron type in any session (Figure 3.15 A, B) (session 5, main effect of action number, $F_{(3,150)} = 0.877$, $p = 0.454$; main effect of neuron type, $F_{(1,50)} = 1.479$, $p = 0.230$; interaction, $F_{(3,150)} = 0.513$, $p = 0.674$; session 6, main effect of action number, $F_{(3,144)} = 2.185$, $p = 0.092$; main effect of neuron type, $F_{(1,48)} = 0.804$, $p = 0.374$; interaction, $F_{(3,144)} = 0.598$, $p = 0.618$; session 7, main effect of action number, $F_{(3,156)} = 2.163$, $p = 0.095$; main effect of neuron type, $F_{(1,52)} = 0.006$, $p = 0.937$; interaction, $F_{(3,156)} = 0.326$, $p = 0.806$). Dopaminergic neuronal activity was uncorrelated with action number (Session 5, $r = 0.01$, $p = 0.797$; Session 6, $r = 0.06$, $p = 0.250$; Session 7, $r = 0.05$, $p = 0.253$). Similar results were obtained with non-dopamine neurons (Session 5, $r = 0.05$, $p = 0.244$; Session 6, $r = 0.05$, $p = 0.258$; Session 7, $r = 0.02$, $p = 0.633$). Thus, action evoked neuronal activity patterns yielded a population average function that is invariant with action number. See Table 3.1 for a description of what proportion of units were significantly activated or suppressed by each action bin. The results detailed so far lead to the conclusion that actions modulate the activity of individual neurons in a unique fashion, and the population average is not likely to convey information about ongoing action number.

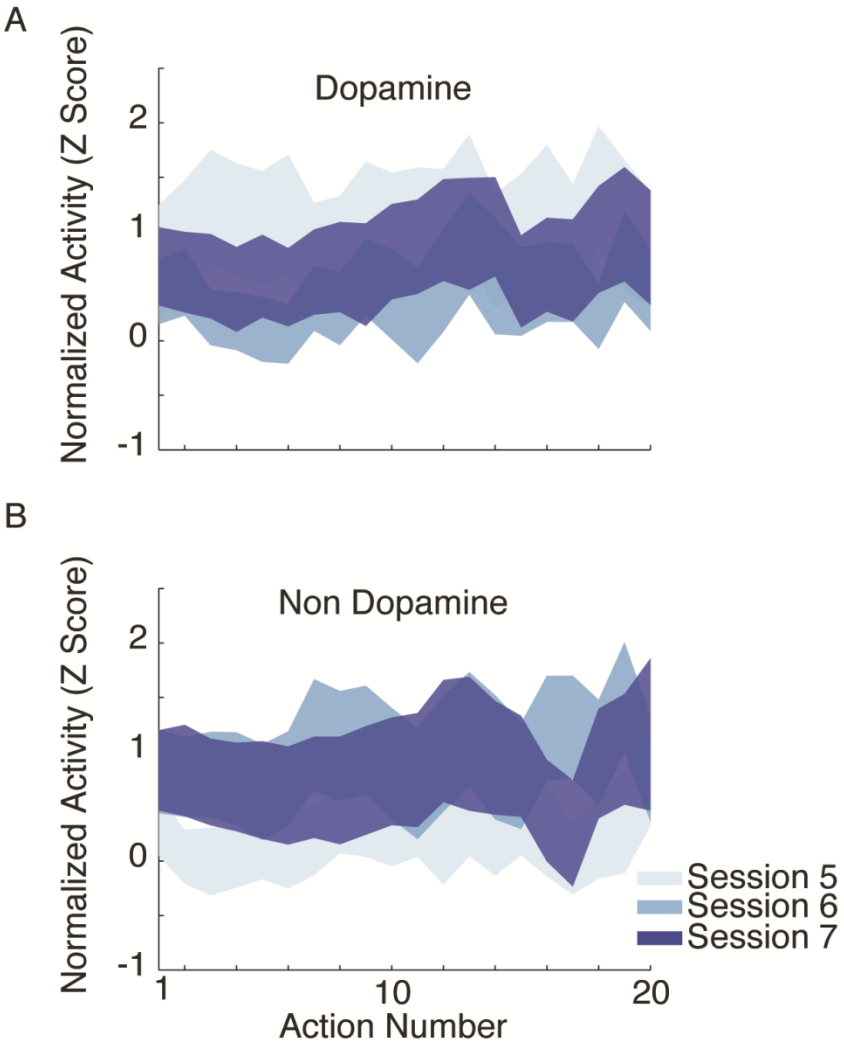

Figure 3.15. Population Averaged Activity Evoked by Actions 1 – 20.

(A) The mean dopaminergic neuronal response as a function of action number. Data are depicted in all RR10 sessions (sessions 5- 7), for actions 1 – 20 (0.25 sec window centered on action execution). Data are depicted as mean ± SE. Note that the population average is invariant with action number. These data suggest averaged activity is not likely to encode action number. (B) Non-dopamine neuronal responses to serial actions. Data plotted identically to (A). Note that similar results were obtained in non-dopamine neurons.

			Activated						Suppressed	
Actions		**1-5**	**6-10**	**11-15**	**16-20**		**1-5**	**6-10**	**11-15**	**16-20**
Session 6		0.17	0.19	0.25	0.23		0.12	0.10	0.13	0.12
Session 7		0.22	0.22	0.20	0.28		0.06	0.06	0.06	0.06
Session 8		0.24	0.17	0.24	0.20		0.06	0.09	0.07	0.06

Table 3.1. Proportion of Neurons Activated by Different Action Numbers.

The proportion of neurons that were activated (left of table) or suppressed (right of table) by each bin of 5 consecutive action numbers is listed for each of the RR10 sessions. Note that for each of the 4 action bins, approximately one quarter to one third of the neurons had significantly activated or suppressed action-evoked activity levels. Though it is not depicted explicitly in this table, please note that there is some overlap between neurons that were significantly modulated by each action bin. For example, some neurons were significantly activated by more than one bin. However, there are a large number of neurons that are selectively modulated for each bin of actions. Also note that activation was the most common activity pattern evoked by a group of actions.

ACTION NUMBER IS ACCURATELY DECODED FROM VTA ENSEMBLE ACTIVITY, BUT NOT THE POPULATION AVERAGE

The heterogeneity of neuronal activity evoked by action execution suggested VTA could encode information about actions via the collective activity of ensembles of neurons. Actions were reinforced randomly, so animals could not predict which action would be reinforced. Each action was equally likely to be reinforced. Thus, individual actions did not differ in terms of value, but instead differed only by action number. These observations prompted us to quantify the degree to which action number could be decoded from the activity of many differently tuned neurons. As a point of comparison the population average was used to decode action number as well.

The same basic approach was utilized in all decoding analyses. A basic description of these analyses is detailed here, and some aspects of the approach are omitted for the sake of clarity. See the Methods Section for a more detailed account of the approach. Two different decoders were utilized in the current work (discussed below). For these analyses, four categories that corresponded to five consecutive action numbers were created (actions 1 - 5, 6 - 10, 11 - 15, and 16 - 20). Each decoder classified observed action-evoked spike counts as belonging to one of the aforementioned four possible categories. Thus, decoder performance depended on how well action number corresponded to firing rate. All the observed firing rates were divided into two datasets. One dataset, the training dataset, was used to build expectations about how firing rate related to action number. The decoder classified the remaining data, which is referred to as the test dataset. Each decoder estimated the probability that an observed firing rate belonged to the

four classes of actions, based on the relationship between firing rates and category in the training data. The decoder classified an each observation of a firing rate as the most likely category of action to have evoked that response.

The first decoder, the population average decoder, used the averaged activity of VTA neurons to classify actions (Figure 3.16 A). This decoder does not take advantage of the fact the VTA neurons were diversely tuned for action number, because it utilizes the average evoked response. A second decoder, the naïve Bayesian ensemble decoder, simply considered each neuron as independent (Figure 3.16 B). This decoder did not assume any structure between neurons, and did not average or otherwise collapse together the activity of multiple neurons into a single quantity. This decoder capitalizes on the fact that neurons are differently tuned to action number. By doing so, this decoder takes into account the collective activity of ensembles of VTA neurons. In the case of the naïve Bayesian ensemble decoder, the process of assessing the probability that an observed firing rate belonged to each category was repeated for every neuron. Within each class, these probabilities were multiplied, and the decoder estimated the observed action class as the class with the highest resulting product of these probabilities. The population average decoder (Figure 3.16 A) serves as a useful point of comparison against the naïve Bayesian ensemble decoder (Figure 3.16 B), as it does not capture the collective activity of the ensemble.

The population average decoder did not decode action number accurately (Figure 3.16 C), as most actions (12 of 20) were correctly decoded at or below chance levels (Table 2.1). Only actions 1-6, and 13-14, were correctly decoded more frequently than chance levels (Table 2.1). These results are consistent with the interpretation that that the population average obscures the unique tuning of each neuron. By averaging each neuron's activity, unique action-evoked

patterns of activity are mostly cancelled out, so that the population average does not contain substantial information about action number. In sharp contrast to the performance of the population average decoder, the naïve Bayesian ensemble decoder correctly classified 14 of 20 action numbers, actions 1-4 and 6-15, above chance levels (Figure 3.16 C, Table 2.1). Only the highest numbered actions were classified below chance levels (Table 2.1). The naïve Bayesian ensemble decoder performed significantly better than the population average decoder (Figure 3.16 C; $\hat{\beta} = 0.914$, $t_{(19)} = 26.620$, $p < 0.001$). Thus, simple conceptualizations of VTA ensemble activity, independent activity between dopaminergic and non-dopaminergic neurons, signal information about ongoing action number. These data suggest that, unlike the population average, the collective activity of VTA ensembles is a viable signal for post-synaptic networks to decode action number from. The current data demonstrate a novel and surprising form of information processing by VTA neurons. It should also be noted that both decoders, and especially the naïve Bayesian decoder, performed better at low numbered versus high numbered actions. This is because the variance in spike counts evoked by higher numbered actions was greater, possibly due to the decreased sample size associated with higher numbered actions (Figure 3.17).

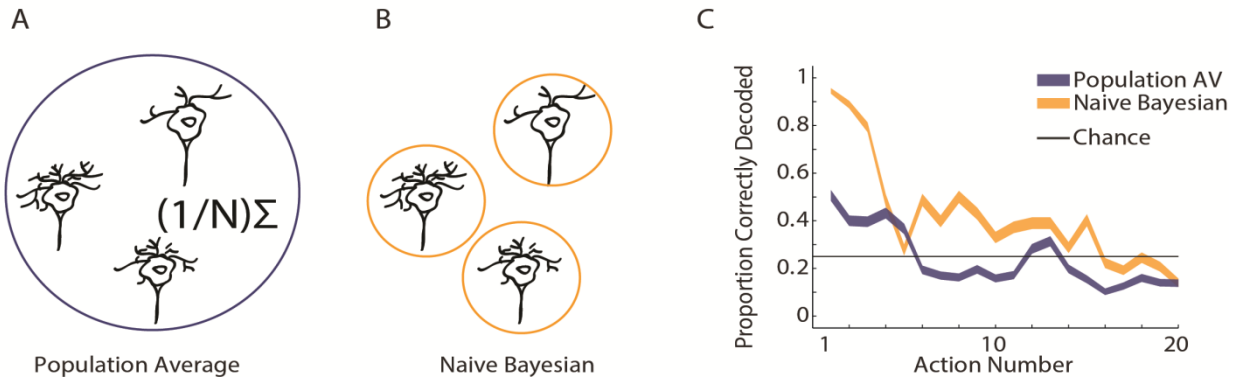

Figure 3.16. Decoding Analyses.

(A) The population average decoder classified activity that was averaged across neurons, and does not take advantage of the unique patterns of activity evoked in each neuron. (B) The naïve Bayesian ensemble decoder used each neuron's action-evoked activity to classify test data (e.g. no averaging), and the decoder is able to take advantage of the diverse tunings of each neuron. (C) Decoding accuracy for actions 1-20. Data are depicted as the mean ± SEM proportion of correction classifications for each decoder. Chance levels (0.25) of correctly classifying actions are depicted as the solid black line. Note that the naïve Bayesian ensemble decoder was significantly more accurate than the population decoder ($\widehat{\boldsymbol{\beta}} = \mathbf{0.914}$, $t_{(19)} = 26.620$, $p < 0.001$).

Decoder		Population Average	Naïve Bayesian
Action 1		P < 0.001, Above Chance	P < 0.001, Above Chance
Action 2		P < 0.001, Above Chance	P < 0.001, Above Chance
Action 3		P < 0.001, Above Chance	P < 0.001 , Above Chance
Action 4		P < 0.001, Above Chance	P < 0.001, Above Chance
Action 5		P < 0.001, Above Chance	P = 0.113, NS
Action 6		P = 0.002, Above Chance	P < 0.001, Above Chance
Action 7		P < 0.001, Below Chance	P < 0.001, Above Chance
Action 8		P < 0.001, Below Chance	P < 0.001, Above Chance
Action 9		P = 0.003, Below Chance	P < 0.001, Above Chance
Action 10		P < 0.001, Below Chance	P < 0.001, Above Chance
Action 11		P < 0.001, Below Chance	P < 0.001, Above Chance
Action 12		P = 0.056, NS	P < 0.001, Above Chance
Action 13		P < 0.001, Above Chance	P < 0.001, Above Chance
Action 14		P = 0.004, Above Chance	P = 0.049, Above Chance
Action 15		P < 0.001, Below Chance	P < 0.001, Above Chance
Action 16		P < 0.001, Below Chance	P = 0.112, NS
Action 17		P < 0.001, Below Chance	P = 0.004, Below Chance
Action 18		P < 0.001, Below Chance	P = 0.435, NS
Action 19		P < 0.001, Below Chance	P = 0.027, Below Chance
Action 20		P < 0.001, Below Chance	P < 0.001, Below Chance
Totals		Above (8/20) Below (11/20)	Above (14/20) Below (3/20)

Table 3.2. Decoder Performance Compared to Chance Levels.

Each decoder's correct classification rate was assessed against chance levels. For each decoder and action number, the resulting probability of obtaining a value as extreme, or more extreme, than chance levels (0.25) is listed in the cells of the table. If probability of correct classification is statistically above or below chance, this is listed in each cell, as are non-significant differences from chance (NS). Summed totals are listed in the bottom row. Note that the population decoder performs below chance levels for the majority of actions. In contrast the naïve Bayesian ensemble decoder performed above chance levels for the majority of actions.

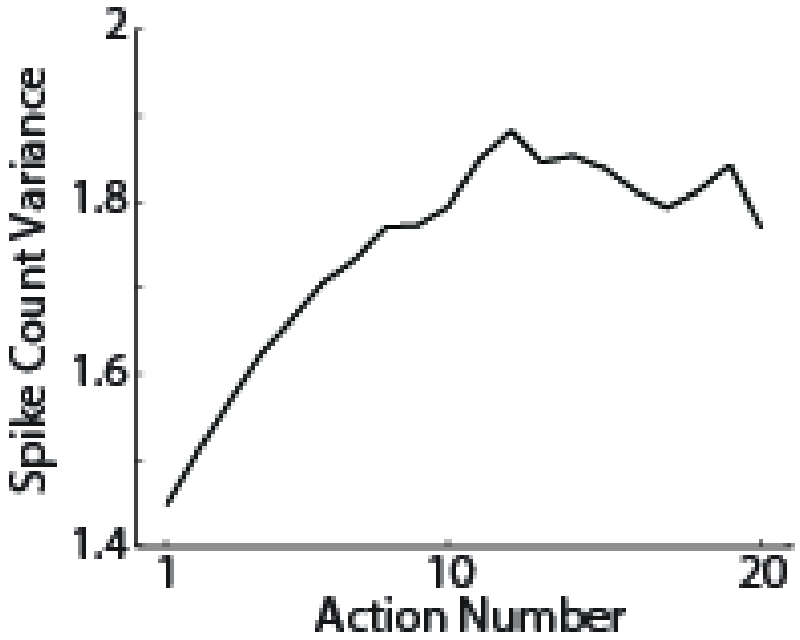

Figure 3.17. Spike Count Variance as a Function of Action Number.

Data depicted as the variability in spike count evoked by differently numbered actions. Each neuron's variability in action evoked spike counts was averaged across neurons to yield population averaged variability. Note the increase in variability as action number increases, which likely accounts for the decreased decoding accuracy at higher action numbers.

SUMMARY OF ACTION EVOKED NEURONAL DATA

In general, actions modestly increased neuronal activity, and individual neurons were tuned to limited and divergent subsets of actions. Action number was accurately decoded by considering the unique action-evoked activity of each neuron, without modeling any interaction between these neurons, and treating them as independent. The population-averaged activity was invariant with action number and offered only poor decoding of action number. These results suggest the collective activity of groups of VTA neurons signals real-time information about action number.

CUE-EVOKED NEURONAL ACTIVITY ABSENT IN RANDOM RATIO SESSIONS

To determine how environmental stimuli that predict outcome availability modulate neuronal activity in the task, neuronal responses that were evoked by the onset of the cue light at the start of each trial were examined. In the FR01 session (session 1), cue onset evoked a strong phasic increase in neuronal activity (Figure 3.18 A, B). The phasic population average response evoked by cue onset significantly diminished in all subsequent recording sessions (Figure 3.18 A, B; main effect of session, $F_{(6,361)} = 2.667$, $p = .015$; all post hoc tests versus session 1 $p < 0.05$). Cue evoked responses in sessions 2 – 7 did not differ from each other (all post hoc tests $p > 0.05$). There was no difference in response magnitudes between dopamine neurons and non-dopamine neurons, and no interaction between neuron type and session (main effect of neuron type, $F_{(1,361)} = 1.543$, $p = .215$; interaction, $F_{(6,361)} = .493$, $p = .814$). Activation was the most common pattern of activity that was evoked by cue onset, and suppressed firing rates were rarely evoked (Figure 3.18 C). The proportion of dopaminergic neurons activated by cue onset decreased across

recording sessions (Figure 3.18 C; $X^2_{(6)} = 20.109$, $p = .003$). There was no change in the number of non-dopamine neurons activated by cue light onset across sessions (Figure 3.18 C; $X^2_{(6)} = 10.636$, $p = .100$). This may be due to a floor effect as a smaller proportion of non-dopaminergic neurons were activated by cue onset in the initial session. Thus, in the initial FR01 recording session, cue onset evoked a phasic response that diminished in subsequent random ratio reinforcement schedule sessions, and was present in decreasing number of dopaminergic neurons. Because minimal numbers of VTA neurons responded to cue onset in random ratio sessions, cue evoked neuronal responses are not likely to guide the execution of serial actions. These data further highlight the importance of the previously mentioned action-evoked neuronal responses, in terms of encoding information about ongoing behaviors.

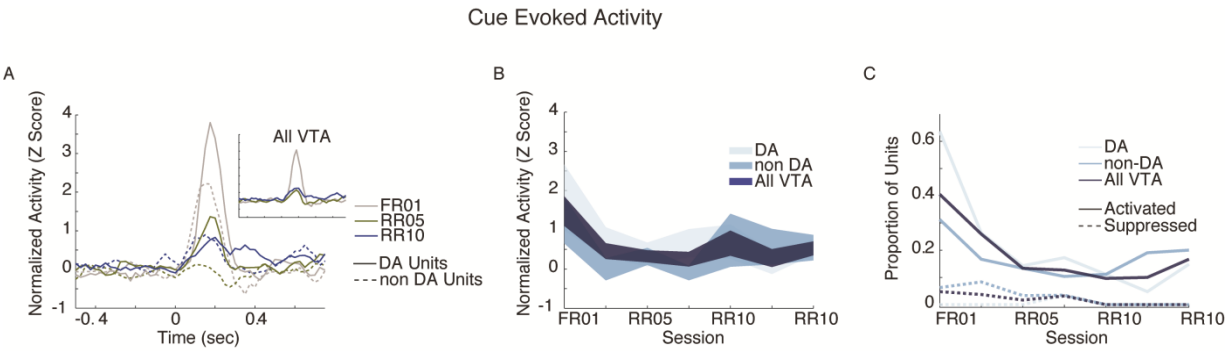

Figure 3.18. Cue-Evoked Neuronal Responses.

(A) The mean population response evoked by cue light onset is depicted for the FR01 session (session 1) the final RR05 session (session 4) and the final RR10 session (session 7). The main figure depicts the normalized population response for all dopamine neurons (solid lines) and non-dopamine neurons (dashed lines), aligned to the time of cue light onset. Inset depicts data plotted identically, for all neurons grouped together. (B) Mean ± SEM neuronal response evoked by cue onset, across all sessions. Each neuron's data were averaged across a time window +0.05 - +0.3 sec, relative to cue onset. Data depict responses of putative dopamine and non-dopamine neurons, as well as all VTA units grouped together. Note that in each grouping of units, the evoked population response was strongest in session 1 and significantly declined in subsequent sessions (main effect of session, $F_{(6,361)} = 2.667$, $p = .015$). There was no difference in response magnitudes between dopamine units and non-dopamine units, and no interaction between neuron type and session (main effect of neuron type, $F_{(1,361)} = 1.543$, $p = .215$; interaction, $F_{(6,361)} = .493$, $p = .814$). (C) The proportion of units classified as either significantly activated (solid lines) or suppressed (dashed lines) by cue light onset are depicted across all sessions, for putative dopamine, non-dopamine, and all VTA neurons. Note that suppression was rarely evoked. The proportion of activated neurons decreased in later sessions (All VTA units, $X^2_{(6)} = 23.844$, $p = .001$; dopamine neurons, $X^2_{(6)} = 20.109$, $p = .003$). There was no change in the number of non-dopamine neurons activated by cue light onset across sessions ($X^2_{(6)} = 10.636$, $p = .100$).

PERI-OUTCOME NEURONAL DATA

In each trial, once animals performed a reinforced action, the cue light was immediately extinguished, there was a half second delay, and then the outcome was delivered to the animal. Dopaminergic and non-dopaminergic population averaged activity was examined around these events in all sessions. For clarity, the mean population activity from selected sessions is plotted in Figure 3.19.

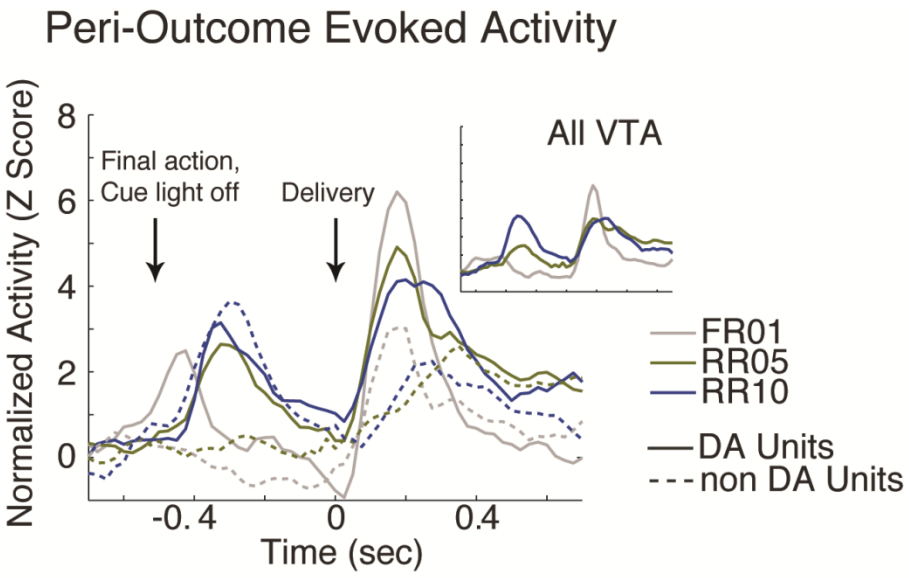

Figure 3.19. Neuronal Responses Aligned to Outcome Delivery.

The normalized population response aligned to outcome delivery (time zero). The final action in each trial (left arrow), which was reinforced, occurred 0.5 sec prior to outcome delivery (right arrow). Cue light offset was simultaneous with execution of the final action, which signaled the completion of the trial and pending outcome delivery. Thus, the animal executed the last action in the trial (left arrow), there was a delay period, and then the outcome was delivered (right arrow). Data are depicted for the FR01 session (session 1) the final RR05 session (session 4) and the final RR10 session (session 7). The main figure depicts the normalized population response for all dopamine units (solid lines) and non-dopamine units (dashed lines), aligned to the time of cue light onset. Inset depicts data plotted identically, for all neurons grouped together.

STABLE OUTCOME EVOKED POPULATION ACTIVITY ACROSS SESSIONS

Neuronal responses evoked by outcome delivery (+0.050 – 0.300 sec) were examined in each session (Figure 3.20 A). The magnitude of the outcome delivery evoked response did not change across sessions and, dopaminergic neurons had larger evoked responses than non-dopamine neurons (Figure 3.20 A; main effect of neuronal type, $F_{(1,361)} = 9.159$, $p = .003$; main effect of session type, $F_{(6,361)} = 1.352$, $p = .233$; interaction, $F_{(6,361)} = 0.276$, $p = .960$). The most common response that was evoked by outcome delivery was activation, and suppressed firing rates were only evoked in a very small number of cases (Figure 3.20 B). Non-systematically varying proportions of non-dopaminergic neurons were activated by outcome delivery (Figure 3.20 B). This variation between sessions reached significance, but the proportion of activated non-dopaminergic neurons does not relate well to changes in reinforcement schedule (Figure 3.20 B; $X^2_{(6)} = 13.234$, $p = .039$). These fluctuations may relate to some aspect of the experimental design that is currently elusive, or could be a spurious effect. The proportion of dopaminergic neurons activated by outcome delivery did not differ between sessions (Figure 3.20 B; $X^2_{(6)} = 3.664$, $p = .722$). These data suggest that outcome delivery evokes robust and stable increases in dopaminergic neuronal activity. Further, these data are compatible with previous reports that outcome delivery evokes dopaminergic activation well into learning (Fiorillo *et al.*, 2003).

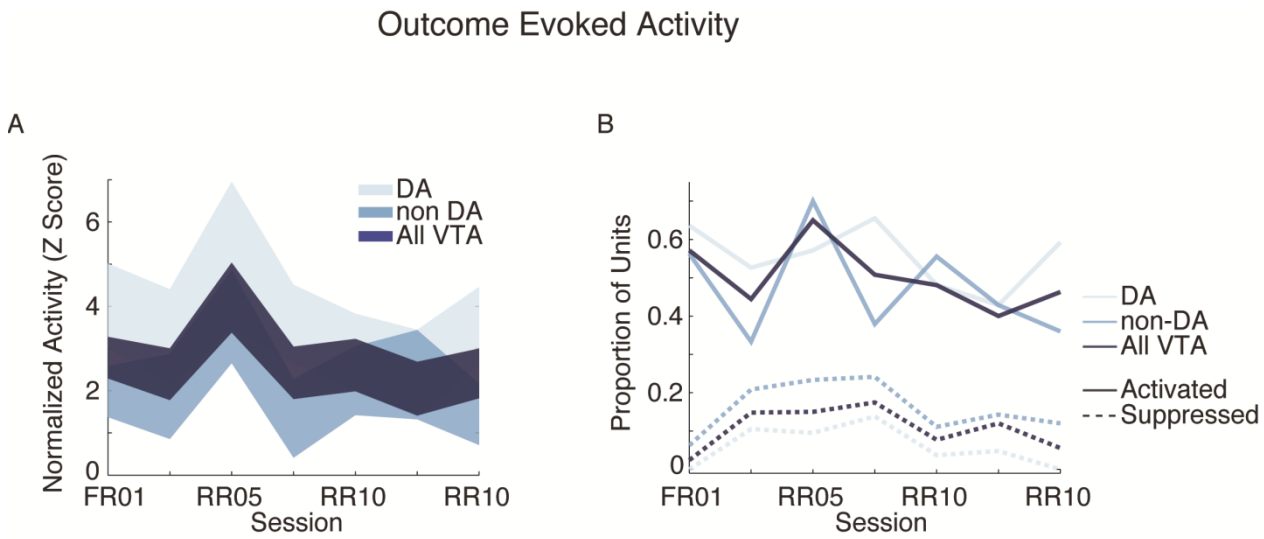

Figure 3.20. Outcome Delivery Evoked Neuronal Activity.

(A) Mean ± SEM neuronal responses evoked by outcome delivery (+0.050 - +0.300 sec) in each session. Data depict responses of putative dopamine and non-dopamine neurons, as well as all VTA units grouped together. Note that dopamine neurons had greater magnitude outcome evoked responses than non-dopamine neurons (main effect of neuronal type, $F_{(1,361)} = 9.159$, $p = .003$; main effect of session type, $F_{(6,361)} = 1.352$, $p = .233$; interaction, $F_{(6,361)} = 0.276$, $p = .960$). (B) The proportion of units classified as either significantly activated (solid lines) or suppressed (dashed lines) by outcome delivery. Data are depicted across all sessions, for putative dopamine, non-dopamine and all VTA neurons. Note that the proportion of dopaminergic neurons that were activated by outcome delivery did not change across sessions (dopamine neurons, $X^2_{(6)} = 3.664$, $p = .722$). There were modest, but significant, fluctuations in the proportion of non-dopamine neurons that were activated in each session (non-dopamine neurons, $X^2_{(6)} = 13.234$, $p = .039$).

DELAY PERIOD NEURONAL ACTIVATION INCREASES WITH LEARNING

Reward prediction errors should be generated by stimuli that precede and are informative of the value of an outcome. In the FR01 session (session 1), only 1 action was required of the animal. In all subsequent sessions, the animal could not predict which action would be the final action, and therefore could not predict when outcomes would be delivered. Thus, immediately following execution of the final action (signaled by cue light offset), or in the delay between the final action and outcome delivery, prediction errors could occur. These responses should emerge with learning. Final action evoked and delay period neuronal responses were analyzed separately.

To begin, each neuron's data were averaged in a 0.250 sec window centered on the final action. The evoked population response did not adapt across sessions (Figure 3.21 A; main effect of session, $F_{(5,321)} = 0.621$, $p = .684$). Dopaminergic and non-dopaminergic neurons had similar responses in all sessions (Figure 3.21 A; main effect of neuronal type, $F_{(1,321)} = 0.230$, $p = .632$; interaction, $F_{(5,321)} = 1.677$, $p = .140$). Suppressed neuronal activity was seldom observed when the final action in a trial was executed (Figure 3.21 B). More, but still relatively modest numbers of neurons, were activated by the final action. The proportion of dopaminergic and non-dopaminergic neurons that were activated by this event did not change across sessions (Figure 3.21 B; dopaminergic neurons, $X^2_{(5)} = 0.575$, $p = .989$; non-dopaminergic neurons, $X^2_{(5)} = 3.526$, $p = .619$). Taken together, these data suggest that the final action in a trial activated a small numbers of neurons, which did not change across recording sessions.

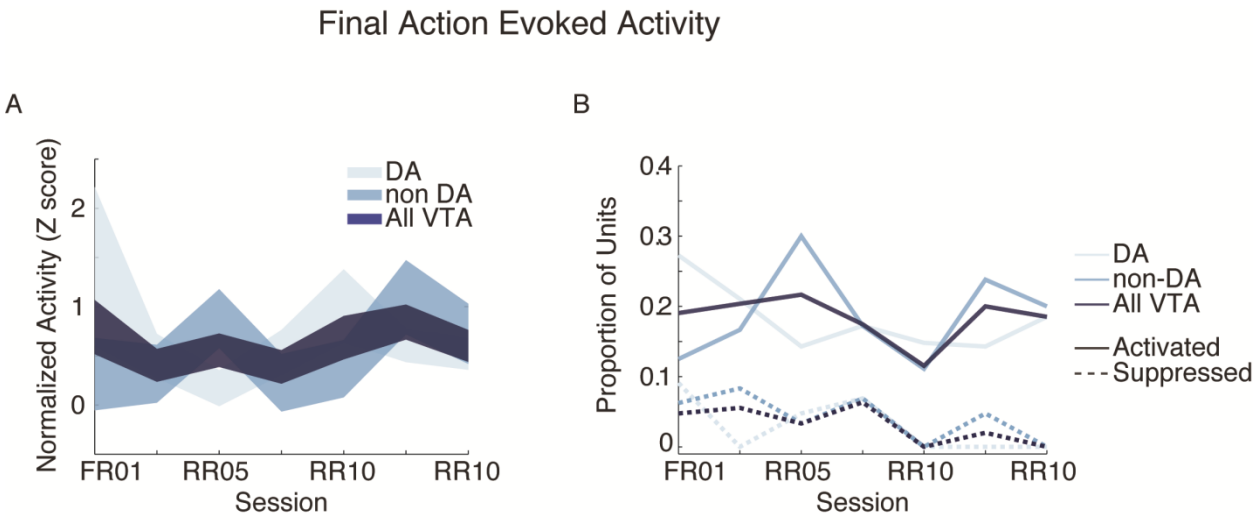

Figure 3.21. Neuronal Activity Evoked by the Final Action in Each Trial.

(A) Mean ± SEM neuronal response evoked by execution of the final action in each trial. Each unit's data were averaged across a time window -0.125 - +0.125 sec, relative to action time. This is the same time window as is used in all other analyses of action-evoked neuronal activity. Data are depicted separately for all putative dopamine and non-dopamine neurons, as well as all VTA units pooled together. Note that the evoked population response did not change magnitudes across sessions, nor differ by neuron type (main effect of neuronal type, $F_{(1,321)} = 0.230$, $p = .632$; main effect of session, $F_{(5,321)} = 0.621$, $p = .684$; interaction, $F_{(5,321)} = 1.677$, $p = .140$). (B) The proportion of units classified as either significantly activated (solid lines) or suppressed (dashed lines) by action execution. Data are depicted across all sessions, for putative dopamine, non-dopamine and all VTA neurons. Note that for all groupings of neurons, suppression was rarely evoked and the proportion of neurons, which were activated, did not change across sessions (dopamine neurons, $X^2_{(5)} = 0.575$, $p = .989$; non-dopamine neurons, $X^2_{(5)} = 3.526$, $p = .619$).

Population averaged activity during the delay period between the final action and outcome delivery was measured in a 0.250 sec window (Figure 3.22 A; +0.150 - +0.400 sec, relative to final action). Note that this window begins 0.050 sec after the window utilized to examine the final action in each trial, and ends 0.100 sec before the outcome is delivered. The evoked population response increased across sessions, with all random ratio 10 sessions being significantly greater than sessions 1 and 2 (Figure 3.22A; main effect of session, $F_{(6,361)} = 4.776$, $p < .001$; post hoc tests $p < 0.05$). This response did not differ between non-dopaminergic or dopaminergic neurons (Figure 3.22 A; main effect of neuronal type, $F_{(1,361)} = 1.585$, $p = .209$ interaction, $F_{(6,361)} = 1.131$, $p = .343$). Similar to neuronal activity around other task events, suppression was rarely evoked at a single unit level, and occurred at an insufficient frequency for reliable statistical analysis (Figure 3.22 B). The proportion of both dopaminergic and non-dopaminergic neurons that were activated during the delay period significantly increased across recording sessions (Figure 3.22 B; dopaminergic neurons, $X^2_{(6)} = 22.093$, $p = .001$; non-dopamine neurons, $X^2_{(6)} = 29.744$, $p < .001$). Taken together, delay period population averaged activity and the proportion of neurons activated during this period, increased across sessions. It is also important to note that this effect is confined to a narrow window of time. Neither the action, which immediately preceded the delay period, nor outcome delivery, which immediately followed the delay, were associated with population activity that changed magnitudes across recording sessions. Further, these data are consistent with a large body of evidence that suggests that environmental stimuli that offer new information about outcome value (including outcome magnitude and the delay until outcomes are received) evoke reward prediction error signals from dopaminergic neurons. The current data suggest that this signal is also carried by non-dopaminergic neurons.

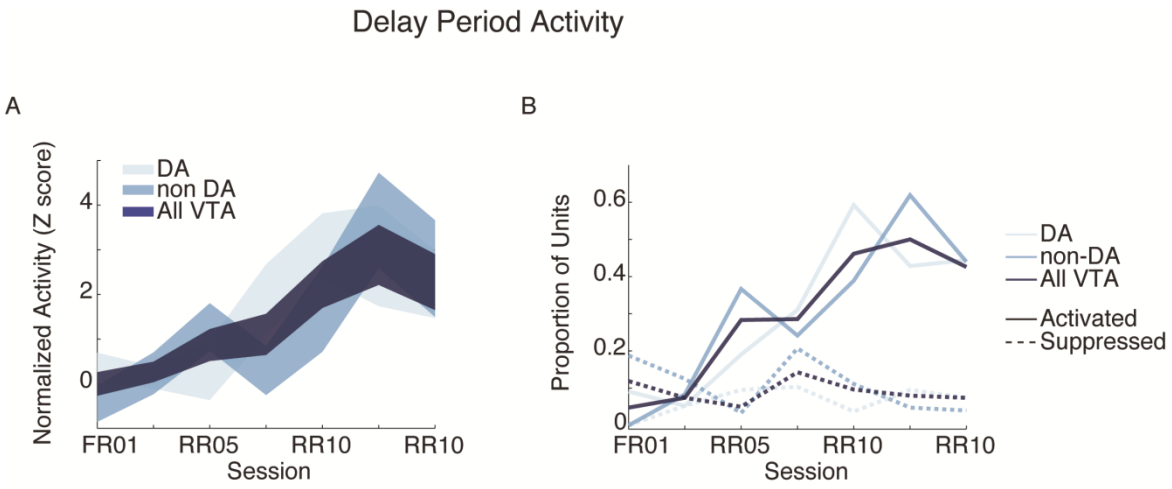

Figure 3.22. Neuronal Activity in the Delay Between the Final Action in Each Trial and the Subsequent Delivery of the Outcome.

(A) Mean ± SEM neuronal response during the delay period (+0.150 - +0.400 sec, relative to final action). Data are depicted for all groups of VTA neurons. Note that in each groups of neurons, the evoked population response increased across sessions (main effect of session, $F_{(6,361)} = 4.776$, $p < .001$). The neuronal response did not differ between non-dopamine and dopaminergic neurons (main effect of neuronal type, $F_{(1,361)} = 1.585$, $p = .209$; interaction, $F_{(6,361)} = 1.131$, $p = .343$). (B) The proportion of units classified as either significantly activated (solid lines) or suppressed (dashed lines) during the delay. Data are depicted across all sessions, for putative dopamine neurons, non-dopamine neurons, and all VTA neurons. Note that for all groupings of neurons, suppression was rarely evoked and the proportion of neurons that were activated increased across sessions (all units, $X^2_{(6)} = 45.949$, $p < .001$; dopamine units, $X^2_{(6)} = 22.093$, $p = .001$; non dopamine units, $X^2_{(6)} = 29.744$, $p < .001$).

CUE AND OUTCOME DELIVERY EVOKED DATA SUMMARY

In the initial session, cue light onset evoked a strong phasic increase in population activity that diminished in magnitude during the random ratio sessions. This suggests that information about a series of actions is most likely processed by VTA neurons during, but not prior to, serial behaviors. Outcome delivery evoked phasic activations were greater in dopamine neurons than non-dopamine neurons and the delay period was associated with increased population activity in all groups of neurons as learning progressed. Taken together, the delay period (signaled by cue offset) and outcome delivery-evoked neuronal responses are in agreement with previous work suggesting that VTA neurons, especially dopamine neurons, encode reward prediction errors.

4.0 DISCUSSION

SUMMARY OF THE CURRENT WORK

In the current work, animals were trained to execute a series of repetitive, self-organized, instrumental actions that were randomly reinforced with a sugar pellet. At the start of each trial a cue light was illuminated and then extinguished once reinforcement was earned. Serial actions evoked heterogeneous patterns of activity in both dopaminergic and non-dopaminergic VTA neurons. Different neurons preferred unique subsets of actions, and each action was preferred by a subset of VTA neurons. Averaging the activity of all neurons across action number obfuscated this pattern, and the resulting population average was uncorrelated with action number. Because VTA neurons produced dozens of distinct and complementary patterns of activation during action execution, a naïve Bayesian decoder was able to classify ensemble neuronal activity above chance levels and more accurately than the population averaged activity could be classified. These ensembles are comprised of dopaminergic and non-dopaminergic VTA neurons that function as a network in which each individual neuron is tuned to a specific subset of actions. This real-time representation of ongoing action number is likely decoded by post-synaptic regions for flexible adaptations of behavior and future decision-making.

Cue light onset did not strongly modulate neuronal activity during random ratio sessions, and is unlikely to influence serial action performance. These data further demonstrate the

importance of the aforementioned action-evoked data in encoding information for real-time behavioral organization. Outcome delivery increased activity in both populations of neurons, though more substantially in dopaminergic neurons. With training, activation also occurred at earlier predictors of outcome delivery (the delay following cue light offset and preceding outcome delivery). These data are qualitatively consistent with the predictions of TD models of dopaminergic activity. They also suggest that non-dopaminergic VTA neurons may process complementary information (Nishino *et al.*, 1987; Seamans & Yang, 2004; Kim *et al.*, 2010; Kim *et al.*, 2012).

It stands to reason that action number, in the current experimental design, could be reflective of a number of variables. This includes, but is not limited to, progress toward completion of a goal, energy or effort expenditure, time since trial start, or an explicit representation of action number. It is difficult, if not impossible, to completely disentangle these explanations, as all of the aforementioned variables are correlated. Explanations based upon time elapsed in a trial are somewhat weakened by the fact that this variable, by itself, only explained the action evoked activity of a small number of neurons. While dopamine is important for various aspects of timing or time perception, time is not a critical element of the current experimental design. It stands to reason that if designs which require animals to process timing were utilized, a different result may have been obtained.

Dopamine is also necessary for effortful behavioral output, though it is not presently clear how dopamine release or firing patterns may modulate effort (Salamone & Correa, 2002; Salamone *et al.*, 2007; Day *et al.*, 2010; Gan *et al.*, 2010; Wanat *et al.*, 2010; Wassum *et al.*, 2012; Pasquereau & Turner, 2013). A series of instrumental behaviors likely requires consistent or sustained levels of effort to complete. This is particularly true when high numbers of

instrumental actions are required of the animal. In addition to providing a signal that contains information about action number for behavioral organization, the ensemble code for action number may also subserve some aspect of motivation, potentially through connections with the nucleus accumbens (Salamone *et al.*, 2007). The ensemble signal may be necessary for sustaining motivation in the face of multiple unrewarded actions or perhaps serve as a surrogate reinforcement signal to engender serial behavioral responding via phasic dopamine release events

All discussion of the concept of 'action number' in this thesis acknowledges the fact that action number may either involve an explicit representation of number, or abstract concepts related to number, such as, goal completion or effort expenditure. Given the role of the VTA in motivated behavioral output, and the diverse patterns of projections from the VTA, it seems likely that the action evoked ensemble signal in the current work could be employed for multiple purposes. A strict role in numerosity seems outside of the scope of the many functions the VTA subserves. Through representations of action number, the VTA most likely represents task completion, progress toward a goal, effort expended, or some other aspect of action number, rather than numerosity, per se. All of these concepts may be generally thought of as behavioral organization-related forms of cognition, and are further discussed below.

DECODING ACTION NUMBER FROM VTA ENSEMBLE ACTIVITY

A pair of decoding approaches was utilized to estimate action number from neuronal activity. Population averaged activity was largely invariant with action number. This is an important finding because most previous attempts to understand how dopaminergic neurons or VTA

neurons encode information have focused on population averaging (Schultz, 1998), and ensemble encoding of information has seldom been suggested in the VTA. A burgeoning body of evidence implicates dopamine and the VTA in an animal's estimates of action number (Gallistel & Gibbon, 2000; Allman *et al.*, 2011; Lustig, 2011). Thus, the fact that the population average transmitted little information about action number is surprising, and suggests that VTA neurons may encode this information using alternative coding regimes.

The naïve Bayesian ensemble decoder estimated action number from the activity of each individual neuron. In this sense, the decoder is analogous to a post-synaptic network that receives equally weighted input from the entire ensemble of VTA neurons. This decoder represents a simple solution for decoding neuronal output, and it makes minimal assumptions about how neuronal activity may be combined into an ensemble signal. By merely reading out the activity of each VTA neuron's firing rate, and assessing what action number would most likely correspond to that level of activity, action number could be accurately decoded. VTA dopaminergic and non-dopaminergic neurons project in parallel to target brain regions with a high degree of convergence on target neurons (Swanson, 1982; Van Bockstaele & Pickel, 1995; Carr & Sesack, 2000b; Sesack & Grace, 2010). While the naïve Bayesian ensemble decoder was not biologically inspired per se, it is reasonable to suggest that many neurons that receive VTA inputs would be capable of analogous computations.

It is also important to point out that the naïve Bayesian ensemble decoder treated each neuron equivalently, though this need not be the case. By weighting the contribution of different neurons to decoding action number, according to various parameters (e.g. how strongly actions modulate firing rate), it may be possible to achieve more accurate decoding. The current approach represents the simplest solution, which is to include all neurons into the analysis. This

approach is preferable, as little is known about VTA ensemble encoding and it may be imprudent to exclude or discount the activity of some neurons. It is reasonable to suggest, however, that *in vivo* networks could decode the activity of select neurons, and potentially improve decoding. This outcome would only strengthen the argument that action number is encoded via an ensemble signal.

Taken together, the current data suggest an entirely novel means by which VTA neurons encode information. Through this previously undescribed ensemble signal, VTA neurons process information that has not been previously observed in electrophysiological studies of the region. The current data directly implicate VTA neurons in processing information that subserves behavioral organization. While dopamine has long been linked to executive function, dopaminergic neurons have seldom been observed to encode information that is exclusively linked with traditional executive processes (e.g. working memory, cognitive flexibility, etc.). Rather, these neurons have been largely associated with encoding reward prediction errors. Thus, the current data also demonstrate a complementary role for these neurons in encoding information that is conceptually distinct from that observed in most previous electrophysiological experiments.

Non-dopaminergic and dopaminergic neurons were incorporated into the current analyses of VTA ensembles, though this is not the case in most previous analyses of VTA activity. Similar patterns of activity were found in both types of neurons, and it stands to reason that both populations of VTA neurons serve similar roles in encoding action number. For this reason, their activity was combined, and this approach is not without precedent. Previous work has suggested that dopaminergic and non-dopaminergic VTA neurons may encode similar or complementary information (Nishino *et al.*, 1987; Seamans & Yang, 2004; Kim *et al.*, 2010;

Cohen *et al.*, 2012; Kim *et al.*, 2012), though there have been very few direct demonstrations of this phenomenon. Thus, the current data are important because they explicitly demonstrate that multiple types of VTA neurons may work together to encode information. The majority of the non-dopaminergic cells in the VTA are GABAergic. These neurons are implicated in slowing the rate of conditioning and reducing reward consumption (Tan *et al.*, 2012; van Zessen *et al.*, 2012), though there is still a great deal to learn about the role of GABAergic VTA neurons in cognition (Sesack & Grace, 2010; Creed *et al.*, 2014). GABAergic neurons represent approximately one third of the neurons in the VTA and can be either long range projection neurons, which run parallel to dopaminergic projection neurons, or make local inhibitory connections which modulate dopaminergic output (Swanson, 1982; Johnson & North, 1992; Carr & Sesack, 2000b; Nair-Roberts *et al.*, 2008; Omelchenko & Sesack, 2009; Dobi *et al.*, 2010; Sesack & Grace, 2010; Creed *et al.*, 2014). Thus, in addition to serving a local modulatory role in the VTA, non-dopaminergic neurons may encode and transmit information to post-synaptic networks in a similar fashion to dopaminergic neurons. The current dataset point to a previously unrecognized role in cognition for this population of VTA neurons, and suggest that non-dopaminergic neurons transmit this action number information to other brain regions similarly to dopaminergic neurons. Further, these data suggest that both dopamine and non-dopamine signals may be combined together to form a multi-neurochemical signal.

The current analyses provide insight into how neuronal networks may use different approaches to decode action number from VTA ensemble activity. Verifying that any network in the brain decodes neuronal activity in a manner analogous to these approaches is difficult. These approaches were meant to highlight some possibilities for how the brain may interpret the VTA signal, but were not an exhaustive account of the possibilities. These statistical algorithms were

not meant to replicate a biological process, but rather highlighted the relationship that exists between action number and neuronal activity. These decoders demonstrated that population average activity is not strongly related to ongoing action number. In contrast, there was a strong relationship between ensemble activity and action number, and neuronal networks may leverage this relationship in a myriad of ways. These analyses highlight ensemble encoding as a key feature of VTA neuronal activity, and demonstrate how this is related to encoding information about ongoing behaviors. Traditional approaches to analyzing VTA neuronal data do not capture this aspect of neuronal activity, and future work should continue to develop novel ways of decoding information from VTA activity, as well as investigate the biological feasibility of these analyses.

It should be noted that the current experimental design did not require the animals to be aware of how many actions they had performed in each trial, nor is there explicit evidence that the animals were tracking this information throughout the task. However, the rate that actions were performed did increase as the reinforcement schedule was changed, which suggests that the animals had some sense of average action number requirement. Further, inter-action intervals increased from lower to higher numbered actions. This finding demonstrates behavioral sensitivity to ongoing action number. It could be the case that the VTA signal is necessary for this calculation, though additional work is necessary to substantiate this claim. Lesion, inactivation, or optogenetic studies could be utilized to demonstrate the necessity of the VTA for processing this information (see Future Directions, below).

Naturalistic settings are marked by uncertainty, and in such an environment an animal would need to track ongoing action number (via task completion, trial progress, effort expended, etc.) to update behavioral policies. Perhaps the VTA processes action number information in the

current task, despite the fact that it does not confer an advantage to the animal, because this is information that the region processes in naturalistic settings. Regardless of how, or if, animals employed the information about action number encoded by VTA neurons, the fact remains that these neurons clearly encode something related to ongoing action number. Because each neuron responded to a unique set of action numbers, the ensemble signal, by definition, contains information about action number. It stands to reason that if a signal containing information about action number exists in the brain, that a downstream brain region may have evolved to decode this information

VTA ENSEMBLE ACTIVITY

Instead of focusing on network phenomena, most VTA electrophysiology experiments have focused on averaged activity, with several notable exceptions. VTA firing rate correlations increase when stimuli predict rewarding outcomes and decrease when stimuli predict negative outcomes (Kim *et al.*, 2012). Similarly, substantia nigra dopamine neurons become increasingly correlated following rewarding events (Joshua *et al.*, 2009). Additionally, several groups have observed synchronous firing between VTA or substantia nigra dopaminergic neurons (Hyland *et al.*, 2002; Joshua *et al.*, 2009; Li *et al.*, 2011). Together, these data suggest that coordinated activity is a prominent feature of the VTA, or of the dopamine systems, and that network-wide interactions or information processing may contribute to information processing in the region.

In a previous experiment, animals learned two sets of Pavlovian associations (Kim *et al.*, 2012). Information about stimulus identity encoded by pairs of VTA neurons was equivalent to the summed information encoded by each neuron separately. In that study, like many others,

stimulus identity was most likely encoded by the population average. These results were generally consistent with existing views of VTA information processing. In contrast, the current data point to a different conclusion: VTA neurons can also encode information in a heterogeneous fashion with unique patterns of activity that can be decoded as an ensemble for an accurate estimate of action number.

One key difference between action number in the current study, and stimulus identity in previous work (Kim *et al.*, 2010), is the dimensionality of these variables. While only two stimuli were used in the previous work, action number can take on many values. In situations where state space (the set of values of a variable can take on) is limited (e.g. variables with only 2 values), redundant encoding may be unavoidable, because limited numbers of neuronal activity patterns are possible in the small state space. In contrast, heterogeneous responses may emerge from larger state spaces, in which there are additional opportunities for different neuronal responses to occur. Heterogeneity scaled to dimensionality may emerge naturally, without the need for the system to switch between encoding modes. If this is the case, then previous observations of homogenous and redundant encoding in VTA may be limited to low dimensional variables. Importantly, this suggestion highlights the compatibility of the current results with a great deal of previous work. Indeed, the overwhelming majority of recordings from dopaminergic neurons utilize designs, in which only low dimensional variables are built into the task. Taken together, these data suggest VTA ensembles encode information when more complex information needs to be processed.

ANATOMICAL AND PHYSIOLOGICAL CONTRIBUTIONS TO ENSEMBLE ACTIVITY PATTERNS

Heterogeneous tuning to subsets of action numbers may arise from extrinsic or intrinsic sources (Paladini & Roeper, 2014). Intrinsic characteristics like pacemaker currents, auto-inhibition, baseline firing rate, or resting potential, may modulate each neuron's tuning and are easily observed phenomena in VTA dopamine and non-dopamine neurons (Grace & Bunney, 1983a; b; Grace & Bunney, 1984a; Grace & Bunney, 1984b; Steffensen *et al.*, 1998; Lee *et al.*, 2001). Differences in these characteristics could contribute to unique action-evoked response patterns, though it is unclear how this may occur. Divergent afferent inputs may also give rise to heterogeneous patterns of activation. Many regions innervate the VTA in a sparse and intermingled fashion (for representative illustrations see Geisler and Zahm, 2005). Therefore, small pockets of VTA (spatial scales similar to those recorded in the current work) may receive inputs arising from many brain regions, and individual neurons within these microdomains may have unique sets of inputs. It seems likely that many VTA neurons have some common inputs as well, because correlated firing rates are common in the VTA and indicative of shared inputs (Cohen & Kohn, 2011). The confluence of shared and unique connections could contribute to neurons in close proximity being uniquely tuned to different subsets of actions. Together, these anatomical and physiological attributes could ultimately engender heterogeneous action-evoked neuronal responses amongst different VTA neurons. Recently, differing cognitive functions have been related to distinct patterns of connectivity and intrinsic properties in subsets of VTA neurons (Lammel *et al.*, 2008; Matsumoto & Hikosaka, 2009; Lammel *et al.*, 2011; Lammel *et al.*, 2012; Roeper, 2013; Volman *et al.*, 2013; Lammel *et al.*, 2014). The current work is part of an emerging field focusing on the heterogeneity inherent to VTA neuronal networks, but is

unique because it shows how this heterogeneity can subserve cooperative information processing.

At first glance, real-time information about ongoing action number could be conceptualized as a fairly complex computation, perhaps involving cortical circuitry. While it is unclear if this is an accurate assumption, if this is indeed the case, it raises an interesting point. The projection from the prefrontal cortex to the VTA is the only identified source of cortical input to the VTA (Sesack & Grace, 2010). If VTA representations of action number rely upon cortical input, then it stands to reason that the prefrontal cortex may provide some aspect of this information to the VTA. If the prefrontal cortex, however, needs the information about action number for executive cognition (as discussed above), and already possesses some information about action number, it is unclear why the region would need to rely upon the VTA for this computation. The prefrontal cortex may need action number representations to reflect some other information that is encoded by VTA neurons (e.g. reward prediction errors or motivation related information). This information may be multiplexed together with action number representations or modulate action number representations. In the current work, synthesis of such information was not identifiable, but future experimentation may uncover these factors with the proper experimental designs. Another possibility is that the prefrontal cortex conveys action number information to the VTA, so the VTA can act as a relay station and transmit this information to a large number of additional brain regions. VTA has a diverse series of projections, and may be ideally suited for this role. A third possibility is that brain may require the action number signal to be encoded by neurotransmitters contained within the VTA, such as dopamine and GABA. In such a configuration, prefrontal cortex would be incapable of encoding this signal, as all corticofugal fibers are glutamatergic. Thus, in this scheme, VTA could convert a glutamatergic,

cortical signal into a dopaminergic and GABAergic signal. It is presently unclear if this is the case.

DIVERGENT VIEWS OF THE DOPAMINERGIC SYSTEMS

The most influential views of the function of the dopamine system can be divided into several camps. The first emphasizes the role that dopamine plays in incentive salience, motivation, effort, locomotion and movement. This view notes that dopaminergic manipulations or diseases alter energy expenditure, motoric output, and reward seeking (Salamone & Correa, 2002; Wise, 2004; Salamone *et al.*, 2005; Berridge, 2007). A second camp emphasizes the role that dopamine plays in stress responses, because of strong increases in dopaminergic neurotransmission and firing during stress (Abercrombie *et al.*, 1989; Finlay *et al.*, 1995; Anstrom & Woodward, 2005).

This dissertation largely focuses on two additional frameworks that implicate the dopaminergic systems in different aspects of cognition- executive functions and reward prediction error signaling. The current discussion of these data will appear as two parallel and compatible interpretations based on these frameworks. The first concerns the role of the dopamine system in high-order aspects of cognition, which subserves behavioral organization. These functions are often thought of as 'executive functions' and are dependent upon dopaminergic neurotransmission from the VTA to the prefrontal cortex (Seamans & Yang, 2004; Robbins & Arnsten, 2009). The discussion below concerns how the current action-evoked neuronal data relate to this powerful framework for understanding the function of the dopaminergic system. Cue and outcome evoked neuronal responses may play a role in behavioral

organization as well, but a clear role for these VTA neuronal responses is lacking in the current data.

While these data are discussed in the context of the VTA projection to the prefrontal cortex, the projection targets of VTA neurons cannot be discerned using only chronic recording techniques. Moreover, VTA neurons with differing projection targets are intermingled (Swanson, 1982), and a neuron's projection targets cannot be assumed based upon intra-VTA localization. In general, other populations of neurons may carry a similar signal, other signals may be encoded by mesocortical neurons, and the neurons in the current dataset could project to non-cortical regions. The prefrontal cortex is discussed only as an archetypal region that may decode action number for behavioral organization, although this does not preclude the involvement of other regions in decoding this information.

Another equally important concept is that dopaminergic neurons encode reward prediction errors (Schultz, 1998). This phenomenon is well approximated as the error signal in TD models, as discussed in the Introduction section of this dissertation. Below, the cue and outcome delivery evoked data are discussed in relation to this framework. Following this is a discussion of how to integrate the action-evoked neuronal data into this framework, and the implications of these notions.

BEHAVIORAL ORGANIZATION FRAMEWORK

Planning behavior and processing ongoing information about behavior, such as action number, is the central role of the prefrontal cortex (Goldman-Rakic, 1988; Fuster, 1991; Constantinidis & Goldman-Rakic, 2002), which also requires dopaminergic input for these functions. The VTA

dopamine projection to the prefrontal cortex regulates working memory, cognitive flexibility, decision-making, attention, goal directedness, and response inhibition (Goldman-Rakic, 1998; Seamans & Yang, 2004; Floresco & Magyar, 2006; Naneix *et al.*, 2009; Robbins & Arnsten, 2009; Sesack & Grace, 2010). To understand how the current VTA data may contribute to behavioral organization, it is first necessary to understand how dopamine contributes to other aspects of executive function.

Dopaminergic regulation of working memory has been extensively studied. Endogenous dopamine levels in the prefrontal cortex increase during working memory tasks (Watanabe *et al.*, 1997; Phillips *et al.*, 2004), which suggests that dopamine release is critical for proper working memory function. Dopamine affects working memory via an "inverted-U shaped" function of D1 receptor activation (Arnsten & Li, 2005; Robbins & Arnsten, 2009). Thus, excessive activation or blockade of prefrontal D1 receptors disrupts working memory (Brozoski *et al.*, 1979; Sawaguchi & Goldman-Rakic, 1991; Sawaguchi & Goldman-Rakic, 1994; Arnsten & Li, 2005), and low concentrations of dopamine agonists improve working memory performance (Cai & Arnsten, 1997; Aultman & Moghaddam, 2001). Dopamine's effects on working memory can be traced to the cellular level. During working memory delays, some primate prefrontal cortex neurons represent information across the delay period, via sustained elevations in firing rate. Direct application of D1 agonists to these cells bi-directionally modulates this signal (Sawaguchi *et al.*, 1988; Williams & Goldman-Rakic, 1995). Dopaminergic neurons, however, do not appear to represent information in working memory, as these neurons only fire phasically at the start of working memory trials (Schultz *et al.*, 1993). As discussed in the Introduction section and again below, these firing patterns also occur in tasks without a strong working memory component.

Thus, prefrontal dopamine modulates working memory, but neuronal firing patterns do not appear to encode information specific to working memory.

Prefrontal dopamine release (or dopamine metabolite concentration) also increases during periods of cognitive flexibility, rule learning or decision-making (Stefani & Moghaddam, 2006; Winstanley *et al.*, 2006). Low doses of dopamine agonists can improve these aspects of cognition, while dopaminergic lesions or antagonists have detrimental effects (Granon *et al.*, 2000; Crofts *et al.*, 2001; Ragozzino, 2002; Chudasama & Robbins, 2004; Floresco *et al.*, 2006; Robbins & Roberts, 2007). Dopaminergic drugs also modulate decision-making, albeit in a more complex fashion (Evenden & Ryan, 1996; Cardinal *et al.*, 2000; Floresco & Magyar, 2006; Setlow *et al.*, 2009). Despite clear roles for dopamine in these cognitive processes, definitive dopaminergic neuronal correlates remain elusive. There is no known dopaminergic signal exclusive to cognitive flexibility. Very little work has documented dopaminergic signals correlated with attention (Bromberg-Martin *et al.*, 2010b; Totah *et al.*, 2013). Further, behavioral choices are actually decoupled from VTA dopamine responses (Roesch *et al.*, 2007). Instead, dopaminergic activity is consistent with models that use the dopaminergic signal for value learning, but not directly selecting behaviors (Morris *et al.*, 2006; Niv *et al.*, 2006; Roesch *et al.*, 2007). Taken together, prefrontal dopamine is necessary for multiple aspects of behavioral organization, but how this relates to the firing of dopamine neurons remains unknown.

Instead of hypothesizing that dopamine encodes information exclusively for behavioral organization, dopamine is often thought to "…curtail or prolong, augment, or diminish effects of fast signaling in neuronal networks" (Robbins & Arnsten, 2009). This type of neuromodulation may alter the lability of prefrontal networks and allow cortical representations to be strengthened or weakened (Seamans & Yang, 2004; Robbins, 2005; Floresco & Magyar, 2006; Floresco *et al.*,

2006). According to this view, dopamine adjusts the messages encoded by other neuronal systems, but does not encode information directly. Extracellular dopamine is cleared slowly and diffuses widely in prefrontal cortex, and these spatiotemporal aspects of dopamine signaling are consistent with a neuromodulatory role (Seamans & Yang, 2004; Schultz, 2007). Thus, conceiving of dopamine as a neuromodulator is compatible with basic neurochemical aspects of dopamine signaling and may partially explain how dopamine contributes to behavioral organization.

Previous emphasis on a neuromodulatory role of dopamine in behavioral organization may also arise from a paucity of evidence correlating the firing of dopamine neurons with these cognitive processes. Electrophysiological experiments suggest dopamine neurons encode a temporally precise and informative signal, reward prediction errors, which does not seem consistent with a neuromodulatory signal (Seamans & Yang, 2004). Instead of being used for behavioral organization, this signal is related to value prediction and association formation (Schultz, 1998). It should be noted, however, that there are no strict definitions for what form a neuromodulatory signal may take. A signal may function as either a direct encoder of information or as a neuromodulator, depending on how the post-synaptic neuron receives and acts upon that signal. It is nearly impossible to determine how any neuron in the brain extracts information from a signal. For these reasons, it is difficult to truly assess to what extent the dopamine signal resembles either a direct encoder of information or a neuromodulator. It should be noted that most speculation that dopamine acts as a neuromodulator is technically speculative, and based upon the aforementioned time course of dopamine clearance and spatial diffusion of dopamine. This view is not based upon firing patterns of dopamine neurons.

The current data suggest that VTA ensembles, including dopaminergic neurons, may directly encode a detailed and informative account of ongoing behaviors. Thus, the current data suggest VTA ensembles may encode a signal that could be decoded by the prefrontal cortex for behavioral organization, as opposed to being strictly utilized as a neuromodulator. This does not preclude the dopaminergic component of the ensemble signal from also acting in a neuromodulatory fashion. Distributed activation of subsets of dopamine neurons during a series of behaviors may optimize prefrontal dopaminergic tone in a traditional, neuromodulatory role. The current data, however, also suggest that action number is encoded directly by VTA ensembles. Moving forward, it is important to consider ensemble activity in models of dopamine and VTA function, especially signals for behavioral organization. These ideas can incorporate neuromodulation and direct encoding of information.

VTA ensemble signals may utilize several neurotransmitters that cooperatively encode information. The mesocortical projection contains a sizeable number of dopamine neurons, but is mostly GABAergic (Carr & Sesack, 2000b). Dopamine release exerts complex effects on prefrontal cortex physiology, and is capable of exciting or inhibiting pyramidal cells (Seamans & Yang, 2004). These effects are likely mediated by multiple types of dopaminergic receptors on pyramidal cells, local interneurons, and the GABAergic component of the mesocortical projection (Sesack & Bunney, 1989; Pirot *et al.*, 1992; Zhou & Hablitz, 1999; Carr & Sesack, 2000a; Lewis & O'Donnell, 2000; Gorelova *et al.*, 2002; Gao & Goldman-Rakic, 2003; Trantham-Davidson *et al.*, 2004; Tseng *et al.*, 2006; Tierney *et al.*, 2008). The effects of dopamine release on prefrontal cortex physiology are dependent upon a myriad of factors, such as background activity level, dopamine release pattern, and dopaminergic tone. For an excellent review, see (Seamans & Yang, 2004). The spatially and temporally imprecise nature of the

dopaminergic signal may be offset if combined with non-dopaminergic VTA signals (Seamans & Yang, 2004; Kim *et al.*, 2010; Kim *et al.*, 2012; Tritsch *et al.*, 2012; Tritsch *et al.*, 2014). This arrangement would allow fast, focal neurotransmission to complement dopamine release. The current data support this notion by demonstrating that dopaminergic and non-dopaminergic neurons in the VTA encode similar signals, as has been previously suggested (Kim *et al.*, 2010; Kim *et al.*, 2012).

Sustaining or adapting behavior until goals are met is a central theme of prefrontal cortex's role in executive function (Goldman-Rakic, 1988; Fuster, 1991; 2001). In an excellent review, Fuster (2001) wrote, "Here are two critical questions to be resolved: (1) How are the components of an executive cortical network timely and selectively activated in the execution of a goal-directed sequence of behavior? (2) How is a cortical network maintained active in the process of bridging temporally separate components of the sequence?"

In other words, it is unknown what input patterns activate, or silence, prefrontal neurons during critical moments in a behavioral sequence. The VTA ensemble signal in the current dataset is a prime candidate. The ensemble signal in the current dataset could evoke or silence the activity of networks in the prefrontal cortex. This, in turn, could sculpt the flow of information through these networks in service of executive function. These input patterns could transmit a real-time account of how many actions have been executed could be sent to the prefrontal cortex. By comparing the current action number with predictions about action requirement, an animal could evaluate and adapt the current behavioral policy. Further, the current action number signal "bridges" the actions within a behavioral series, and dopaminergic lesions disrupt performance of behavioral sequences (Veeneman *et al.*, 2012). Thus, ensemble patterns encoding action number may partially answer the pair of questions posed above. Both

humans and non-human animals possess a sense of numerosity, which is dopamine-dependent (Gallistel & Gibbon, 2000; Allman *et al.*, 2011; Lustig, 2011). This sense of numerosity could underlie multiple aspects of executive function. It remains unknown how this information is processed by the dopamine system. The current study casts new light on this idea by suggesting that VTA ensembles encode a running account of the current action number.

In summary of the ideas above, VTA ensembles encode a real-time account of action number within a trial. This information is consistent with a well-accepted role for the VTA dopamine system in executive functions and behavioral organization. This role may be extended to the non-dopaminergic populations of the VTA as well. The prefrontal cortex, amongst other regions, could decode ongoing action number to alter or sustain behavior until goals are met.

REWARD PREDICTION ERROR SIGNALING FRAMEWORK

As discussed in the Introduction section, dopaminergic neurons encode reward prediction errors, which are modeled by the TD error term. These prediction error signals represent actual and estimated future reward values, minus the predicted reward value. The prediction error is generated at the earliest predictor of future value, often a stimulus that predicts impending rewards. Because this signal predicts future value, it is implicated in associative learning.

Cue-Evoked Neuronal Data:

Reward prediction errors are often associated with cues that predict the consequences of actions or impending rewards (Miller *et al.*, 1981; Schultz, 1986; Romo & Schultz, 1990; Mirenowicz & Schultz, 1996; Schultz, 1998; Nakahara *et al.*, 2004; Matsumoto & Hikosaka, 2007; Roesch *et al.*, 2007; Bromberg-Martin & Hikosaka, 2009; Kim *et al.*, 2010; Takahashi *et al.*, 2011; Cohen *et al.*, 2012; Totah *et al.*, 2013). In instrumental behavior tasks, these cues almost always precede actions, and instruct an animal to begin responding.

In the current study, during the FR01 session (session 1), a sizeable proportion of neurons responded to cue light onset with phasic increases in firing rate. This activation was observed across all populations of VTA neurons. These data are generally consistent with previous work demonstrating that cues predicting outcomes generate reward prediction errors. Approximately two thirds of dopaminergic neurons were activated by cue onset in the current dataset, and it has previously been suggested that roughly 70% of dopaminergic neurons encode prediction error signals (Schultz, 1998; 2010). Thus, the current FR01 data are also generally consistent with the prevalence of reward prediction error signaling amongst dopaminergic neurons. The current data further suggest that non-dopaminergic neurons, albeit only a third of the non-dopamine neurons, also encode this information. This observation is consistent with prior suggestions that dopaminergic and non-dopaminergic neurons may encode similar or complementary signals (Seamans & Yang, 2004; Kim *et al.*, 2010; Kim *et al.*, 2012).

The cue-evoked population response magnitude decreased in random ratio reinforcement sessions (sessions 2-7). At first glance, this may seem incompatible with most previous research, although cue evoked responses have been observed to disappear or diminish with training in some previous studies (Schultz & Romo, 1990; Ljungberg *et al.*, 1992; Schultz *et al.*, 1993).

Schultz and colleagues briefly discussed a compelling explanation for this phenomenon. The authors explained that when predictive cues are placeholders (e.g. do not lead directly to reward or strongly predict when reward will be delivered), they do not generate phasic dopamine responses (Schultz *et al.*, 1993). Moreover, cues that are temporally distant from outcome delivery are known to evoke smaller magnitude neuronal activations (Roesch *et al.*, 2007; Fiorillo *et al.*, 2008; Kobayashi & Schultz, 2008). Thus, it is not enough to state that cues that provide new information about impending outcome value should generate reward prediction errors. Instead, the reward prediction error response also appears to depend upon the temporal contiguity and contingency of the cue and the outcome.

In the current random ratio reinforcement sessions, the cue was often separated from outcomes by many actions and lengthy temporal delays. Further, there was no contingency between cue and outcome. This contrasts sharply with FR01 reinforcement schedules, in which cue onset is separated from outcome delivery by minimal behavioral output. Therefore, it is not surprising that the cue did not continue to evoke phasic responses in any group of VTA neurons during random ratio reinforcement. Similar results were also obtained in the only other recording of VTA neurons during serial actions (Nishino *et al.*, 1987). Taken together, diminished cue-evoked VTA responses are consistent with a number of previous reports. Previous recordings have suggested that dopaminergic neurons have high learning rates (Hollerman & Schultz, 1998) but see (Pan *et al.*, 2005). This theoretical consideration could explain why cue onset evoked a neuronal response within the first recording session. It is less clear how to model the disappearance of this neuronal response. While it has been observed several times, and even mentioned explicitly in the first formalization of dopaminergic responses as TD errors (Montague *et al.*, 1996), this aspect of neuronal activity has not been well studied in the context

of TD models. Future TD models could potentially weight stimulus value by contiguity and contingency to produce diminished cue responses.

Outcome Delivery Evoked Neuronal Data:

VTA neuronal activity increased during outcome delivery. In all populations of VTA neurons, outcome evoked population activity was stable across all sessions, and was larger in dopaminergic neurons than non-dopaminergic neurons. These data are generally consistent with previous reports that dopaminergic neurons are responsive to the delivery of reinforcing outcomes (see below). Qualitatively similar responses between dopaminergic and non-dopaminergic neurons suggest that these neurons process this information similarly.

Historically important work, largely from Wolfram Schultz's group, suggests that predicted rewards should not evoke a dopaminergic response (Ljungberg *et al.*, 1992; Schultz *et al.*, 1993; Mirenowicz & Schultz, 1994; Hollerman & Schultz, 1998; Waelti *et al.*, 2001). Though predicted rewards evoking dopaminergic responses has also been reported by Schultz and colleagues (Fiorillo *et al.*, 2003), as well as other groups (Morris *et al.*, 2004; Pan *et al.*, 2005; Morris *et al.*, 2006; Roesch *et al.*, 2007; Joshua *et al.*, 2008; Kim *et al.*, 2010; Takahashi *et al.*, 2011; Cohen *et al.*, 2012; Pasquereau & Turner, 2013; Totah *et al.*, 2013). It should be noted that in some of these experiments multiple outcomes were delivered or available. In this case, interpreting outcome-evoked phasic activation becomes more complicated, as the response may partially reflect errors in predicting which reward would be delivered (Tobler *et al.*, 2005). Taken together, it is not surprising to continue to see reward evoked neuronal responses, even after thousands of trials have been completed.

Even though rewards were delivered at the conclusion of each trial, there may still have been a reward prediction error. Reward prediction errors incorporate the timing of reward delivery (Fiorillo *et al.*, 2008), and the delay between the final action and outcome in each trial could have created larger prediction errors. Similarly, requiring animals to execute random numbers of actions prevented animals from predicting how many actions would be required per trial. This uncertainty may ultimately create a larger outcome delivery evoked prediction error. Thus, it is not clear to what extent reward delivery was fully predictable by the animals. Under these circumstances, it seems reasonable to suggest outcome delivery would evoke neuronal responses in all sessions.

Delay Period Neuronal Activity:

Population averaged activity increased during the delay between the final action in a trial and outcome delivery. Previous reports have suggested that pre-delivery delay period activity may be specific to non-dopaminergic VTA neurons (Cohen *et al.*, 2012). In the current dataset, this pattern occurred in all groups of VTA neurons and increased in strength across sessions. The beginning of the delay period, at cue offset, predicted when outcomes would be delivered. In this sense, the delay strongly resembles important aspects of a traditional Pavlovian conditioned stimulus or discriminative stimuli from other instrumental tasks. Once this predictive relationship is learned, cue offset may evoke prediction errors, and the delay period data are potentially consistent with TD models. It is also interesting that cue-evoked responses developed during the first session, when delay period responses were negligible. Thus, events more distal to the outcome generated phasic responses before the delay period began modulating neuronal activity.

Summary of Cue and Outcome Evoked Neuronal Data:

Taken together, the current cue evoked, delay period and outcome delivery evoked dopamine data are generally consistent with previous findings. Weak cue evoked dopaminergic responses after training have been observed previously. Likewise, persisting outcome evoked dopaminergic activity has also been reported numerous times, and may be related to difficulty predicting when rewards would be earned. The emergence of delay period activity across learning may also be explained within the context of TD models, and the late development of this response relative to the cue evoked response is quite interesting. The current data extend the general patterns of activity seen in dopaminergic neurons to non-dopaminergic neurons, as has been noted previously (Kim *et al.*, 2010; Kim *et al.*, 2012). The TD error signal is the most widely used theoretical model of dopaminergic phasic firing patterns, and the aforementioned data seem qualitatively consistent with many aspects of the model's predictions. This information is likely sent to many brain regions (Glimcher, 2011). Phasic dopamine signals are particularly prominent in ventral striatum, and neurochemical studies suggest that prediction errors are transmitted to this region (Robinson *et al.*, 2003; Heien & Wightman, 2006; Hart *et al.*, 2014). These signals may facilitate action sequence learning (Suri & Schultz, 1998; Wassum *et al.*, 2012), as well as acquisition of appetitive behaviors (Phillips *et al.*, 2003; Day *et al.*, 2007). Further, cortical regions receiving VTA inputs have also been implicated in utilizing traditional reward prediction error signals to update outcome values (Takahashi *et al.*, 2009).

Serial Action Evoked Neuronal Data:

In the current dataset, neuronal activity was only weakly modulated when averaged over all actions. Actions are not generally associated with generating reward prediction errors in many instrumental designs. With all actions averaged together, the current data are similar to previous reports in this sense. Different neurons fired preferentially during the execution of unique subsets of actions, and it is unclear if an underlying process that mimics the TD error signal produces this phenomenon, and if so, how this might work. It is common to suggest that the error signal should back-propagate to the earliest predictor of the outcome. This could be the first action within a trial, and there were a sizeable number of neurons tuned to prefer the first few actions in a trial. Many neurons, however, preferred higher numbered actions. It does not seem likely that reward prediction errors had incompletely back propagated to the first action in these neurons, because these responses were present even after hundreds of trials had been completed. Perhaps most importantly, the TD error can only take on one value at each time step, which is inconsistent with an ensemble signal comprised of heterogeneous responses. It is conceivable that a TD algorithm could reproduce the population average activity as a function of action number. This, however, would not reflect the most interesting aspect of this data: the ensemble nature of the signal. Likewise, the information about action number that is encoded by the ensemble would be lost.

The fact that other events, such as the delay period and outcome delivery, are generally consistent with a reward prediction error framework is critical. It confirms a widely replicated observation, and suggests the current dataset conforms to prior assumptions about VTA function. In light of this, the difficulty to explain serial action evoked neuronal activity in the context of TD error signaling is, therefore, all the more striking. The earliest event in random ratio trials

that generates a response mimicking the TD error is the delay period between the final action and outcome delivery. A great deal of theoretical and experimental considerations suggests reward prediction errors should occur before the animal begins responding (Montague *et al.*, 1996; Schultz *et al.*, 1997; Dayan & Abbott, 2001; Morris *et al.*, 2006; Roesch *et al.*, 2007).

The subset of neurons responsive to any single action may be encoding a TD error-like response forecasting the value of each action. Because several actions were often executed over the course of just a few seconds, this would necessitate multiple reward prediction errors in a short amount of time. In most prior work, it is rare to observe more than 2 reward prediction errors in the course of a trial (e.g. one at a predictive cue and another at reward delivery). There may be an upper limit on how many of these responses each neuron can produce in a short period of time. Potentially, reward prediction errors may be spread across subsets of neurons to minimize the activity of individual neurons. Thus, the ensemble signal could function as a traditional reward prediction error signal that is distributed across multiple subsets of neurons. Under such a scheme, reward prediction error representations from activated neurons may be utilized in the standard way, and ensemble encoding of action number could be utilized for behavioral organization. Thus, different networks could decode the signal for divergent purposes.

The notion that large numbers of reward prediction errors in a trial would need to be encoded by distinct subsets of neurons, in a revolving fashion, raises the intriguing possibility that the heterogeneity of VTA neuronal activity patterns scales to the complexity of behavioral requirements. If a neuron is tuned to specialize in encoding information about a limited subset of action numbers, metabolic energy expenditure is limited. The VTA could therefore encode information about a large state-space with ensembles of neurons that are energetically cheap encoders. Most prior experiments have utilized tasks with far fewer events per trial (Dayan &

Niv, 2008; Niv & Schoenbaum, 2008), and would not be expected to uncover this previously hidden aspect of VTA activity.

PREVIOUS VTA RECORDINGS DURING SERIAL ACTIONS

In the only previous VTA recording in a task that required animals to execute large numbers of actions per trial, Sasaki and colleagues (1987) observed irregularly fluctuating firing rates of dopaminergic and non-dopaminergic neurons throughout an action sequence. In that experiment, monkeys bar pressed multiple times per trial for rewards. Thus, the animals executed serial actions for rewards, similar to the current work. There are several notable differences between their work and the current study (Nishino *et al.*, 1987), but the potential similarities in the data are striking. In both datasets, action evoked activity occurred in dopaminergic and non-dopaminergic VTA neurons. It appears that different neurons may be tuned to different actions in their data, and that each neuron preferred subsets of the action numbers (Nishino *et al.*, 1987). Finally, the cue at trial start failed to elicit a phasic response in both cases (Nishino *et al.*, 1987). This suggests that the current findings are reliable and replicable. In the case of both datasets there is one critical point, stated previously in the introduction as well, that bears repeating here. The heterogeneity of activity patterns suggests there is underlying structure in the responses of VTA neurons that has not been incorporated into a theoretical framework.

SUMMARY

To summarize, the current data suggest that VTA neurons may support multiple aspects of cognition, through several information processing regimes. A conceptual schematic is presented in Figure 4.1, which details some hypothesized aspects of these phenomena. The diverse response patterns evoked by action execution may ultimately be derived from sources both inside and afferent to the VTA. For the sake of clear presentation, VTA intrinsic properties are not depicted in Figure 4.1. Unique afferent inputs may partially produce divergent tunings to action number and may originate in distinct brain regions, or from parallel projections from the same region. These inputs likely synapse on intermingled dopaminergic and non-dopaminergic projection neurons. In the schematic, action numbers are conceptualized as high, medium, or low, though action number may be decoded at a different resolution in the brain. Neurons with differing action number preferences likely project in parallel to target brain regions, including the prefrontal cortex. Prefrontal cortex could potentially decode these parallel signals by simply assessing each neuron's current activity in relation to its previously learned tuning. By considering each neuron independently, and not averaging neuronal activity across neurons with differing tunings, action number could be accurately decoded on an action-by-action basis. While many regions could potentially utilize this information, prefrontal cortex stands out in particular, because of the well-described role of this region in behavioral organization. Theoretically, the prefrontal cortex could utilize information about action number to adapt, sustain, and plan behavior on the basis of estimated task completion, proximity to a goal state, or optimality of the behavioral policy.

In addition to information about actions, VTA neurons responded prominently to outcome delivery, and during the delay preceding outcome delivery. For simplicity, in the

schematic these events are represented simply as “outcome delivery”. These events evoked responses in approximately 2/3 of the dopaminergic and non-dopaminergic neurons, which are intermingled with non-responsive neurons. The same neurons that were diversely tuned to action number were homogeneously activated by outcome delivery or during the delay. Theoretically, these neurons could be encoding reward prediction errors related to outcome delivery. Target regions, such as ventral striatum, could decode the population averaged activity level. The ventral striatum, or other regions that receive this signal, could utilize the information for associative, value-based, learning.

It should be noted, that delay period responses developed with learning, while outcome evoked responses were present in the earliest sessions. Delay period input strength likely changed across sessions, and delay period and outcome delivery information could be transmitted to VTA neurons by separate pools of afferent input. Additionally, in all recordings, the projection targets of recorded neurons are not known. Instead, the ideas expressed in the schematic are purely for the sake of illustration.

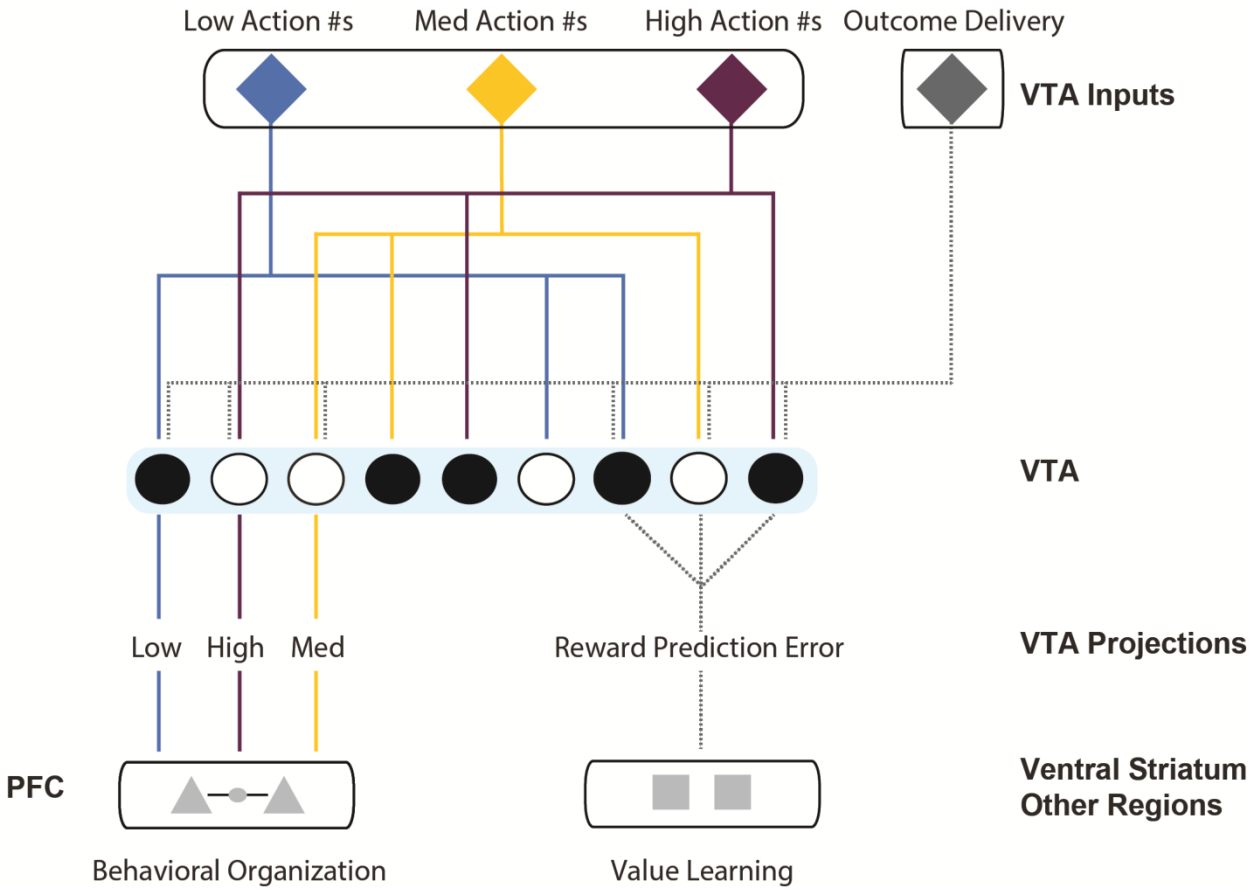

Figure 4.1. Theoretical Schematic.

Heterogeneous representations of action number could arise from distinct pools of afferent inputs (represented as blue, gold, maroon diamonds at the top level of the schematic). These inputs could arise from one or from multiple brain regions. Many dopamine (white circles) and non-dopamine (black circles) neurons represented subsets of the actions in the current data. The schematic depicts action numbers grouped into low, medium, and high bins of action number. On the left of the middle level of the schematic, a group of neurons is depicted projecting in parallel to both pyramidal neurons and interneurons (light grey triangles and oval in the lower left level of the schematic) in the prefrontal cortex (PFC). Because each neuron is tuned to prefer a subset of action numbers, prefrontal cortex could decode the current action number by reading out each input separately. This information could theoretically be utilized to update ongoing behavioral strategies, plan new behavioral strategies, and keep a running account of ongoing behaviors (behavioral organization).

Many VTA neurons responded to outcome delivery or the preceding delay period after learning had occurred (both notated as outcome delivery). Thus, the majority of dopaminergic and non-dopaminergic neurons likely receive information about these events. This is depicted as inputs arising from the dark grey diamond (top level of schematic). The lines representing these inputs are dashed to add visual clarity and, are not meant to be reflective of input strength. The averaged activity of these neurons could represent reward prediction errors (right set of merged inputs to the bottom level of the schematic). Regions such as ventral striatum could utilize this signal for value-based associative learning (right bottom level of schematic).

FUTURE DIRECTIONS

The current work suggests several avenues of future research: modifications to theoretical models of VTA neuronal activity, and demonstrating a causal role for the current ensemble patterns of activity in encoding action number. These future research directions are discussed below.

TD error signals are not modeled as an ensemble with heterogeneous tunings, though the current data suggest this may be critical to explore. One possibility is that the standard approach to calculating the TD error, $\delta(t)$, could be used with few modifications. As detailed in the Introduction and Appendix A, the TD error is calculated in Equation 1 (from Introduction):

$$\delta(t) = [r(t) + v(t+1)] - v(t).$$

This equation could be potentially modified to weight each neuron's response by action number preference with a single term, θ_{na}, where θ is the scaled tuning of each n^{th} neuron's activity at the a^{th} action (this is conceptually identical to the manner in which tuning curves were scaled in Figure 3.7 A and Figure 3.8). Values of θ_{na} would range from 0 to 1 (0 = least preferred action and 1 = most preferred action). Action number would weight each neuronal response as follows (Equation 7):

$$\delta_{na}(t) = \theta_{na} \times ([r(t) + v(t+1)] - v(t))$$

By doing so, the TD error that each neuron generates is scaled to its relative preference for that action. When actions that are highly preferred occur a typical TD error is produced. When actions that are less preferred occurred, weights approaching zero minimize the other terms and the result is a TD error for that neuron that is close to zero. This would qualitatively

replicate some aspects of the current data. Namely, each neuron would produce phasic TD error like activation for only a subset of actions. This is not an exhaustive account of all possibilities. For instance, θ_{na} may also be weighted with negative and positive values to capture bidirectional firing rate modulations. Other modeling approaches have suggested that only the first action of a series should generate the prediction error, though these have used very different experimental designs than the current experiments (Suri & Schultz, 1998). While it remains unsettled which approach is the optimal manner to model the current data, future work should rigorously explore the question of when and how TD errors are generated during trials requiring serial actions.

Optogenetic or electrical modulations of VTA activity could provide one way to test the necessity of VTA ensemble signaling for estimating ongoing action number. Briefly, rats could be trained to execute a given number of actions before rewards were made available. Once this action requirement had been met, animals could then execute a different action to cause outcome delivery. This could be accomplished with the use of two operant levers in a simple behavioral chamber. If both levers were available during the response period, and no cues were given when the outcome was available for delivery, then the animal would need to keep track of ongoing action number to perform this task correctly. If the second lever was pressed before or after the precise action requirement was met it could be penalized, inducing motivated animals to keep track of action number as they performed the behavior.

Optical or electrical stimulation could be utilized to test the necessity of VTA signaling for an animal to estimate how many actions have been executed in a trial. Briefly, animals could perform the task as described above in a subset of trials to demonstrate that they understand task rules and are able correctly estimate action number. In another subset of trials, VTA neurons could be optogenetically inhibited during or just after action execution. This would

prevent or decrease neuronal activation and potentially prevent the animal from accurately estimating action number. This approach would target those neurons that were activated by the current action. On the other hand, electrical or optogenetic stimulation could activate those neurons that were normally silent during the execution of that action. This would also potentially disrupt ongoing estimates of action number. In either case, perturbations of the VTA should disrupt the animal's ability to correctly track ongoing action number, and demonstrate the necessity of intact VTA signaling for this cognitive function. More elaborate designs may target specific VTA projection systems with retrograde transportation of optogenetic constructs, demonstrating the involvement of VTA subsystems in this phenomenon.

IMPACT OF THE CURRENT WORK

The mosaic of neuronal activity patterns as a function of action number represents an unexpected level of heterogeneity. There is a general consensus that the dopamine system, and by extension the VTA, lacks diversity, complexity, or synergy of information processing (Schultz, 1998; Glimcher, 2011; Schultz, 2013) but see (Roeper, 2013; Volman *et al.*, 2013). The current data suggest that this notion falters when VTA activity is examined through the lens of more complex behavioral paradigms, and following different approaches to analysis of neuronal data. Ultimately, the current data could serve as a catalyst for challenging popular notions of the capabilities of the VTA and how it contributes to behavior. The current work has the capacity to invigorate future research and offer new insight into VTA function by demonstrating how VTA neurons process previously unexpected information. Many excellent electrophysiological studies of the VTA, and specifically the dopaminergic component of the region, have myopically

focused on population-averaged activity and reward prediction error signaling by dopamine neurons. The result is that sub populations of the system (non-dopaminergic neurons) and information encoding regimes (ensemble encoding) that may be critical to the function of the system are neglected. Thus, the current data can contribute to a burgeoning reconceptualization of ideas about the VTA (Volman *et al.*, 2013), as well drive future research.

The current work may also be extended to disease treatments, insofar as it suggests a novel direction in therapeutic research. For instance, the dopamine system was amongst the first neuronal systems to be linked to schizophrenia, and every approved pharmacological treatment for the disease has affinity for dopaminergic receptors (Seeman & Lee, 1975). Patients with schizophrenia, however, do not consistently demonstrate coarse disruptions of the dopaminergic system, such as cell death or altered levels of dopamine metabolites (Davis *et al.*, 1991; Howes & Kapur, 2009; Moghaddam & Wood, 2014). Further, approaches to understanding schizophrenia through unidirectional dysregulation of the dopamine system, such as global hyperdopaminergic neurotransmission, have been rejected in favor of more subtle interpretations (Davis *et al.*, 1991; Howes & Kapur, 2009). In short, cognitive impairment most likely does not result from simple, unidirectional, or static alterations in neurotransmission.

Instead, the integrity of neuronal networks, and the resulting information processing of the network, may be the key to understanding cognitive impairments in a number of diseases such as addiction disorders or schizophrenia (Moghaddam & Homayoun, 2008; Bassett & Bullmore, 2009; Akil *et al.*, 2010; Malsburg *et al.*, 2010; Uhlhaas & Singer, 2010; Wood *et al.*, 2012; Moghaddam & Wood, 2014). These data provide one of the first direct observations of VTA information processing that emphasize the integrity of the network. Disease

processes may corrupt ensemble signaling, though this has not been explored in the context of disease previously. It stands to reason that while gross disruptions of the VTA are not as prominent in these diseases, impaired ensemble signaling may underlie much of the symptomology. Indeed, impaired behavioral organization is a hallmark of these diseases (Goldman-Rakic, 1999; Jentsch *et al.*, 2000; Nuechterlein *et al.*, 2004; Everitt & Robbins, 2005; Kalivas & Volkow, 2005; Robbins, 2005; Schoenbaum *et al.*, 2006).

Taken together, these data suggest it is fundamentally important to study network signaling in diseased and healthy brains. While these concepts are often applied to other brain regions, such as the cortex, they have seldom been applied to subcortical regions, such as the VTA. The intricate patterns of activity in the current dataset suggest that the VTA accomplishes far more than has been previously suggested. In fact, it is exciting to note that heterogeneous patterns of action-evoked activation may balance specialization and diversity, as this type of scheme leads to optimal information encoding (Tripathy *et al.*, 2013). Moving forward, the VTA should be conceptualized as a complex network, which is capable of encoding diverse types of information through multiple encoding regimes.

Appendix A

SUPPLEMENTAL EXPLANATION OF TD ALGORITHM

The TD model predicts the value of each future moment in a trial (Dayan & Abbott, 2001). Equation 8:

$$v(t) = \sum_{\tau=0}^{t} w(\tau)u(t-\tau)$$

Where value, v, is simply reflective of a weighted representation of u, a single time dependent stimulus (u = 1 if the stimulus is present, u = 0 if the stimulus is absent), and w is the weight for that stimulus.

The value of the weight, w, is updated by δ, *the difference between actual and expected future rewards*. This is a reward prediction error. The parameter, ϵ, describes a learning rate, and controls how much weight is given to the prediction error signal in terms of updating the weight, w. Equation 9:

$$w(\tau) \rightarrow w(\tau) + \epsilon\delta(t)u(t-\tau)$$

The TD prediction error signal cannot be calculated directly. The signal represents the difference between actual and expected future rewards. But because future rewards have yet to occur, this quantity must be estimated. The key to accomplishing this comes from the notion that the estimated value at the next moment in time, plus the current reward, is a good estimator of future value (Dayan & Abbott, 2001). Equation 10:

$$\sum_{\tau=0}^{T-t} r(t+\tau) \approx r(t) + v(t+1)$$

The term δ can be calculated from Equation 10. Equation 1 (from Introduction section): $\delta(t) = r(t) + v(t+1) - v(t)$. Thus, the error signal represents an estimate of the actual and current reward value plus estimated value of future rewards, minus the prediction of the current value. The final two terms, $v(t+1) - v(t)$, are where the name 'temporal difference' is derived from. This term represents the difference between estimates at successive time steps (Dayan & Abbott, 2001).

BIBLIOGRAPHY

Abercrombie, E.D., Keefe, K.A., DiFrischia, D.S. & Zigmond, M.S. (1989) Differential effect of stress on in vivo dopamine release in striatum, nucleus accumbens, and medial frontal cortex. *Journal of Neurochemistry*, **52**, 1655-1658.

Akil, H., Brenner, S., Kandel, E., Kendler, K.S., King, M.C., Scolnick, E., Watson, J.D. & Zoghbi, H.Y. (2010) Medicine. The future of psychiatric research: genomes and neural circuits. *Science*, **327**, 1580-1581.

Allison, D.W., Ohran, A.J., Stobbs, S.H., Mameli, M., Valenzuela, C.F., Sudweeks, S.N., Ray, A.P., Henriksen, S.J. & Steffensen, S.C. (2006) Connexin-36 gap junctions mediate electrical coupling between ventral tegmental area GABA neurons. *Synapse*, **60**, 20-31.

Allman, M.J., Pelphrey, K.A. & Meck, W.H. (2011) Developmental neuroscience of time and number: implications for autism and other neurodevelopmental disabilities. *Frontiers in integrative neuroscience*, **6**, 7.

Anstrom, K.K., Cromwell, H.C. & Woodward, D.J. (2007) Effects of restraint and haloperidol on sensory gating in the midbrain of awake rats. *Neuroscience*, **146**, 515-524.

Anstrom, K.K. & Woodward, D.J. (2005) Restraint increases dopaminergic burst firing in awake rats. *Neuropsychopharmacology*, **30**, 1832-1840.

Arnsten, A.F. & Li, B.M. (2005) Neurobiology of executive functions: catecholamine influences on prefrontal cortical functions. *Biol Psychiatry*, **57**, 1377-1384.

Aultman, J.M. & Moghaddam, B. (2001) Distinct contributions of glutamate and dopamine receptors to temporal aspects of rodent working memory using a clinically relevant task. *Psychopharmacologia*, **153**, 353-364.

Bassett, D.S. & Bullmore, E.T. (2009) Human brain networks in health and disease. *Current opinion in neurology*, **22**, 340-347.

Bayer, H.M. & Glimcher, P.W. (2005) Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron*, **47**, 129-141.

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

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

Bromberg-Martin, E.S. & Hikosaka, O. (2009) Midbrain dopamine neurons signal preference for advance information about upcoming rewards. *Neuron*, **63**, 119-126.

Bromberg-Martin, E.S., Matsumoto, M. & Hikosaka, O. (2010a) Distinct tonic and phasic anticipatory activity in lateral habenula and dopamine neurons. *Neuron*, **67**, 144-155.

Bromberg-Martin, E.S., Matsumoto, M. & Hikosaka, O. (2010b) Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron*, **68**, 815-834.

Brozoski, T.J., Brown, R.M., Rosvold, H.E. & Goldman, P.S. (1979) Cognitive deficit caused by regional depletion of dopamine in prefrontal cortex of rhesus monkey. *Science*, **205**, 929-932.

Bunney, B.S., Aghajanian, G.K. & Roth, R.H. (1973) Comparison of effects of L-dopa, amphetamine and apomorphine on firing rate of rat dopaminergic neurones. *Nature: New biology*, **245**, 123-125.

Bush, R.R. & Mosteller, F. (1951) A mathematical model for simple learning. *Psychol Rev*, **58**, 313-323.

Cai, J.X. & Arnsten, A.F. (1997) Dose-dependent effects of the dopamine D1 receptor agonists A77636 or SKF81297 on spatial working memory in aged monkeys. *J Pharmacol Exp Ther*, **283**, 183-189.

Cardinal, R.N., Robbins, T.W. & Everitt, B.J. (2000) The effects of *d*-amphetamine, chloradiazepoxide, a-flupenthixol and behavioural manipulations on choice of signalled and unsignalled delayed reinforcement in rats. *Psychopharmacology,* **152**, 362-375.

Carr, D. & Sesack, S. (2000a) Projections from the rat prefrontal cortex to the ventral tegmental area: target specificity in the synaptic associations with mesoaccumbens and mesocortical neurons. *Journal of Neuroscience,* **20**, 3864-3873.

Carr, D.B. & Sesack, S.R. (2000b) GABA-containing neurons in the rat ventral tegmental area project to the prefrontal cortex. *Synapse,* **38**, 114-123.

Chudasama, Y. & Robbins, T.W. (2004) Dopaminergic modulation of visual attention and working memory in the rodent prefrontal cortex. *Neuropsychopharmacology*, **29**, 1628-1636.

Cohen, J.Y., Haesler, S., Vong, L., Lowell, B.B. & Uchida, N. (2012) Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature,* **482**, 85-88.

Cohen, M.R. & Kohn, A. (2011) Measuring and interpreting neuronal correlations. *Nat Neurosci*, **14**, 811-819.

Constantinidis, C. & Goldman-Rakic, P. (2002) Correlated discharges among putative pyramidal neurons and interneurons in the primate prefrontal cortex. *Journal of Neurophysiology*, **88**, 3487-3497.

Creed, M.C., Ntamati, N.R. & Tan, K.R. (2014) VTA GABA neurons modulate specific learning behaviors through the control of dopamine and cholinergic systems. *Front Behav Neurosci*, **8**, 8.

Crofts, H.S., Dalley, J.W., Collins, P., Van Denderen, J.C., Everitt, B.J., Robbins, T.W. & Roberts, A.C. (2001) Differential effects of 6-OHDA lesions of the frontal cortex and caudate nucleus on the ability to acquire an attentional set. *Cereb Cortex*, **11**, 1015-1026.

Davis, K.L., Kahn, R.S., Ko, G. & Davidson, M. (1991) Dopamine in schizophrenia: a review and reconceptualization. *Am J Psychiatry*, **148**, 1474-1486.

Day, J.J., Jones, J.L., Wightman, R.M. & Carelli, R.M. (2010) Phasic nucleus accumbens dopamine release encodes effort- and delay-related costs. *Biol Psychiatry*, **68**, 306-309.

Day, J.J., Roitman, M.F., Wightman, R.M. & Carelli, R.M. (2007) Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nat Neurosci*, **10**, 1020-1028.

Dayan, P. & Abbott, L.F. (2001) *Theoretical neuroscience : computational and mathematical modeling of neural systems*. Massachusetts Institute of Technology Press, Cambridge, Mass.

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

Dobi, A., Margolis, E.B., Wang, H.L., Harvey, B.K. & Morales, M. (2010) Glutamatergic and nonglutamatergic neurons of the ventral tegmental area establish local synaptic contacts with dopaminergic and nondopaminergic neurons. *J Neurosci*, **30**, 218-229.

Evenden, J.L. & Ryan, C.N. (1996) The pharmacology of impulsive behaviour in rats: the effects of drugs on response choice with varying delays of reinforcement. *Psychopharmacology (Berl)*, **128**, 161-170.

Everitt, B.J. & Robbins, T.W. (2005) Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat Neurosci*, **8**, 1481-1489.

Finlay, J.M., Zigmond, M.J. & Abercrombie, E.D. (1995) Increased dopamine and norepinephrine release in medial prefrontal cortex induced by acute and chronic stress: effects of diazepam. *Neuroscience*, **64**, 619-628.

Fiorillo, C.D. (2013) Two dimensions of value: dopamine neurons represent reward but not aversiveness. *Science*, **341**, 546-549.

Fiorillo, C.D., Newsome, W.T. & Schultz, W. (2008) The temporal precision of reward prediction in dopamine neurons. *Nat Neurosci*.

Fiorillo, C.D., Song, M.R. & Yun, S.R. (2013a) Multiphasic temporal dynamics in responses of midbrain dopamine neurons to appetitive and aversive stimuli. *J Neurosci*, **33**, 4710-4725.

Fiorillo, C.D., Tobler, P.N. & Schultz, W. (2003) Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*, **299**, 1898-1902.

Fiorillo, C.D., Yun, S.R. & Song, M.R. (2013b) Diversity and homogeneity in responses of midbrain dopamine neurons. *J Neurosci*, **33**, 4693-4709.

Floresco, S.B. & Magyar, O. (2006) Mesocortical dopamine modulation of executive functions: beyond working memory. *Psychopharmacology (Berl)*, **188**, 567-585.

Floresco, S.B., Magyar, O., Ghods-Sharifi, S., Vexelman, C. & Tse, M.T. (2006) Multiple dopamine receptor subtypes in the medial prefrontal cortex of the rat regulate set-shifting. *Neuropsychopharmacology*, **31**, 297-309.

Freeman, A.S. & Bunney, B.S. (1987) Activity of A9 and A10 dopaminergic neurons in unrestrained rats: further characterization and effects of apomorphine and cholecystokinin. *Brain Res*, **405**, 46-55.

Fuster, J.M. (1991) The prefrontal cortex and its relation to behavior. *Progress in brain research*, **87**, 201-211.

Fuster, J.M. (2001) The prefrontal cortex--an update: time is of the essence. *Neuron*, **30**, 319-333.

Gallistel, C.R. & Gibbon, J. (2000) Time, rate, and conditioning. *Psychol Rev*, **107**, 289-344.

Gan, J.O., Walton, M.E. & Phillips, P.E. (2010) Dissociable cost and benefit encoding of future rewards by mesolimbic dopamine. *Nat Neurosci*, **13**, 25-27.

Gao, W. & Goldman-Rakic, P. (2003) Selective modulation of excitatory and inhibitory microcircuits by dopamine. *Proceedings of the National Academy of Sciences of U.S.A.*, **100**, 2836-2841.

Geisler, S. & 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 U S A*, **108 Suppl 3**, 15647-15654.

Goldman-Rakic, P. (1998) The cortical dopamine system: Role in memory and cognition. *Advances in Pharmacology*, **42**, 707-711.

Goldman-Rakic, P. (1999) The physiological approach: Functional architecture of working memory and disordered cognition in schizophrenia. *Biological Psychiatry*, **46**, 650-661.

Goldman-Rakic, P.S. (1988) Topography of cognition: parallel distributed networks in primate association cortex. *Annual Reviews in Neurosciences*, **11**, 137-156.

Gorelova, N., Seamans, J.K. & Yang, C.R. (2002) Mechanisms of dopamine activation of fast-spiking interneurons that exert inhibition in rat prefrontal cortex. *J Neurophysiol*, **88**, 3150-3166.

Grace, A.A. & Bunney, B.S. (1983a) Intracellular and extracellular electrophysiology of nigral dopaminergic neurons--1. Identification and characterization. *Neuroscience*, **10**, 301-315.

Grace, A.A. & Bunney, B.S. (1983b) Intracellular and extracellular electrophysiology of nigral dopaminergic neurons--2. Action potential generating mechanisms and morphological correlates. *Neuroscience*, **10**, 317-331.

Grace, A.A. & Bunney, B.S. (1984a) The control of firing pattern in nigral dopamine neurons: burst firing. *Journal of Neuroscience*, **4**, 2877-2890.

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

Granon, S., Passetti, F., Thomas, K.L., Dalley, J.W., Everitt, B.J. & Robbins, T.W. (2000) Enhanced and impaired attentional performance after infusion of D1 dopaminergic receptor agents into rat prefrontal cortex. *J Neurosci*, **20**, 1208-1215.

Gruber, A.J., Hussain, R.J. & O'Donnell, P. (2009) The nucleus accumbens: a switchboard for goal-directed behaviors. *PLoS One*, **4**, e5062.

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

Heien, M.L. & Wightman, R.M. (2006) Phasic dopamine signaling during behavior, reward, and disease states. *CNS & neurological disorders drug targets*, **5**, 99-108.

Hollerman, J.R. & Schultz, W. (1998) Dopamine neurons report an error in the temporal prediction of reward during learning. *Nature Neuroscience*, **1**, 304-309.

Houk, J.C., Adams, J.L. & Barto, A.G. (eds) (1995) *A model of how the basal ganglia generate and use neural signals that predict reinforcement.* The MIT Press.

Howes, O.D. & Kapur, S. (2009) The dopamine hypothesis of schizophrenia: version III--the final common pathway. *Schizophr Bull*, **35**, 549-562.

Hyland, B.I., Reynolds, J.N., Hay, J., Perk, C.G. & Miller, R. (2002) Firing modes of midbrain dopamine cells in the freely moving rat. *Neuroscience*, **114**, 475-492.

Jentsch, J.D., Roth, R.H. & Taylor, J.R. (2000) Role for dopamine in the behavioral functions of the prefrontal corticostriatal system: implications for mental disorders and psychotropic drug action. *Progress in brain research*, **126**, 433-453.

Joel, D., Niv, Y. & Ruppin, E. (2002) Actor-critic models of the basal ganglia: new anatomical and computational perspectives. *Neural Netw*, **15**, 535-547.

Johnson, S.W. & North, R.A. (1992) Two types of neurone in the rat ventral tegmental area and their synaptic inputs. *J Physiol*, **450**, 455-468.

Joshua, M., Adler, A., Mitelman, R., Vaadia, E. & Bergman, H. (2008) Midbrain dopaminergic neurons and striatal cholinergic interneurons encode the difference between reward and aversive events at different epochs of probabilistic classical conditioning trials. *J Neurosci*, **28**, 11673-11684.

Joshua, M., Adler, A., Prut, Y., Vaadia, E., Wickens, J.R. & Bergman, H. (2009) Synchronization of midbrain dopaminergic neurons is enhanced by rewarding events. *Neuron*, **62**, 695-704.

Kalivas, P.W. & Volkow, N.D. (2005) The neural basis of addiction: a pathology of motivation and choice. *Am J Psychiatry*, **162**, 1403-1413.

Kass, R.E., Eden, U.T. & Brown, E.N. (2014) *Analysis of neural data*.

Kim, Y., Wood, J. & Moghaddam, B. (2012) Coordinated activity of ventral tegmental neurons adapts to appetitive and aversive learning. *PLoS One*, **7**, e29766.

Kim, Y.B., Matthews, M. & Moghaddam, B. (2010) Putative gamma-aminobutyric acid neurons in the ventral tegmental area have a similar pattern of plasticity as dopamine neurons during appetitive and aversive learning. *The European journal of neuroscience*, **32**, 1564-1572.

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

Krivanek, J.A. & McGaugh, J.L. (1969) Facilitating effects of pre- and posttrial amphetamine administration on discrimination learning in mice. *Agents and actions*, **1**, 36-42.

Lammel, S., Hetzel, A., Hackel, O., Jones, I., Liss, B. & Roeper, J. (2008) Unique properties of mesoprefrontal neurons within a dual mesocorticolimbic dopamine system. *Neuron*, **57**, 760-773.

Lammel, S., Ion, D.I., Roeper, J. & Malenka, R.C. (2011) Projection-specific modulation of dopamine neuron synapses by aversive and rewarding stimuli. *Neuron*, **70**, 855-862.

Lammel, S., Lim, B.K. & Malenka, R.C. (2014) Reward and aversion in a heterogeneous midbrain dopamine system. *Neuropharmacology*, **76 Pt B**, 351-359.

Lammel, S., Lim, B.K., Ran, C., Huang, K.W., Betley, M.J., Tye, K.M., Deisseroth, K. & Malenka, R.C. (2012) Input-specific control of reward and aversion in the ventral tegmental area. *Nature*, **491**, 212-217.

Lassen, M.B., Brown, J.E., Stobbs, S.H., Gunderson, S.H., Maes, L., Valenzuela, C.F., Ray, A.P., Henriksen, S.J. & Steffensen, S.C. (2007) Brain stimulation reward is integrated by a network of electrically coupled GABA neurons. *Brain Res*, **1156**, 46-58.

Lee, R.S., Steffensen, S.C. & Henriksen, S.J. (2001) Discharge profiles of ventral tegmental area GABA neurons during movement, anesthesia, and the sleep-wake cycle. *J Neurosci*, **21**, 1757-1766.

Lewis, B.L. & O'Donnell, P. (2000) Ventral tegmental area afferents to the prefrontal cortex maintain membrane potential 'up' states in pyramidal neurons via D(1) dopamine receptors. *Cereb Cortex*, **10**, 1168-1175.

Li, W., Doyon, W.M. & Dani, J.A. (2011) Acute in vivo nicotine administration enhances synchrony among dopamine neurons. *Biochem Pharmacol*, **82**, 977-983.

Ljungberg, T., Apicella, P. & Schultz, W. (1992) Responses of monkey dopamine neurons during learning of behavioral reactions. *J Neurophysiol*, **67**, 145-163.

Lustig, C. (2011) The neuroscience of time and number: untying the gordian knot. *Frontiers in integrative neuroscience*, **5**, 47.

Malsburg, C.v.d., Phillips, W.A. & Singer, W. (2010) *Dynamic coordination in the brain : from neurons to mind*. MIT Press, Cambridge, Mass.

Margolis, E.B., Lock, H., Hjelmstad, G.O. & Fields, H.L. (2006) The ventral tegmental area revisited: is there an electrophysiological marker for dopaminergic neurons? *J Physiol*, **577**, 907-924.

Matsumoto, M. & Hikosaka, O. (2007) Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature*, **447**, 1111-1115.

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

McGaugh, J.L. (2000) Memory--a century of consolidation. *Science*, **287**, 248-251.

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

Miller, J.D., Sanghera, M.K. & German, D.C. (1981) Mesencephalic dopaminergic unit activity in the behaviorally conditioned rat. *Life Sci*, **29**, 1255-1263.

Mirenowicz, J. & Schultz, W. (1994) Importance of unpredictability for reward responses in primate dopamine neurons. *J Neurophysiol*, **72**, 1024-1027.

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

Moghaddam, B. & Homayoun, H. (2008) Divergent plasticity of prefrontal cortex networks. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology*, **33**, 42-55.

Moghaddam, B. & Wood, J. (2014) Team work matters: coordinated neuronal activity in brain systems relevant to psychiatric disorders. *JAMA psychiarty*.

Montague, P.R., Dayan, P. & Sejnowski, T.J. (1996) A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci*, **16**, 1936-1947.

Morris, G., Arkadir, D., Nevet, A., Vaadia, E. & Bergman, H. (2004) Coincident but distinct messages of midbrain dopamine and striatal tonically active neurons. *Neuron*, **43**, 133-143.

Morris, G., Nevet, A., Arkadir, D., Vaadia, E. & Bergman, H. (2006) Midbrain dopamine neurons encode decisions for future action. *Nat Neurosci*, **9**, 1057-1063.

Nair-Roberts, R.G., Chatelain-Badie, S.D., Benson, E., White-Cooper, H., Bolam, J.P. & Ungless, M.A. (2008) Stereological estimates of dopaminergic, GABAergic and glutamatergic neurons in the ventral tegmental area, substantia nigra and retrorubral field in the rat. *Neuroscience*, **152**, 1024-1031.

Nakahara, H., Itoh, H., Kawagoe, R., Takikawa, Y. & Hikosaka, O. (2004) Dopamine neurons can represent context-dependent prediction error. *Neuron*, **41**, 269-280.

Naneix, F., Marchand, A.R., Di Scala, G., Pape, J.R. & Coutureau, E. (2009) A role for medial prefrontal dopaminergic innervation in instrumental conditioning. *J Neurosci*, **29**, 6599-6606.

Nishino, H., Ono, T., Muramoto, K., Fukuda, M. & Sasaki, K. (1987) Neuronal activity in the ventral tegmental area (VTA) during motivated bar press feeding in the monkey. *Brain Res*, **413**, 302-313.

Niv, Y., Daw, N.D. & Dayan, P. (2006) Choice values. *Nat Neurosci*, **9**, 987-988.

Niv, Y. & Schoenbaum, G. (2008) Dialogues on prediction errors. *Trends Cogn Sci*, **12**, 265-272.

Nuechterlein, K.H., Barch, D.M., Gold, J.M., Goldberg, T.E., Green, M.F. & Heaton, R.K. (2004) Identification of separable cognitive factors in schizophrenia. *Schizophr Res*, **72**, 29-39.

Oades, R.D., Taghzouti, K., Rivet, J.M., Simon, H. & Le Moal, M. (1986) Locomotor activity in relation to dopamine and noradrenaline in the nucleus accumbens, septal and frontal areas: a 6-hydroxydopamine study. *Neuropsychobiology*, **16**, 37-42.

Omelchenko, N. & Sesack, S.R. (2007) Glutamate synaptic inputs to ventral tegmental area neurons in the rat derive primarily from subcortical sources. *Neuroscience*, **146**, 1259-1274.

Omelchenko, N. & Sesack, S.R. (2009) Ultrastructural analysis of local collaterals of rat ventral tegmental area neurons: GABA phenotype and synapses onto dopamine and GABA cells. *Synapse*, **63**, 895-906.

Paladini, C.A. & Roeper, J. (2014) Generating bursts (and pauses) in the dopamine midbrain neurons. *Neuroscience*, **282C**, 109-121.

Pan, W.X., Schmidt, R., Wickens, J.R. & Hyland, B.I. (2005) Dopamine cells respond to predicted events during classical conditioning: evidence for eligibility traces in the reward-learning network. *J Neurosci*, **25**, 6235-6242.

Pasquereau, B. & Turner, R.S. (2013) Limited encoding of effort by dopamine neurons in a cost-benefit trade-off task. *J Neurosci*, **33**, 8288-8300.

Phillips, A.G., Ahn, S. & Floresco, S.B. (2004) Magnitude of dopamine release in medial prefrontal cortex predicts accuracy of memory on a delayed response task. *J Neurosci*, **24**, 547-553.

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

Pirot, S., Godbout, R., Mantz, J., Tassin, J.P., Glowinski, J. & Thierry, A.M. (1992) Inhibitory effects of ventral tegmental area stimulation on the activity of prefrontal cortical neurons: evidence for the involvement of both dopaminergic and GABAergic components. *Neuroscience*, **49**, 857-865.

Ragozzino, M.E. (2002) The effects of dopamine D(1) receptor blockade in the prelimbic-infralimbic areas on behavioral flexibility. *Learn Mem*, **9**, 18-28.

Ramon-Moliner, E. & Nauta, W.J. (1966) The isodendritic core of the brain stem. *J Comp Neurol*, **126**, 311-335.

Rescorla, R.A. & Wagner, A.R. (1972) *Classical conditioning II: current research and theory*. Appleton-Century-Crofts, New York,.

Reynolds, G.S. (1975) *A primer of operant conditioning*. Scott, Foresman, Glenview, Ill.

Reynolds, J.N. & Wickens, J.R. (2002) Dopamine-dependent plasticity of corticostriatal synapses. *Neural Netw*, **15**, 507-521.

Robbins, T.W. (2005) Chemistry of the mind: neurochemical modulation of prefrontal cortical function. *Journal of Comparative Neurology*, **493**, 146-146.

Robbins, T.W. & Arnsten, A.F. (2009) The neuropsychopharmacology of fronto-executive function: monoaminergic modulation. *Annu Rev Neurosci*, **32**, 267-287.

Robbins, T.W. & Roberts, A.C. (2007) Differential regulation of fronto-executive function by the monoamines and acetylcholine. *Cereb Cortex*, **17 Suppl 1**, i151-160.

Robbins, T.W., Roberts, D.C. & Koob, G.F. (1983) Effects of d-amphetamine and apomorphine upon operant behavior and schedule-induced licking in rats with 6-hydroxydopamine-induced lesions of the nucleus accumbens. *J Pharmacol Exp Ther*, **224**, 662-673.

Robinson, D.L., Venton, B.J., Heien, M.L. & Wightman, R.M. (2003) Detecting subsecond dopamine release with fast-scan cyclic voltammetry in vivo. *Clinical chemistry*, **49**, 1763-1773.

Roeper, J. (2013) Dissecting the diversity of midbrain dopamine neurons. *Trends Neurosci*, **36**, 336-342.

Roesch, M.R., Calu, D.J. & Schoenbaum, G. (2007) Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nat Neurosci*, **10**, 1615-1624.

Romo, R. & 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. (2002) Motivational views of reinforcement: implications for understanding the behavioral functions of nucleus accumbens dopamine. *Behav Brain Res*, **137**, 3-25.

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

Salamone, J.D., Correa, M., Mingote, S.M. & Weber, S.M. (2005) Beyond the reward hypothesis: alternative functions of nucleus accumbens dopamine. *Current opinion in pharmacology*, **5**, 34-41.

Sawaguchi, T. & Goldman-Rakic, P. (1991) D1 dopamine receptors in prefrontal cortex: involvement in working memory. *Science*, **251**, 947-950.

Sawaguchi, T. & Goldman-Rakic, P.S. (1994) The role of D1-dopamine receptor in working memory: local injections of dopamine antagonists into the prefrontal cortex of rhesus monkeys performing an oculomotor delayed-response task. *J Neurophysiol*, **71**, 515-528.

Sawaguchi, T., Matsumura, M. & Kubota, K. (1988) Delayed response deficit in monkeys by locally disturbed prefrontal neuronal activity by bicuculline. *Behav Brain Res*, **31**, 193-198.

Schoenbaum, G., Roesch, M.R. & Stalnaker, T.A. (2006) Orbitofrontal cortex, decision-making and drug addiction. *Trends Neurosci*, **29**, 116-124.

Schultz, W. (1986) Responses of midbrain dopamine neurons to behavioral trigger stimuli in the monkey. *J Neurophysiol*, **56**, 1439-1461.

Schultz, W. (1998) Predictive reward signal of dopamine neurons. *Journal of Neurophysiology*, **80**, 1-27.

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

Schultz, W. (2007) Multiple dopamine functions at different time courses. *Annu Rev Neurosci*, **30**, 259-288.

Schultz, W. (2010) Dopamine signals for reward value and risk: basic and recent data. *Behavioral and brain functions : BBF*, **6**, 24.

Schultz, W. (2013) Updating dopamine reward signals. *Curr Opin Neurobiol*, **23**, 229-238.

Schultz, W., Apicella, P. & Ljungberg, T. (1993) Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *Journal of Neuroscience*, **13**, 900-913.

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

Schultz, W. & Romo, R. (1987) Responses of nigrostriatal dopamine neurons to high-intensity somatosensory stimulation in the anesthetized monkey. *J Neurophysiol*, **57**, 201-217.

Schultz, W. & Romo, R. (1990) Dopamine neurons of the monkey midbrain: contingencies of responses to stimuli eliciting immediate behavioral reactions. *J Neurophysiol*, **63**, 607-624.

Seamans, J.K. & Yang, C.R. (2004) The principal features and mechanisms of dopamine modulation in the prefrontal cortex. *Prog Neurobiol*, **74**, 1-58.

Seeman, P. & Lee, T. (1975) Antipsychotic drugs: direct correlation between clinical potency and presynaptic action on dopamine neurons. *Science*, **188**, 1217-1219.

Sesack, S. & Bunney, B. (1989) Pharmacological characterization of the receptor mediating electrophysiological responses to dopamine in the rat medial prefrontal cortex: a microionophoretic study. *Journal of Pharmacology & Experimental Therapeutics*, **248**, 1323-1333.

Sesack, S.R. & Grace, A.A. (2010) Cortico-Basal Ganglia reward network: microcircuitry. *Neuropsychopharmacology*, **35**, 27-47.

Setlow, B. & McGaugh, J.L. (2000) D2 dopamine receptor blockade immediately post-training enhances retention in hidden and visible platform versions of the water maze. *Learn Mem*, **7**, 187-191.

Setlow, B., Mendez, I.A., Mitchell, M.R. & Simon, N.W. (2009) Effects of chronic administration of drugs of abuse on impulsive choice (delay discounting) in animal models. *Behav Pharmacol*, **20**, 380-389.

Simon, N.W. & Setlow, B. (2006) Post-training amphetamine administration enhances memory consolidation in appetitive Pavlovian conditioning: Implications for drug addiction. *Neurobiology of learning and memory*, **86**, 305-310.

Stefani, M.R. & Moghaddam, B. (2006) Rule learning and reward contingency are associated with dissociable patterns of dopamine activation in the rat prefrontal cortex, nucleus accumbens, and dorsal striatum. *J Neurosci*, **26**, 8810-8818.

Steffensen, S.C., Bradley, K.D., Hansen, D.M., Wilcox, J.D., Wilcox, R.S., Allison, D.W., Merrill, C.B. & Edwards, J.G. (2011) The role of connexin-36 gap junctions in alcohol intoxication and consumption. *Synapse*, **65**, 695-707.

Steffensen, S.C., Lee, R.S., Stobbs, S.H. & Henriksen, S.J. (2001) Responses of ventral tegmental area GABA neurons to brain stimulation reward. *Brain Res*, **906**, 190-197.

Steffensen, S.C., Svingos, A.L., Pickel, V.M. & Henriksen, S.J. (1998) Electrophysiological characterization of GABAergic neurons in the ventral tegmental area. *J Neurosci*, **18**, 8003-8015.

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*, **16**, 966-973.

Steinfels, G.F., Heym, J., Strecker, R.E. & Jacobs, B.L. (1983a) Behavioral correlates of dopaminergic unit activity in freely moving cats. *Brain Res*, **258**, 217-228.

Steinfels, G.F., Heym, J., Strecker, R.E. & Jacobs, B.L. (1983b) Response of dopaminergic neurons in cat to auditory stimuli presented across the sleep-waking cycle. *Brain Res*, **277**, 150-154.

Strecker, R.E. & Jacobs, B.L. (1985) Substantia nigra dopaminergic unit activity in behaving cats: effect of arousal on spontaneous discharge and sensory evoked activity. *Brain Res*, **361**, 339-350.

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

Suri, R.E. & Schultz, W. (1998) Learning of sequential movements by neural network model with dopamine-like reinforcement signal. *Exp Brain Res*, **121**, 350-354.

Swanson, L. (1982) The projection of the ventral tegmental area and adjacent regions: a combined fluorescent retrograde tracer and immunofluorescence study in the rat. *Brain Research Bulletin*, **9**, 321-353.

Swanson, L.W. (2004) *Brain maps III : structure of the rat brain : an atlas with printed and electronic templates for data, models, and schematics*. Elsevier, Academic Press, Amsterdam ; Boston.

Takahashi, Y.K., Roesch, M.R., Stalnaker, T.A., Haney, R.Z., Calu, D.J., Taylor, A.R., Burke, K.A. & Schoenbaum, G. (2009) The orbitofrontal cortex and ventral tegmental area are necessary for learning from unexpected outcomes. *Neuron*, **62**, 269-280.

Takahashi, Y.K., Roesch, M.R., Wilson, R.C., Toreson, K., O'Donnell, P., Niv, Y. & Schoenbaum, G. (2011) Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex. *Nat Neurosci*, **14**, 1590-1597.

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

Thierry, A.M., Stinus, L., Blanc, G. & Glowinski, J. (1973) Some evidence for the existence of dopaminergic neurons in the rat cortex. *Brain Res*, **50**, 230-234.

Tierney, P.L., Thierry, A.M., Glowinski, J., Deniau, J.M. & Gioanni, Y. (2008) Dopamine modulates temporal dynamics of feedforward inhibition in rat prefrontal cortex in vivo. *Cereb Cortex*, **18**, 2251-2262.

Tobler, P.N., Fiorillo, C.D. & Schultz, W. (2005) Adaptive coding of reward value by dopamine neurons. *Science*, **307**, 1642-1645.

Totah, N.K., Kim, Y. & Moghaddam, B. (2013) Distinct prestimulus and poststimulus activation of VTA neurons correlates with stimulus detection. *J Neurophysiol*, **110**, 75-85.

Trantham-Davidson, H., Neely, L., Lavin, A. & Seamans, J. (2004) Mechanisms underlying differential D1 versus D2 dopamine receptor regulation of inhibition in prefrontal cortex. *Journal of Neuroscience*, **24**, 10652-10659.

Tripathy, S.J., Padmanabhan, K., Gerkin, R.C. & Urban, N.N. (2013) Intermediate intrinsic diversity enhances neural population coding. *Proc Natl Acad Sci U S A*, **110**, 8248-8253.

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

Tritsch, N.X., Oh, W.J., Gu, C. & Sabatini, B.L. (2014) Midbrain dopamine neurons sustain inhibitory transmission using plasma membrane uptake of GABA, not synthesis. *eLife*, **3**, e01936.

Tseng, K.Y., Mallet, N., Toreson, K.L., Le Moine, C., Gonon, F. & O'Donnell, P. (2006) Excitatory response of prefrontal cortical fast-spiking interneurons to ventral tegmental area stimulation in vivo. *Synapse*, **59**, 412-417.

Uhlhaas, P.J. & Singer, W. (2010) Abnormal neural oscillations and synchrony in schizophrenia. *Nat Rev Neurosci*, **11**, 100-113.

Ungless, M.A. & Grace, A.A. (2012) Are you or aren't you? Challenges associated with physiologically identifying dopamine neurons. *Trends Neurosci*, **35**, 422-430.

Ungless, M.A., Magill, P.J. & Bolam, J.P. (2004) Uniform inhibition of dopamine neurons in the ventral tegmental area by aversive stimuli. *Science*, **303**, 2040-2042.

Van Bockstaele, E.J. & Pickel, V.M. (1995) GABA-containing neurons in the ventral tegmental area project to the nucleus accumbens in rat brain. *Brain Res*, **682**, 215-221.

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

Vandecasteele, M., Glowinski, J., Deniau, J.M. & Venance, L. (2008) Chemical transmission between dopaminergic neuron pairs. *Proc Natl Acad Sci U S A*, **105**, 4904-4909.

Vandecasteele, M., Glowinski, J. & Venance, L. (2005) Electrical synapses between dopaminergic neurons of the substantia nigra pars compacta. *J Neurosci*, **25**, 291-298.

Veeneman, M.M., van Ast, M., Broekhoven, M.H., Limpens, J.H. & Vanderschuren, L.J. (2012) Seeking-taking chain schedules of cocaine and sucrose self-administration: effects of reward size, reward omission, and alpha-flupenthixol. *Psychopharmacology (Berl)*, **220**, 771-785.

Vijayraghavan, S., Wang, M., Birnbaum, S.G., Williams, G.V. & Arnsten, A.F. (2007) Inverted-U dopamine D1 receptor actions on prefrontal neurons engaged in working memory. *Nat Neurosci*, **10**, 376-384.

Volman, S.F., Lammel, S., Margolis, E.B., Kim, Y., Richard, J.M., Roitman, M.F. & Lobo, M.K. (2013) New insights into the specificity and plasticity of reward and aversion encoding in the mesolimbic system. *J Neurosci*, **33**, 17569-17576.

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

Wanat, M.J., Kuhnen, C.M. & Phillips, P.E. (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. & Maidment, N.T. (2012) Phasic mesolimbic dopamine signaling precedes and predicts performance of a self-initiated action sequence task. *Biol Psychiatry*, **71**, 846-854.

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

Watanabe, M., Kodama, T. & Hikosaka, K. (1997) Increase of extracellular dopamine in primate prefrontal cortex during a working memory task. *J Neurophysiol*, **78**, 2795-2798.

Williams, G.V. & Goldman-Rakic, P.S. (1995) Modulation of memory fields by dopamine D1 receptors in prefrontal cortex. *Nature*, **376**, 393-401.

Winstanley, C.A., Theobald, D.E.H., Dalley, J.W., Cardinal, R.N. & Robbins, T.W. (2006) Double dissociation between serotonergic and dopaminergic modulation of medial prefrontal and orbitofrontal cortex during a test of impulsive choice. *Cereb Cortex*, **16**, 106-114.

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

Wood, J., Kim, Y. & Moghaddam, B. (2012) Disruption of prefrontal cortex large scale neuronal activity by different classes of psychotomimetic drugs. *J Neurosci*, **32**, 3022-3031.

Yamaguchi, T., Sheen, W. & Morales, M. (2007) Glutamatergic neurons are present in the rat ventral tegmental area. *Eur J Neurosci*, **25**, 106-118.

Zhou, F.M. & Hablitz, J.J. (1999) Dopamine modulation of membrane and synaptic properties of interneurons in rat cerebral cortex. *J Neurophysiol*, **81**, 967-976.