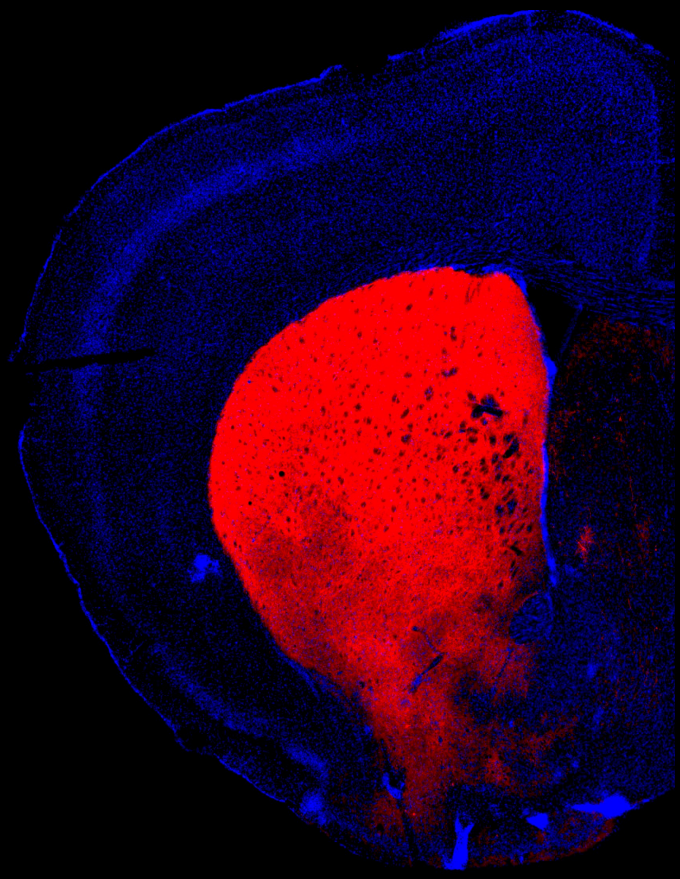
Time in the basal ganglia

The contributions of striatal and midbrain dopamine neurons to timing behavior

Sofia Lima da Silva Soares



Dissertation presented to obtain the Ph.D degree
in Biology | Neuroscience

Instituto de Tecnologia Química e Biológica António Xavier | Universidade Nova de Lisboa

Oeiras,
December, 2017

UNIVERSIDADE NOVA DE LISBOA

Time in the basal ganglia

The contributions of striatal and midbrain dopamine neurons to timing behavior

Sofia Lima da Silva Soares

Dissertation presented to obtain the Ph.D degree
in Biology | Neuroscience

Instituto de Tecnologia Química e Biológica António Xavier | Universidade Nova de Lisboa

Research work coordinated by:

Champalimaud Foundation

FCT
Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA EDUCAÇÃO E CIÊNCIA

FUNDAÇÃO CALOUSTE GULBENKIAN
Instituto Gulbenkian de Ciência

ITQB UNIVERSIDADE NOVA DE LISBOA

Oeiras, December, 2017

Cover image A coronal slice of mouse brain showing the left hemisphere. Midbrain dopamine neurons are expressing tdTomato (red) and in the image we can see dopaminergic projections in the striatum. Nuclei are stained with DAPI (blue).

Time in the Basal Ganglia

the contributions of striatal and midbrain dopamine neurons to timing behavior

Sofia Lima da Silva Soares

A Dissertation presented to the
Instituto de Tecnologia Química e Biológica
of the Universidade Nova de Lisboa
in Candidacy for the Degree
of Doctor of Philosophy

Supervisor: Joseph James Paton

2017

Time is not something objective and real, neither a substance, nor an accident, nor a relation. It is the subjective condition necessary by the nature of the human mind for coordinating any sensible objects among themselves by a certain law; time is a pure intuition.

Immanuel Kant, 1770

Acknowledgments

This work represents the end of an unforgettable journey that was only possible thanks to the incredible group of people that I was so lucky to have by my side. In this all too short page, I would like to give a huge 'thank you' to:

Toda a minha família, obrigada pelo apoio incondicional e por serem sempre um porto seguro. À Eva em especial, porque ser irmã não tem preço;

Joe Paton for all the guidance, mentoring, trust and contagious excitement throughout this journey. My thesis committee members, Susana Lima and Alfonso Renart, for fruitful discussions. All members of the Paton Lab, an incredibly creative, critical and supportive group, that made this experience so much richer. Bassam Atallah, for spending countless hours teaching me many of the tools that made this work possible, and whose persistence and dedication are inspiring;

Gustavo Mello (Guga), um colega caloroso que se tornou um irmão, pelo carinho, alegria, sabedoria e inteligência com que marcou este processo, dentro e fora do laboratório; Gonçalo Lopes, por seres um génio generoso e uma fonte de inspiração; Thiago Gouvêa e Tiago Monteiro, por nunca negarem um pedido de ajuda e por criarem um ambiente de companheirismo do qual já sinto saudades;

All my amazing friends in the INDP2011 class: Banas, Jacques, Jens, João, Luís, Pedro, Roberto and Zaca, for sharing this crazy adventure, the good and the bad, and for never taking ourselves too seriously; Banas levas com outro agradecimento porque para mim és família e quando for grande quero ser como tu;

My friends in Lisbon, for creating the most fantastic experience I could ever ask for. You don't fit all in here, but I must mention Vivek, Diogo, Catarina, Bruno, Tiago, Cindy, Ricardo, Susana, Mert, Kirsten, Claire, Paco, Mattia and Danbee for our meals, games, parties, music, trips and conversations - you are a bright light in my world. Especially to Ana, Antonia and Madalena, three beautiful, inspiring and precious friends who are always there;

Susana, Cláudia e Emanuel, aqueles amigos fora da bolha, que me mostram sempre o que é importante na vida;

And last but never least, to Sam, for being my best friend and the best person I know.

Título

A estimativa de tempo nos gânglios da base: as contribuições dos neurónios dopaminérgicos e do corpo estriado para o comportamento de cronometragem.

Resumo

A capacidade de estabelecer relações temporais entre eventos, assim como de prever as consequências das próprias acções, é essencial para que um animal consiga adaptar-se ao ambiente que o rodeia. Assim sendo, é fundamental que o cérebro tenha acesso a informação temporal. Como é que o cérebro representa a passagem do tempo? E quais são os mecanismos cerebrais que contribuem para a subjectividade das estimavas temporais? Esta dissertação aborda as contribuições do corpo estriado e dos neurónios de dopamina (DA) do mesencéfalo para comportamentos de cronometragem. Começámos por monitorizar a actividade de neurónios do corpo estriado enquanto ratazanas desempenhavam uma tarefa comportamental que exigia estimativas temporais. Observámos que diferentes neurónios respondem em alturas distintas do intervalo a ser estimado. Descobrimos ainda que quando os intervalos a estimar são mais curtos ou mais longos, cada neurónio contrai ou dilata o seu perfil de resposta, respectivamente, revelando um código temporal em termos relativos. De seguida, monitorizámos e manipulámos a actividade dos neurónios de DA da substância negra compacta (SNc) enquanto murganhos categorizavam intervalos de tempo. Descobrimos que estes neurónios reflectem estimativas de tempo, e que a activação ou inibição transiente destes neurónios é suficiente para desacelerar ou acelerar essas estimativas, respectivamente. No mesencéfalo, a relação entre a actividade dopaminérgica e as estimativas temporais mostrou-se especifica à SNc, já que tal relação não foi observada em neurónios de DA da área tegmental ventral. Para além da representação de estimativas temporais, descobrimos ainda que os neurónios do corpo estriado, assim como os neurónios de DA do mesencéfalo, reflectem o nível de empenho do animal na tarefa comportamental que está a ser realizada, isto é, reflectem um estado que se caracteriza por níveis baixos de escolhas prematuras e níveis altos de escolhas correctas. Por fim, observámos que estas representações relacionadas com o empenho na tarefa são dissociáveis das representações relacionadas com estimativas temporais. Estes resultados sugerem que o corpo estriado possui um código ajustável, ao nível populacional, que representa a passagem de tempo, e que os neurónios de DA da SNc reflectem e controlam decisões baseadas em estimativas temporais. Estas representações são, muito provavelmente, cruciais para que os animais consigam prever eventos e consequências e, assim, agir da forma mais apropriada em cada situação.

Abstract

To behave adaptively, animals must learn the temporal structure of events in their environment, and they must also predict the sometimes delayed consequences of their own actions. Therefore, to produce adaptive behavior, it is essential that the brain maintains a representation of time. How does the brain represent elapsed time in a manner that supports adaptive behavior, and what are the mechanisms that contribute to variability in subjective time estimates? In this monograph I address the roles of striatal and midbrain dopamine (DA) neurons in timing behavior. First, we recorded the activity of striatal neurons in rats performing an interval production task. We found that these neurons responded at different delays spanning the interval being timed. In addition, individual neurons rescaled their responses in time when intervals changed, indicating that relative time can be decoded from striatal populations. Next, we both measured and manipulated the activity of DA neurons in the substantia nigra pars compacta (SNc) while mice performed a temporal categorization task. We found that SNc-DA neurons reflected time judgments, and that transiently activating or inhibiting SNc-DA neurons was sufficient to alter these judgments, slowing or speeding time estimation, respectively. Within the midbrain DAergic system, the relationship between DAergic activity and time judgments was specific to the SNc, as it was not present when we targeted DA neurons in the ventral tegmental area (VTA). In addition to timing signals, we found that both striatal and midbrain DA neurons reflected a state of task engagement (i.e., a state of general adherence to task rules, characterized by low levels of premature choices and high probability of correct choices). Importantly, task engagement signals were dissociable from signals related to temporal judgments. These results suggest that the striatum exhibits a scalable population code for time, and that variable time judgments are reflected and controlled by SNc-DA neurons. Such timing signals are likely crucial for animals to generate predictions and to guide their actions in order to produce adaptive behavior.

Author Contributions

Experiments were designed by Sofia Soares, Bassam Atallah, Gustavo Mello and Joseph Paton. Data were acquired by Sofia Soares, Bassam Atallah and Gustavo Mello, with contributions from Alessandro Braga, Tiago Monteiro, Thiago Gouvêa and Bruno Cruz in Chapters 3 and 4. Data were analyzed by Sofia Soares, Bassam Atallah, Gustavo Mello and Joseph Paton with contributions from Asma Motiwala, Thiago Gouvêa, Tiago Monteiro and Bruno Cruz in Chapter 4. Chapter 2 and part of Chapter 3 are adapted from the manuscripts reprinted in Appendix A (Mello, Soares, & Paton, 2015; Soares, Atallah, & Paton, 2016) which were written by all authors. Chapters 1, 4, 5 as well as part of Chapter 3 contain unpublished material and were written by Sofia Soares with input from Joseph Paton.

Financial Support

This work was funded by the Portuguese FCT (Fundação para a Ciência e a Tecnologia), Fundação Champalimaud, Fundação Bial, Simons Foundation, European Molecular Biology Organization and Marie Curie Actions.

Overview

This monograph is structured into five Chapters. Chapter 1 consists of a general introduction, in which literature of relevance for the topics discussed in the following Chapters is presented. In particular, Chapter 1 covers the importance of time for adaptive behavior and briefly describes the history of time estimation research. Furthermore, I introduce the main behavioral paradigms and theoretical models of interval timing, while highlighting their possible neural basis. Finally, I introduce the idea that a common neural (and theoretical) basis is shared between interval timing and reinforcement learning (RL). In this initial Chapter, I opted to only reference studies published until the year of 2014, since this was the work that informed the conception of the projects described in the three subsequent Chapters. These three Chapters correspond to the body of work produced during my doctoral studies. Chapter 2 describes a scalable population code for time found in the rat striatum. Additionally, consistent with the role of the basal ganglia (BG) in action selection, Chapter 2 describes populations of striatal neurons that multiplexed information about action and time. Chapter 3 investigates the role of midbrain dopamine (DA) neurons in mice judging duration. In this Chapter I describe how DA neurons in the substantia nigra pars compacta (SNc), but not in the ventral tegmental area (VTA), reflect and control temporal judgments in a manner consistent with their encoding of a reward prediction error (RPE) signal. In Chapter 4, I present an additional signal in the BG, in a period before trial initiation, that reflects task engagement levels in both rats and mice. This signal, found in both SNc-DA and VTA-DA neurons, is dissociable from the timing signals described in Chapter 3. Furthermore, I describe how the activity of different striatal neuronal subtypes also reflects task engagement signals. Chapter 5 consists of a general discussion where I summarize and contextualize our findings. Additionally, I address possible contradictions and questions that arise from our work, while arguing in favor of a unified view of interval timing and RL. Lastly, I discuss why it might be adaptive for a system that is linked to reward processing to also impact temporal processing.

Contents

List of Figures

Acronyms

2AFC two alternative forced-choice. 7, 8

AAV adeno-associated virus. 61, 106, 108

ADHD attention deficit hyperactivity disorder. 18, 99, 118

AP anterior-posterior. 131

auROC area under the ROC curve. 71

BeT behavioral theory of timing. 12, 13

BG basal ganglia. xiii, xix, xxii, 17–19, 21, 24, 26, 28, 47, 48, 50, 97, 100, 115, 118, 119, 129, 131, 136, 140–142, 144

BIC bayesian information criterion. 71, 72, 79, 94, 106, 116

CI confidence interval. 71, 75, 79, 82, 106, 116

CNO clozapine N-oxide. 60, 90

COM center of mass. xix, 35, 36

CPu caudate-putamen. 19, 131, 133

CR conditioned response. 4, 5

Cre cyclic recombination enzyme. 63, 87, 94, 108, 109, 112, 120

CS conditioned stimulus. 4, 136, 140

CSC complete serial compound. 25, 26

DA dopamine. xi, xiii, xxii, 17–22, 25, 26, 28, 57–61, 63–75, 77–81, 84–88, 91–94, 97, 99, 100, 104–108, 111–120, 124, 127, 129, 131, 134, 135, 137–142, 144

DAT dopamine transporter. 63, 87, 94, 120

DL dorsal-lateral. 131, 132, 140, 141

DM dorsal-medial. 132

dMSN direct pathway MSN. 19, 108–113, 115, 116, 120, 124

DS dorsal striatum. 59, 63, 85, 105, 106, 108, 111, 112, 124, 129, 133, 140, 141

DV dorsal-ventral. 132

FI fixed interval. xix, xx, 6, 7, 10, 15, 22, 29–47, 49–55

fMRI functional magnetic resonance imaging. 15, 22, 28

FSI parvalbumin positive, fast spiking interneuron. 116, 125

GABA γ-aminobutyric acid. 19, 108

GABAa $GABA_a$ receptor. 41, 53

GFP green fluorescent protein. 94

GLM generalized linear model. 97, 100, 102–104, 111, 112, 114, 121–123

GP globus pallidus. 17, 108

GPe external globus pallidus. 19

GPi internal globus pallidus. 19

iMSN indirect pathway MSN. 19, 108–113, 115, 116, 120, 124

ITI inter-rial interval. 4, 86, 100

L-dopa L-3,4-dihydroxyphenylalanine. 20, 21

LeT learning to time. 13

MFC dorsomedial frontal cortex. 133, 134

ML medial-lateral. 132

MSN medium spiny neuron. 19, 108, 111, 112, 114–119, 127, 133

MTS multiple time scale. 13

MVT marginal value theory. 143

NAc nucleus accumbens. 19, 59, 85, 105, 106, 124, 131, 139, 140

PCA principal component analysis. 51

PFC prefrontal cortex. 15, 16

PI peak-interval. 6, 7, 20–22, 142

pInt putative interneuron. 115–117, 125, 126

pMSN putative MSN. 115, 125

PPC posterior parietal cortex. 16

PPN pedunculopontine nucleus. 19

PSE point of subjective equivalence. 8, 9

PST pressing start time. xix, 30–32, 40, 41, 43, 45, 47, 49–51, 53, 54

PSTH peri-stimulus time histogram. 34, 42, 43, 47, 52

R^2 coefficient of determination. 71

RL reinforcement learning. xiii, 2, 23–26, 48, 58, 130, 142–144

RNN recurrent neural network. 14, 16, 133, 135

ROC receiver operating characteristic. 71, 91

RPE reward prediction error. xiii, 23, 25, 26, 57, 58, 65–67, 70, 72, 78, 79, 85, 86, 91, 118, 119, 139, 140

RT response time. xxi, 64, 65, 75, 77, 80–82, 84, 100, 101, 103, 115, 121, 122

SC superior colliculus. 19

SD standard deviation. 10, 11, 52, 56, 60, 66, 91

SDF spike density function. xx, 32–36, 47, 51

SEM standard error of the mean. 30, 31, 60, 64, 65, 67, 71, 75, 79, 84, 101, 102, 105, 107

SET scalar expectancy theory. 11–13

SFI serial fixed interval. 28–30, 32, 34, 43, 45, 46, 49, 52–54

SMA supplementary motor area. 15, 16

SNc substantia nigra pars compacta. xi, xiii, 17, 19, 20, 57–59, 61, 63–75, 77–81, 83–86, 88, 89, 91–94, 100, 104–106, 118, 119, 124, 129, 131, 137–142

SNr substantia nigra pars reticulata. 17, 19, 108

STM spectral timing model. 13

STN subthalamic nucleus. 17, 19, 59, 108

TD temporal-difference. 23–26

TH tyrosine hydroxylase. 59, 63, 74, 87, 94, 120

US unconditioned stimulus. 4, 5, 136, 140

VM ventral-medial. 131, 132, 141

VTA ventral tegmental area. xi, xiii, 17, 18, 57–59, 77–81, 83–86, 89, 92, 93, 100, 104–106, 118, 119, 124, 131, 138, 139, 141

Chapter 1

Introduction

> "Timing is everything – whether in making shots, in making love or in making dinner. Indeed, it is difficult to conceive of an action that does not require temporal control."
>
> (Gibbon & Malapani, 2006)

1.1 Why time?

One incredible feature in the animal kingdom is the vast range of complex behaviors that are executed with exquisite temporal precision in the appropriate context. From a hungry bird foraging for food, to a lioness waiting for the perfect moment to attack a zebra, virtually all behaviors are executed in complex sequences that are temporally coordinated. But animals don't just exhibit these behaviors at random. What makes most animal behavior astounding is that it is *adaptive*. In a given environmental context, an animal will learn to produce a particular behavior at a specific time in order to obtain a desired outcome. In a different context, however, the selected behavior, as well as its timing, might be completely different. How do animals achieve this remarkable ability? The challenge of producing adaptive behavior can be split into four parts: i) animals must sense their environment and identify relevant features of the current context; ii) from the vast range of behaviors at their disposal, animals must select the behavior most likely to lead to the best possible outcome given the context; iii) animals must perform the selected behavior at the appropriate time; and finally, iv) animals must observe the outcome of the behavior they selected and use that information to guide future behavior.

For all the above, a representation of time is crucial. Firstly, any environmental context is composed of stimuli distributed over space and time. Secondly, the process of selecting the optimal behavior requires that animals *predict* the delayed outcome of each possible behavior. They must do so by evaluating prior outcomes experienced under a context-behavior contingency and then selecting the behavior that they expect will lead to the best possible outcome. To infer these causal links and generate predictions, a temporal representation of when contexts-behaviors-outcomes occurred relative to each other is required. Additionally, a representation of elapsed time is essential for animals to exhibit the selected behavior in the appropriate moment. Finally, once animals experience the outcome of their selected behavior, a representation of time is necessary for a causal relationship to be formed between context-action and outcome, so that animals update their predictions about what outcomes to expect next time they encounter the same context and select the same action.

These same problems that govern adaptive behavior are at the heart of reinforcement learning (RL) theory, where an agent's goal is to learn which actions to take in a given state in order to maximize expected future rewards (Sutton & Barto, 1998). A key issue that RL tries to solve is the credit-assignment problem: how to distribute credit appropriately between the many actions that an agent may perform until a delayed reward occurs (Minsky, 1961; Barto, Sutton, & Anderson, 1983). Therefore, in essence, understanding how animals represent temporal information is key to understanding how they learn to behave adaptively.

Although animals are able to represent many relevant timescales [there is extensive evidence for molecular clocks that regulate circadian rhythms of about 24 hours (Reppert & Weaver, 2002), and for representations in the microsecond to millisecond range that are crucial for fine motor control (Braitenberg, 1967; Ivry & Keele, 1989)], adaptive behavior relies on temporal representations in the range of seconds to minutes to hours (interval timing, e.g., Rakitin et al., 1998; Staddon & Higa, 1999; Buhusi & Meck, 2005). Virtually all animal species rely on their estimates of time to guide adaptive behavior. Throughout evolution the behavioral and neural mechanisms that allow animals to estimate elapsed time in the absence of mechanical devices such as modern clocks have been selected. What do we know about how animals estimate the passage of time?

1.2 A brief history of time estimation research

Before experimental psychologists and neuroscientists began to systematically study timing behavior and its neural correlates, the notion of time as a crucial dimension guiding behavior had been recognized and discussed by philosophers for centuries [e.g., Augustine of Hippo (Jordan, 1955); Immanuel Kant (Kant, 1770); William James (James, 1886)]. Although throughout the XIXth century there was an effort among experimental psychologists to study temporal perception in humans, this endeavor did not prove fruitful from the get go. According to Judd et al., 1899:

> "The investigations of time-perception are among the most difficult undertaken in experimental psychology. If one overlooks the earlier, rather crude experiments, the number of valuable treatments of this subject reduces to five or six. Unfortunately, the results thus far obtained are frequently of such a conflicting character that one feels that the whole ground must be gone over anew before any interpretation can be commenced. The articles have come to take on a polemical and too often personal character that does not tend to stimulate unqualified acceptance of the statements of either party. The personal equation doubtless plays a very large part in estimation of time intervals, and perhaps it would be well for investigators to recognize this fact once for all, and not to be too hasty in generalizing from their own individual observations or from those of their own 'school'. The theories, too, are hypothetical in the last degree, and must be so regarded."

One of such "valuable treatments" was an 1892 study by Heim[1], who addressed the relationship between temporal perception and near-death experiences in human subjects that had suffered almost fatal falls in the Alps. For Heim's subjects:

> "Time became greatly expanded. (...) Mental activity became enormous, rising to a hundred-fold velocity(...)."

In other words, Heim's subjects reported that external time appeared to slow down. This apparent slowing down of external time was accompanied by accounts of *"mental quickness"*. This description exposes the sometimes counterintuitive relationship between external and internal time: if one's internal representation of time is sped up, external time seems to slow down. We'll come back to the topic of speeding or slowing down of time estimates in Section 1.5, but note that just as Judd et al. described, the

[1] This work was translated to English and expanded upon by Noyes Jr and Kletti in 1972, 80 years after Heim.

subjective nature of temporal estimation posed real obstacles to the study of time perception. Until then, introspective studies such as Heim's were the norm, where human subjects were asked questions and verbally reported their experiences. Controlled behavioral tasks, theories of learning and mathematical models of timing behavior would only flourish in the XXth century[2].

1.3 Behavioral paradigms for the study of learning and interval timing

Classical conditioning The XXth century saw a proliferation of controlled experimental paradigms to study learning and timing behavior. Between the late 1890's and the 1920's, Ivan Pavlov (Pavlov, 1927; Pavlov & Anrep, 2003) studied dogs' salivary responses to auditory stimuli that preceded the delivery of food, introducing the paradigm of classical (or Pavlovian) conditioning. In classical conditioning, a conditioned stimulus (CS), such as an auditory stimulus, is repeatedly paired with an unconditioned stimulus (US), such as food. Initially, only the US elicits a response (such as salivation), but after enough repetitions, the same response will be elicited by the CS, and therefore is termed conditioned response (CR).

Although the initial theory behind classical conditioning posited that animals learn these CS-US associations because they occurred together in time (in other words, due to temporal *contiguity*, Pavlov, 1927), a series of experiments from Kamin and Rescorla highlighted *surprise* and *contingency* as key factors guiding conditioning. Firstly, Kamin (1969) used a so-called *blocking* procedure to show that learning only takes place if a reward is unpredicted. In other words, after an animal learns a given CS-US pairing, if a second CS is introduced together with the first CS such that both are paired with the same US, it will fail to elicit a CR when it is later presented alone. The explanation for this is that, despite the clear temporal contiguity, the initial CS-US training *blocks* any future learning to the second CS. Secondly, a set of experiments by Rescorla (1966, 1967) showed that in fact *contingency* between CS and US was essential for learning to occur. Rescorla found that adding enough extra USs randomly during the inter-rial interval (ITI) without a preceding CS abolished the CR to the CS.

Together, these experiments showed how classical conditioning paradigms require a combination of contingency and unpredictability between a CS and US to drive learning.

[2]Although early work from Thorndike in 1898 laid ground for his "Law of effect", this principle was only formulated into the XXth century.

Because the vast majority of CRs are reflexive reactions to the US to begin with, the classical conditioning framework does not address the issue of how animals learn to perform non-reflexive behaviors that produce rewarding outcomes.

Operant conditioning Edward Thorndike (1898, 1911) was a pioneer in the study of how animals learn to produce behaviors that are followed by positive outcomes. In one of his best known experiments, Thorndike placed hungry cats individually inside a box from which they could escape by operating latches (Thorndike, 1911). If they succeeded in escaping the box they received food and then were placed inside the box for a new trial at escaping. By recording how long cats would take to escape on single trials, Thorndike found that cats escaped the box more quickly as the number of trials increased. Thorndike introduced the idea of learning by trial and error: if an animal performs an action that is followed by an outcome that produces "*satisfaction to the animal*", then that action will "*be more firmly connected with the situation*". Therefore, if the same situation occurs, that same action will also be more likely to occur. Furthermore, Thorndike suggested that the opposite relationship was true for actions that were followed by outcomes that caused "*discomfort to the animal*". Thorndike named his theory the "*Law of effect*", because it described the effect of outcomes on behavior.

These ideas planted the seeds for the work of Skinner (1938), although Skinner moved away from Thorndike's reference to animals' *sensations*. Instead, Skinner measured behavior and manipulated it using *reinforcement*: the strengthening of behavior by the presentation of an event (positive reinforcement) or by the removal/avoidance of an aversive event (negative reinforcement). These ideas defined operant conditioning, and the use of reinforcers to control behavior was accompanied by the design of a number of paradigms. These paradigms differed in how the reinforcers were delivered over time (in other words, they had distinct *schedules of reinforcement*). It soon became apparent that subjects trained in such schedules of reinforcement exhibited response patterns with clear temporal regularities, leading to the application of these paradigms to the study of timing behavior. Since then, a vast range of paradigms have been developed to study interval timing. Below is a short selection of some of the most common interval production, interval reproduction and interval discrimination paradigms (more detailed reviews on interval timing paradigms can be found in Grondin, 2010 or Merchant, Pérez, Zarco, & Gámez, 2013).

1.3.1 Interval production

Initial operant conditioning paradigms became crucial to the study of timing behavior and fall in the category of interval production paradigms. In interval production paradigms, an animal's subjective temporal estimate is correlated with the timing of a measurable self-paced response that occurs relative to a particular event, usually the last reinforcer presentation. Two of the most popular interval production paradigms are described below.

Fixed interval schedules of reinforcement Operant conditioning under fixed interval (FI) schedules of reinforcement is one of the earliest and most common paradigms in the study of interval timing behavior (Ferster & Skinner, 1957; Cumming & Schoenfeld, 1958; Grossmann, 1973; Dews, 1978; Jaldow, Oakley, & Davey, 1989; Lejeune & Wearden, 1991). FI paradigms reinforce a response (e.g., pressing a lever) that occurs after a fixed elapsed time since the last reinforcer. After a response is reinforced, a new FI begins and only when it has elapsed does a response produce a new reinforcer, and so on. The temporal regularities in animals' patterns of responding during FI paradigms have been described extensively (Ferster & Skinner, 1957; Catania & Reynolds, 1968; Lejeune & Wearden, 1991). Subjects typically respond very little or not at all immediately following the delivery of the reinforcer. This period is often referred to as the *post-reinforcement pause*. Then, approximately two-thirds into the FI, animals begin responding with high frequency without any external cue signaling how much time has elapsed. This moment is used as a correlate of the animal's estimate of the end of the FI and of the availability of a new reinforcer. When animals gradually increase their rate of responding, this pattern is referred to as *scalop* (Dews, 1978). When animals respond at a constant rate from the end of the post-reinforcement pause until the end of the FI, they are said to display a *break and run* pattern (Ferster & Skinner, 1957; Cumming & Schoenfeld, 1958)[3]. By the end of the FI, responses stop and animals collect the reinforcer. In FI schedules, the post-reinforcement pause is taken as an estimate of the subject's temporal expectation of reinforcement. There are a number of variations of the FI paradigm; one of the most widely used is the peak-interval (PI) procedure.

Peak-interval The PI procedure is essentially a modification of the FI paradigm where the reinforcer is omitted in a small proportion of probe (or *peak*) trials (Catania, 1970; Roberts, 1981). On average, responses in these probe trials begin as in the normal

[3]Ferster and Skinner (1957) referred to this pattern as *break-throughs*

trials (with the same patterns as in the FI paradigm) peaking at around the time of the FI, after which the average rate of responding slowly decreases. Therefore, the peak of the average response rate on probe trials (also referred to as peak-rate time) allows the experimenter to infer the subject's estimate of when the reinforcer should have been delivered. FI and PI paradigms have been successfully implemented in a vast range of species, from bees (e.g., Grossmann, 1973) to humans (e.g., Rakitin et al., 1998). One disadvantage of these classic paradigms, however, is that only one interval duration is probed across all trials. Other disadvantages will be discussed in Chapter 2.

1.3.2 Interval reproduction

As the name suggests, interval reproduction paradigms rely on a subject's ability to reproduce an interval duration that has been previously presented to them. The subject needs to hold the presented interval in memory for some period of time before reproducing their estimate of that interval. The exact duration of the interval to be reproduced can vary from trial to trial, granting this class of tasks more flexibility. The reproduction can be achieved verbally or by behaviors such as tapping at the beginning and end of the estimated interval or key-pressing for the entire estimated duration (Woodrow, 1930; Treisman, 1963; Jazayeri & Shadlen, 2010). Interval reproduction is therefore more complex than interval production or discrimination, and therefore has not been as successfully applied to non-primate animals.

1.3.3 Interval discrimination

The principles of psychophysics developed by Fechner (1966, but originally published in 1860) began to be applied to timing behavior in laboratory animals with the development of temporal discrimination tasks. Psychophysics studies the behavioral effects of parametric variations along some physical dimension of a stimulus, and has been extensively applied in the context of perceptual decision-making (Parker & Newsome, 1998). A common approach in psychophysics is to train animals in tasks where one of two choices is reinforced depending on a previously presented context. The context can be manipulated by varying any one physical dimension of a stimulus - for example, the vibration frequency of somatosensory stimuli (Mountcastle, Talbot, Sakata, & Hyvarinen, 1969), or the contrast/coherence in the motion direction of visual stimuli (Newsome & Pare, 1988). This class of paradigms is commonly referred to as two alternative forced-choice (2AFC) tasks.

In the case of 2AFC tasks based on temporal discrimination, the context can be defined, for example, by stimuli of different durations. Typically, subjects are presented with a stimulus with a given duration and discrimination is evident if the probability of performing one of two behaviors is dependent on the duration that was presented. Initial interval discrimination paradigms reinforced behavior when a given stimulus duration was presented (e.g., a 1.5 sec stimulus) but not when a different duration (e.g., a 4 sec stimulus) was presented (Woodrow, 1928; Reynolds & Catania, 1962). In another study, Cowles and Finan (1941) trained rats to run through different arms of T-maze depending on the duration of time for which they had been restrained before being released into the maze. These initial studies led to the development of several paradigms with better stimulus control. Some of the most widely used types of temporal discrimination tasks are described below.

Temporal bisection In temporal bisection paradigms (Stubbs, 1968; Church & Deluty, 1977), subjects are trained to perform one of two actions depending on the duration of a stimulus. Reinforcers are delivered following correct responses to the *short* or *long* stimulus durations. Typically, these responses occur at different spatial locations in a training box, and are referred to as the *short* and *long* choices (or locations). Once subjects have learned to perform accurately under these contingencies, intermediate interval durations are presented and the frequency of responding at either of the two locations is recorded. These intermediate durations are either reinforced randomly or not reinforced, and they allow the experimenter to obtain the duration at which subjects respond with equal probability at the locations associated with the extreme *short* or *long* durations. This point is called the point of subjective equivalence (PSE), and it has been observed that the PSE in this class of paradigms lies close to the geometric mean of the two extreme initial durations.

Both Stubbs (1968) and Church and Deluty (1977) designed their paradigms in such a way that their pigeons and rats, respectively, had to wait until the end of stimulus presentation before reporting their choice. In 1983, Platt and Davis published results under a modified version of the bisection paradigm, allowing pigeons to respond at either the *short* or *long* location at any point in the trial. Although anticipatory responses were allowed, they were not reinforced, and as in the classic bisection paradigm, the first response after the stimulus elapsed was recorded as the actual choice, and was rewarded if performed at the correct location. During a trial, Platt and Davis observed that pigeons began by responding at the *short* location, and if the *short* delay elapsed without a reinforcer being delivered, pigeons would then eventually *switch* to the *long*

location. Because of this *switching* behavior, this version is sometimes called a switching paradigm, and has been applied to other species such as the mouse (Balcı et al., 2008). Behavioral data in this paradigm (Platt & Davis, 1983; Balcı et al., 2008) builds on the knowledge from the classic bisection paradigms in two aspects: i) like the PSE, switching also occurred at the geometric mean of the two intervals but, unlike the PSE, it can be seen as a continuous metric over trials; ii) this paradigm is useful to infer bisection points only if the ratio between the short and long delays is not larger than 1:4[4].

Interval categorization Creelman (1962) was one of the first to implement interval categorization paradigms. Interval categorization can be seen as a modified version of the temporal bisection paradigm where the experimenter controls the boundary of categorization by reinforcing correct responses to intermediate durations (Creelman, 1962; Leon & Shadlen, 2003; Gouvêa, Monteiro, Soares, Atallah, & Paton, 2014). This modification confers an advantage to interval categorization over temporal bisection, because animals have an increased incentive to perform all trials as accurately as possible. As in the temporal bisection paradigm, one can infer the PSE in temporal categorization tasks.

1.4 Theoretical models of interval timing behavior

Until the middle of the XXth century, despite the growing number of behavioral studies, there was no consensus in the field regarding a mathematical model of timing behavior. As early as 1886, James suggested that:

> "It is but dates and events, *representing* time; their abundance *symbolizing* its length. I am sure that this is so, even where the times compared are no more than an hour or so in length."

This idea that time could be represented by the quantity of events that occur in a given period would become the basis for one of the most influential classes of timing behavior models, pacemaker-accumulator-based models. From the early days of interval production and discrimination tasks, two features of timing behavior were prevalent across studies and animal species. Firstly, the average post-reinforcement-pause

[4]If the ratio between the two durations is larger than 1:4, animals would stop responding and the resulting switching or PSE was ambiguous (Platt & Davis, 1983; Balcı et al., 2008).

or peak-rate times were shown to be proportional to the FI duration. This feature has been referred to as *proportional timing* (Staddon & Higa, 1999) and *flexible accuracy* (Gibbon & Malapani, 2006). Secondly, there was a linear relationship between the standard deviation (SD) of either post-reinforcement-pause or peak-rate times and their average values [in other words, their coefficient of variation (SD over mean) is constant]. This second feature was interpreted as a reflection of Weber's law[5] applied to the timing domain, which was later referred to as *scalar timing/scalar property/scalar variability/Weber timing* (Gibbon, 1977; Gibbon & Church, 1984; Staddon & Higa, 1999; Gallistel & Gibbon, 2000). Below is a brief account of some mathematical models of timing behavior (a review of interval timing models can be found in Matell, Meck, et al., 2000; Ivry & Schlerf, 2008; Buhusi & Meck, 2005).

1.4.1 Pacemaker-accumulator models

The classic version of the pacemaker-accumulator model, as well some of its successors, are described in this subsection.

The classic pacemaker-accumulator model In 1963, Treisman proposed the classic pacemaker-accumulator model, giving rise to arguably the most influential class of interval timing behavior models. Treisman's ideas were influenced by the work of Hoagland (1933), who suggested that subjective time estimates could result from a "chemical clock" that is capable of generating temperature-dependent rhythmic activity. Treisman's 1963 pacemaker-accumulator model consists of an *internal clock* unit (or a pacemaker) that emits pulses, a counter unit that accumulates the number of emitted pulses, a memory unit that can store the pulse count for a given duration, and a decision unit that compares the current pulse count with a stored value and can produce a decision (e.g., deciding whether two stimulus durations are different, or deciding when to begin responding in an interval production task). The pacemaker emits pulses at a given inter-pulse-interval that is described to be relatively constant within each trial but variable across trials, varying around a fixed mean depending on the value of a parameter that reflects the subject's level of arousal. Due to this feature, and without assuming *a priori* Weber's law, the classic version of the pacemaker-accumulator model

[5]Weber's law states that the just noticeable difference in intensity between two stimuli is a constant proportion of the initial stimulus intensity. Another way to interpret Weber's law is to say that errors in discrimination should be constant if the two stimulus intensities are proportionally increased or decreased. This suggests that the distribution of intensity estimates over several trials should be scale-invariant: their coefficient of variation should be constant.

accounts for the timing analogue of Weber's law observed in interval production and discrimination paradigms (Treisman, 1966).

Scalar expectancy theory The scalar expectancy theory (SET) of timing behavior (Gibbon, 1977) and its information processing implementation (Gibbon & Church, 1984) are heavily inspired by Treisman's 1963 pacemaker-accumulator model, but scalar timing was one of the main aspects that the theory was explicitly devised to address. Gibbon's 1977 work further supported that Weber's law was applicable to the timing domain by highlighting that, across a wide range of studies, dividing the distributions of response times for each schedule of reinforcement by their mean values produced almost perfectly overlapping distributions. To account for this scalar timing pattern, Gibbon (1977) devised a theory in which interval timing behavior is a result of a ratio comparison between two *expectancies*: i) an instantaneous measure of reinforcement rate (calculated as the inverse of the current estimate of time until reinforcement); and ii) the rate of reinforcement estimated over a whole session. Gibbon tested four models to generate temporal estimates but only his "scalar timing" model (where estimates are normally distributed with mean and SD proportional to a reference interval) was consistent with the timescale invariance of behavior; hence the *scalar expectancy* name of his theory. This approach was successful at explaining the data from the behavioral studies mentioned in the previous Section. For example, in temporal production paradigms, "break-and-run" behaviors are generated when the ratio between the two expectancies crosses a fixed arbitrary threshold.

In 1984, Gibbon and Church formalized an information processing implementation of SET. In this implementation, a Poisson-variable pacemaker generates independent and exponentially distributed inter-pulse intervals (Gibbon & Church, 1984; Gibbon, Church, & Meck, 1984). These pulses are accumulated over the course of a trial and serially stored in memory after reinforcement. A consequence of the Poisson-variable inter-pulse-interval is that the variance in estimates increases with the square of the mean of the interval, which results in a decrease of the coefficient of variation in proportion to the square root of the interval being estimated. In other words, this within-trial Poisson variability should cause estimates to be relatively more precise for longer intervals than for shorter ones, which is inconsistent with behavioral data. Therefore, to account for scalar timing in their model, Gibbon and Church (1984) explored a number of modifications, including ratio-based memory comparisons. Because, in each trial, accumulated values are stored in memory, after a few trials the memory unit contains a distribution of values of memorized reinforcement times. SET randomly selects one

sample from memory and performs a comparison between the current estimate and the selected memorized estimate. This comparison is performed in the form of a ratio as described above for the temporal production case. For the temporal discrimination case between categories A and B, SET requires an extra memory unit and a slightly more complex decision rule. One memory unit sample is required for each category (M_A and M_B) and a comparison needs to be performed between the ratios of the current accumulated sample (X_t) and the sample drawn from each memory unit. For example, if $X_t/M_A < M_B/X_t$, then the model favors the category A choice. This noisy memory ratio-discrimination rule allows for scalar timing in SET, but its plausibility has been challenged since the basis for this noisy-memory rule is unclear (Staddon & Higa, 1999), particularly since noisy-memory ratio comparisons are not the only way to account for scalar timing in pacemaker-accumulator models (Simen, Rivest, Ludvig, Balci, & Killeen, 2013). For example, Simen et al. (2013) developed a model that combines concepts from drift-diffusion and pacemaker-accumulator models, where the rate of a Poisson-variable pacemaker is adapted depending on the interval being estimated while maintaining the accumulator threshold fixed. Simen et al.'s model, like Treisman's, accounts for scalar timing without assuming Weber's law *a priori*. In fact, Gibbon and Church (1984) suggested another variation to SET: removing within-trial Poisson-variance in the pacemaker and including across-trial pulse rate variance that is normally distributed would reproduce the scalar property. This produces a much simpler model that is similar to Treisman's and where scalar timing can be achieved without a ratio comparison rule.

1.4.2 Time as a sequence of states

From the work described in Section 1.3 to more recent studies (e.g., Balcı et al., 2008), we know that animals exhibit sequences of temporally structured behaviors if, for example, there are strong temporal regularities in reinforcer delivery. Killeen and Fetterman (1988) proposed the idea that these *"adjunctive behaviors"* that lead up to the time of reinforcement could be used as *"conditional discriminations of the passage of time"*. In the behavioral theory of timing (BeT), Killeen and Fetterman posited that transitions between adjunctive behaviors are well approximated by a Poisson process similar to the pacemaker described in SET. Furthermore, the rate of transition between behaviors is proportional to the rate of reinforcement, allowing BeT to account for the scalar property in a different way than SET[6].

[6]This approach has been challenged by a number of studies from the authors of SET (e.g., Gibbon & Church, 1992 and Church, Meck, & Gibbon, 1994)

1.4.3 Beat-frequency model

The notion that representations of elapsed time in mathematical models of interval timing behavior should be biologically plausible became increasingly prominent during the XX^{th} century. In 1989, Miall suggested a model based on the idea that groups of simultaneously entrained neurons may oscillate at slightly different frequencies and therefore the *beat frequency* (the frequency at which any oscillator pair spikes simultaneously) between neurons is longer than their individual oscillation frequencies. Using this model, Miall suggested that the brain could in principle encode a wide range of durations with the help of a detector mechanism that keeps track of these beat frequencies and stores them in memory for future recall. More recently, Matell and Meck (2004) suggested that distinct brain areas could serve as oscillators and detectors (cortex and striatum, respectively). The literature on the neural basis of interval timing that gave rise to Matell and Meck's striatal beat-frequency model will be introduced in Section 1.5.4.

1.4.4 Time distributed over elements

In another class of timing models, temporal estimates in interval timing behavior are suggested to be represented in a distributed manner across multiple elements (akin to the sequences of behavioral states in BeT). In other words, time estimates are viewed as a result of the linear combination of time-varying (or *basis*) functions. One of the most influential models of this class was introduced by Staddon and Higa in 1999. Like SET, Staddon and Higa propose a model where the animal's current estimate of elapsed time and the memorized times of reinforcement are calculated/stored separately and ultimately compared before a behavior is produced. Yet, in Staddon and Higa's Multiple time scale (MTS) model, there is no pacemaker unit. Instead, an input stimulus triggers basis functions that decrease monotonically at different rates, similar to a cascade of leaky integrators with different time constants. Also based on BeT, Machado (1997) proposed the learning to time (LeT) model, where behavioral states are linked in a feed-forward manner during the interval period. Additionally, an associative process links those behavioral states and the final operant response, and the strength of that association is mapped onto the rate (or probability) of responding.

Similar non-linear time-series representations of elapsed time were introduced before Staddon and Higa and Machado. One such example is the spectral timing model (STM) proposed by Grossberg & Schmajuk (1989) that successfully reproduces key aspects of behavior in both classical conditioning and the peak-interval paradigm. The application of temporal representations distributed over elements in the context of reinforcement

learning theories (e.g.,Grossberg & Merrill, 1992; Suri & Schultz, 1999; Ludvig, Sutton, & Kehoe, 2008) will be introduced in Section 1.6.

1.5 The neural basis of interval timing

> "In previous decades, the idea that animals can represent time was resisted, partly because it was difficult to imagine a plausible physiological basis for this representational capacity."

As Gibbon and Malapani (2006) emphasized, scientists studying interval timing encountered a difficult problem: to understand how biological brains are capable of representing time in the range of seconds to minutes, a timescale much larger than the milliseconds range time constants of individual neurons. The lack of a dedicated sensory system for time meant an added difficulty. Where in the brain should one even begin to look for the so called *internal clock* (i.e., a dedicated system that generates a time-varying signal)? Or does the brain represent time through the temporal statistics of time-varying sensory and motor signals? Particularly following the development of behavioral and theoretical models of interval timing, such as those described in the previous Sections, the field moved towards exploring these questions and probing the neural basis of timing behavior.

The neurobiological implementation of models mentioned in Section 1.4 might be possible under a number of different activity patterns and neural architectures. Initial theories proposed the existence of an *internal clock*, but this centralized view of temporal representations has been challenged by a number of studies (see Section "No Central Clock" in Grondin, 2010 for a review on this topic). In recent years, the view that temporal representations that guide timing behavior might be the result of a distributed system has gained increasing support, in part due to studies highlighting that neural signals able to support temporal representations at timescales longer than the time constants of single neurons can be found in a vast range of brain areas.

Across the brain, connectivity patterns among neurons are complex and recurrent. Some theoretical approaches have taken this complexity into consideration when attempting to model dynamics that can support temporal representations. For example, theoretical work using recurrent neural networks (RNNs) has been successful in reproducing *in vivo* activity patterns that support temporal representations at long timescales (e.g., Buonomano, Merzenich, et al., 1995; Wang, 2001). This Section briefly reviews both experimental and theoretical work that has illuminated the neural basis for temporal representations that are able to support timing behavior (reviews on the neural

basis of timing behavior can be found in Buhusi & Meck, 2005; Gibbon & Malapani, 2006).

1.5.1 The cortex

The involvement of cortical areas in timing behavior has been suggested by a number of studies. Initial attempts at probing a causal relationship between cortical activity and interval timing using decortication methods (i.e., removal of the entire cortex) proved hard to interpret. For example, Jaldow et al. (1989) found that, although decorticated rats exhibited later post-reinforcement pause times than control animals in an FI paradigm, these results were most likely due to motor-related impairments. In fact, the same group of rats was able to adjust post-reinforcement pause times when presented with a new FI duration. In a more recent study, however, reversible inactivation of the rat prefrontal cortex (PFC) led to impairment in performance in a temporal bisection paradigm (Kim, Jung, Byun, Jo, & Jung, 2009). Additionally, functional magnetic resonance imaging (fMRI) studies have suggested that cortical areas such as supplementary motor area (SMA) and PFC may be recruited during interval timing tasks (Wiener, Turkeltaub, & Coslett, 2010).

The neural implementation of different interval timing models can be achieved with a number of distinct activity patterns, most of which have been found in cortical areas. One such implementation might be achieved through **oscillating patterns** (e.g., Miall, 1989; Treisman, Faulkner, Naish, & Brogan, 1990). Such sinusoidal oscillating patterns have been found in several brain areas including the cortex and can be evoked by sensory stimuli (e.g., Freeman & Skarda, 1985; Freeman & van Dijk, 1987; Treisman, Cook, Naish, & MacCrone, 1994) or spontaneously generated (e.g., Silva, Amitai, & Connors, 1991). Matell and Meck's (2004) extension of Miall's Beat-frequency model proposed the cortex as the source of the oscillation patterns at the basis of their model that reproduced some of the basic behavioral outputs of the peak-interval paradigm.

A second neural implementation of interval timing models might be achieved through neural integration (Simen, Balci, Cohen, Holmes, et al., 2011), relating to **ramping activity patterns** that have also been described in cortical areas. For example, during an interval discrimination task, neurons in the parietal cortex reflect the stimulus category that monkeys are likely to choose irrespective of outcome (Leon & Shadlen, 2003). A representation of elapsed time was also found in parietal cortex in the context of a delayed eye movement task (Janssen & Shadlen, 2005). In addition, ramping activity

leading up to movement onset was observed in the preSMA area of monkeys performing a timing task (Mita, Mushiake, Shima, Matsuzaka, & Tanji, 2009).

A third example of a possible neural implementation of interval timing models can be achieved through the **intrinsic dynamics of neural networks**. In this architecture type, temporal representations are distributed over a set of elements that can be represented in a number of ways. Several feed-forward types of network models have been proposed (e.g., serial activation of behavioral states in Machado, 1997; leaky integrators such as the ones described by Staddon & Higa, 1999). A perhaps more biologically plausible architecture that resembles the recurrent connectivity of cortical neurons is that of RNNs (e.g., Wang, 2001; Maass, Natschläger, & Markram, 2002), which have been directly applied to the problem of representing time in cortical neurons (e.g., Karmarkar & Buonomano, 2007; Buonomano & Maass, 2009; Buonomano & Laje, 2010). These studies suggest that temporal representations depend on changes in the spatial-temporal patterns of activity in a neural network. Therefore, decoding temporal information depends on the ability of a system to identify those spatial-temporal patterns. Through the activity dynamics in these state-dependent RNN models, one can observe persistent representations of stimulus identity and time for timescales of hundreds of milliseconds to seconds (Buonomano, 2014). Corroborating this work, activity in the mouse posterior parietal cortex (PPC) exhibits a sequential pattern capable of encoding stimulus identity and elapsed time (Harvey, Coen, & Tank, 2012), and neurons in the PFC of both primates (Machens, Romo, & Brody, 2010) and rats (Kim, Ghim, Lee, & Jung, 2013) have been shown to exhibit complex temporal patterns from which time can be decoded.

1.5.2 The hippocampus and time cells

The hippocampus plays an essential role in encoding spatial information (O'Keefe & Dostrovsky, 1971). Inspired by Tolman's (1948) work, in 1978 O'Keefe and Nadel proposed a broader role for the hippocampus as a cognitive map. In fact, before O'Keefe and Dostrovsky's (1971) work, the hippocampus had been linked to episodic memory through the famous case study of patient H.M. (Scoville & Milner, 1957). Although direct evidence for a role of the hippocampus in interval timing behavior is still missing, a number of studies in both laboratory animals and humans have linked the hippocampus to the organization of memories *in time* (Manns, Howard, & Eichenbaum, 2007). This temporal coding has been observed in situations where it can not be attributed to external events (e.g., Pastalkova, Itskov, Amarasingham, & Buzsáki, 2008; MacDon-

ald, Lepage, Eden, & Eichenbaum, 2011; see also Eichenbaum, 2014 for a review on this topic) and cells exhibiting such activity patterns have been dubbed *time cells*. A mathematical framework proposed by Howard et al. (2014) reproduces the patterns exhibited by time cells (MacDonald et al., 2011) using leaky integrators, sharing a similar kind of temporal representation as models of interval timing mentioned in Section 1.4 (Staddon & Higa, 1999; Grossberg & Merrill, 1992). These studies highlight the concept that hippocampal neurons exhibit sequential activity patterns, similar to those observed in cortical areas, that could potentially be used to guide timing behavior.

1.5.3 The cerebellum

The relationship between the cerebellum and timing on a millisecond timescale has been widely studied (Braitenberg, 1967; Ivry & Keele, 1989; Clarke, Ivry, Grinband, Roberts, & Shimizu, 1996; Buonomano & Mauk, 1994), particularly in the context of timed motor actions (Buonomano & Laje, 2010). Yet, whether the cerebellum plays a role at the longer timescales associated with interval timing is still to be determined. In 2010, Gooch et al. published a study suggesting that cerebellar lesions impair reproduction and production of intervals in the suprasecond timescale. The same authors simultaneously found that temporal discrimination in the same timescale is unaffected (Gooch et al., 2010). Other studies have reported conflicting results, some claiming no evidence that cerebellar lesions impair timekeeping in similar tasks (Harrington, Lee, Boyd, Rapcsak, & Knight, 2004), and others reporting that cerebellar lesions in humans result in increased (but scalar) variability in time estimates but do not affect accuracy (Malapani, Dubois, Rancurel, & Gibbon, 1998). This last study is particularly relevant since dysfunctions of the basal ganglia (BG) induce accuracy distortions as well as increases in the variability of time estimates that deviate from the scalar property when the same time ranges are probed (Malapani, Rakitin, et al., 1998). These results suggest that there are marked differences between the functional roles played by cerebellar and BG circuits in timing behavior.

1.5.4 The striatum and the dopamine clock hypothesis

The striatum is a key component of the BG, a group of subcortical structures that also includes the globus pallidus (GP), the subthalamic nucleus (STN) and both substantia nigra pars compacta (SNc) and substantia nigra pars reticulata (SNr) (see Figure 1.1, adapted from Gerfen & Bolam, 2010). Midbrain dopamine (DA) neurons in the SNc and ventral tegmental area (VTA) are a crucial part of BG circuitry due to their recip-

rocal connectivity with the striatum (Gerfen & Bolam, 2010), although the VTA is not considered part of the canonical organization of the BG.

The BG have been implicated in a vast range of functions, such as motor control (Graybiel, Aosaki, Flaherty, Kimura, et al., 1994), action selection (Mink, 1996), novelty, and motivation (Wise, 2004), all of which are functions that have been linked to timing behavior (Pariyadath & Eagleman, 2007; Cools, 2008; Gable & Poole, 2012). Additionally, interval timing behavior has been shown to be altered in a number of neurological and neuropsychiatric disorders that affect the BG, particularly the striatum and the midbrain DAergic system. Examples of such disorders include Huntington's (Rowe et al., 2010) and Parkinson's (Artieda, Pastor, Lacruz, & Obeso, 1992; Pastor, Artieda, Jahanshahi, & Obeso, 1992) diseases, as well as attention deficit hyperactivity disorder (ADHD) (Noreika, Falter, & Rubia, 2013), schizophrenia (Wahl & Sieg, 1980; Rammsayer, 1990) and substance abuse (Wittmann, Leland, Churan, & Paulus, 2007; Lüthi & Lüscher, 2014). Regarding the use of substances, the impact of cannabinoid drugs such as hashish on time perception has been described, for example, by James (1886):

> "In hashish-intoxication there is a curious increase in the apparent time-perspective. We utter a sentence, and ere the end is reached the beginning seems already to date from indefinitely long ago. We enter a short street, and it is as if we should never get to the end of it."

Since then, a wide range of studies have probed the effects of different drugs that target the DAergic system on interval timing behavior, as well as on striatal function. These studies have led to the proposition that the DAergic system may be involved in regulating the speed of a timekeeping process. The data in support of this idea is presented below.

Dopamine neurons and the dopamine clock hypothesis A long-standing theory in the field of interval timing research posits that levels of DA may impact the speed of an *internal clock* (Maricq, Roberts, & Church, 1981; Maricq & Church, 1983; Meck, 1983, 1986; Rammsayer, 1993; Matell, King, & Meck, 2004). Perhaps the root of this idea dates back to the late 1950's and 1960's and to the experiments that illuminated the role of DA in Parkinson's disease. Parkinson's disease is a neurodegenerative disorder that mainly affects motor control, where patients exhibit slowness of movement. In the late 1950's, Carlsson and colleagues performed a series of key experiments linking the levels of DA in the central nervous system with Parkinson's disease (Carlsson, Lindqvist, & Magnusson, 1957; Carlsson, Lindqvist, Magnusson, &

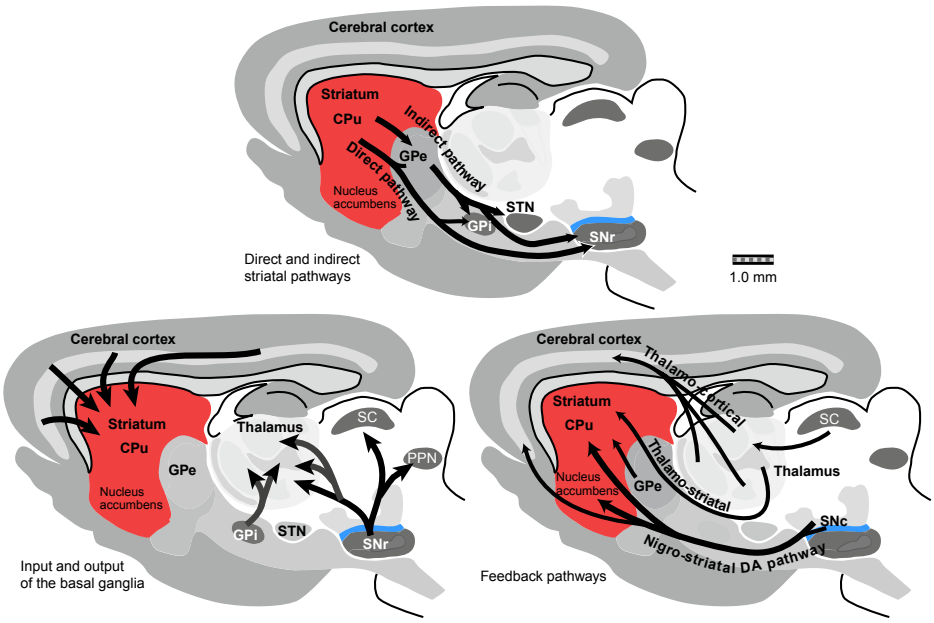


Figure 1.1. **Illustration of basal ganglia connectivity**. Sagittal views of the rat brain illustrate distinct connections between BG nuclei, their inputs and outputs, as well as feedback pathways. This figure is adapted from Gerfen & Bolam, 2010. The striatum is highlighted in red and the substantia nigra pars compacta (SNc) is highlighted in blue. *Top* - The striatum is comprised of two main projection pathways which arise from two distinct types of striatal medium spiny neurons (MSNs). Direct pathway MSNs (dMSNs) project to the external globus pallidus (GPe) as well as to the main output nuclei of the BG, the internal globus pallidus (GPi) and substantia nigra pars reticulata (SNr). Indirect pathway MSNs (iMSNs) project to the GPe, and therefore only connect to the output nuclei of the BG indirectly through the subthalamic nucleus (STN). *Bottom left* - Similar scheme as *top* panel but depicting the major input and output targets of the BG. Most cortical areas provide input to the BG through their connections to the caudate-putamen (CPu) and nucleus accumbens (NAc), which together constitute the rodent striatum. γ-aminobutyric acid (GABA) neurons in the GPi and SNr constitute the output of the BG, and provide inhibitory input to the superior colliculus (SC) and pedunculopontine nucleus (PPN), as well as to thalamic nuclei. *Bottom right* - There are three main feedback pathways to the BG, namely the nigro-striatal pathway, constituted by SNc-DA neuron inputs to the striatum, thalamo-striatal and thalamo-cortical pathways.

Waldeck, 1958; Bertler & Rosengren, 1959; Carlsson, 1959). Carlsson and colleagues discovered that DA was present in the brain (and mainly in the striatum) at concentrations comparable to those of noradrenaline, suggesting that DA was not merely a precursor in the biosynthesis pathway of other catecholamines as was believed until then. The authors also found that DA depletion (mainly in the striatum) was the basis of a reserpine-induced parkinsonism phenotype in mice. Next, they discovered that the administration of L-3,4-dihydroxyphenylalanine (L-dopa), a DA precursor, could restore both normal DA levels and normal motor function to animals treated with reserpine. Inspired by the work from Carlsson's group, Hornykiewicz investigated the brains of Parkinson's disease patients and control subjects (*post mortem*), in search of altered DA levels. Hornykiewicz found that DA levels in the striatum were lower in Parkinson's disease patients than controls (Ehringer & Hornykiewicz, 1960), and he speculated that this depletion might be due to neurodegeneration in the SNc after discovering that levels of DA were also reduced in the SNc of Parkinson's patients (Hornykiewicz, 1963). These findings led to the first administration of L-dopa in humans to treat Parkinson's disease symptoms (Cotzias, Van Woert, & Schiffer, 1967; Cotzias, Papavasiliou, & Gellene, 1969), a treatment still in use today.

The ability of DA to restore movement to Parkinson's patients and the effect of DAergic agonists in effectively *speeding up* movement (Carlsson et al., 1957) suggested that the DAergic system was a good candidate as the speed controller of an *internal clock* for interval timing. This theory suggested that an increase in DA levels should increase the speed of timekeeping, whereas decreasing DA levels should slow down that speed. Initial experiments produced compelling results in favor of this hypothesis: pharmacological manipulations using DA agonists and antagonists led to impaired timing behavior, consistent with an apparent speeding up or slowing down, respectively, of timekeeping across temporal production and discrimination paradigms (Maricq et al., 1981; Maricq & Church, 1983; Meck, 1983, 1986; Matell et al., 2004; MacDonald & Meck, 2005; Balcı et al., 2008). These pharmacological results have been supported by more recent genetic studies where the overexpression of D2-type DA receptors seems to cause a shift in response curves to later times during PI paradigms (e.g., Drew et al., 2007). Also in the context of a PI paradigm, Meck (2006) has shown that a selective lesion of midbrain DAergic neurons impairs the temporal dynamics of timing behavior.

In some conditions, however, data from similar pharmacological and genetic manipulations are inconsistent with this specific hypothesis. For example, the results from Drew et al. (2007) can be explained by a decrease in motivation (Ward et al., 2009). Additionally, while some studies have found that administration of DA agonists can cause

a global timing impairment with no specific apparent control of timing speed (Odum, Lieving, & Schaal, 2002; Balcı et al., 2008), other studies have not found an effect of manipulating the DAergic system on timekeeping (Balcı et al., 2008, 2010) or have even suggested an effect in the opposite direction (i.e., increased DA might decrease the speed of timekeeping) (Lake & Meck, 2013). Finally, pharmacological studies have reported that DA modulates attentional mechanisms during timing behavior (Buhusi & Meck, 2002; Ward et al., 2009), arguing against a specialized role of DA activity in controlling only timekeeping speed.

The idea that there is a direct relationship between DA levels and timekeeping speed has also been challenged by studies in Parkinson's disease patients. Some studies report overall timing accuracy deficits in Parkinson's patients (Artieda et al., 1992; Malapani, Rakitin, et al., 1998). Malapani, Rakitin, et al. (1998) tested Parkinson's disease patients in a PI paradigm where target times of 8 or 21s were presented in blocks of trials. When under the effect of L-dopa medication, patients' performance was comparable to age-matched controls. When patients performed in the absence of L-dopa, estimates were impaired both in accuracy (the 8s interval was over-estimated and the 21s interval was under-estimated) and in precision (increased variance that violates the scalar property). However, the predicted slowing down of timekeeping in patients tested in the absence of L-dopa medication (and therefore with low levels of DA) has been harder to observe. While Pastor et al. (1992) found evidence in favor of the DA clock hypothesis in Parkinson's disease patients, other studies have shown normal timing behavior in such patients (Spencer & Ivry, 2005; Wearden et al., 2008).

These conflicting results may, in part, reflect the diversity of functions in which DA neurons have been implicated. In fact, some studies have suggested that pharmacologically-induced changes in timing behavior may be explained by changes in motivation (Odum et al., 2002; Balcı et al., 2010) and attention (Buhusi & Meck, 2002; Lake & Meck, 2013). Therefore, studying the specific activity patterns that DA neurons might exhibit during interval timing tasks, as well as understanding how DA neurons mediate these distinct functions, may require measurements/manipulations with higher spatial and temporal resolution than those outlined above.

The striatum The striatum is the main input nucleus of the BG (see Figure 1.1) and receives inputs from most cortical areas as well as from the hippocampus, amygdala, thalamus and brainstem (Gerfen & Bolam, 2010). This input diversity may explain the wide range of functions in which the striatum, and the BG in general, have been implicated, including motor control, reinforcement learning and cognitive functions.

The ability to encode sensory inputs, ongoing behaviors and reward outcomes makes the striatum an ideal structure to represent elapsed time. In fact, a number of studies indicate that the striatum may play a key role in timing behavior. Firstly, fMRI studies have reported striatal activation during tasks that require the processing of durations (Tanaka et al., 2004) as well as specifically during the performance of PI paradigms (Hinton & Meck, 2004). Secondly, the striatum is one of the main targets of the midbrain DAergic system (Bertler & Rosengren, 1959; Watabe-Uchida, Zhu, Ogawa, Vamanrao, & Uchida, 2012), and it is therefore possible that the impact of manipulations and disease states affecting the DAergic system on timing behavior are a result of modulation of striatal activity. Consistent with this idea, Meck (2006) has shown that the lack of temporally structured responses following selective lesioning of midbrain DA neurons was reproduced if the DAergic terminals in the striatum alone were lesioned. Lastly, there is evidence that the striatum, much like the cortex and the hippocampus, exhibits sequential activity patterns that may support timing behavior. Jin, Fujii, and Graybiel (2009) found that the dynamics of a subset of striatal neurons seem to be modulated during fixed delay periods leading up to reward delivery in monkeys, with different neurons displaying maximum activity at different points in time (Jin et al., 2009). Therefore, elapsed time during the delay period could be decoded from this neural activity that spanned the delay period until reward delivery. Furthermore, work from Adler et al. (2012) showed similar results in monkeys trained in a classical conditioning paradigm. Adler et al. also found groups of striatal neurons that exhibited peak responses at different times, with long latency response neurons exhibiting greater variability.

It remains unclear, however, if these representations are also present in animals performing tasks that require estimation of elapsed time, and whether such representations of time can be used to guide behavior. At least one study (Matell, Meck, & Nicolelis, 2003) suggests that this scenario is likely, since striatal (and cortical) neurons exhibit distinct firing rate modulations depending on whether a lever press leading up to reward occurs early or later within an FI, allowing the experimenters to decode where in the FI a lever press occurred.

Interestingly, the sequential activity patterns mentioned above (Jin et al., 2009; Adler et al., 2012) bear close resemblance to the non-linear time-series representations of elapsed time that have been proposed in the context of reinforcement learning theories (Grossberg & Merrill, 1992; Suri & Schultz, 1999; Ludvig et al., 2008).

1.6 The basal ganglia: where interval timing meets reinforcement learning

The same "Law of effect" principles that motivated early work in operant conditioning paradigms (see Section 1.3) also inspired the development of an initial thread of RL theory[7] based on trial-and-error learning (Sutton & Barto, 1998). In RL, animals are modeled as *agents* whose goal is to maximize future rewards in an environment represented by distinct *states*. To achieve this goal, agents must learn the appropriate mappings between the states they encounter and a set of actions they may take (e.g., which state to visit next). To learn these mappings, agents must predict the (discounted) amount of future rewards expected from being in a state (i.e., *the state value function*) or from taking a certain action in a given state (i.e., *the action value function*) and select actions according to a *policy* that maximizes future rewards.

Temporal-difference learning In a simple scenario akin to classical conditioning, the agent's estimate of the value function of a given state can be expressed as a weighted sum of the features that describe that state (Sutton & Barto, 1990, 1998). These weights can be updated by a simple temporal-difference (TD) learning rule at each time step, where a difference between the current prediction the agent has about future rewards and the prediction generated in the previous timestep is calculated. This *temporal-difference* in reward prediction is compared with the actual experienced reward, and if they are different, a reward prediction error (RPE) is generated. TD learning shares some basic principles with a model from Rescorla and Wagner (1972) while overcoming some of the limitations of the latter. In the Rescorla–Wagner model, on each trial, the difference between the reward the animal is expecting to receive and the actual reward experienced is calculated: if there is a difference, an RPE signal is generated. Therefore, the TD model can be seen as an extension of the Rescorla–Wagner model with two important distinctions that grant TD learning a better fit with some conditioning data: i) the Rescorla–Wagner model is a *"trial-level"* model (Sutton & Barto, 1990), because it represents trials as a whole and therefore its predictions are insensitive to the temporal relationships between stimuli, responses and rewards; TD updates, in contrast, are performed on each time-step, a more relevant time-frame for an animal, and are able to account for these temporal relationships within each trial; and ii) the

[7]There were mainly two initial threads in RL: one concerning trial-and-error learning and another concerning the problem of optimal control. The modern theory of RL as of the late 1980's is the result of the combination of these two threads (Sutton & Barto, 1998).

temporal-difference in reward predictions in the TD algorithm means that the strength with which stimuli are associated with rewards at a given time is not only predictive of the immediate reward outcome but also of future ones. Extended comparisons between these learning rules have been described by several authors (Sutton & Barto, 1990; Niv, 2009; Ludvig, Bellemare, & Pearson, 2011).

Actor-Critic architecture The framework described so far applies to cases in which the transitions between states are fixed, as in classical conditioning. However, this framework is insufficient to model interval timing paradigms, where an animal must not only predict the timing of a reward but also learn when to perform a given action that maximizes future rewards. This process of selecting the optimal action is more difficult when the consequences of that action are delayed, as is the case for many natural behaviors, as well as some interval timing tasks. A major reason for this difficulty is known as the *credit assignment problem*: how to assign the appropriate credit to one or more actions that preceded the current outcome (e.g., Minsky, 1961; Barto et al., 1983). Selecting which actions to perform is a problem that can be solved in several ways by RL algorithms (Sutton & Barto, 1998). One example, and in fact one of the earliest model architectures designed to perform action selection, is the Actor-Critic architecture (Barto et al., 1983). As the name suggests, the Actor-Critic architecture is composed of two separate structures, the *actor* and the *critic*. The *actor* is represented by the policy structure that selects an *action* given a certain state, independently of the value function. The estimation of the value function is done by the *critic*, who must learn about the policy that the actor is following and provide the *actor* with a *critique* on how to improve its current policy in the form of a TD error.

What is the evidence that the brain may implement an RL-like algorithm? Furthermore, how are features (and consequently time) represented in RL models to appropriately assign credit? The following Sections will briefly address these questions (see Ludvig et al., 2011; Daw, Courville, & Touretzky, 2006 and Gershman, Moustafa, & Ludvig, 2014 for extensive reviews on this topic).

1.6.1 The neural basis of reinforcement learning

In the last three decades, a number of studies have accumulated evidence in support of the theory that the BG implements aspects of RL (Daw et al., 2006; Samejima & Doya, 2007). Firstly, state features and value functions are thought to be represented in the cortex and striatum, respectively (Doya, 1999; Samejima, Ueda, Doya, & Kimura, 2005; Lau & Glimcher, 2008; Kim, Sul, Huh, Lee, & Jung, 2009). Secondly, and par-

ticularly important for RL, the firing rate of midbrain DA neurons following rewards, and reward-predicting stimuli, resembles the TD learning RPE signal mentioned above, both qualitatively (Schultz, Apicella, & Ljungberg, 1993; Schultz, Dayan, & Montague, 1997) and quantitatively (Waelti, Dickinson, & Schultz, 2001; Bayer & Glimcher, 2005). For example, DA neurons increase their firing rate when an unexpected reward is presented. Conversely, the same neurons exhibit a pause in firing rate when an expected reward is omitted. At least in the context of classical conditioning, this error signal can be found in the vast majority of DA neurons (Cohen, Haesler, Vong, Lowell, & Uchida, 2012). This teaching signal is thought to mediate learning at the level of corticostriatal synapses through DA-dependent plasticity (Reynolds, Hyland, & Wickens, 2001; Reynolds & Wickens, 2002; Steinberg et al., 2013). Indeed, stimulation of DA neurons suggests that RPE signals are able to elicit learned behaviors (Witten et al., 2011; Steinberg et al., 2013), in agreement with TD learning theories.

1.6.2 Temporal representations in reinforcement learning

Complete serial compound As mentioned before, a central aspect in RL models involves choosing what form the temporal representation of features should take. In the simple case of TD learning, a representation named complete serial compound (CSC) has been widely applied (Sutton & Barto, 1990; Montague, Dayan, & Sejnowski, 1996; Schultz et al., 1997). In CSC, every single time step in a stimulus is represented as an individual non-overlapping binary pulse (Sutton & Barto, 1990). Although a CSC representation has proven to be able to reproduce some key aspects of DA neuron responses in the context of classical conditioning, its implementation has suffered criticism (Daw et al., 2006; Ludvig et al., 2008; Ludvig, Sutton, & Kehoe, 2012; Gershman et al., 2014). One concern regarding the CSC representation is that it functions as a perfect digital clock with no uncertainty, a representation that lacks generalization and is biologically improbable, since it is inconsistent with animal behavior (see Sections 1.3 and 1.4) and does not accurately reproduce the firing patterns of DA neurons observed when the *timing* of rewards is altered.

For example, a CSC-TD learning algorithm predicts a large and temporally precise negative RPE when expected rewards are omitted (Ludvig et al., 2008), unlike the modest and temporally broad decrease in midbrain DA neuron activity (Schultz et al., 1997). On the other hand, if a reward is delivered earlier than expected, a CSC-TD learning algorithm will produce a positive RPE signal at the time of reward delivery but will also exhibit a large, temporally precise, negative RPE at the time of expected

reward. Yet, experimental evidence suggests that there is no such abrupt decrease in DA neuron activity (although perhaps a small, temporally smeared effect) at the expected reward time (Hollerman & Schultz, 1998).

Microstimulus representation Alternative temporal representations to the CSC have been shown to overcome at least some of the limitations mentioned above. One of these alternatives was introduced by Ludvig et al. (2008), and proposes that events (sensory cues or rewards) leave behind a memory trace that decays over time. A series of basis functions (in this case, temporal receptive fields) encode this memory trace and the resulting microstimuli become weaker and wider over time, providing a measure of confidence of the trace hight (Ludvig et al., 2008). TD learning algorithms that use microstimuli to predict upcoming rewards gain temporal generalization compared to those employing CSC (Ludvig et al., 2008, 2012) and, as a result, exhibit RPE signals that match the experimental data of DA neuron recordings under variable reward timing conditions mentioned earlier (e.g., reward omission and reward anticipation, Schultz et al., 1997; Hollerman & Schultz, 1998).

A number of studies have proposed similar temporal representations to Ludvig's microstimuli, both in the context of RL (Suri & Schultz, 1999) as well as in the context of interval timing models (see Section 1.4.4; Grossberg & Schmajuk, 1989; Machado, 1997; Staddon & Higa, 1999). Still, for the most part, interval timing and RL have been addressed separately, although studies mentioned in this Chapter highlight important common features between the two frameworks (see Gershman et al., 2014 for an attempt at reconciling the two). RL and interval timing models share not only similar forms of theoretical temporal representations, but also common brain areas (in particular the striatum and the midbrain DA system) thought to implement these models. Additionally, there is evidence that DA neurons have access to temporal information in the context of RPE signals because, in the presence of varying temporal expectation, DAergic responses are smaller when rewards and reward-predicting cues are expected (Fiorillo, Newsome, & Schultz, 2008). However, an important missing piece is the understanding of what signals are carried by BG nuclei (particularly in the striatum and midbrain DAergic neurons) during interval timing tasks. How do those potential signals relate to these predicted by RL algorithms? The following Chapters will explore these questions, in an attempt to provide experimental support for a more unified view of interval timing and RL.

Chapter 2

A scalable population code for time in the striatum

To guide behavior and learn from its consequences, the brain must represent time over many scales. Yet, the neural signals used to encode time in the range of seconds to minutes to hours (i.e., interval timing) are not known. The striatum is a major input area of the basal ganglia associated with learning and motor function. Previous studies have also shown that the striatum is necessary for normal timing behavior. To address how striatal signals might be involved in timing, we recorded from striatal neurons in rats performing an interval timing task. We found that neurons fired at delays spanning tens of seconds and that this pattern of responding reflected the interaction between time and the animals' ongoing sensorimotor state. Surprisingly, cells rescaled their responses in time when intervals changed, indicating that striatal populations encoded relative time. Moreover, time estimates decoded from activity predicted timing behavior as animals adjusted to new intervals, and disrupting striatal function led to a decrease in timing performance. The results in this Chapter suggest that striatal activity forms a scalable population code for time, providing timing signals that animals use to guide their actions.

2.1 Introduction

To behave adaptively in complex, ever-changing environments, animals must learn which actions to take in a particular context based on their past experience. However, to learn about the sometimes-delayed consequences of actions and to guide future behavior, it is absolutely necessary that the brain represent not only actions and consequences but also temporal information about when those actions and consequences occur (Schultz et al., 1997).

Multiple lines of evidence implicate the basal ganglia (BG) as a locus for the representation of such temporal information. Lesions of the striatum in rats (Meck, 2006), disease states that affect the BG such as Parkinson's (Malapani, Rakitin, et al., 1998) and Huntington's diseases (Rowe et al., 2010), drugs that affect dopamine (DA) signaling (Maricq & Church, 1983), and genetic manipulations that affect the DA system in the BG (Ward et al., 2009) all result in interval timing dysfunction. Furthermore, human functional magnetic resonance imaging (fMRI) studies have found that the striatum, a main input area of the BG, is activated by tasks that involve the processing of interval information (Hinton & Meck, 2004; Tanaka et al., 2004). In addition, many theoretical models have been proposed to explain timing behavior. These models can be grouped into at least three categories. Pacemaker-accumulator models integrate pulses emitted from a central pacemaker to measure elapsed time (Gibbon, 1977; Simen et al., 2011). Beat frequency models detect patterns of activation across resettable oscillatory processes at different frequencies to encode time delays from a resetting event (Meck, Penney, & Pouthas, 2008). Sequential state models contain orderly transitions between different activity states that can be used to encode time (Killeen & Fetterman, 1988; Buonomano et al., 1995; Machado, Malheiro, & Erlhagen, 2009). These theories reproduce various aspects of timing behavior in many interval timing tasks. However, neural data providing evidence for or against these various theories are lacking.

To understand how time is encoded in neural circuits, we recorded the spiking activity of neurons as rats performed a serial fixed interval (SFI) timing task. Specifically, given the apparent localization of timing function in striatal tissue, we asked whether striatal neural activity could encode elapsed time over durations of tens of seconds to one minute while we measured behavior that reflected animals' estimates of time. We found that different striatal neurons fired maximally at different delays from reward receipt and that information about animals' time estimates could be extracted from striatal populations by simply treating neurons as tuned for time. Importantly, this tuning for time, while affected by sensorimotor event-related neural responses, could not be fully

explained by ongoing behavior, as even cells that displayed responses locked to a specific behavior varied their responses depending on when that behavior was executed within a given interval. Strikingly, we found that temporal tuning stretched or contracted, rescaling with the interval being timed. Thus, striatal populations encoded relative time, flexibly adapting to the immediate demands of the environment. Finally, we ran a simple simulation of the SFI task and showed that neural responses resembling those we observe in the striatum are suitable as a basis for timing behavior. The results in this Chapter provide important biological insight into how a major brain system encodes time during behavior.

2.2 Results

2.2.1 Lever pressing start time in SFI task is a behavioral measure of rats' expectation of time until reward

To elicit robust time-dependent behavior over a broad range of timescales, we employed operant conditioning procedures under SFI schedules of reinforcement

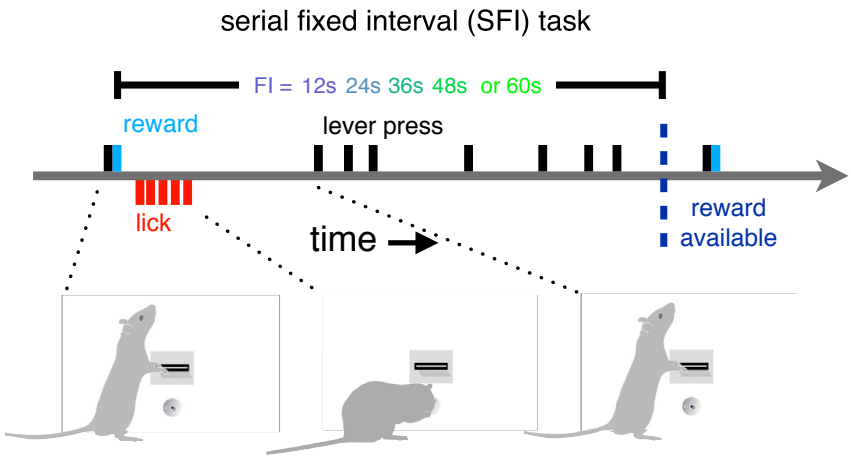


Figure 2.1. **Schematic of the serial fixed interval task**. The following color code will be commonly used: blue represents short fixed intervals (FIs), and green represents longer FIs.

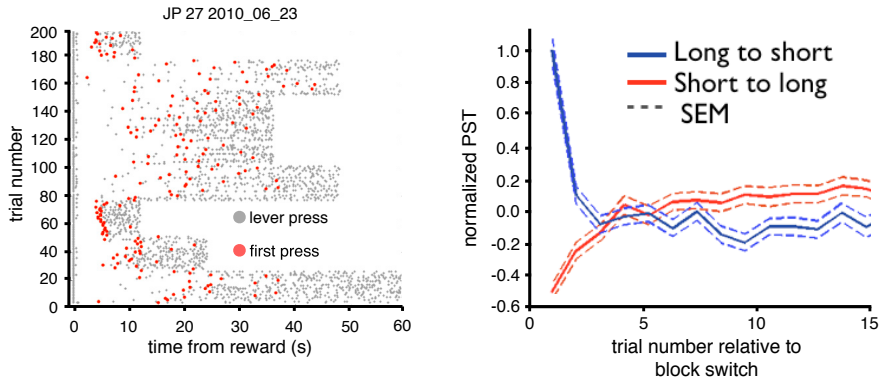


Figure 2.2. **Pressing behavior during serial fixed interval task**. *Left* - Example of lever pressing behavior in a single session of the SFI task. Gray markers indicate a lever press; red markers indicate the pressing start time (PST). *Right* - Normalized PST for the first 15 trials after block switch. Solid blue trace represents block switches from any FI to the 12s FI block (long to short). The solid red trace represents block switches from any FI to the 60s FI (short to long). Dashed lines show the standard error of the mean (SEM). PSTs were normalized by subtracting the mean within-FI PST, and dividing by the within-FI standard deviation, across all sessions.

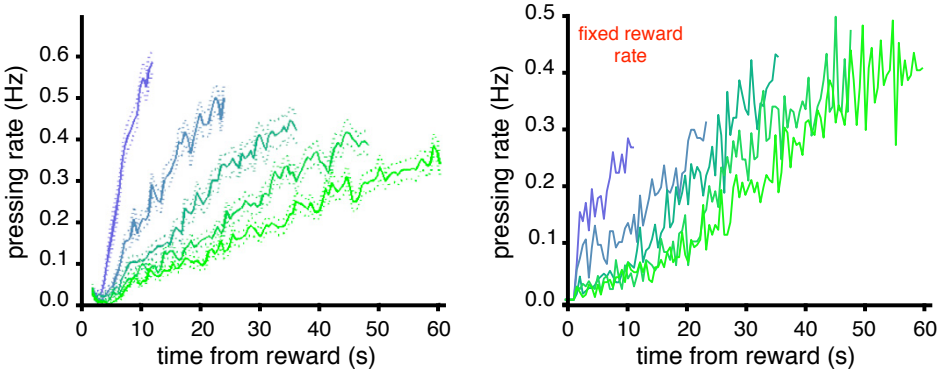


Figure 2.3. **Lever pressing rate aligned on reward**. *Left* - Average lever pressing rate in each of the five FIs, aligned on preceding reward. Dashed lines represent SEM. *Right* - Conventions as in (*left*). Average lever pressing rate aligned on preceding reward, for 12 sessions from two control animals where reward amount varied in proportion to the current FI, holding reward volume over time constant during the session.

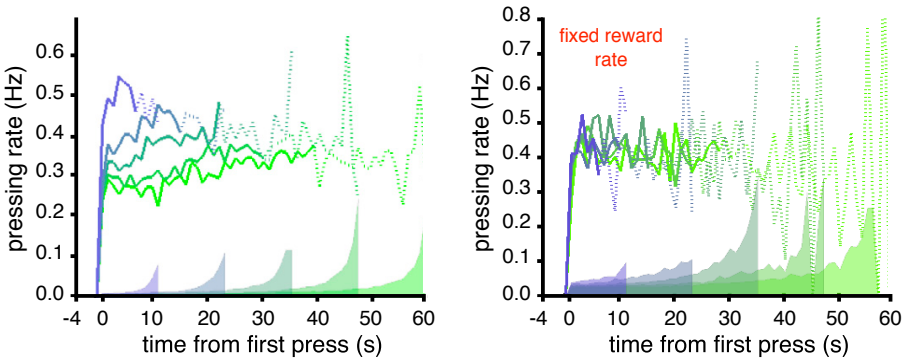


Figure 2.4. **Lever pressing rate aligned on first press**. Same as Figure 2.3 but aligned on PST. Traces are plotted on a solid line for the period for which more than half of the trials contribute data and on a dotted line after that point. Shaded patches along the horizontal axis represent SEM.

(Figure 2.1). Briefly, rats were placed in a behavioral box containing a lever positioned over a liquid delivery port and were trained to press the lever to receive water reward. Reward delivery triggered a timer, and reward became available again only after the timer exceeded an FI ranging from 12s to 60s in multiples of 12s. Lever presses occurring after reward delivery but before the FI had elapsed were not reinforced. An FI was maintained for between 18 and 40 rewards before changing to another FI, randomly chosen from the interval set. In single sessions, rats tended to distribute lever pressing toward the latter portion of the FI, shifting the timing of responding as FI changes occurred (Figure 2.2). This pattern of responding produced ramps in block-wise averaged pressing as a function of time that varied in slope in relation to the FI (Figure 2.3). However, this did not reflect the pattern of responding in single trials. We asked

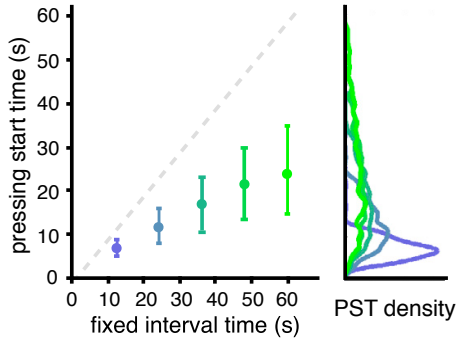


Figure 2.5. **Median and interquartile range of PST**. *Left* - Median and interquartile range of PST for each of the five FIs. *Right* - Smoothed density functions depicting the full distributions of PST.

how pressing evolved after pressing onset (pressing start times, PSTs) in each trial by aligning on the PST and averaging lever press rates across trials and within blocks of the same FI (Figure 2.4, left panel). This PST is equivalent to the post-reinforcement pause described in Section 1.3. Rats pressed at a relatively constant rate after the first press in each trial, with a rate determined by the experienced reward rate (Figure 2.4, right panel). The ramps in the reward-aligned pressing as a function of time largely result from changing distributions of PSTs (Figure 2.5), as these vary systematically with FI, and averaging a group of step functions with onset times drawn from these distributions will produce ramps of varying slope. This SFI lever pressing task produced systematic variation in the distributions of PSTs of bouts of anticipatory pressing, consistent with previous timing studies employing FI schedules of reinforcement (Gibbon, 1977). These bouts were of a relatively constant rate (see *break and run* pattern in Section 1.3) that varied with reward rate over time (Figure 2.4). The PST thus provided a behavioral metric that covaried with the animals' changing expectation about time until the next available reward, which we compared to the activity of neurons recorded in the striatum during performance of the task as described below.

2.2.2 Striatal neurons display temporal tuning

In the SFI task, reward delivery is both the timing cue and the reinforcer. Since animals reported knowledge of time between reward availability by when they began to press a lever, we asked whether neuronal responses in the striatum aligned on reward might reveal a signal that animals could use to guide the decision of when to begin pressing. We recorded broadly in the dorsal striatum so as to sample neurons from regions previously shown to be important for interval timing behavior (Meck, 2006; Figure 2.6, inset), and the vast majority of units we recorded exhibited average firing rates lower than five spikes per second, consistent with a population made up of mostly medium spiny projection neurons (Gage, Stoetzner, Wiltschko, & Berke, 2010; Figure 2.6). Aligned on reward delivery, the population of recorded cells exhibited a broad distribution of activity patterns, as reflected in the normalized spike density functions (SDFs; see Materials & Methods in section 2.4 for details) shown in Figure 2.7. Some cells fired just after reward delivery, others fired in the middle of the delay, and others fired leading up to the next reward (Figures 2.7 and 2.8). This produced a slow-moving *bump* of activity that traversed the population during each FI. In theory, reading out the location of this bump in the population could provide an estimate of time within the FI.

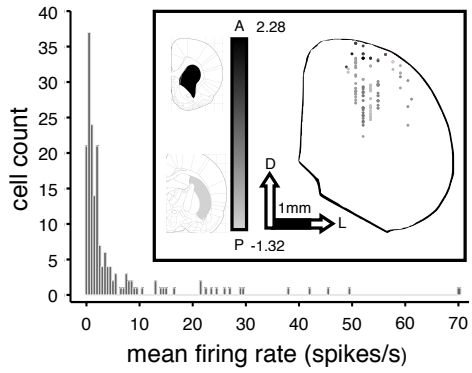


Figure 2.6. **Neuron firing rates and reconstruction of recording sites**. Distribution of average firing rates (bin size of 1spike/s), calculated using all spikes from each neuron. *Inset* - Recording sites projected onto one coronal silhouette of the striatum ($n = 5$ rats). Each data point represents one recording site. Black-gray color coding indicates position along the anterior-posterior axis (respectively). Coronal slices at the the extreme anterior and posterior positions are shown. A - Anterior, P - Posterior, D - Dorsal, L - Lateral.

However, a core feature of interval timing behavior is that timing accuracy decreases with the magnitude of the interval being timed (Gibbon, 1977). Two features of the neural data could potentially contribute to this phenomenon: an increased spread of each neuron's responses as a function of their peak latency and a decreasing density of neurons displaying peak firing rates as time progresses. We found that the widths of responses were indeed correlated with their latencies to peak firing within each FI (left panel in figure 2.9, linear regression, FI 12s, $R = 0.4443, p < 0.001$; FI 24s, $R = 0.7563, p < 0.001$; FI 36s, $R = 0.7188, p < 0.001$; FI 48s,$R = 0.5910, p < 0.001$;

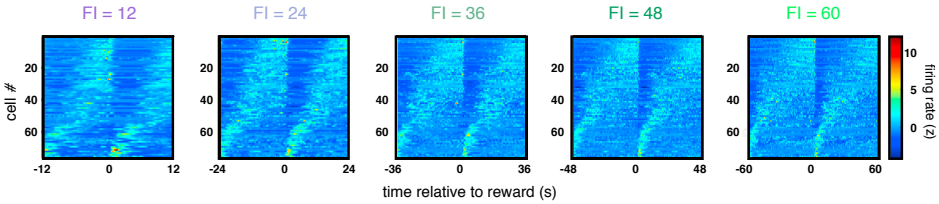


Figure 2.7. **Striatal neuron responses tile tens of seconds to one minute**. SDFs of neurons that maintained their relative ordinal position in time within the population across all five FIs, aligned on reward.

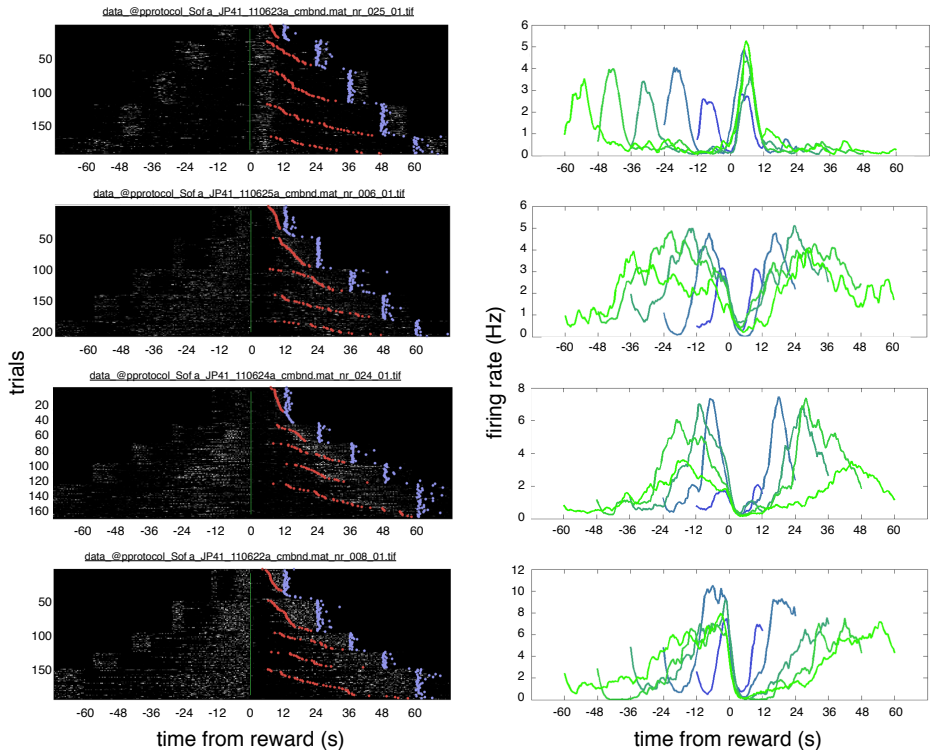


Figure 2.8. **Single neuron examples with different time courses of response that rescale with FI**. Each row corresponds to a single neuron example recorded from the striatum during performance in the SFI task. *Left* - Single trial peri-stimulus time histograms (PSTHs), aligned on reward delivery (green line), with a bin size of 20ms. Red tick marks indicate the onset of pressing and purple tick marks indicate reward delivery. Trials are ordered by FI, and within each FI by pressing onset time. *Right* - SDFs for each FI block, aligned on reward delivery.

FI 60s, $R = 0.4733, p < 0.001$; see Materials & Methods in section 2.4 for details). In addition, the density of peak firing rate latencies in our population decreased over time within the FI (Figure 2.9, right). Thus, the bump in activity within the striatum population moved progressively slower as the FI wore on. Strikingly, the overall time taken by this bump to traverse the population appeared to scale with the FI (Figures 2.7 and 2.10). To begin to assess apparent scaling of response times, we first selected cells that we had recorded in all five FIs and that maintained their ordinal position within the population when responses within each FI were ordered by firing dynamics (Geffen,

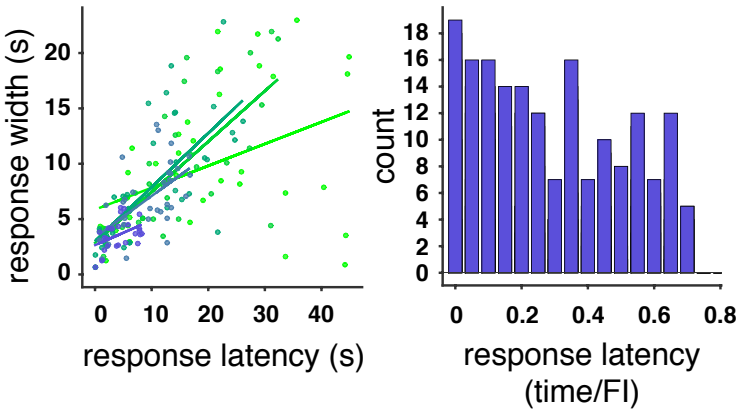


Figure 2.9. **Response width correlates with latencies to peak firing within each FI**. *Left* - Width of each cell's response within each FI as a function of latency to peak firing. Colored lines represent the best linear fit to the data. *Right* - Histogram of relative peak latencies pooled over all FIs, using the data in the *left* panel.

Broome, Laurent, & Meister, 2009). Of the 112 neurons recorded in all FIs, we found that 76 neurons (68%) maintained their ordinal position in time across the population (see Materials & Methods in section 2.4 for details). The responses of these neurons can be observed in Figure 2.7, wherein the position of cells along the y axis is the same across the panels displaying average responses in each of the FIs (for all recorded cells, see Figure 2.10). To quantify to what degree responses rescaled, we computed a scale factor for each neuron as the ratio of the center of mass (COM) of the SDF in the 12s FI over the COM of the SDF in each of the other four FIs (Figure 2.11). The distributions of these scale factors were sharper than and significantly different from null distributions generated by shuffling cell identity across FIs and recomputing

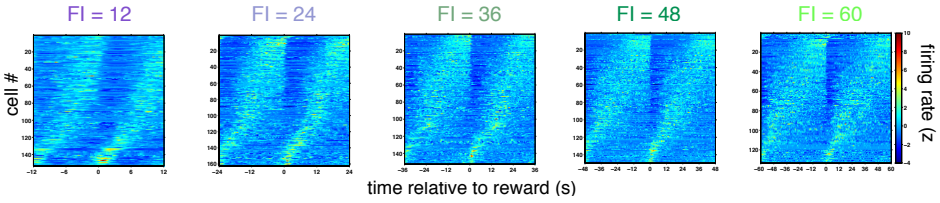


Figure 2.10. **All recorded neuron responses per FI**. SDFs of all recorded striatal neurons across all five FIs, aligned on reward.

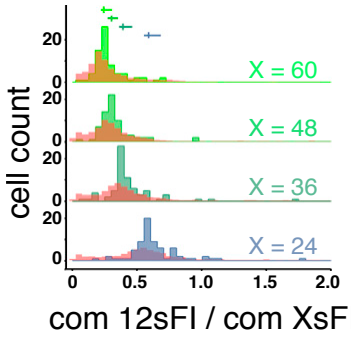


Figure 2.11. **Distributions of scale factors per FI**. Distributions of scale factors obtained by calculating the ratio of the center of mass (COM) of the SDF between the 12s FI and each of the other FIs (24s, 36s, 48s, and 60s, respectively, from blue to green) for each cell. For each distribution of scale factors, a null distribution in red was generated by shuffling cell identity across FIs and recomputing the scale factors.

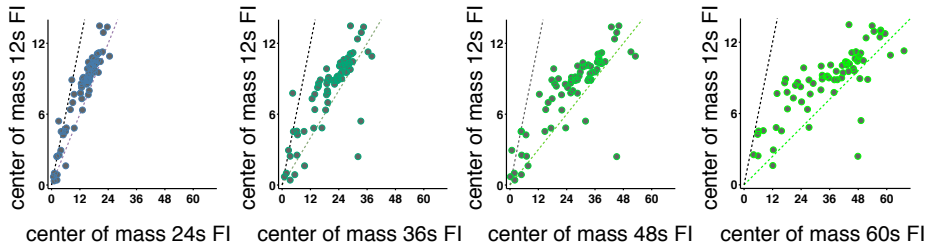


Figure 2.12. **Center of mass comparisons per FI**. COM of each cell's SDF in the 12s FI against each of the other FIs. The black dotted line signifies no change in COM from block to block. The colored dotted line signifies a change in COM that is proportional to the change in FI relative to the 12s FI.

the scale factors (red distributions in Figure 2.11, Kolmogorov-Smirnov test, $p < 0.001$ for all pairwise comparisons). Were the population to have rescaled its responses in direct proportion to the FI, the medians of these distributions should lie at 1/2, 1/3, 1/4, and 1/5 for the scale factors corresponding to 12/24s, 12/36s, 12/48s, and 12/60s FIs, respectively. We observed median values of 0.59, 0.39, 0.30, 0.24 for the corresponding distributions, indicating near-proportional rescaling of response times across the recorded striatal population. A more complete description of the relative scale of responses can be seen in Figure 2.12, where the COM of each cell's SDF in the 12s FI

against each of the other FIs are displayed. These data demonstrate a strong tendency for rescaling of neural responses across the population, suggesting that the state of striatal populations may convey relative elapsed time information scaled to the animal's estimate of the current behaviorally relevant timescale in the environment. We explore this hypothesis in greater detail below.

2.2.3 Striatal populations encode information about timing behavior

The above analyses of striatal neural responses indicate a gross correspondence between striatal activity and timing behavior across blocks of trials, suggesting that striatal activity patterns might guide decisions about when to begin pressing the lever during each FI. To test this hypothesis, we applied a decoding approach to data collected from single trials near block transitions, wherein animals systematically changed the time that they began to press the lever. Specifically, we asked three questions. First, we asked whether decoded time estimates covaried with true time. Second, we asked whether systematic errors in estimated time as compared to true time occurred at these block transitions. Lastly, we asked whether any observed errors in time encoding correlated with timing behavior. We first built a probabilistic decoder to derive an estimate of elapsed time from reward in single trials given the observed spiking response of the population. We focused on the first trials of the 12s and 60s FI blocks because these blocks were the shortest and longest FIs employed, respectively. Thus, animals consistently overestimated and underestimated the amount of time remaining until reward as they entered 12s and 60s blocks. Briefly, our decoder was constructed as follows. In each of the first seven trials of a block, we counted spikes within defined time bins and asked how likely we were to have observed that number of spikes at each time given the observed distributions of spike counts in trials 8 onward of the corresponding block. This generated a likelihood function for current time, given an observed spike count in each bin, for each individual cell. To derive a measure of the population's estimate of the likelihood for current time, we multiplied together the individual cells' likelihood functions. We then took the mean of this likelihood function as our estimate for current time (Dayan & Abbott, 2005). In Figure 2.13, we display decoded estimates as a function of time for the first seven trials of 12s and 60s FI blocks. We found that decoded estimates tracked true time but that systematic errors between estimates and true time were present in the first few trials of the 12s and 60s FI blocks. This feature can be observed more readily when estimates derived from multiple trials are plotted on the

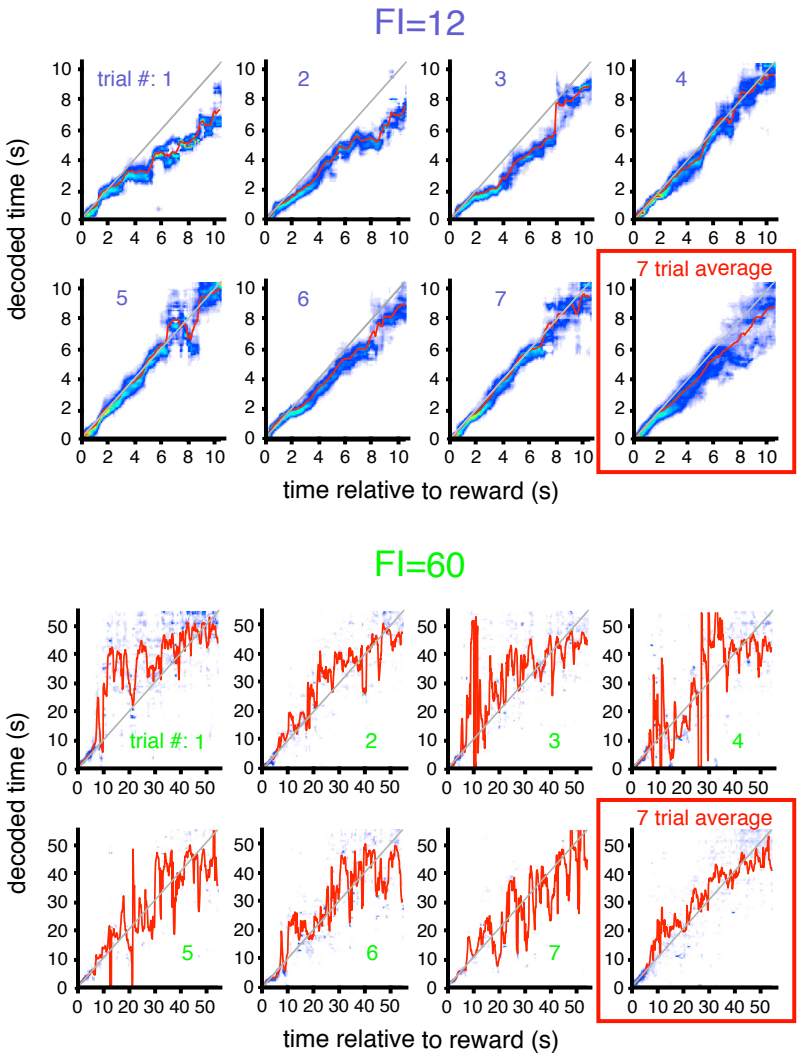


Figure 2.13. **Decoder estimates of elapsed time**. Decoded population estimates of elapsed time from reward in single trials, for the first seven trials of the 12s and 60s FIs block plotted against true time. Red traces indicate the mean of the population likelihood function, and the underlying heatmap indicates the population likelihood function. The last panel shows a seven-trial average likelihood function using the first seven trials of the respective FI block.

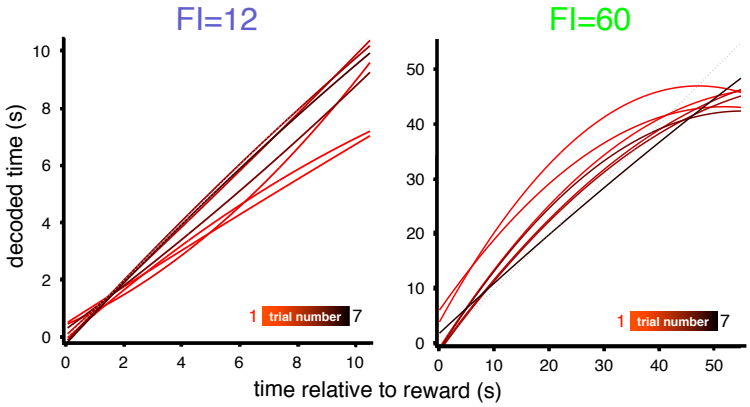


Figure 2.14. **Decoder estimates run slow or fast depending on FI transition**. *Left* - Decoded estimates of elapsed time for the first seven trials of the 12s FI block plotted on the same axis. Curves are quadratic fits to the mean likelihood function of each individual trial (red lines in Figure 2.13). Red curves represent early trials, and black curves represent later trials. *Right* - Same as *left* but for the 60s FI block.

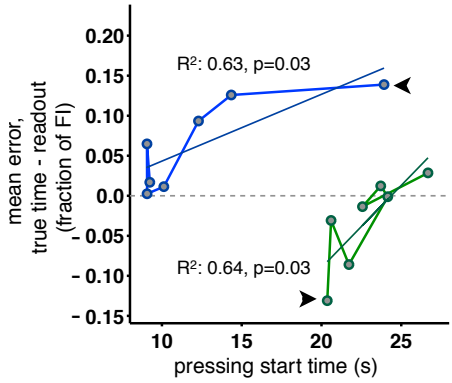


Figure 2.15. **Errors in decoded time predicted timing behavior**. Mean error between true time and the decoded population estimate in the first seven trials of the 12s (blue) and 60s (green) FI blocks. Contiguous trials are connected by solid lines to display the trajectory of the data over trials, and the first trial on each block is indicated by the black arrow. Dashed horizontal gray line represents zero error average decoding as compared to true time.

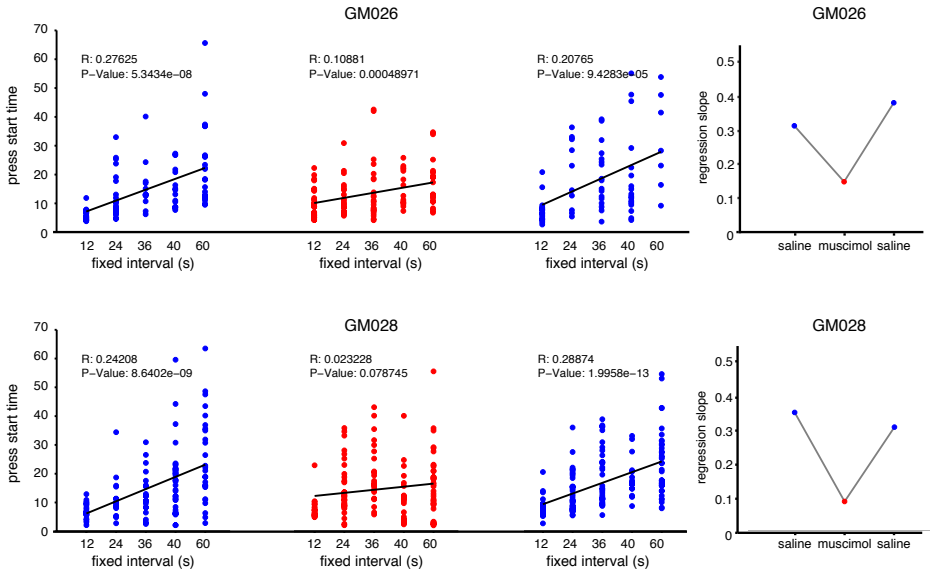


Figure 2.16. **Striatal muscimol infusion diminishes PST relationship to the FI**. Each row corresponds to one of two rats that underwent muscimol infusion. *Left* - Pressing start times per FI of three successive sessions (from left to right) and their respective regression slopes (black line) for two rats (top and bottom). Muscimol treatment session is depicted in red and saline treatment in blue. *Right* - Regression slopes for the same three sessions depicted on the *left*, same color code is applied.

same axes (Figures 2.14, quadratic fits). Initial estimates were relatively slow and fast in the first trials of the 12s and 60s FI blocks, respectively, and became more accurate after the first few trials.

Next we asked whether such timing signals may be used by animals to guide timing behavior. We first asked whether errors in decoded time estimates over the first trials of blocks were correlated with timing behavior. We found that the mean PST was significantly correlated with the errors in time estimates derived from the population over the first seven trials of 12s and 60s FI blocks (Figure 2.15; FI 12s, $R^2 = 0.63, p = 0.03$; FI 60, $R^2 = 0.64, p = 0.03$). In the initial trials of the 12s FI block, rats began pressing late relative to subsequent trials, and likewise, the decoded estimate of time relative to reward ran slow (left panel in Figure 2.14 and blue trace in Figure 2.15). The first trials of the 60s FI block showed a similar relationship, yet opposite in direction: the decoded estimate ran quickly in early trials, and rats were early to press (right panel

in Figure 2.14 and green trace in Figure 2.15). We then tested in two control animals whether manipulating striatal circuitry via bilateral infusions of the $GABA_a$ receptor (GABAa) agonist muscimol produced deficits in timing behavior (Figure 2.16). Indeed, at a dose that rendered rats able to perform the task, muscimol reversibly and significantly diminished the relationship between PST and FI (linear regression, likelihood ratio test, significant effect of treatment, $p < 0.001$), showing that a normally functioning striatum is critical for normal timing behavior. The consistency between time estimates decoded from striatal populations and trial-by-trial variations in timing behavior at block transitions, together with observed dependence of a normally functioning striatum for normal timing behavior, suggests that the brain uses a population code for time that samples broadly from striatal neurons to guide decisions about when to act.

2.2.4 Striatal neurons multiplex information about action and time

Based on previous studies (Mink, 1996; Kim et al., 2009; Jin & Costa, 2010), we expected that striatal neurons would display significant modulation by behaviors during the FI. Could behaviors that accompany task performance fully explain the sequential neural responses we observed? Several features of the data argue against this possibility. Rats consistently licked at the reward port from 0.5s to 5.5s after reward delivery (Figure 2.17), and yet, our ability to decode time was unaffected by the animal being engaged in a fixed behavior over this time (see initial ~5s of decoded time estimates in Figure 2.13). After departing from the reward port, however, it is possible that observed dynamics in neural responses are accounted for by ongoing behaviors. Were this the case, responses related to a particular behavior should not vary depending on when in a trial the rat engaged in that behavior. To identify neurons that were significantly modulated

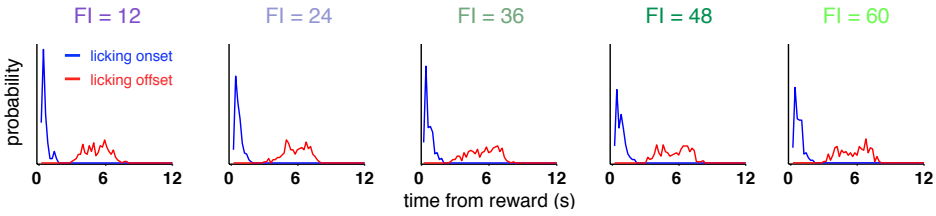


Figure 2.17. **Licking behavior across FIs**. Probability of licking sequence onset (blue) and licking sequence offset (red) in the first 12s of all five FIs.

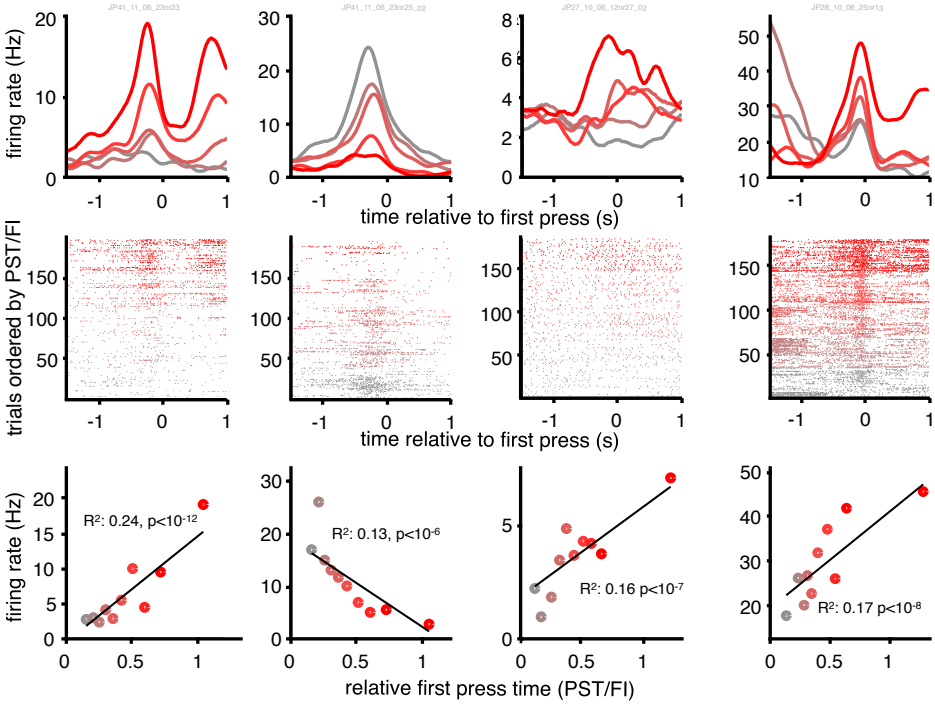


Figure 2.18. **Pressing onset responsive neurons display sensitivity to the time relative to the FIs**. Four single-neuron PSTHs (*top*) and raster plots (*middle*) of 2.5 s epochs aligned on pressing onset event (from three animals; the two first columns display data from two neurons recorded in the same animal and same session). Trials were sorted in ascending fashion from bottom to top on the vertical axis by the pressing onset time relative to the FI (*middle*) and grouped into quintiles. Here, the colors from gray to red represent the first to the fifth quintile, respectively (middle and top panels). *Bottom* - Correlation between the firing rate of the respective neuron and the time of the pressing onset relative to FI. Each data point is color coded from gray to red for the first to the tenth decile of the relative pressing onset time. Firing rates were extracted from the most modulated 500 ms bin of the four bins surrounding the pressing onset event.

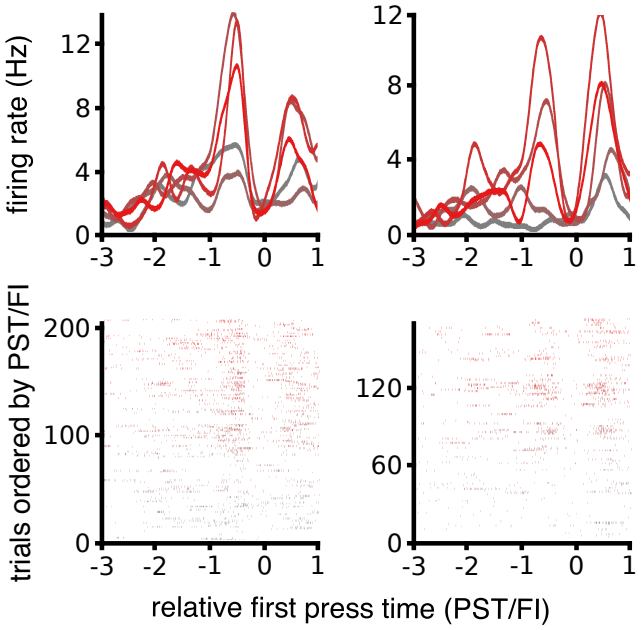


Figure 2.19. **Non-monotonic time-dependent modulations in pressing onset response**. Two additional single-neuron PSTHs (*top*) and raster plots (*middle*) of 4 s epochs aligned on pressing onset event. Trial sorting, alignment and color code as in Figure 2.18.

by a measured behavior in the SFI task, we focused on a 2.5s epoch centered on the PST in each trial. We found that of the 76 neurons displayed in Figure 2.7, 31 exhibited significant modulations around the onset of lever pressing. Next, we asked whether spiking observed in time bins aligned on the PST was additionally correlated with the time, relative to the FI, that pressing onset occurred. More than half of pressing onset-modulated neurons (16/31, 52%) displayed a significant correlation between spiking around each press initiation and the relative time that press onset occurred within the FI (Pearson's linear regression, $p < 0.01$). Figure 2.18 shows examples of four such neurons from three different animals, all of which vary in their responses around the PST, from none at all to robust firing.

The regression approach described above is only expected to identify neurons that display a monotonic relationship between pressing onset response and the relative time

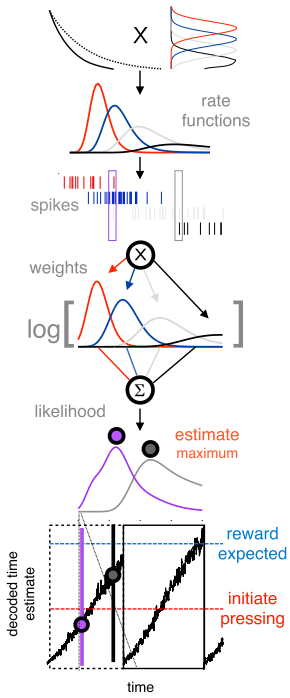


Figure 2.20. **Schematic of timing behavior simulation**. Firing of striatal neurons was modeled based on receptive fields for the height of a decaying trace that is reset in each trial by reward delivery (top left). This trace can decay faster (solid line) or slower (dotted line) by adjusting the parameter γ. The Gaussian functions (top right) represent receptive fields evenly spaced along the height of the trace function. The trace function was multiplied by the receptive fields to generate rate functions, the levels of which vary across time as the memory trace decays. Spike counts observed within defined time bins were then multiplied by the logarithm of their respective rate functions and summed to compute the population log likelihood function for current time given the population response, from $t = 0$ to $t =$ FI. The maximum of this likelihood function was used to derive our estimate for current time relative to reward, for each time bin. Decoded time estimates can run faster or slower depending on whether the trace function decays quickly or slowly. For each trial, when the decoded time estimate reached a given threshold (red dotted line), we simulated a probabilistic pressing process. If the decoded estimate runs too slowly, it fails to reach the threshold value for expected reward (blue dotted line) before the current FI elapses, and the reward happens before it was expected (dotted black box), generating a large prediction error that drives appropriate updating of γ in the next trial. If the decoded estimate runs more accurately (solid black box), a small prediction error is generated, and γ is minimally adjusted in the next trial.

of pressing onset. Other cells may have displayed significant time-dependent modulations in pressing onset response that were not monotonic (for example, see Figure 2.19). To identify such cells, we asked whether the median of distributions of spikes counts, collected around pressing onsets and falling into each of five quintiles of relative PSTs, differed from each other. We found that 53 out of 76 neurons (70%) displayed in Figure 2.7 exhibited significantly different median spike counts across relative time within the FI ($p < 0.01$, Kruskal-Wallis). Of these, nine cells were significantly modulated by the onset of lever pressing and were not identified in the linear regression analysis. Overall, only six cells that displayed response modulation around PST did not exhibit additional modulation by relative time in the FI as assessed by linear regression and/or nonparametric testing for median difference in spike count.

These results suggest that striatal neurons multiplex information about time and immediate sensorimotor state of the animal and argue strongly against the possibility that the striatal population responses we observed can be explained by purely non-time-related responses to specific sensory or motor components of ongoing behavior.

2.2.5 A simple simulation of timing behavior

In order to understand the relationship between the recorded striatal signals and rats' behavior, we ran a simple simulation that performed the SFI task (Figure 2.20). The core of this simulation is comprised of a set of temporal basis functions that were inspired by the diverse single-neuron responses observed in our striatal dataset as well as existing timing and learning models, some of which were introduced in Chapter 1 (Grossberg & Schmajuk, 1989; Suri & Schultz, 1999; Ludvig et al., 2008; Howard et al., 2014).

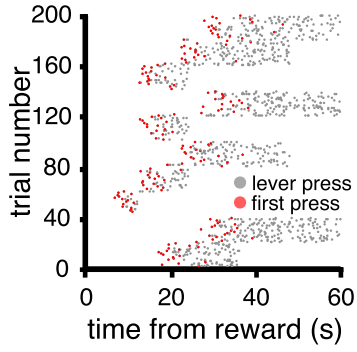


Figure 2.21. **Simulated lever pressing behavior**. Example of simulated lever pressing behavior on the SFI task. Gray markers indicate a lever press and red markers indicate the PST.

We used the method described in Ludvig et al. (2008) to generate temporal basis (see also Section 1.6.2). Each function was used as a rate function for generating inhomogeneous Poisson spike trains from which time was read out during task performance. Whenever this time readout

passed a threshold, presses were produced at a fixed rate. In order to adapt to the changing FIs, we implemented a simple learning rule to update a temporal scale factor for the basis functions depending on the difference between expected time of reward and encoded time at the time of reward delivery. Lastly, to account for our observation that many striatal neurons multiplexed information about action and time, each press produced a response in the temporal basis that was proportional to the product of the original time-dependent rate function at the time of the press and a rate function generated by the press itself. With these elements, we ran the simulation under the conditions contained in the SFI task.

The simulation produced qualitatively similar behavior to that of rats (Figures 2.21 and 2.22) and reproduced the three main features that we observed in striatal neurons: temporal tuning, rescaling of neural responses (Figure 2.23), and multiplexing of information about action and time (Figure 2.24). Although simple, the simulation serves as proof of principle that neural activity with the properties that we observe in this Chapter can serve as a basis for timing behavior and suggests candidate computational elements such as a scale factor and temporal error signal for which there might exist functional analogs in the brain.

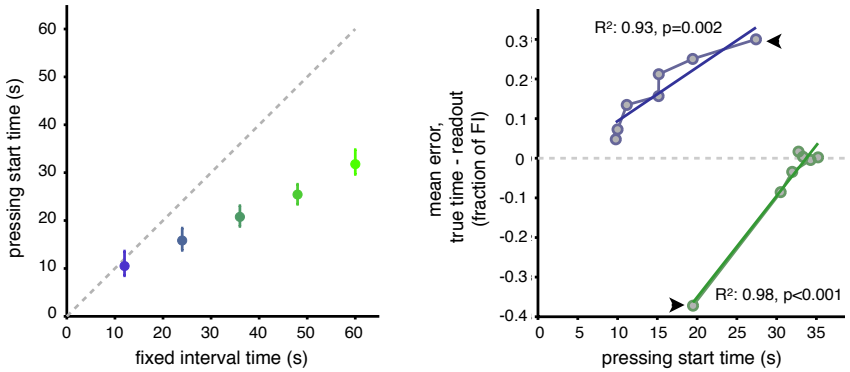


Figure 2.22. **Model accuracy and precision across FIs**. *Left* - Median and interquartile range of modeled PST for each of the five FIs. *Right* - Mean error between true time and the decoded estimate from model units on the first seven trials of the 12s (blue) and 60s (green) FI blocks as in Figure 2.15.

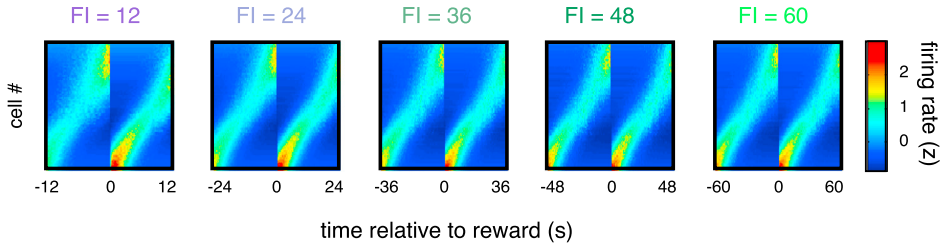


Figure 2.23. **SDFs of simulated units ordered by response profile**. Each panel corresponds to one FI as in Figure 2.7.

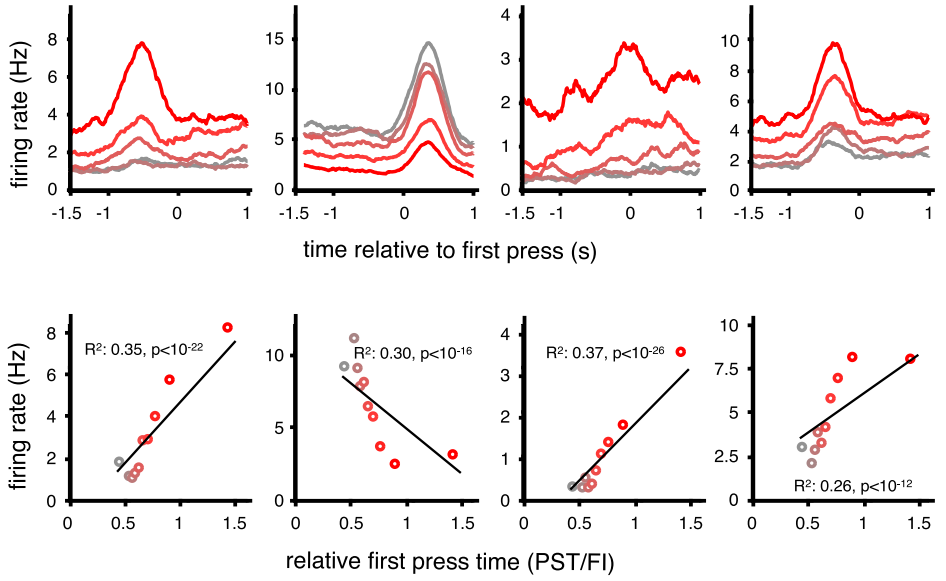


Figure 2.24. **Time dependent modulation of press onset responses in simulated units**. Four simulated single-unit PSTHs of 2.5 s epochs aligned on pressing onset event (*top*). Trials were grouped, aligned and color coded as in Figure 2.18. *Bottom* - Correlation between the firing rate of the correspondent unit on the *top* panel and the PST relative to FI. Each data point is color coded as in Figure 2.18.

2.3 Discussion

Time is a fundamental dimension of animals' experience in the world. As such, it plays an integral role in many brain processes, from perception to motor control to learning and memory formation. What is the role of temporal representation within the BG?

A dominant view supported by a wide range of neurobiological data posits that the BG implements aspects of reinforcement learning (RL) (Schultz et al., 1997; Kim et al., 2009; Doya, 1999; Lauwereyns, Watanabe, Coe, & Hikosaka, 2002; Samejima et al., 2005; Lau & Glimcher, 2008), learning how an organism ought to act in order to maximize reward. However, to learn about the sometimes-delayed consequences of actions and to guide future behavior toward rewarding outcomes, it is absolutely necessary that the brain represent situations and actions through time (Schultz et al., 1997; Sutton & Barto, 1998). Indeed, temporal relations among actions and events contain the causal information that learning systems have evolved to detect through a process sometimes referred to as credit assignment (Balsam & Gallistel, 2009). Once credit for the occurrence of predictable events has been assigned, this information must be used to profitably guide the course and timing of action as situations arise. This continuous learning-behaving cycle is what RL algorithms naturally account for (Sutton & Barto, 1998). Yet, it is not known how the BG, the brain system most often associated with RL, represents temporal relationships over the durations necessary to explain its purported role in animal learning and behavior.

The sequential neural states that we describe in the striatum during timing behavior can provide a unifying view of the BG's role in timing and RL. These signals are strikingly similar to temporal basis functions proposed in existing learning models as more neurally plausible and efficient representations of time (Grossberg & Schmajuk, 1989; Suri & Schultz, 1999; Ludvig et al., 2008), which we show can be used to generate timing behavior similar to what we observed experimentally. Such models operate by learning a set of weights used in a weighted sum of the temporal basis to construct a moment-by-moment prediction about future events such as expected reward. In theory, a weighted combination of activity patterns in the cortical or thalamic inputs to the striatum could act as such temporal basis and modulate the responses of striatal neurons that we observed.

An important question for future studies concerns the mechanism that generates the striatal dynamics we observed. We find it unlikely, given the duration of the intervals we examined, that striatal dynamics were purely locally generated, although several modeling studies suggest mechanisms for generating sequential activity states using striatum-like circuitry over shorter timescales (Ponzi & Wickens, 2010; Berns & Sejnowski, 1998). Indeed, the signals we use to decode time were affected, but not fully explained by, the ongoing sensorimotor state of the animal. Thus, our decoding approach implicitly endorses a number of prominent interval timing theories, positing that animals may use behavioral (Killeen & Fetterman, 1988; Machado et al., 2009)

or sensory state (Ahrens & Sahani, 2011) transitions to learn to time events in the environment and their own behavior.

Our behavioral data did not strictly adhere to the scalar property, as has been shown to be the case in other timing studies (Staddon & Higa, 1999), particularly when subjects estimate long timescales such as the ones presented in this Chapter. Although the variance of the PST distributions increased with the mean PST across the different FIs, the dynamic nature of the SFI task might promote a strategy where animals begin pressing earlier on longer intervals, so as to not avoid missing a possible block transition to a shorter FI.

The data presented in this Chapter appear most consistent with theoretical models that suggest distributed representations of time encoded by the joint activity of populations of neurons (Buonomano et al., 1995). Indeed, the decoder used in this Chapter assumes that time information may be present in many different neurons. However, we cannot rule out that upstream of the population we recorded in the striatum, other forms of temporal representations may exist. For instance, an accumulating process such as that contained within pacemaker-accumulator models (e.g., Treisman, 1963; Gibbon, 1977; see Section 1.4.1) might act to trigger neurons to become active at different delays as the accumulator passes a series of thresholds.

We show that sequential neural activation in the striatum can be used to encode time on a scale of tens of seconds up to one minute. These results add to a growing list of studies that demonstrate sequential activation of neurons over multi-second time-scales in other brain areas, such as the hippocampus (Pastalkova et al., 2008; MacDonald et al., 2011), the cerebellum (Buonomano & Mauk, 1994), the parietal cortex (Harvey et al., 2012), and the prefrontal cortex (Machens et al., 2010; Shinomoto et al., 2011; Kim et al., 2013). Unlike previous studies, we found that many individual striatal neurons exhibited responses that dynamically rescaled with the timing of events in the environment and that this scaling of responses produced changes in time encoding by the population that correlated with timing behavior. Combined with previous studies highlighting the importance of a normally functioning striatum for timing behavior (Meck, 2006; Malapani, Rakitin, et al., 1998; Rowe et al., 2010; Ward et al., 2009), the effect of striatal inactivation shown in Figure 2.16, and other work that demonstrated time encoding by striatal populations over shorter timescales (Jin et al., 2009), the results presented in this Chapter suggest that information about where in time a subject finds itself relative to anticipated events in the environment is present in populations of striatal neurons and is used to guide behavior.

Similar timing signals observed in areas other than the striatum are viewed within the larger context of the functional role of those areas where they were recorded. Timing signals in the hippocampus might endow explicit memories with accurate information about the order and temporal context of events (Howard et al., 2014), and timing signals in the cerebellum might coordinate learned actions at a fine timescale (Buonomano & Mauk, 1994), while timing signals in premotor cortex might enable accurate timing of movement in general (Merchant et al., 2013). The striatal neurons we observed appear to multiplex temporal information with other, non-temporal types of information, such as signals related to the ongoing sensorimotor state of the animal and likely other previously identified striatal signals related to actions, motor sequences, or reinforcement (Lauwereyns et al., 2002; Samejima et al., 2005; Lau & Glimcher, 2008; Jin & Costa, 2010). Such multiplexing of temporal and other information in populations of striatal neurons as observed in this Chapter is consistent with previous observations (Matell et al., 2003) and is likely to be critical to the previously ascribed and well-studied function of the BG in learning and action selection.

2.4 Materials & Methods

All experiments were approved by the Champalimaud Centre for the Unknown Bioethics Committee and the Portuguese Direção Geral Veterinária.

2.4.1 Behavior

Five male Long-Evans hooded rats were trained in an operant lever pressing paradigm reinforced with 0.015ml of water on a FI reinforcement schedule (Figure 2.1). The FI was varied randomly in blocks of ≥ 18 trials among five intervals ranging in multiples of 12 from 12s to 60s . A real-time LINUX state machine directed by custom software in MATLAB (bcontrol, Mathworks) controlled the task. To remove incidential lever depressions that sometimes occurred as animals explored the box, PST was computed as the first press on each trial where the interval until the next press fell below the 85th percentile of the inter-press interval distribution of the entire session.

2.4.2 Neurophysiology

Movable arrays of 32 tungsten microwires (CD neural systems) were implanted unilaterally in the striatum under isoflurane anesthesia. Neural signals recorded during

behavior were amplified and high pass filtered at 100Hz or 250Hz (i2smicro), and waveforms corresponding to action potentials from single neurons were sorted offline using principal component analysis (PCA) (offline sorter, Plexon). All isolated units (179 total from 5 rats, 25 R1, 9 R2, 21 R3, 28 R4, 96 R5) recorded for at least three blocks in sessions in which PSTs correlated significantly with FI ($p < 0.05$) were included in subsequent analyses. To construct SDF, spikes were counted in bins of 20ms. This histogram was then divided by the number of trials, smoothed with a half gaussian kernel with a standard deviation of 500ms, and z-scored by subtracting the mean and dividing by the standard deviation of the time series. We applied PCA to the SDFs of all recorded neurons within each FI separately (Figure 2.10), or the concatenated SDFs of a subset of neurons across all FIs together (Figure 2.7). SDFs were aligned on reward, including activity from reward delivery minus one FI to reward delivery plus one FI. As has been done previously, we applied an ordering procedure wherein SDFs were ordered by their angular position within a plane defined by the contributions made by the first two PCs to SDFs, rotating around the origin (Geffen et al., 2009). This method has the advantage that it orders cells with respect to their dynamics while taking into account the full firing profile of each neuron over time, as opposed to methods that order by peak response time that only take into account one moment in the cell's average firing profile.

2.4.3 Selection for cells with consistent relative response profiles

To identify cells that maintained their position in the population across the FIs we performed the following selection. First, The PCA-based ordering process was applied to all neurons that were recorded in all five FIs ($n = 112$). Importantly, for this analysis PCA was run separately on data collected in each FI so that each cell's ordered position within the population was free to change across FI. Next, each within-FI position was converted to a unit vector, at an angle determined by its ordered position. These unit vectors were then averaged, and those average vectors with a length of 0.75 or greater ($n = 76/112$, 68%) were identified as having consistent ordinal positions within the population across FIs.

2.4.4 Scale factors

To quantify the temporal rescaling observed in striatal neurons across the different FIs, we computed a scale factor for each neuron in the 24s, 36s, 48s, and 60s FIs as

the ratio of the center of mass of neural firing in the 12s FI, over the the center of mass of neural firing in each of the four other FIs (Figure 2.12 and 2.11). We then generated null distributions by recomputing scale factors using the same data but where scale factors were repeatedly computed across cells instead of within single cells (100 repetitions). We then tested for significant differences between the data distributions and the null distributions for each of the four sets of ratios using a Kolmogorov-Smirnov test ($p < 0.001$).

2.4.5 Latency and width of responses

To quantify the latency and the spread of the observed striatal responses (Figure 2.9), we estimated the time of the peak firing rate in each FI and the width at its half-height from from a smoothed PSTH aligned on reward delivery. 10ms bins were used to build the PSTH, smoothed with a gaussian kernel with a standard deviation (SD) that was inversely proportional to the median firing rate of the neuron (SD = 11/median firing rate). We selected the firing mode as the time of the peak of the smoothed PSTH, and the width of the peak at the half height between the trough (minimum) and the peak (maximum) of the smoothed PSTH as our measure of spread. To prevent edge effects, peaks occurring later than three quarters of the FI were removed from the analysis. We asked whether each cell's spread was correlated with it's delay to peak firing by linear regression (FI 12s, $R = 0.4443, p < 0.001$; FI 24s, $R = 0.7563, p < 0.001$; FI 36s, $R = 0.7188, p < 0.001$; FI 48s, $R = 0.5910, p < 0.001$; FI 60s, $R = 0.4733, p < 0.001$).

2.4.6 Decoding methods

We built a maximum likelihood decoder to estimate current time within the SFI task given the pattern of activity across the population. To best control for conditions across sessions, we focus on the the first seven trials of the 12s and 60s FI blocks. In 12s FI trials, spikes were counted in 1.5s bins beginning with the earlier edge aligned to reward delivery and moved in 100ms steps until the later edge reached 12s. In 60s FI trials,spikes were counted in 5s bins beginning with the earlier edge aligned to reward delivery and moved in 100ms steps until the later edge reached 60s. Time labels were placed at the earlier edge of the bins. To build distributions of spikes counts used to generate the single cell likelihood functions, $P(t \mid r)$, meaning the probability of it being any moment in time given the observed spike count, we did the following: first we counted the number of spikes observed for trials 8 onward since a block switch, and for all time bins. To estimate the underlying probability density we smoothed the resulting

histograms using local regression within a window of 30 spikes. We then counted spikes within the same time bins during trials 1 through 7 since the block switch. For each spike count, we determined the likelihood of observing that many spikes in each time bin, given the data from the later trials. These single cell likelihood functions from every cell recorded ($n = 179$) were multiplied together to derive a population estimate of current time given the number of spikes observed from each cell in the population. To derive a single estimate from this likelihood function derived from the population, we took the mean (Dayan & Abbott, 2005).

2.4.7 Muscimol infusions

We implanted 24-gauge stainless steel cannulas bilaterally into the striatum (coordinates from bregma: 2mm anterior; ± 2.7mm lateral; 4mm ventral) of two rats under isoflurane anesthesia. After one week of recovery from surgery, rats were allowed to perform the SFI task. Once the rats performed well (5 block switches with significant regression of PST vs. FI) in the SFI task, we injected saline (PBS, 1x), muscimol (GABAa agonist, SigmaTM) and saline in three successive days. To perform the injections, rats were anaesthetised using 1.5-2.5% isoflurane (v/v). Muscimol or saline solution was delivered using a 1 µL (Hamilton) syringe attached to a injection pump (Harvard ApparatusTM, HA11D 702209) through a 24-gauge injector. The injector extended 1.5mm beyond the tip of the guide cannulas. We used a muscimol concentration of $22.2\,\mathrm{ng\,\mu L^{-1}}$ in saline, injecting 0.6 µL during 2.5 min. The injector was left in place for an additional 1.5 min and the rats were placed for a 45 min recovery period in a home-cage before starting the test session. Both rats received the saline vehicle injections using the same parameters as the muscimol treatment.

2.4.8 Identification of pressing onset related neurons

To identify neurons modulated by the onset of lever pressing, we compared distributions of "test" spike counts in time bins of 500ms during two seconds centered on lever pressing onset, to distributions of "baseline" spike counts from 500ms time bins situated between 2.5 and 1s before lever pressing onset (Jin & Costa, 2010). Baseline time bins consisted of 3 non-overlapping 500ms time bins, while test time bins consisted of 500ms bins that were moved in 10ms steps starting with the trailing edge at 1s before lever pressing onset, until the leading edge met 1s after lever pressing onset. Cells were considered significantly modulated around lever press onset if at least 50 consecutive test bins were

significantly different from a baseline distribution constructed by pooling spike counts in the three non-overlapping baseline bins Kruskal-Wallis test, $p < 0.05$).

2.4.9 Identification of press start time modulated neurons

To identify neurons displaying press onset responses that varied additionally with time, we compared observed spike counts in four bins of 500ms centered on the PST in each trial. Given the observed rescaling in striatal neurons during the task, we first normalized PSTs by dividing each PST by the FI of the block in which it occurred, resulting in fractional PSTs. We then used two methods to identify neurons that exhibited responses aligned on the PST that varied with fractional PST. First we performed linear regression, using fractional PST to predict spike counts in each time bin. We performed an additional analysis to test for differences in median spike count observed across trials that had been binned with respect to fractional PST. Fractional PSTs were separated into five bins, the edges of which corresponded to 0, 20, 40, 60, 80 and 100^{th} percentiles. Responses were considered not to be uniform with respect to fractional PST if trials in at least one bin displayed a median spike count that was significantly different from the rest (Kruskal-Wallis, $p < 0.01$).

2.4.10 Simulation of timing behavior

We ran a simulation of timing behavior in the SFI task that takes into account the temporal tuning, rescaling and multiplexing properties of striatal neurons presented in this Chapter. The firing of striatal neurons was simulated based on receptive fields for the height of a decaying memory trace of the reward (Figure 2.20,top panel) (Ludvig et al., 2008).

$$y_\tau = \gamma^\tau \tag{2.1}$$

The trace function y (equation 2.1), was reset to a value of 1 at each reward delivery; Its exponential decay rate γ was initialized as 0.999. The trace function decays over an arbitrary time unit τ that varies from 0 to 10000. The receptive fields f had equidistant means μ on the height of the trace function y, and standard deviation σ fixed at 0.1.

$$f(y, \mu, \sigma) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2\sigma^2}} \tag{2.2}$$

Next, we generated 75 temporal basis functions that served as rate functions for generating spike trains. The basis functions were computed as follows:

$$b_t(i) = f\left(y_t, \frac{i}{m}, \sigma\right) y_t \tag{2.3}$$

The level of the i^{th} basis function b out of m total basis at time t, was determined by the product of the corresponding height of the trace function y_t and the receptive field f. The result was a set of temporal basis functions that became progressively wider, delayed and with a lower peak (Figure 2.20, second panel from top). We then generated spikes for each unit using these basis functions rate functions in an inhomogeneous poisson spiking process. To compute internal estimates of time and simulate behavior, we first counted spikes bins of 10ms and at each time step we multiplied the number of observed spikes by the logarithm of its' respective unit's basis function, it's tuning curve for time. This produced a log likelihood for time given the observed spikes for each unit. We then summed the individual unit log-likelihood functions (Jazayeri & Movshon, 2006). We used the maximum of this likelihood function as our estimate of current time relative to reward for each time bin. Two fixed thresholds were established to trigger changes in the simulated behavior. Crossing the first threshold set at three eighths (3/8) of the maximally possible internal time estimate triggered press initiation. The second threshold, set at twice the level of the pressing threshold, was the expected reward time. Once the decoded time crossed the pressing threshold, we simulated pressing at a fixed rate. Pressing persisted at that fixed rate until the end of the FI was reached and reward delivered. The time of reward receipt was then compared to the expected time of reward (the second threshold). Differences between real and expected time generated a type of prediction error that was used to update the time constant γ of the exponential trace that controlled scaling of the basis functions. The update proceeded as the equation below:

$$\gamma := \gamma + \alpha \cdot \left(\frac{1}{\gamma^{t-t_r}}\right) \cdot \left(\frac{l(t_r) - \tau_r}{max(l_\tau)}\right) \tag{2.4}$$

Where α was the learning rate set to 0.00007, t was the current time, and t_r was the time of the reward. The expected time of reward τ_r was fixed as three fourths of the length of the τ vector ($\tau \approx 7500$) and $l(t_r)$ is the current estimation of time at the moment that the reward is received. This adjustment of the parameter γ in the trace function (equation 2.1) led encoded time to change faster or slower, allowing for the simulation to exhibit an adaptive timing of neural and behavioral responses that resembled the

experimental data. Information about each press fed back onto the temporal basis in the following manner. Each unit possessed a press locked contribution to its rate function. This press contribution to firing was a gaussian rate function with a standard deviation of 200ms, and a mean precedence relative to the press that was drawn from another gaussian distribution with a mean of -1500ms and a SD of 300ms. However, the contribution of each press to the firing of a given unit was computed as the product of the instantaneous, time dependent rate at the time of the press, and the press locked contribution to firing.

Acknowledgments

The experiments presented in this Chapter were designed by Gustavo Mello and Joseph Paton. Gustavo Mello and Sofia Soares acquired the data. Gustavo Mello, Sofia Soares and Joseph Paton analyzed the data We thank Bassam Atallah, Brian Lau, Kenway Louie, Christian Machens, Zachary Mainen, Thiago Gouvêa, Eric DeWitt, Alfonso Renart, and Masayoshi Murakami for discussions and critical comments on versions of the manuscript on which this Chapter is based (Mello et al., 2015). We thank the histopathology and vivarium staff from the Champalimaud Scientific and Technological Platforms for support. This work was supported by Champalimaud and Gulbenkian Foundations as well as by fellowships from the Portuguese FCT (Fundação para a Ciência e a Tecnologia) to Gustavo Mello and Sofia Soares.

Chapter 3

Midbrain dopamine neurons control judgment of time

Our sense of time is far from constant. For instance, time flies when we are having fun, and it slows to a trickle when we are bored. Midbrain dopamine (DA) neurons have been implicated in variable time estimation. However, a direct link between signals carried by DA neurons and temporal judgments is lacking. Additionally, it is unclear whether distinct DA neuron subpopulations in the substantia nigra pars compacta (SNc) and ventral tegmental area (VTA) carry distinct signals during timing behavior. In this Chapter, we measured and manipulated the activity of DA neurons in both SNc and VTA as mice judged the duration of time intervals. We found that pharmacogenetic suppression of DA neurons decreased behavioral sensitivity to time. Next, we measured DA neuron activity in the SNc and found that these neurons encode a reward prediction error signal that reflects both temporal expectation and the expected probability of reward. Furthermore, we observed a horizontal shift of the psychometric curve toward long choices when SNc-DA neuron activity was low, and a horizontal shift in the opposite direction when activity was high. Additionally, we show that transient activation or inhibition of DA neurons in the SNc was sufficient to slow or speed time estimation, respectively. Interestingly, DA neurons in the VTA did not reflect nor control temporal judgments. Therefore, activity of DA neurons in the SNc, but not in the VTA, reflects and can directly control the judgment of time.

3.1 Introduction

Our ability to accurately estimate and reproduce time intervals is variable and depends on many factors, including motivation (Gable & Poole, 2012), attention (Coull, Vidal, Nazarian, & Macar, 2004), sensory change (Ahrens & Sahani, 2011), novelty (Pariyadath & Eagleman, 2007), and emotions (Droit-Volet & Meck, 2007). In addition, several neurological and neuropsychiatric disorders (Pastor et al., 1992; Wittmann et al., 2007; Noreika et al., 2013; Wahl & Sieg, 1980) are accompanied by changes in timing behavior. Midbrain dopamine (DA) neurons are implicated in many of the psychological factors (Cools, 2008) and disorders (Pastor et al., 1992; Noreika et al., 2013; Lüthi & Lüscher, 2014) associated with changes in time estimation.

Midbrain DA neurons also encode reward prediction errors (RPEs; Schultz et al., 1997; Bayer & Glimcher, 2005; Eshel, Tian, Bukwich, & Uchida, 2016; Steinberg et al., 2013), an important teaching signal in reinforcement learning (RL; Sutton & Barto, 1998). Phasic DA responses to reward-predicting cues reflect the magnitude of (Tobler, Fiorillo, & Schultz, 2005; Cohen et al., 2012), probability of (Fiorillo, Tobler, & Schultz, 2003), and expected time delay until the reward (Kobayashi & Schultz, 2008; Fiorillo et al., 2008). When expectation varies over time, DA neuron responses are smaller at times when rewards and reward-predicting cues are more expected (Fiorillo et al., 2008; Pasquereau & Turner, 2015), indicating that DA neurons receive temporal information. Manipulations of the DAergic system by pharmacological (Maricq & Church, 1983) or genetic (Drew et al., 2007) approaches disrupt timing behavior, suggesting that DA neurons may directly modulate timing. However, the data from pharmacological and genetic manipulations are inconsistent: in some cases, DA seems to speed up timekeeping (Maricq & Church, 1983; Drew et al., 2007; Buhusi & Meck, 2002), and in others, DA seems to slow down or not affect timekeeping (Lake & Meck, 2013; Balcı et al., 2010). Additionally, there are distinct subpopulations of midbrain DA neurons, that may carry distinct signals in the context of timing behavior and RL, adding another layer of complexity to the study of the functional role of these neurons. However, most of the manipulation studies mentioned above did not distinguish between the two main subpopulations of midbrain DA neurons: those in the substantia nigra pars compacta (SNc) and those in the ventral tegmental area (VTA).

SNc and VTA are located in anatomical proximity to each other and have classically been treated as a relatively homogeneous neuronal population. Many SNc and VTA neurons share electrophysiological properties such as broad waveforms, two modes of activity patterns (a tonic mode between 2 and 5Hz and a phasic mode characterized by

brief discharges >15Hz) and inhibition mediated by the D2-type autoreceptor (Sulzer, Zhang, Benoit-Marand, & Gonon, 2010). These properties have been used to identify these neurons *in vivo* (e.g., Schultz, 1986; Schultz & Romo, 1987; Schultz et al., 1993; Hollerman & Schultz, 1998; Fiorillo et al., 2003). Additionally, one traditional tool used to identify both SNc-DA and VTA-DA neurons histologically is immunostaining brain slices for tyrosine hydroxylase (TH), the rate-limiting enzyme for DA synthesis (e.g., Watabe-Uchida et al., 2012; Stamatakis et al., 2013).

Despite their common features, there are also profound differences between SNc-DA and VTA-DA neurons, suggesting that this seemingly homogeneous group of neurons is in fact quite heterogeneous. One major difference between the two is their distinct input and output connectivity patterns with both cortical and subcortical areas. SNc-DA neurons have been shown to receive inputs from areas such as motor and somatosensory cortices, subthalamic nucleus (STN) and dorsal striatum (DS) (Watabe-Uchida et al., 2012). VTA-DA neurons, in contrast, receive inputs from the lateral hypothalamus and nucleus accumbens (NAc), among others (Watabe-Uchida et al., 2012). Many of these connections are reciprocal: for example, SNc-DA neurons are known to send the vast majority of their axons to the DS (nigrostriatal pathway) while VTA-DA neurons mainly project to the NAc (mesolimbic pathway) (Malenka, Nestler, & Hyman, 2009).

How these distinct anatomical connectivity patterns might translate to differences in the response of SNc-DA or VTA-DA neurons is currently unclear. Although some studies suggest that distinct DA neuron connectivity patterns relate to different responses in the presence of aversive stimuli (Lammel, Ion, Roeper, & Malenka, 2011; Lammel et al., 2012), whether these anatomical differences relate to distinct signals in the context of timing behavior is unknown. To determine (i) what signals are encoded by midbrain DA neurons during timing behavior and (ii) how DA neurons contribute to variability in temporal judgments, in this Chapter we measured and manipulated the activity of SNc-DA and VTA-DA neurons in mice as they performed categorical decisions about duration (Gouvêa et al., 2014, 2015).

3.2 Results

3.2.1 An intact DAergic system is required during performance of a temporal discrimination task in mice

We first trained mice to perform a temporal discrimination task (Figure 3.1, left panel). Mice initiated trials at a central nose port, immediately triggering the delivery of two

identical tones separated by a variable delay. Mice reported the delay between tones as shorter or longer than 1.5s at one of two lateral nose ports for water reward. Incorrect choices were not rewarded. Performance was nearly perfect for the easiest intervals but more variable for intervals near 1.5s (the boundary between the "short" and "long" categories) and was well described by a sigmoidal psychometric function (Figure 3.1, middle panel).

We then pharmacogenetically suppressed DAergic neuronal activity and observed impaired temporal judgments on treatment days as compared with adjacent nontreatment days ($p < 0.004$, $n = 3$ mice; Figure 3.1, right panel). We also observed a tendency to perform fewer trials [control group, 177 ± 15 trials; CNO–treated group, 115 ± 54 trials; mean ± standard deviation (SD); $p = 0.05$], suggesting that the animals' motivation was affected by DAergic suppression.

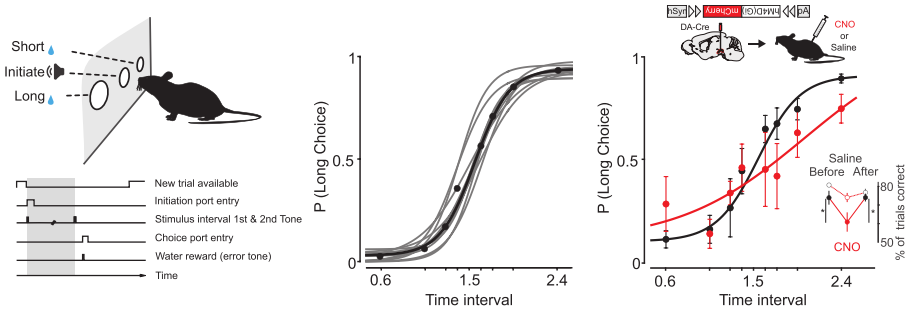


Figure 3.1. **Performance of temporal discrimination task requires DAergic activity**.
Left - Task schematic and order of events (circles in the upper panel, nose-ports; gray shading in the lower panel, interval period). *Middle* - Logistic function fit to the daily (gray) and average (black) performance of an example mouse (10 sessions). *Right* - Pharmacogenetic suppression (hM4D) was targeted to midbrain DAergic neurons, and mice were injected with either clozapine N-oxide (CNO) or saline on consecutive days (top panel); the main panels shows the mean psychometric performance on days with saline or CNO treatment (black or red, respectively; $n = 3$ mice). Error bars, standard error of the mean (SEM). The inset shows the percent of correct trials on days before and after CNO treatment in mice expressing hM4D (filled circles, n = 3; * $p < 0.005$) or non–hM4D-expressing controls (open circles, $n = 4$). Error bars, SEM.

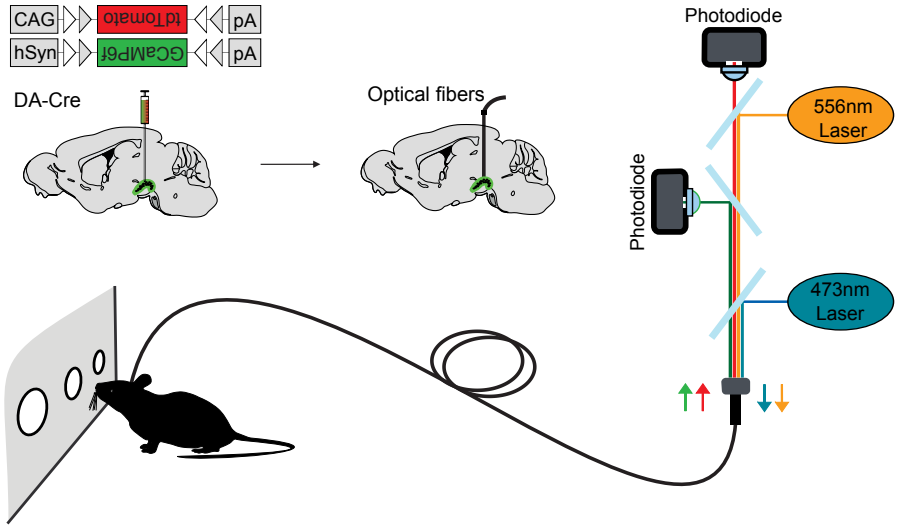


Figure 3.2. **Schematic of the photometry apparatus and surgical procedure**. A single surgical procedure was performed for adeno-associated virus (AAV) injections and optical fiber implantation. After 1-2 weeks of recovery, mice were retrained on the temporal discrimination task and the Ca^{2+} activity of DAergic neurons was recorded during task performance.

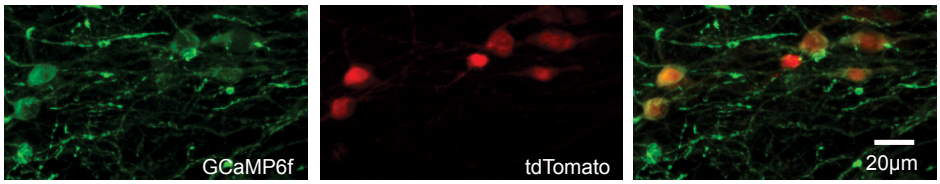


Figure 3.3. **Image of the SNc histology**. Confocal image of SNc neurons displaying GCaMP6f (*left*) and tdTomato (*middle*) positive neurons. A merged image is shown on the *right*.

3.2.2 SNc-DAergic activity is precisely aligned to temporal cues, not movement

To test whether fluctuations in endogenous DA neuron activity predicted systematic changes in temporal judgments, we used fiber photometry (Matias, Lottem, Dugue, & Mainen, 2017) to measure Ca^{2+} activity in DAergic neurons (Figure 3.2). Briefly, we injected a combination of two adeno-associated viruses (AAVs) containing double-

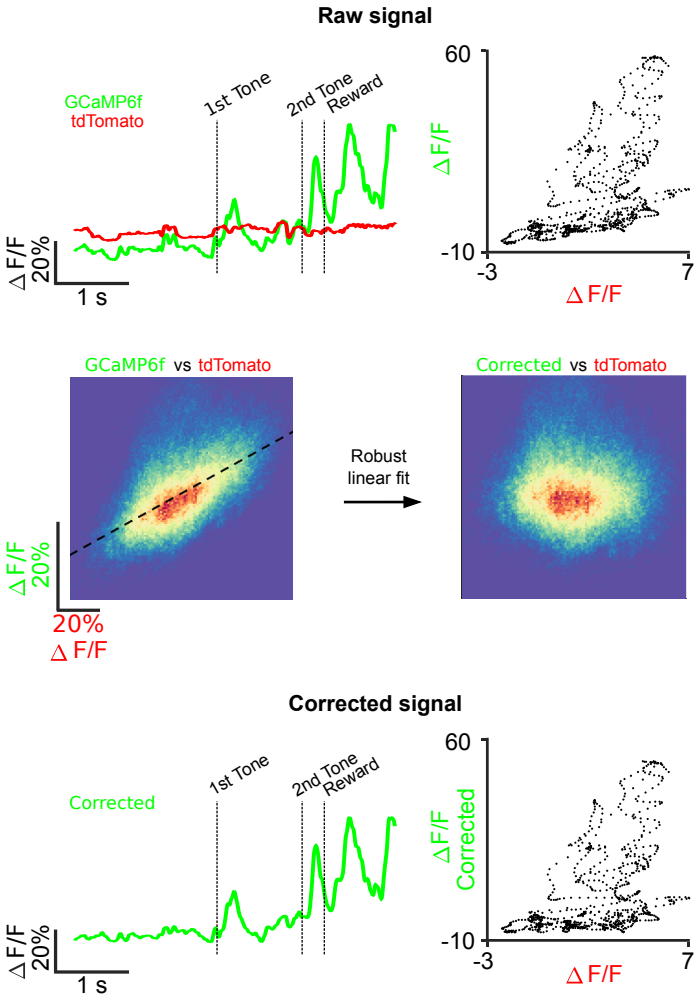


Figure 3.4. **Common mode correction of photometric recordings**. *Top* - Simultaneously recorded time course ΔF/F in green and red channels (*left*), and correlation between these two channels (*right*). *Middle* - Color density plot of correlation between green and red channels before (*left*) and after (*right*) correcting for common mode noise by removing correlations. *Bottom* - Same as *top* after common mode correction.

floxed inverted open reading frames encoding either the Ca^{2+} indicator GCaMP6f or the activity-independent fluorescent protein tdTomato, targeting the SNc of transgenic mice expressing cyclic recombination enzyme (Cre) under the control of the promoter of either TH or dopamine transporter (DAT) (DA-Cre, Figure 3.2 and Figure 3.3). The tdTomato signal was used to correct for small changes in GCaMP6f signal due to motion (Figure 3.4, see Materials & Methods in Section 3.4 for details).

We began by targeting SNc-DA neurons since they densely innervate the DS (Watabe-Uchida et al., 2012), and removal of DA input to the DS can cause a selective deficit in timing (Meck, 2006). Additionally, the DS has been shown to encode relative time as well as temporal judgments (see data in Chapter 2; Gouvêa et al., 2015). The reconstruction of the location of optical fiber placement is shown in Figure 3.5. Next, we analyzed single trial traces and observed that SNc-DAergic responses were locked to the three main task events on single trials: the first tone, the second tone, and reward delivery (or omission thereof, see Figure 3.6). Average traces of these signals per interval duration across five mice are shown in Figure 3.7. Activity increased after reward delivery and decreased when the reward was omitted in the case of incorrect choices (Figure 3.8, *left* panel; Schultz et al., 1993). SNc-DAergic signaling has also been implicated in movement. In our task, DAergic responses following a tone could in principle reflect the tone itself or a subject's movement, as subjects reported their choice shortly after the second tone.

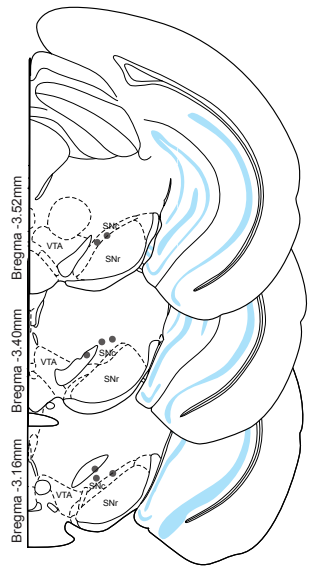


Figure 3.5. **Optical fiber placement for photometry experiment**. Circles represent the tips of optical fibers.

However, as can be seen in Figure 3.8, SNc-DA neuron activity in this task was consistently locked to the presentation of the second tone and to reward delivery, but not to incorrect choices. Indeed, responses to the first tone, which coincided with animals' entry into a noseport, and responses to the second tone, which was not tightly aligned with animals' entry into a noseport, were similar to each other (*middle* panel in Figure 3.8, red and gray traces). In contrast, we observed no increase in activity aligned on unrewarded entries into the choice noseport (*middle* panel in Figure 3.8, black trace) or to entries into the initiation port in the absence of a tone (*right* panel in Figure 3.8,

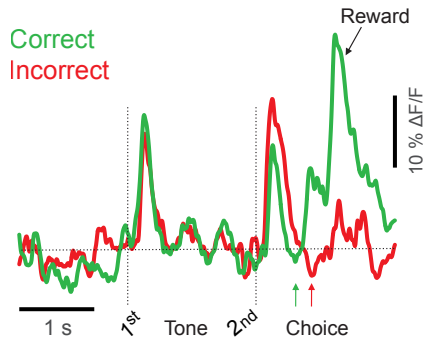


Figure 3.6. **Single trial example photometric traces**. Two example traces recorded during a single correct and incorrect trial of the 1.74s interval.

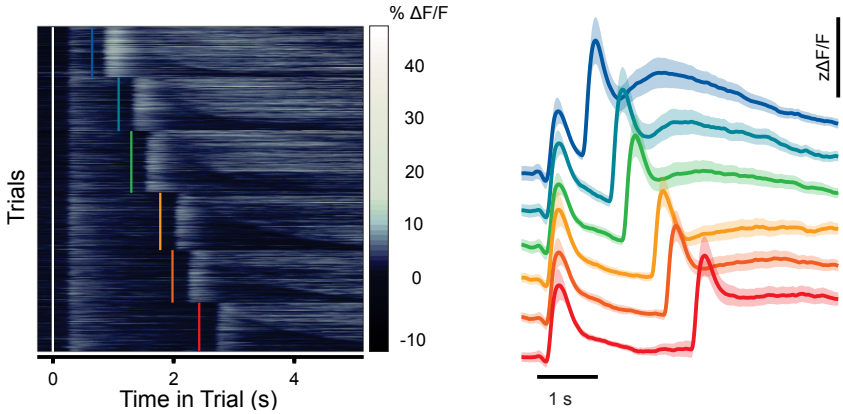


Figure 3.7. **Single trial and average DA neuron activity grouped by interval duration**. *Left* - All trials of DA neuronal activity recorded from a single subject are shown, split by interval duration and aligned on trial initiation (first tone delivery; white vertical line). Each row represents a trial, and within each interval, trials are sorted from fast (*top*) to slow (*bottom*) response time (RT, time from the second tone to choice; 3759 trials). *Right* - Mean DAergic neuron responses, split by interval duration ($n = 5$ mice; intervals are color-coded as throughout). Shading, SEM across mice. z, z-score, ΔF/F, see Materials & Methods in Section 3.4.

black trace). Lastly, SNc-DA neuron activity was better aligned to tone presentation than to the movement of leaving the noseport (*right* panel in Figure 3.8, red and blue traces). These results demonstrate that SNc-DA neuron activity in this task did not reflect movement per se.

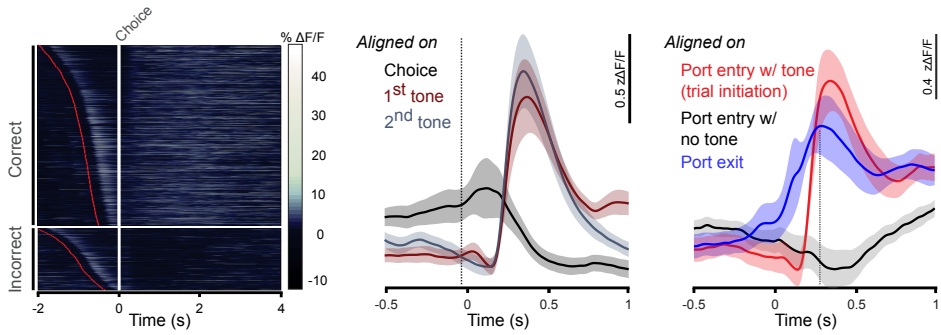


Figure 3.8. **SNc-DA neuron activity does not reflect movement per se**. *Left* - Photometric recordings of SNc-DA neuronal activity from a single subject, split by outcome (correct choices, *top*; incorrect choices, *bottom*) and aligned on choice (white). Within each outcome, trials were sorted by RTs [slow (*top*) to fast (*bottom*)]. Red dots mark the time of second-tone presentation (2426 trials). *Middle* - Mean DAergic responses of incorrect trials aligned on the three main task events [first tone (red), second tone (gray), and choice (black); $n = 5$ mice]. Shading, SEM across mice. *Right* - Mean SNc-DA neuron activity aligned on three events: central port entry followed by the 1st tone that signals trial initiation (red); central port entry in the absence of a tone (black); central port exit after trial initiation (blue). Shade: SEM across mice ($n = 5$ mice).

3.2.3 SNc-DAergic responses correlate with temporal judgments and are explained by a simple model of RPE

In this task, the second tone marks the end of the interval to be categorized and is a sensory cue that predicts reward. The amplitude of a RPE at the time of the second tone should be modulated by two factors: the subject's expectation of reward at tone delivery and their temporal expectation of the second tone itself. First, expectation of reward varies as a function of stimulus difficulty, where the more difficult the interval to be categorized, the lower the probability of reward (Figure 3.9). Second, because delay intervals were randomly selected from the stimulus set on each trial, occurrence of the second tone becomes less surprising with time (Figure 3.9). Indeed, animals were sensitive to changing temporal expectation, as indicated by a systematic decrease in response time (RT, the delay between second tone delivery and choice execution) with increasing interval duration (RT for the shortest interval greater than RT for the longest interval; $p < 0.005$ in each of five mice). To test whether second tone responses reflected a RPE that integrated information about temporal expectation and expected reward,

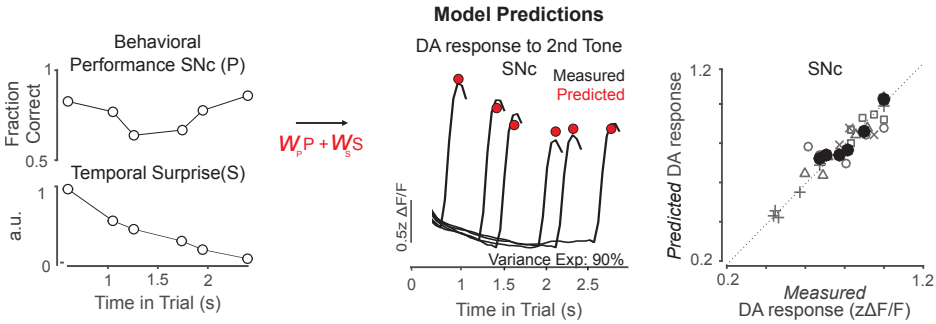


Figure 3.9. **SNc-DA responses evoked by the second tone are explained by a simple model of RPE**. *Left* - Linear model including RPE components: expectation of reward P (subject performance, *top*) and temporal expectation S (surprise, the inverse of the subjective hazard function; *bottom*). W, weight; a.u., arbitrary units. See Materials & Methods in Section 3.4 for details. *Middle* - Measured second tone SNc-DA responses for six time intervals (black traces; $n = 5$ mice) are compared to predicted SNc-DA response (red dots). *Right* - The graph on the right shows model predictions versus measured SNc-DAergic activity (gray symbols, individual mice; mean responses across mice, black filled circles).

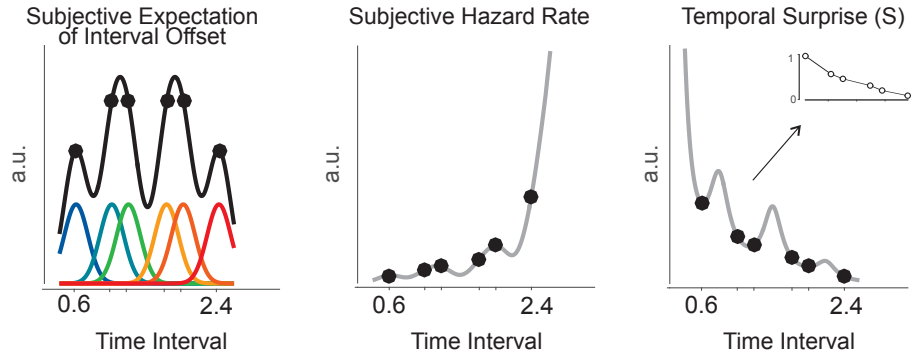


Figure 3.10. **Computation of Temporal Surprise in RPE model**. *Left* - Subjective expectation, E, of interval offset, i.e. second tone (black) computed as the sum of the subjective expectation that each of the 6 intervals (colors) would occur. The expectation of each interval is assumed to be a Gaussian, $SD = 0.15s$. *Middle* - Subjective hazard rate, H, (gray) is the probability of the second tone occurring given that it has not yet occurred. *Right* - Temporal surprise, S, (gray, the inverse of H) at each of the 6 intervals (black dots) and plotted in the inset as in Figure 3.9. See Materials & Methods in Section 3.4 for details.

we asked how well the pattern of average responses to all six second tones could be explained by a linear combination of temporal expectation (i.e., surprise, the inverse of the subjective hazard function; Figure 3.10) and performance (the probability of reward for each stimulus). On average, 90% of variance in mean responses could be explained by a relatively equal contribution of these two factors (range, 58 to 99%; $n = 5$ mice; Figure 3.9). Reward responses were also consistent with RPE coding: within a given choice category, they tended to be larger for intervals that animals miscategorized more often (Figure 3.11). On average, SNc-DA neuron responses to the second tone contained information about elapsed time through their encoding of temporal expectation. Do

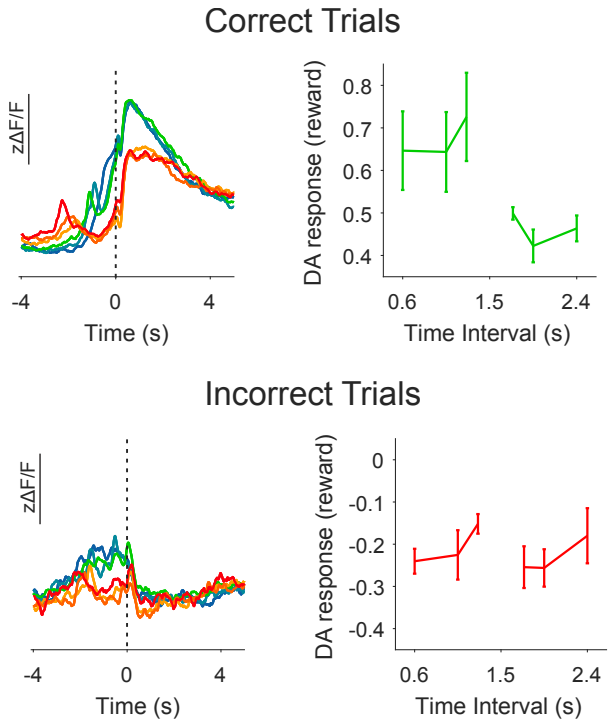


Figure 3.11. **Reward responses reflect performance**. *Top left* - Mean photometry signals aligned on choice (reward) during correct trials. Colors as in Figure 3.7 indicate the different interval durations ($n = 5$ mice). *Top right* - Mean DAergic neuron response for correct trials, window 0.2 - 5s after choice, as a function of interval duration. Error bars indicate SEM ($n = 5$ mice). *Bottom* - Same as *top* but for incorrect (no reward) trials.

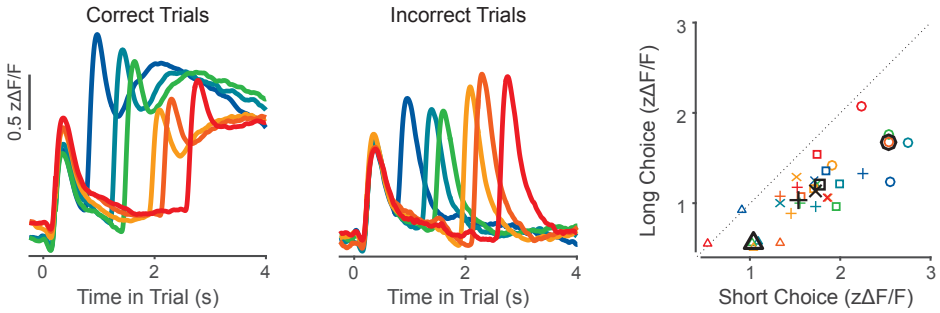


Figure 3.12. **SNc-DAergic responses correlate with choice behavior**. *Left* - Mean photometry signals aligned on trial initiation for all intervals during correct trials ($n = 5$ mice). *Middle* - Same as *left* but for incorrect trials. *Right* - Mean SNc-DA response to the second tone when an interval was judged as long versus short. Each shape represents a different mouse. Black symbols represent responses averaged across all interval stimuli.

these responses relate to variations in judgments of time? When animals correctly judged intervals, the response to the second tone was, on average, larger for intervals in the short category (Figure 3.12, *left* panel). However, on incorrect trials, the pattern was reversed: The response to the second tone was larger for intervals in the long category (Figure 3.12, *middle* panel). Thus, SNc-DA response magnitude reflected the animals' assessment of the interval duration, not the actual interval duration. Over all intervals, the second-tone response for a given interval was significantly larger when that interval was judged as short ($p < 0.001$; Figure 3.12, *right* panel). How do these results relate to the underlying decision and motor processes that guide choice during the task?

3.2.4 Changes in a time-dependent component of choice behavior are predicted by SNc-DAergic activity

In principle, the trial-to-trial variations in SNc-DA neuron activity could be related to a time-dependent component of the decision, such as the speed of internal timekeeping or the location of the decision boundary in time. Alternatively, variations in SNc-DA activity might reflect a time-independent component of the behavior, such as a constant action bias.

To quantitatively evaluate these two possibilities, we performed a logistic regression to assess the degree to which the magnitude of the SNc-DA neuron response to the second tone predicted animals' choices on single trials. We found that activity predicted

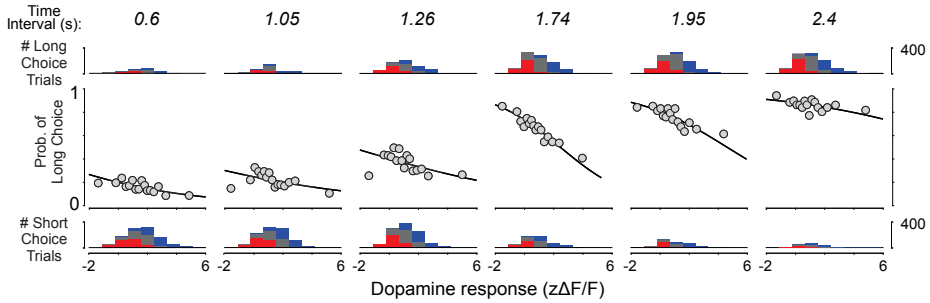


Figure 3.13. **Logistic regression predicting choice using DA activity**. *Middle* - Trial-by-trial logistic regression (black) that predicts choice from the amplitude of the second-tone DA response (gray), for each of the six time intervals (left to right). *Top and bottom* - Histograms that illustrate the number of trials, as a function of DA response, in which subjects made long and short choices, respectively ($n = 8533$ trials, 5 mice). For each session and interval, DA responses are grouped into terciles: high (blue), medium (gray), and low (red). This color code will be used throughout this Chapter.

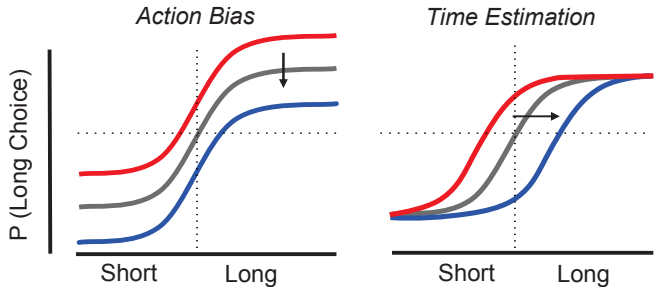


Figure 3.14. **Illustration of possible relationships between DA activity and judgments**. Distinct patterns of temporal judgments are expected depending on the nature of the relationship between DA response and choice.

choice to a lesser extent in the case of easy stimuli than in the case of difficult stimuli (Figure 3.13).

These data suggest that the SNc-DA neuron response was systematically related to the horizontal position of the psychometric curve along the time axis and not the vertical position along the choice axis (Figure 3.14). To test this, we split trials into high, medium, and low terciles of the distribution of responses to the second tone [Figures 3.13 (histograms) and 3.15]. While the second tone response amplitude was

used to group trials, the systematic ordering of SNc-DA neuron responses emerged toward the beginning of the trial and persisted throughout an interval (Figure 3.16, *top left* and *bottom* panels). We next constructed psychometric curves for trials in each tercile and compared a range of models for the psychometric curve.

The model that best explained the behavioral data collected from high-, medium-, and low-tercile trials consisted of three sigmoid curves that differed only in their horizontal location along the time axis (Figure 3.16, *top right* panel). We observed a shift toward long choices when SNc-DAergic activity was low, and the opposite shift when activity was high. Specifically, as SNc-DAergic activity varied from the lower to the upper tercile, the psychometric threshold shifted by ∼340ms (i.e., 20% of the 1.5s category boundary; range, 90 to 620ms; 6 to 42%; $n = 5$ mice). The relationship between SNc-DAergic response and psychometric shift was observed for recordings in either hemisphere (Figure 3.17), thus ruling out an explanation based on the laterality of short versus long choices.

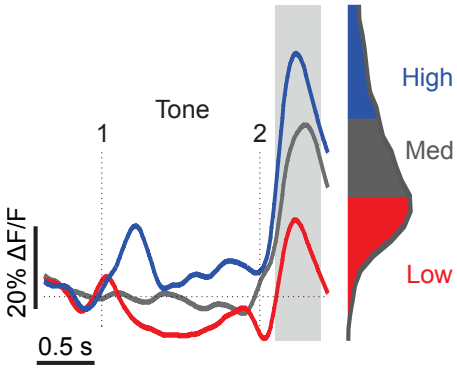


Figure 3.15. **Illustration of trial-by-trial tercile analysis**. Three individual trials illustrating low, medium, and high second-tone DA responses (quantified as the mean response in the gray-shaded box) and grouped by tercile within the entire second tone response distribution, depicted at right.

Instead, these results indicate that higher or lower SNc-DAergic activity is correlated with a change in a time-dependent component of the decision. How might this correlation between SNc-DA neuron activity and the location of the psychometric curve along the time axis relate to our initial finding that temporal expectation contributed to the average second-tone response? The theory of DAergic RPE coding predicts that slower (or faster) timekeeping, by stretching (contracting) temporal surprise along the time axis, should increase (decrease) DAergic responses to the second tone (Figure 3.18). We observed a pattern of SNc-DAergic response to the second tone that was consistent with this (Figure 3.12 and 3.18). Furthermore, if SNc-DAergic activity reflects RPE continuously throughout a trial, differences in activity associated with slower or faster timekeeping (i.e., the separation between low- and high-activity terciles) should also grow continuously over time, and indeed, this is the case in our data (Figure 3.16 and

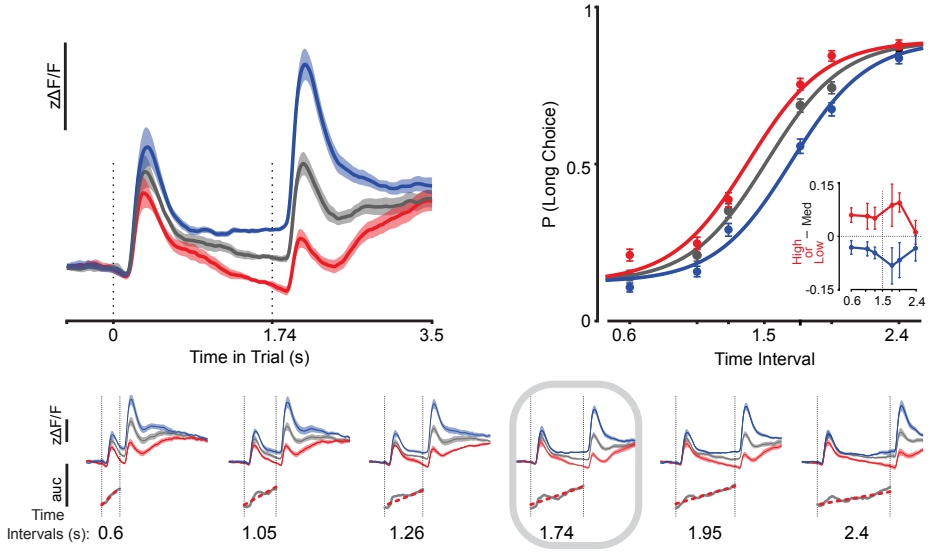


Figure 3.16. **SNc-DAergic activity correlates with a change in a time-dependent component of choice behavior**. *Top left* - Average SNc-DA response in each tercile for the 1.74s interval stimulus ($n = 1868$ trials, 5 mice). Shading, SEM. *Top right* - Psychometric curves constructed using trials from each tercile of SNc-DA response. Curves are the maximum-likelihood fits of logistic functions with the lowest bayesian information criterion scores ($n = 8533$ trials, 5 mice). Error bars, 95% confidence interval (CI). The inset shows the difference in the probability of making a long choice between medium and low or high (red or blue) SNc-DA response trials. Error bars, SEM. *Bottom* - The top row is as in *top left* but for all six interval durations; data shown on the *top left* are outlined in gray. The bottom row shows results from the receiver operating characteristic (ROC) analysis: the area under the ROC curve (auROC), distinguishing high- and low-tercile SNc-DA responses. This difference in DA response increased during the course of the trial (red linear regression; coefficient of determination (R^2) ranging from 0.72 to 0.98; $p < 0.0001$).

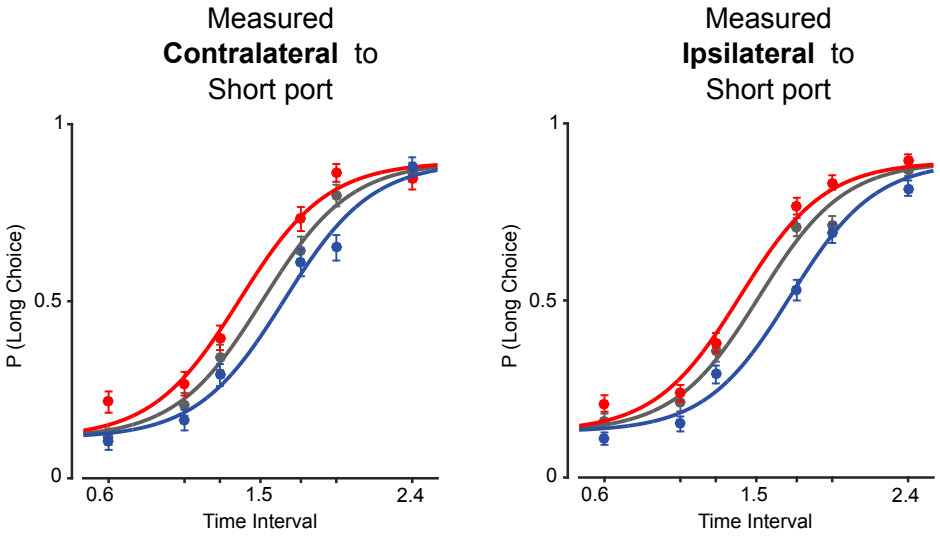


Figure 3.17. **Changes in temporal judgments are similarly predicted by SNc-DA neuron activity in both hemispheres**. Probability of mice making long choices, P(Long Choices), during low (red), medium (gray) and high (blue) tercile DAergic activity recorded contralateral (left) and ipsilateral (right) relative to the short choice port. Sigmoidal curves are the lowest bayesian information criterion (BIC) score maximum likelihood fits of logistic functions ($n = 5$ mice).

3.18). In contrast to the expected impact of variability in the speed of timekeeping on RPE coding, it is not apparent to us how changes in the location of the decision boundary along an animal's internal notion of time should change RPEs arising at the presentation of the second tone. The most parsimonious explanation of the data is that SNc-DA neuron activity reflects variability in the speed of internal timekeeping.

3.2.5 Optogenetic manipulation of SNc-DA neurons is sufficient to change judgments of time

These results demonstrate a correlation between temporal judgments and SNc-DA neuron activity. However, it is unclear whether SNc-DA neuron activity simply reflects, or whether it is sufficient to cause changes in, time judgments. We mimicked the observed variability in SNc-DAergic responses by optogenetically activating or inhibiting SNc-DA neurons (Figure 3.19 and 3.20). We confirmed that SNc-DA neurons could be reliably activated and inhibited during our light protocols by electrophysiologically recording the

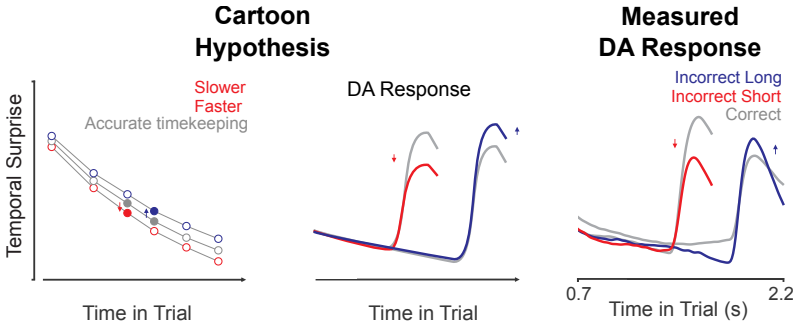


Figure 3.18. **Illustration of how trial-by-trial variation in temporal estimation may impact DAergic response**. *Left and middle* - Illustration of how changes in temporal estimation (i.e. slower or faster decrease of the temporal surprise, *left*) may result in smaller DAergic responses when time estimates run faster than average (red) and larger responses when they run slower (blue) (*middle*). *Right* - Measured DA response for the two time intervals nearest the category boundary (same data as in Figure 3.12). Arrows in all panels illustrate consistency between hypothesized effect of faster (red) and slower (blue) timekeeping on responses to the second tone shown in the *left* and *middle* panels, and experimentally observed changes in response to second tone shown in the *right* panel associated with incorrect "long" (red) and "short" (blue) judgments.

response of optogenetically identified SNc-DA neurons using the same photoactivation or photoinhibition protocol outside of the context of the task (Figure 3.21). Next, we activated or inhibited SNc-DA neurons on a minority of randomly chosen trials during task performance. Notably, we found that increasing or decreasing SNc-DA activity resulted in a horizontal shift in the psychometric curve in the directions predicted by

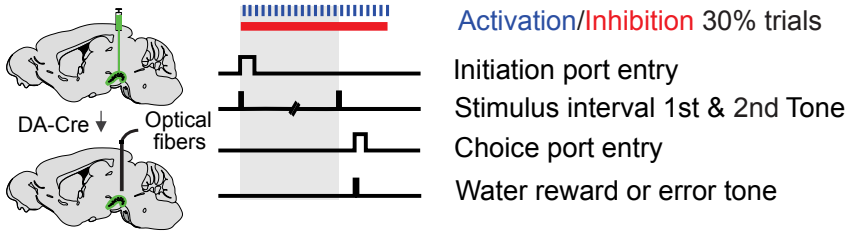


Figure 3.19. **Schematic of optogenetic experiment**. *Left* - Illustration of viral strategy and subsequent fiber implantation. *Right* - Activation (blue) and inhibition (red) protocol schematic.

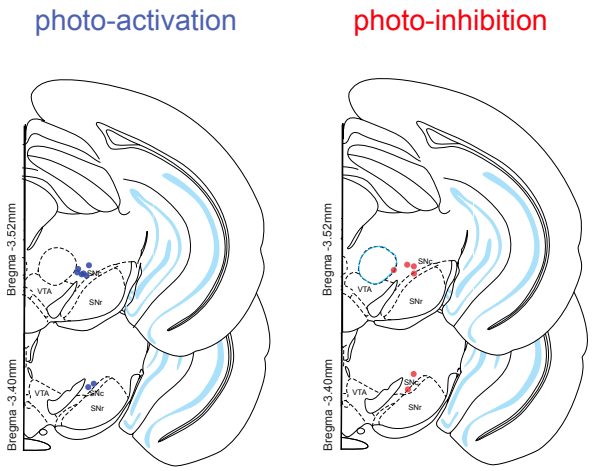


Figure 3.20. **Optical fiber placement for optogenetic experiment**. Circles represent the tips of the optical fibers across the two experiments.

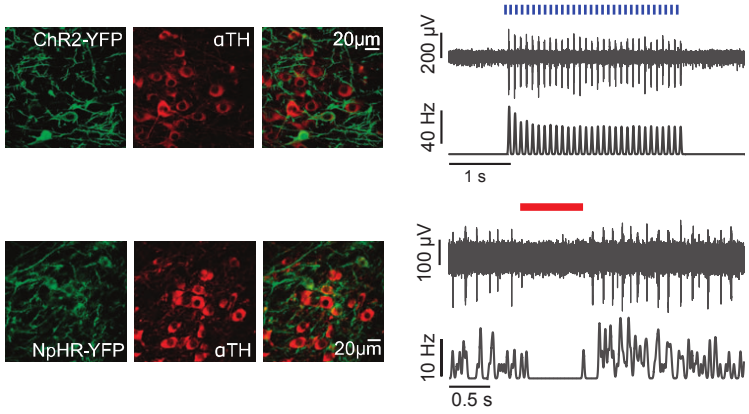


Figure 3.21. **Histology confirmation and validation of light protocols**. *Left* - Histology confirming membrane expression of ChR2-YFP or NpHR-YFP, both in green, in neurons of the SNc expressing TH in red. *Right* - Histology confirming membrane expression of ChR2-YFP or NpHR-YFP, both in green, in neurons of the SNc expressing TH in red. *Right* - Single-trial (*top* panels) and peri-stimulus time histogram (*bottom* panels) of *in vivo* electrophysiological measurement of two DA neurons reliably activated and inactivated by light ($n = 53$ and 8 trials, respectively).

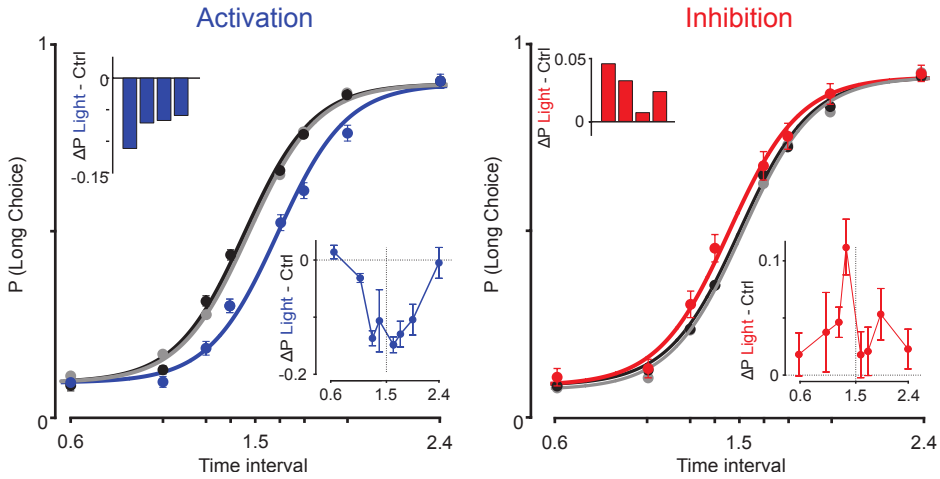


Figure 3.22. **Optogenetic manipulation of SNc-DA neurons is sufficient to change judgment of time**. *Left* - Choice behavior and psychometric curves during control trials (black), photoactivated trials (blue), and unstimulated trials immediately after photoactivation (gray, $n = 4$ mice). Error bars, 95% CI. Insets show the mean difference in the probability of a long choice between photoactivated and control trials (*top*, one bar per animal; *bottom*, one data point per stimulus). Error bars, SEM. *Right* - Same as *left* panel but for animals whose DA neurons were inhibited ($n = 4$ mice).

the photometry data, albeit more modestly in the case of photoinhibition (excitation, 140 ± 20 ms, $n = 4$ mice; inhibition, –68 ± 23 ms, $n = 4$ mice; Figure 3.22 and 3.23). These effects were transient, occurring only on stimulated trials, and thus could not be explained as resulting from learning (gray curves in Figure 3.22), nor were they observed in control animals (Figure 3.24).

In addition, as was the case when sorting trials on the basis of SNc-DA response to the second tone, we observed no systematic effect on RTs, arguing against SNc-DAergic neuron activity affecting the subjects' movement toward or incentive salience of choice options during the task (Figure 3.25).

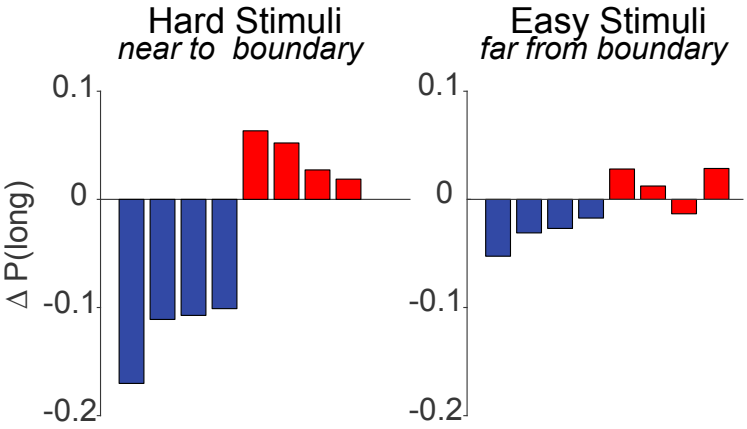


Figure 3.23. **Bidirectional control of temporal judgments by SNc-DA neurons in individual animals**. Mean difference in probability of a long choice [ΔP(long)] between activation (blue) or inhibition (red) trials and control trials. Each bar represents one mouse. Hard stimuli (*left* panel) include the four intervals closest to the 1.5s category boundary, and easy stimuli (*right* panel) include the 4 stimuli farthest from the 1.5s category boundary. Note the larger change for stimuli closer to the category boundary.

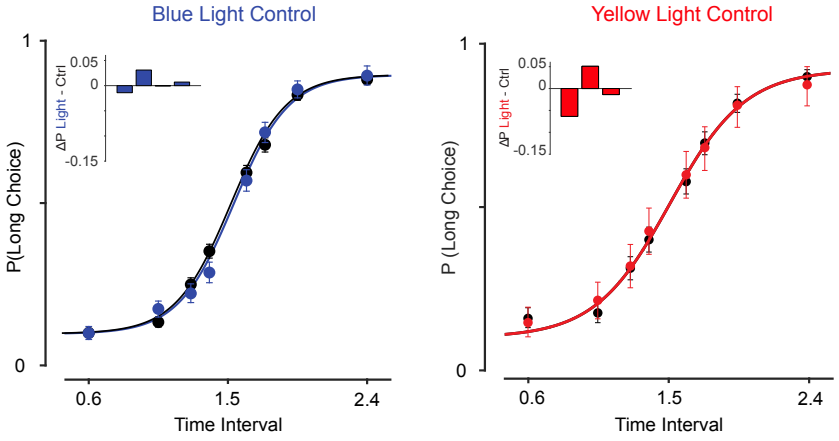


Figure 3.24. **Light alone does not impact temporal judgments**. Choice behavior and psychometric curves for control mice (*left*, $n = 4$ mice; *right*, $n = 3$ mice), related to Figure 3.22.

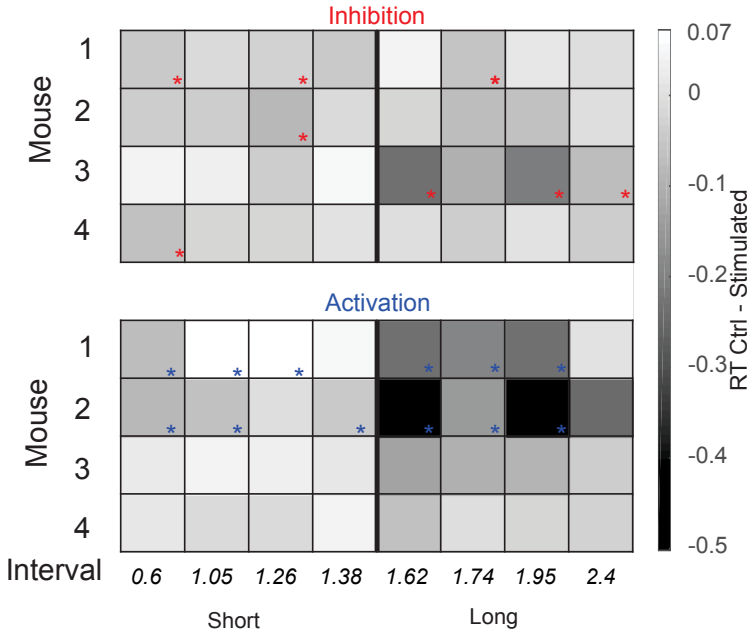


Figure 3.25. **Activation and inhibition of SNc-DA neurons do not systematically impact RTs**. Change in median RT for each interval stimulus (columns) in each stimulated mouse (rows). White and black indicate increases and decreases in RT respectively, star indicates $p < 0.05$ significance. Note that, despite the opposite effects of activation and inhibition on temporal judgments, they have no systematic pattern of impact on RTs.

3.2.6 VTA-DA neuron responses do not correlate with or change temporal judgments

The results shown so far in this Chapter demonstrate a direct link between SNc-DA neuron activity and temporal judgments. Do DA neurons in other regions, such as the VTA, carry similar signals during this task?

To answer this question, we repeated the fiber-photometry experiment illustrated in Figure 3.2, but aimed the optical fibers at the VTA. At first, the responses from VTA-DA neurons appeared very similar to the responses of SNc-DA: VTA-DA neuron activity was also locked to the main task events (first tone, second tone and reward delivery/omission). A closer look at the data, however, highlighted noticeable differences between DAergic activity in the two areas. Figure 3.9 shows that SNc-DA neuron activ-

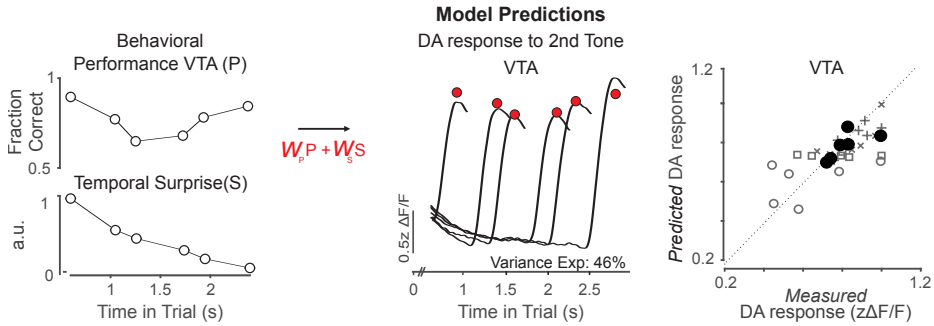


Figure 3.26. **VTA-DA responses evoked by the second tone do not reflect temporal surprise in their RPE**. *Left* - Linear model including RPE components: expectation of reward P (subject performance, *top*) and temporal expectation S (surprise, the inverse of the subjective hazard function; *bottom*). W, weight; a.u., arbitrary units. See Materials & Methods in Section 3.4 for details. *Middle* - Measured second tone VTA-DA responses for six time intervals (black traces; $n = 4$ mice) are compared to predicted SNc-DA response (red dots). *Right* - The graph on the right shows model predictions versus measured VTA-DAergic activity (gray symbols, individual mice; black filled circles, mean responses across mice).

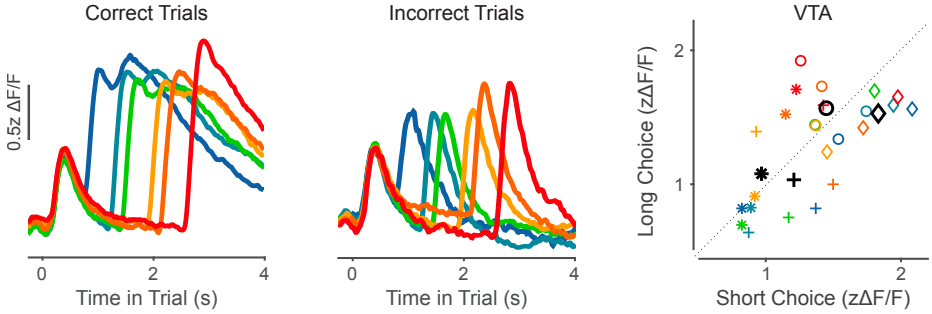


Figure 3.27. **VTA-DAergic responses do not correlate with choice behavior**. *Left* - Mean photometry signals aligned on trial initiation for all intervals during correct trials ($n = 5$ mice). *Middle* - Same as *left* but for incorrect trials. *Right* - Mean VTA-DA response to the second tone when an interval was judged as long versus short. Each shape represents a different mouse. Black symbols represent responses averaged across all interval stimuli.

ity is well explained by a relatively equal contribution of both predictors (performance and temporal surprise) in a simple RPE model. When we applied the same model to the VTA-DA neuron data, however, we observed that the contribution of the temporal

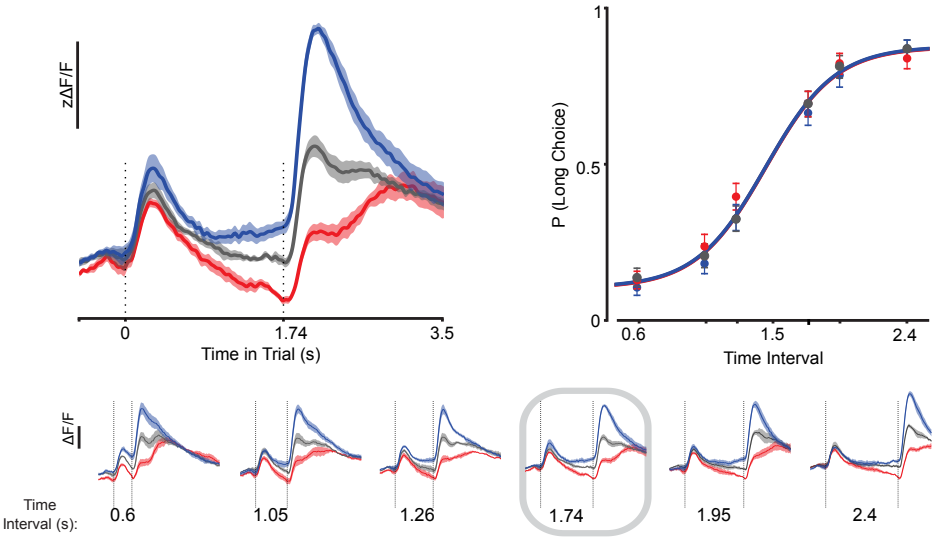


Figure 3.28. **VTA-DAergic activity does not correlate with a change in a time-dependent component of choice behavior**. This Figure follows the structure of Figure 3.16, but using data from VTA-DA neurons. *Top left* - Average VTA-DA response in each tercile for the 1.74s interval stimulus (5 mice). Shading, SEM. *Top right* - Psychometric curves constructed using trials from each tercile of VTA-DA response. Curves are the maximum-likelihood fits of logistic functions with the lowest bayesian information criterion scores. Error bars, 95% CI. *Bottom* - As in *top left* but for all six interval durations; data shown on the *top left* are outlined in gray. Compare this data with data shown in Figure 3.16.

expectation predictor was greatly reduced (Figure 3.26, 0.14 ratio between temporal expectation and performance, see Section 3.4.8 for details). Additionally, we found that splitting average VTA-DA neuron activity by stimulus and judgment did not reveal the same pattern as was observed in SNc-DA neurons (i.e., second tone-evoked VTA-DA neuron activity was not consistently different between short and long judgments; Figure 3.27, compare with Figure 3.12). These result suggest that, in our task, the RPE signal (evoked by the second tone) carried by VTA-DA neurons mainly reflects the average probability of reward the animal expects, whereas the response of SNc-DA neurons additionally reflects the temporal expectation of the second tone event itself.

The results in Figure 3.12 suggest that VTA-DA neuron activity evoked by the second tone might not related to variations in time judgments, unlike SNc-DA activity. First we confirmed that second tone-evoked VTA-DA neuron activity was variable in

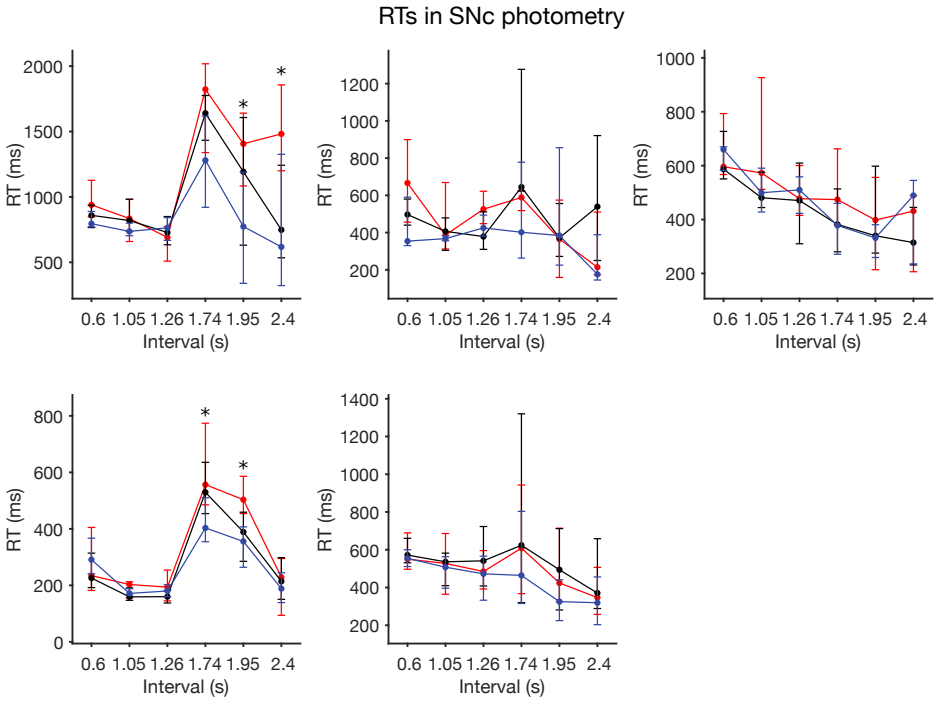


Figure 3.29. **There are no consistent differences between RTs on trials divided by SNc-DA neuron activity tercile**. Median RTs and interquartile range of correct trials are plotted for each interval duration. Trials were split by SNc-DA tercile as described before. Each panel represents a single subject. Stars indicate $p < 0.05$ significance when testing the distributions of RTs of the two extreme terciles using a Wilcoxon rank sum test.

amplitude on a trial-by-trial basis when controlling for stimulus duration. To do this, we again split trials according to the tercile within which they fell on the distributions of second tone-evoked activity (Figure 3.15). This second tone response variability was similar to that previously observed in the SNc (Figure 3.16). We had additionally observed that variability in SNc-DA neuron activity was systematically related to temporal judgments, evidenced by an horizontal shift towards short choices with higher SNc-DA neuron activity (Figure 3.16, *top right* panel). However, unlike SNc-DA neurons, VTA-DA neuron activity evoked by the second tone did not relate to temporal judgments. This was evident since the performance from trials corresponding to the different terciles of VTA-DA activity overlapped (Figure 3.28). We reasoned that perhaps the variability

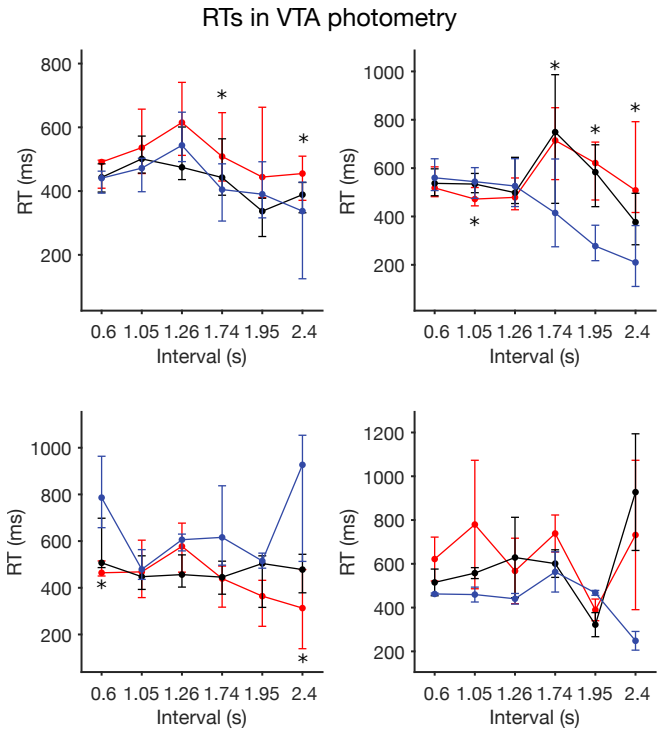


Figure 3.30. **There are no consistent differences between RTs on trials divided by VTA-DA neuron activity tercile**. Median RTs and interquartile range of correct trials are plotted for each interval duration. Trials were split by VTA-DA tercile as described before. Each panel represents an single subject. Stars indicate $p < 0.05$ significance when testing the distributions of RTs of the two extreme terciles using a Wilcoxon rank sum test.

in second tone evoked VTA-DA activity could underlie a vigor or motivational state that could be reflected in the animals' RTs. When we analyzed the RTs for each individual animal split by terciles of DA activity, we did not observe any consistent differences (Figures 3.29 and 3.30), indicating that neither SNc-DA nor VTA-DA neuron second tone evoked activity related to RT.

The data presented so far suggest that the relationship between DA activity and temporal judgments is specific to SNc neurons. We then reasoned that we should not observe an effect on temporal judgments by activating VTA-DA neurons, since we did not observe major temporal expectation signals in VTA-DA neurons, nor did we

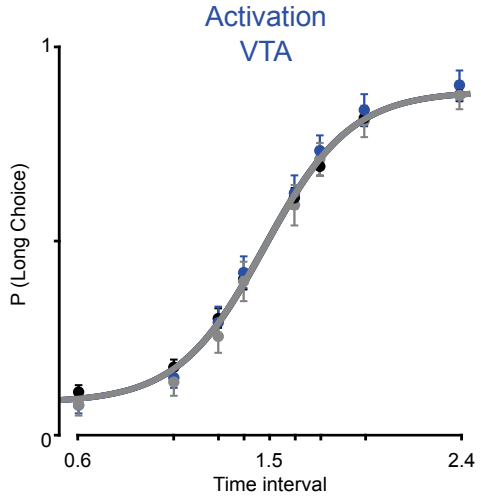


Figure 3.31. **Optogenetic manipulation of VTA-DA neurons does not impact judgment of time**. Choice behavior and psychometric curves during control trials (black), photoactivated trials (blue), and unstimulated trials immediately after photoactivation (gray, $n = 5$ mice). Error bars, 95% CI.

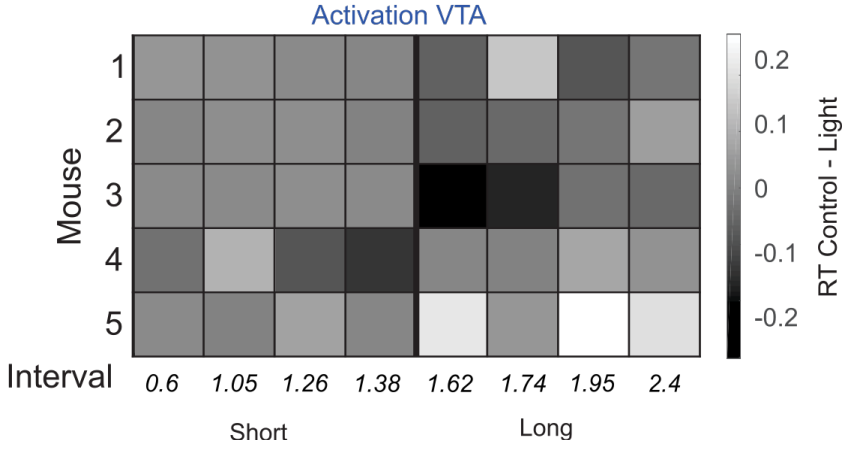


Figure 3.32. **Activation and inhibition of VTA-DA neurons do not systematically impact RTs**. Change in median RT for each interval stimulus (columns) in each stimulated mouse (rows). White and black indicate increases and decreases in RT respectively.

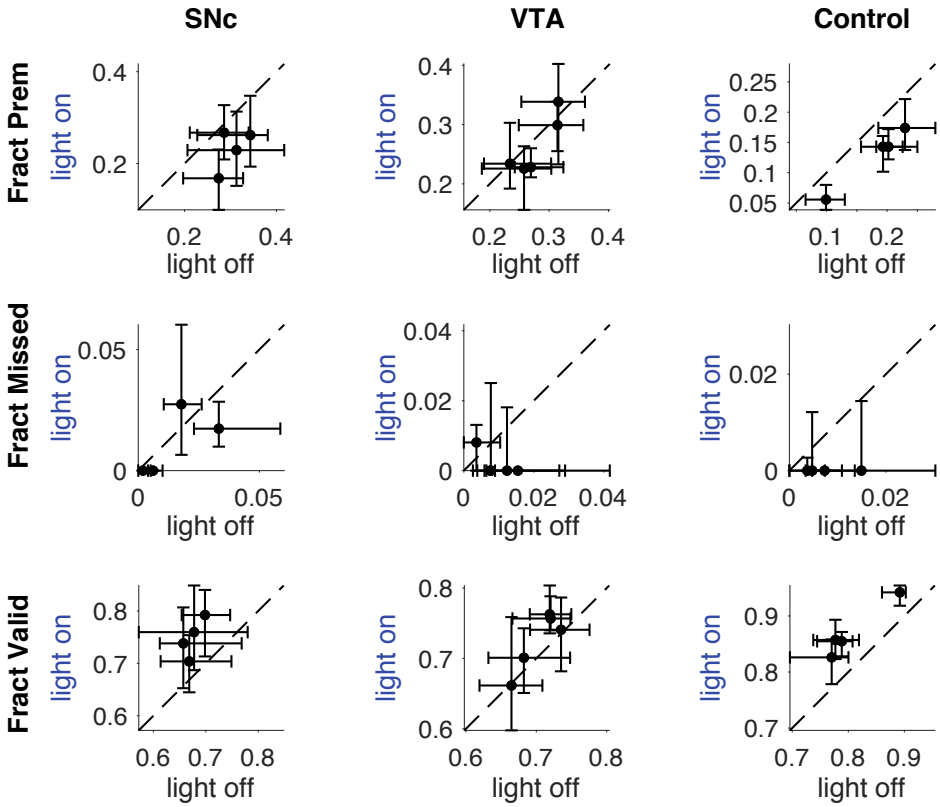


Figure 3.33. **Comparisons of fraction of premature and valid choice types between photoactivated and control trials in SNc, VTA and control groups of mice**. *Top row* - Fraction of premature choices in trials where photoactivation was performed vs when it was not. Mice where the fiber was aimed at the SNc are shown in the left column, the ones targeting the VTA are plotted in the middle column and control animals are plotted in the right column. *Middle row* - Same as *top* row but for fraction of missed choices. *Bottom row* - Same as *top* row but for fraction of valid choices.

observe a relationship between trial-to-trial variability in VTA-DA activity and temporal judgments. Indeed, when we optogenetically activated VTA-DA neurons in an identical manner to the SNc-DA neuron experiments, we did not observe an effect on temporal judgments in the photoactivated trial, nor in the subsequent one (Figure 3.31).

Furthermore, we compared animals' RTs and probability of premature choices, two behavioral variables that could be impacted by DA neuron activation but might not be directly reflected in temporal judgments. We found no systematic differences in RT across stimulation conditions in either SNc or VTA (Figures 3.25 and 3.32). Although we found a decrease in premature choices in photoactivation trials compared to control trials in both SNc and VTA animals, this tendency was also present in light control animals (Figure 3.33). Therefore, we attribute these differences to an artifact caused by the light.

This lack of an effect on temporal judgments by photoactivating VTA-DA neurons was not due to an inability to produce any behavioral effect at the power levels used in the experiment, since we had selected power levels for this experiment with the same criteria used for the SNc experiment (see Section 3.4.11 for details). These criteria involved performing a self-reinforcement assay (Figure 3.34), where light delivery was triggered when mice interrupted one of two available noseports. Mice quickly displayed a

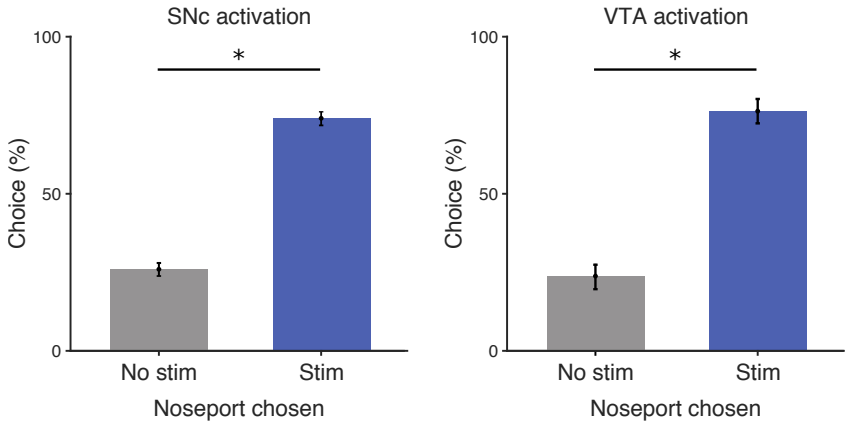


Figure 3.34. **Photoactivation of both SNc-DA and VTA-DA neurons drives self-reinforcement**. *Left* - Mean percentage of choices performed in either the non-stimulated (gray) or the stimulated (blue) noseports for the animals in the SNc group ($n = 4$ mice, $p < 0.005$). Error bars, SEM. *Right* - As in the *left* panel but for mice in the VTA group ($n = 5$ mice, $p < 0.005$).

higher rate of responding at the light-paired noseport compared to the inactive noseport in both SNc-DA and VTA-DA photoactivation experiments. These results indicate that the lack of effect in the temporal discrimination task was not due to an inability to produce a behavioral effect. Rather, these results suggest that VTA-DA neurons, unlike SNc-DA neurons, do not reflect nor control temporal judgments in this timing task, and that the relationship between second tone-evoked activity and temporal judgments is specific to SNc-DA neurons.

3.3 Discussion

Here we demonstrate a direct link between signals carried by midbrain SNc-DA neurons and judgments of elapsed time. Higher or lower levels of SNc-DAergic activity not only correlated with but could directly control timekeeping. Furthermore, the absence both correlative and causal relationships between trial-by-trial variability in VTA-DA neuron activity and temporal judgments suggests that this link between DA neuron activity and temporal judgments in our task is specific to the SNc. We argue that the differences observed here between the RPE signals in SNc-DA and VTA-DA neurons may be due to the distinct input patterns between these two regions. Studies have shown both the encoding of elapsed time in the DS (see results in Chapter 2 and Gouvêa et al., 2015) as well as the requirement of intact DA input to the DS, but not to the NAc, for proper timing behavior (Meck, 2006). Our results are consistent with this work since the DS is an area preferentially and reciprocally connected to DA neurons in the SNc rather than in the VTA (Watabe-Uchida et al., 2012).

The results shown in this Chapter are in agreement with some results of pharmacological manipulations of the DAergic system during timing tasks (Lake & Meck, 2013), but appear at odds with some others that showed accelerated timekeeping with increased DAergic tone (Maricq & Church, 1983; Buhusi & Meck, 2002). However, recent studies demonstrate that many of the pharmacological effects on timing behavior can be explained by the changes in motivation (Balcı et al., 2010; Odum et al., 2002) and attention (Buhusi & Meck, 2002; Ward et al., 2009) that accompany DAergic drug administration (Panigrahi et al., 2015). Indeed, pharmacogenetic DAergic manipulation in our task affected motivated behavior. Variability in the effects of pharmacology on timing may result from its relatively slow time course, which allows for compensation and/or the superposition of multiple distinct behavioral effects. Additionally, many of these pharmacological approaches lacked the spatial resolution to distinguish between the signals carried by VTA-DA and SNc-DA neurons in timing behavior. Our

approach circumvents these issues with genetically targeted, transient manipulations of DA neuron activity.

Our results represent an important addition to the study of differences between signals carried by subpopulations of midbrain DA neurons by identifying a specific relationship between SNc-DA, but not VTA-DA, neuron activity and temporal judgments. Additional studies expanding on the differences between these subpopulations will be fundamental for our understanding of how differential signals are conveyed and transmitted by these neurons during temporal judgments and learning in general.

Recent work has proposed that RPE signals in the mouse VTA reflect a hidden-state inference across time (Starkweather, Babayan, Uchida, & Gershman, 2017). In this study, animals were presented with odor cues that predicted rewards at different delays. A particular odor cue predicted a variable reward delay drawn from a Gaussian distribution, with a 90% probability. In this context, RPE signals in the VTA following reward delivery gradually increased with the delay between cue and reward delivery. The authors propose that animals infer whether they are in a state that can lead to reward delivery or if an uncued transition to an (unrewarded) inter-rial interval (ITI) state occurred. Furthermore, they propose that VTA activity reflects this inference. This hypothesis is unlikely to explain our results for a number of reasons. First, in our task, animals always have direct auditory feedback of the transitions to the ITI, independently of the trial outcome. Additionally, animals in both the SNc and VTA groups are presented with each second tone delay in a relatively uniform manner. For these reasons, a hidden-state inference hypothesis is unlikely to explain the differences between the activity of SNc-DA and VTA-DA neurons observed in this Chapter.

Importantly, we monitored and manipulated the activity of midbrain DA neurons, and not the levels of released DA. The relationship between tonic and phasic firing of DA neurons and DA release is not entirely clear, and is complicated by feedback mechanisms by which released DA can affect the firing of DA neurons (Bunney & Aghajanian, 1978).

Although unexpected, the data presented here may explain existing behavioral data. Situations in which DAergic activity is elevated naturally, such as states of high approach motivation (Bromberg-Martin, Matsumoto, & Hikosaka, 2010), response uncertainty (de Lafuente & Romo, 2011), or cognitive engagement (Fried et al., 2001), are associated with underestimation of time (Gable & Poole, 2012; Coull et al., 2004; Hicks, Miller, & Kinsbourne, 1976). Conversely, situations that decrease DAergic activity, such as when fearful or aversive stimuli are presented (Oleson, Gentry, Chioma, & Cheer, 2012), are associated with overestimation of time (Watts & Sharrock, 1984). These observations, together with our data, suggest that flexibility in time estimation may confer an adaptive

advantage on the individual. For example, underestimating duration in better-than-expected situations may lead to longer engagement in those situations, resulting in even greater reward than if time estimation were not flexible. In other words, there may be a normative explanation for why "time flies when we are having fun" underlying our observation that DA neurons, which are so central to reward processing, exert control over time estimation.

3.4 Materials & Methods

All experiments were approved by the Champalimaud Centre for the Unknown Bioethics Committee and by the Portuguese Veterinary General Board (Direção Geral Veterinária, project approval 014303 - 0420/000/000/2011) and in accordance with the European Union Directive 86/609/EEC.

3.4.1 Animals

Adult (over 2 months old) male and female mice of DAT:Cre and TH:Cre lines (DA-Cre) were used for the experiments presented in this Chapter. We observed no differences in results between the two genotypes and hence combined data across genotypes. Mice were group housed prior to surgical procedures and single housed following surgery. Mice were maintained under water deprivation for all behavioral experiments (>70% body weight from baseline ad libitum period before deprivation).

3.4.2 Behavioral setup

The behavioral box (20 × 17 × 19 cm, model 003102.0001; Island motion corporation), contained 3 nose ports and a speaker. The behavioral box contained 3 front walls (135°angle between the center and the side walls) 2 side walls and a back wall with 90°angle between them. Each of the three front walls contained a nose port equipped with infrared emitter / sensor pair to access port entry and exit times (model 007120.0002; Island motion corporation). The central nose port was set as the trial initiation port, and choices were reported at the lateral nose ports. For correct trials, a 4 μL water reward was delivered using solenoid valves (LHDA1233115H; The Lee Company, Westbrook, CT). Tones were delivered through a speaker mounted on one of the side walls (Neo3 PDRW W/BC; Bohlender-Graebener, Bad Oeynhausen, Germany). All sensors and effectors in the behavioral box were read and controlled using a

microprocessor (Arduino Mega 2560) via a custom circuit board. The task was implemented by the microprocessor, which outputted data via serial a communication port to a desktop running custom Python based software.

3.4.3 Behavioral task

Mice were trained to categorize interval durations as either short or long by performing right and left choices as previously described (Gouvêa et al., 2014, 2015). Briefly, mice self-initiated trials by entering the central nose port, and this action triggered the delivery of a pair of tones (7500 Hz, 150ms) separated by one of 8 randomly selected intervals (0.6, 1.05, 1.26, 1.38, 1.62, 1.74, 1.95 and 2.4s). There was a limited window (∼2-3 weeks) in which to image DA neurons before signal loss. Therefore, with the goal of maximizing data collected for each stimulus, for fiber photometry experiments the 1.38s and 1.62s stimuli were omitted. Trial availability was not signalled to the the animal. Thus, if the animal inserted their snout into the initiation port before the point that a trial became available, nothing happened. After both tones were played, mice reported their judgments by entering one of the two laterally located nose ports. For intervals shorter than a 1.5s category boundary, responses were reinforced at one of the lateral ports. For intervals longer than 1.5s, responses were reinforced at the opposite port. Incorrect responses were followed by a white noise burst (150ms) and a time out (8s). The short/long right/left were counterbalanced across animals. Sessions typically lasted ∼2 hours. Psychometric functions were fit using a 4-parameter logistic function:

$$P(x) = (u - l)\frac{e^{\frac{x-b}{s}}}{1 + e^{\frac{x-b}{s}}} + l \tag{3.1}$$

Where $P(x)$ is the probability of a long choice as a function of stimulus duration, s controls the slope, b the inflection point and l and u control the minimum and maximum values of the curve respectively. Only sessions where performance was $\geq 70\%$ were used for further analysis.

3.4.4 Surgery

Viral injections

All surgeries were performed with mice under isoflurane anesthesia (1-2% at 1 L/min). We stereotaxically targeted the SNc bilaterally for all viral deliveries (from bregma: 3.16 mm posterior, 1.4 mm lateral, 4.2 mm ventral) using an automated microproces-

sor controlled microinjection pipette (Nanoject II; Drummond Scientific Company, Broomall, PA) with micropipettes pulled from borosilicate capillaries (Drummond Scientific Company, Broomall, PA). Injections were performed at 0.2 Hz frequency pulses, with 4.6 nL or 9.2 nL injection volumes per pulse. For all injections, the micropipette was kept at the injection site for 10 min before withdrawal. For experiments targeting the VTA, all viral injections were performed at the following target coordinates (from bregma: 3.3 mm posterior, 0.34 mm lateral, 4.2 mm ventral) with a lateral angle of 10°.

hM4D-mediated pharmacogenetic inhibition For this experiment, we injected 0.5 µL of a AAV2.hSyn.DIO.hM4D(Gi).mCherry virus (titer 4.6 × 1012 vg/mL; University of North Carolina Chapel Hill Vector Core).

Photometry For the photometry experiments, we injected 0.5 µL of a mixture of two virus: AAV1.Syn.Flex.GCaMP6f (titer 1.71 × 1013 gc/mL; University of Pennsylvania Vector Core) and AAV1.CAG.Flex.tdTomato (titer 5.88 × 1012 gc/mL; University of Pennsylvania Vector Core) at a 3:1 ratio.

Photoactivation and photoinhibition For the photoactivation experiments, we injected animals with 1 µL of AAV2/1.EF1a.DIO.ChR2.EYFP (titer 1.11 × 1013 gc/mL; University of Pennsylvania Vector Core). For the photoinhibition experiments, we injected animals with either 1 µL of AAV1.CAG.DIO.ArchT.GFP (titer 4 × 1010 gc/mL; University of North Carolina Chapel Hill Vector Core) or 0.5 µL of AAV2.EF1a.DIO.eNpHR3.0.EYFP (titer 5.7 × 1012 vg/mL; University of North Carolina Chapel Hill Vector Core). Light controls were DA-cre implanted with fibers either with or without 1 µL AAV2/1.EF1a.DIO.EYFP (titer 1.11 × 1013 gc/mL; University of North Carolina Chapel Hill Vector Core) injection. All titers were calculated using qPCR.

Optical fiber implantation

For the photometry and optogenetic experiments, we chronically implanted optical fibers bilaterally in the SNc (from bregma: 3.16 mm posterior, 1.4 mm lateral, 4 mm ventral for photometry, 3.9 mm ventral for optogenetic experiments) or in the VTA (from bregma: 3.3 mm posterior, 0.35 mm lateral, 4.2 mm ventral for photometry and optogenetic experiments) a lateral angle of 10°. Analgesia was administered intraperitoneally following each surgery (buprenorphine, 0.1 mg/kg).

3.4.5 CNO administration for hM4D mediated inactivation

CNO (10 mg/kg, Tocris Bioscience, Bristol, UK) and saline solutions (0.05-0.1 mL volume) were injected intraperitoneally, 20 min prior to session initiation. We tested for a nonspecific CNO effect by repeating the protocol on a separate group of animals that had not been injected with the AAV2-hM4D(Gi) virus.

3.4.6 Fiber photometry setup

Two lasers (473 nm and 556 nm; Changchun New Industries Tech. Co. Ltd., Changchun, China) were coupled to individual patchcords (100 µm core diameter, 0.22 NA; Doric lenses) and connected to an individual collimator adapter each (EFL 4.5 mm, NA 0.50; Doric lenses) and a neutral density filter (NE30B; Thorlabs), mounted on the main setup unit. Three dichroic mirrors (Di01-R561-25x36 and FF552-Di02-25x36, Semrock; T495LP, Chroma) were fixed inside the main unit, allowing for 473 nm and 556 nm light delivery and GCaMP6f and tdTomato fluorescence detection. The 473 nm and 556 nm light was coupled into a patchcord (200 µm core diameter, 0.48 NA; Doric lenses) using a lens (EFL 4.5 mm, NA 0.50) and a rotatory joint (Doric lenses). The patchcord was mated to one of two chronically implanted optical fibers (200 µm core diameter, 0.48 NA; Doric lenses). Laser intensities at the patchcord tip, before mating to the chronically implanted fiber, were ∼20 µW. For detection of GCaMP6f fluorescence, light was collected by the lens, transmitted and reflected by the dichroics before final filtering and focusing (bandpass: ET525/50m; Chroma, lens: EFL 25.4mm, model LA1951 - A; Thorlabs) into a photodetector (PDF10A/M; Thorlabs). For detection of tdTomato fluorescence, light was collected by the lens and transmitted through all dichroics before final filtering and focusing (bandpass filter FF01-641/75-25; Semrock) into a second photodetector. Photodetector output was digitized at 1 kHz (NI PCIe-6351 board via NI BNC-2120) and recorded using custom software in Bonsai (Lopes et al., 2015). For each mouse, the hemisphere being recorded was alternated daily.

3.4.7 Fiber photometry data analysis

All data analysis was performed with custom MATLAB software (Mathworks). Raw photometry traces were filtered in the frequency range of 0.2 Hz - 15 Hz. For each session, $\Delta F/F$ was calculated on a trial by trial basis for both GCaMP6f and tdTomato traces as (F - F0) / F0, where F0 was calculated as the mean across all trials of the mean fluorescence values within a window of -4 s to -2s before each trial initiation. Robust

regression using the GCaMP6f and tdTomato ΔF/F was performed and the coefficient estimates were used to calculate a predicted GCaMP6f ΔF/F based on the observed tdTomato ΔF/F. This predicted GCaMP6f ΔF/F was then subtracted to the observed GCaMP6f ΔF/F to calculate the corrected ΔF/F. When pooling data across animals, each session's corrected ΔF/F was z-scored using the mean and SD across the entire session. For splitting trials according to DAergic activity, for each session and interval duration, distributions were built using each trial's mean signal between 300ms and 500ms after second tone onset, after linearly regressing out correlations with a pre-first tone response baseline window spanning -1 to 0.2s relative to first tone onset. These distributions were then split into terciles. Separability of DA responses between trials falling into the upper and lower terciles of the second tone response were assessed by computing the area under the ROC curve at each time point throughout the trial.

3.4.8 RPE model

A simple linear RPE model was fit to the DA response and predicted the amplitude of a RPE at the time of the second tone:

$$D = _{WP}\ P + _{WS}\ S + const \tag{3.2}$$

Where D is predicted DA response. P is average expectation of reward which was determined experimentally, i.e. the measured performance at each interval. S is temporal surprise, i.e. the inverse subjective hazard rate (Figure 3.10) (Janssen & Shadlen, 2005):

$$S = \frac{1}{H} \tag{3.3}$$

To calculate H, we began by calculating the subjective expectation of each individual stimulus, modeled as a Gaussian with mean at each stimulus duration and $SD = 0.15$s. We then summed these subjective expectation Gaussians which is represented as E, subjective expectation of interval offset. We then calculated the subjective hazard rate, H_t, that is:

$$H_t = \frac{E_t}{1 - E_t} \tag{3.4}$$

The W's are the respective weights, where $w \geq 0$ and was determined using MATLAB's global optimization toolbox. The predicted DA neuron response accounted for 90% of variance in the average data recorded in the SNc and 46% in the average data

recorded in the VTA. The exact ratio of weights $S/(P+S)$ varied across recordings and brain area. This ratio was on average 0.52 for the SNc data set, indicating that S and P were equally represented in the pattern of SNc-DA neuron activity. The ratio of weights $S/(P+S)$ for the VTA data set was on average 0.14 indicating that P was mostly represented in the pattern of SNc-DA, with a lower contribution of S.

3.4.9 Trial-by-trial prediction of choice from the dopamine response

To estimate the probability of animals performing a long choice (P_{Long}) given DAergic neuron response to the second tone, we used logistic regression. The log odds ratio was modeled as:

$$ln\frac{P_{Long}}{1-P_{Long}} = B0 \cdot \mathrm{DA}_{\mathrm{response}} + B1 \tag{3.5}$$

Where $B0$ and $B1$ are the logistic fit parameters (slope and intercept respectively) and DA $_{\mathrm{response}}$ is the response of DAergic neurons to the second tone. The P_{Long} was recovered by exponentiating both sides of the above equation and solving for P_{Long}:

$$P_{Long} = \frac{1}{e^{-t}}, t = B0 \cdot \mathrm{DA}_{\mathrm{response}} + B1 \tag{3.6}$$

3.4.10 Electrophysiological recordings

Mice under isoflurane anesthesia (1-2% at 1 L/min). A 480 nm LED (Doric lenses) or a 473 nm laser (Laserglow Technologies, Toronto, Canada) were used as light source. An 'optrode' was created: optical fiber (200 µm core diameter, 0.48 NA; Thorlabs) was glued to a platinum-iridium microelectrode (catalog #UEPLEESE1N2M; FHC, Bowdoin, ME), the tip of which was extended ∼400 µm from the tip of the fiber. Light power at the tip of the optical fiber was 4 mW for activation and 15 mW for inhibition. The optrode targeted the SNc (from bregma: 3.16 mm posterior, 1.4 mm lateral, 4.2 mm ventral) for simultaneous SNc photoactivation and recording. Square light pulses were delivered at a 10 Hz frequency for 3s, with a 2s interval between each stimulation onset. Extracellular signals were amplified (model 4000, A-M Systems), high-pass filtered at 100 Hz and digitized at 32 kHz. A custom MATLAB (Mathworks) based software was used to record and analyze extracellular signals and to control the LED or the laser via a NI PCIe-6351 board, NI SCB-68 and NI BNC-2090 accessories (National Instruments Corporation, Austin, TX).

3.4.11 Optogenetic manipulations during task performance

A laser (473 nm for activation and 556/596 nm for inhibition (Laserglow Technologies, Toronto, Canada) was coupled to a patchcord (200 nm core, 0.48 NA; Doric lenses), connected to a splitting patchcord (200 nm core, 0.48 NA; Doric lenses) for bilateral light delivery through a rotatory joint (FRJ 1×1; Doric lenses) into two chronically implanted optical fibers (200 nm core diameter, 0.48 NA; Doric lenses). Prior to manipulations using light during timing behavior, we performed a self-reinforcement assay in each of the ChR2 injected animals (both in the SNc and VTA groups). The goals for this assay were: i) have a positive behavioral effect and ii) to define the power levels to use in the timing task manipulation. Animals were placed in a behavioral box similar to the one used for training in the timing task but with only two available nose ports. Interrupting one of the nose ports (active port) would elicit a 500ms pulse train of 20ms individual pulses delivered at 25HZ using the 473 nm laser. Interrupting the other nose port (inactive port) was inconsequential. In both SNc and VTA animals, we defined the minimum power that elicited a preference for the active port versus the inactive port and used that power level for the optogenetic activation during task performance. For optogenetic manipulations in the context of the timing task, laser stimulation was delivered randomly in 30% of trials. Starting upon trial initiation, 20ms duration square pulses were delivered at a 10 Hz frequency for 3s for photoactivation and a continuous square pulse of 3s for inhibition experiments. Laser power at the tip of the fiber was on average 4 mW , 15 mW and 40 mW for the 473 nm, 556 nm and 596 nm lasers, respectively. Control animals were tested with the same protocol with an average laser power of 5 mW for the 473 nm laser and 48 mW for the 596 nm laser.

3.4.12 Immunohistochemistry and microscopy

Histological analysis was performed after all experiments to confirm optical fiber placement. Mice were anesthetized with pentobarbital (Eutasil, 50 mg/kg intraperitoneally; CEVA Sante Animale, Libourne, France) and perfused transcardially with 4% paraformaldehyde (P6148, Sigma-Aldrich). The brains were removed from the skull, stored for 24 hours in 4% paraformaldehyde, and then kept in PBS for ~ 1 week until sectioning. A vibratome (VT1000S, Leica Microsystems, Wetzlar, Germany) was used to section the brain into 50 µm thick slices that were then immunostained with antibodies to tyrosine hydroxylase (22941, 1:5000, ImmunoStar) to visualize DAergic neurons and 49,6-diamidino-2-phenylindole (DAPI, D9542, Sigma-Aldrich) to visualize nuclei. For the fiber photometry experiments, slices were additionally immunostained

with an antibody to green fluorescent protein (GFP) (A-6455, 1:1000, Invitrogen) to amplify the fluorescent signal from cells expressing GCaMP6f. Images were acquired with a confocal microscope (Zeiss LSM 710, Zeiss) using a 20X objective. Using stereological procedures (StereoInvestigator, 100 µm^2 windows within the SNc, spaced by 150 µm), we quantified the number of GCaMP6f expressing neurons in the SNc (480 neurons, $n = 2$ mice). We confirmed high levels of co-localization between GCaMP6f and tdTomato (90% in DAT-Cre, 84% in TH-Cre) as well as high degree of specificity of GCaMP6f expression in DA neurons (82% in DAT-Cre, 75% in TH-Cre).

3.4.13 Statistics and model comparison

To select the model that best explained differences in the pattern of the subjects' judgments between conditions (i.e. photometry: high, medium, low DA; photostimulation: light, control) we fit models based on a 4 parameter logistic psychometric function described in Equation 3.1. We tested models which allowed all possible combinations of parameters to vary. That is, in a given model, between 0 and 4 parameters were allowed to vary between experimental conditions. For each model, we calculated the maximized value of the likelihood function (L) and the *BIC*. Models with the lowest *BIC* value are presented in Section 3.2 as lower *BIC* values are evidence for the model.

$$L(\theta \mid x) = P(x \mid \theta) \tag{3.7}$$

Where x are the observed probabilities of a long choice and Θ are the parameters of the model (slope, bias, offset and amplitude).

$$BIC = -2 \cdot ln(L) + k \cdot ln(n) \tag{3.8}$$

Where k is the number of free parameters to be estimated and n is the number of data points in the experimental data. Statistical significance was determined using Wilcoxon sign rank and rank sum tests where appropriate, unless otherwise stated.

Acknowledgments

The experiments presented in this Chapter were designed by Sofia Soares, Bassam Atallah and Joseph Paton. Sofia Soares and Bassam Atallah acquired and analyzed the data. We thank Alessandro Braga for assistance with behavioral training, Gonçalo Lopes for assistance with Bonsai and Margarida Duarte for assistance with mouse colonies. We

thank Tiago Monteiro, Thiago Gouvêa, other members of the Paton laboratory, Brian Lau, Eran Lottem, Masayoshi Murakami, Cindy Poo, Alfonso Renart and Thomas Akam for discussions and/or comments on versions of the manuscript on which part of this Chapter is based (Soares et al., 2016). We thank Zachary Mainen for support. We thank the histopathology and vivarium staff from the Champalimaud Scientific and Technological Platforms for support and animal care.

V. Jayaraman, R. A. Kerr, D. S. Kim, L. L. Looger, and K. Svoboda from the GENIE (Genetically-Encoded Neuronal Indicator and Effector) Project at the Howard Hughes Medical Institute's Janelia Farm Research Campus for providing the AAV-GCaMP6f through the University of Pennsylvania Vector Core. Viruses for expression of NpHR3.0 and EYFP are available from the University of North Carolina Vector Core under a material transfer agreement with K. Deisseroth. Viruses for expression of GCaMP6f and TdTomato are available from the University of Pennsylvania Vector Core under a material transfer agreement with the trustees of the University of Pennsylvania on behalf of J. Wilson.

This work was supported by the following institutions: Champalimaud Foundation; Portuguese FCT (Fundação para a Ciência e a Tecnologia, SFRH/BD/51895/2012 to Sofia Soares); European Molecular Biology Organization (Advanced Long Term Fellowship 983-2012 to Bassam Atallah); Marie Curie Actions (FP7-PEOPLE-2012-IIF 326398 to Bassam Atallah); Bial Foundation (188/12 to Joseph Paton) and Simons Foundation (Simons Collaboration on the Global Brain award 325476 to Joseph Paton).

Chapter 4

Task engagement signals in the basal ganglia

Animals' choices are a complex function of sensory evidence, prior experience and internal states. Although psychophysics has been crucial to the study of how judgments relate to stimulus difficulty, much less is known about the influence of internal states such as task engagement on animals' performance. Here, we trained mice on the same temporal discrimination task described in Chapter 3 and used a generalized linear model (GLM) to identify behavioral and task-related stimuli that predicted performance and premature choices. We also recorded the activity of midbrain dopamine (DA) neurons during task performance and found that activity in a period prior to trial initiation (pre-trial) was predictive of animals' performance: the higher the pre-trial DA activity, the worse the performance. Remarkably, pre-trial DA activity was also higher when animals responded prematurely than when they waited to complete a trial, indicating that DA pre-trial activity reflected not simply upcoming performance levels, but a more general state of task engagement. Surprisingly, similar signals (albeit of opposite sign), were observed in a number of neuronal types in the striatum, arguing for a global role of the basal ganglia (BG) in task engagement. Adding this pre-trial signal to the GLM improved predictions of both performance level and premature behavior, confirming the relationship between this signal and choice behavior. Together, these results suggest that two major BG components reflect a task engagement state that, together with current sensory evidence, predicts animals' choices.

4.1 Introduction

Animals integrate sensory evidence in their environment to guide behavior. To carefully study the neural circuits that underlie decision-making, experimenters often turn to psychophysics to describe how decisions depend on sensory input (Parker & Newsome, 1998; Carandini & Churchland, 2013). The power of this approach is that changes in psychometric curve parameters can be used to infer how measured or imposed neural activity patterns relate to internal decision variables such as sensitivity or bias (Parker & Newsome, 1998; Carandini & Churchland, 2013). In such studies, experimenters often begin by overtraining animals on a given task to reach a stable performance criterion before any measurement or manipulation. Still, performance in overtrained animals is somewhat variable, sometimes even for relatively 'easy' stimuli far from the category boundary (e.g., Brunton, Botvinick, & Brody, 2013; Znamenskiy & Zador, 2013). Additionally, despite this overtraining, animals sometimes fail to maximize reward by not adhering to task rules, as evidenced by impulsive decisions before the termination of the stimulus to be discriminated or lack of trial completion. Thus, an open question remains as to which factors, other than stimulus difficulty, affect choice behavior on psychophysical tasks.

Beyond task-relevant sensory stimuli, behavior is strongly influenced by an animal's internal state. A range of studies have addressed the relationship between states such as arousal, motivation or engagement, and behavior (Anderson, 2016). Arousal is generally related to an increase in physiological, autonomic and motor activities (such as transitions from sleep to wakefulness). Therefore, arousal can be seen as a relatively global regulator of behavior (Anderson, 2016).

Motivation is described as a combination of homeostatic drive states (such as thirst) and states driven by external stimuli (e.g., states driven by reward-predicting cues) that promote goal-directed behaviors (Berridge, 2004; Liljeholm & O'Doherty, 2012; Anderson, 2016). For example, one factor that impacts animals' motivation is the degree of association between sensory cues and previous rewards. Furthermore, satiety levels contribute to changes in motivation, and consequently in performance (Berditchevskaia, Cazé, & Schultz, 2016) and goal-directed behaviors in general, as evidenced by contingency devaluation paradigms (Balleine & Dickinson, 1998; Dias-Ferreira et al., 2009).

Task engagement is another example of an internal state that relates to task performance (Manly, Robertson, Galloway, & Hawkins, 1999; Robertson, Manly, Andrade, Baddeley, & Yiend, 1997; Smallwood et al., 2004; Smallwood, McSpadden, & Schooler, 2007; Cheyne, Solman, Carriere, & Smilek, 2009). Although many definitions of task

engagement exist (in many cases relating it to sustained attentional processes), they share the idea that a core feature of disengagement consists of *"reduced allocation of attentional resources to environmental task-related stimuli"* (Cheyne et al., 2009). This definition mainly derives from studies using human subjects performing tasks similar to Go/No-Go tasks (Robertson et al., 1997; Manly et al., 1999; Smallwood et al., 2007), where performance can be complemented with verbal reports to identify trials where subjects experience *"task unrelated thoughts"* (Smallwood, Baracaia, Lowe, & Obonsawin, 2003; Smallwood et al., 2004; Cheyne et al., 2009). Additionally, a number of studies have identified behavioral correlates of task engagement, providing a bridge for the study of this internal state in laboratory animals. Subjects experiencing low levels of task engagement exhibit a general decrease in performance, an increase in anticipatory/premature behavior, and an increase in choice omission (i.e., failures to report the occurrence of a stimulus; Manly et al., 1999; Smallwood et al., 2004; Cheyne et al., 2009). In this Chapter, our use of the term task engagement/disengagement is based on these ideas, and takes advantage of these behavioral correlates. Although some neural correlates of arousal (Sakurai, 2007; Taylor et al., 2016) and motivation (Mogenson, Jones, & Yim, 1980; Wise, 2004) have been identified, less is known about the neural mechanisms underlying task engagement. How might task engagement be regulated by the brain to influence behavior?

Midbrain dopamine (DA) neurons have been shown to play an important role in motor control (Carlsson et al., 1957), reinforcement learning (Schultz et al., 1997) and timing behavior (Meck, 2006), but they have also been implicated in a broad range of internal states including arousal (Taylor et al., 2016) and motivation (Mogenson et al., 1980; Wise, 2004). Additionally, evidence from attention deficit hyperactivity disorder (ADHD) suggests that DA neurons may also play a role in states beyond arousal and motivation (Levy, 1991). For example, drugs that target the DAergic system are routinely used to treat ADHD patients (Castle, Aubert, Verbrugge, Khalid, & Epstein, 2007), alleviating the symptoms of distractibility, hyperactivity, impulsivity, poor concentration and poor performance on cognitive tasks (Levy, 1991; Wodka et al., 2007; Noreika et al., 2013; Gmehlin et al., 2014). However, the details of how DA neuron activity might relate to these symptoms, and to a general state of task engagement, remains unknown.

In this Chapter, we investigated the behavioral and neural correlates of performance in a psychophysical task, using an approach we believe is suited to parse the relative contributions of task-relevant stimuli, and behavioral variables that correlate with different internal states, to performance. Throughout this Chapter, the term performance

is used as a measure of behavioral sensitivity, interchangeably with the measure of probability of a correct choice across all valid trials. First, to better understand the variables that may influence behavior in this task, we used a generalized linear model (GLM) to model performance using solely task-relevant stimulus information and behavioral variables. We found that current stimulus difficulty and response time (RT) were the best predictors of performance, with minimal effects of trial history and movement-related variables. We used a similar approach to predict the probability of a premature choice, and found that a history of premature choices was predictive of impulsive behavior in the current trial. We then probed the relationship between DA neuron activity in both the substantia nigra pars compacta (SNc) and the ventral tegmental area (VTA) and behavior (either performance or premature choices) in this task using fiber photometry.

We identified a period preceding trial initiation (pre-trial) where activity in both the cell bodies and axonal projections of DA neurons was predictive of performance. Additionally, we found that premature choices and missed trials were also predicted by this pre-trial DA signal, suggesting that DA neurons contribute to a broader state of task engagement. In addition, we found that a similar relationship between neuronal activity and task engagement was present in different neuronal types in the striatum, arguing for a global role of basal ganglia (BG) circuits in task engagement.

Finally, we found that adding these neural signals to our GLM improved predictions of both performance and premature behavior, confirming that pre-trial activity in striatal and midbrain DA neurons predicts how likely an animal is to adhere to task rules, make correct decisions, and ultimately maximize rewards.

4.2 Results

4.2.1 A generalized linear model predicts correct and premature choices

We began by training mice on the same temporal discrimination task described in Chapter 3 (see *left* and *middle* panels of Figure 3.1), where the delay between two tones was categorized as either shorter or longer than a 1.5s category boundary (Gouvêa et al., 2014, 2015). Trials were self-initiated when the mouse interrupted a central nose port and long/short choices were reported at a left/right located noseport. We imposed an inter-rial interval (ITI) period, after which a new trial became available for initiation. The requirement for mice to self-initiate trials ensured a certain level of engagement in task performance. Even so, we observed that animals' performance was variable on a

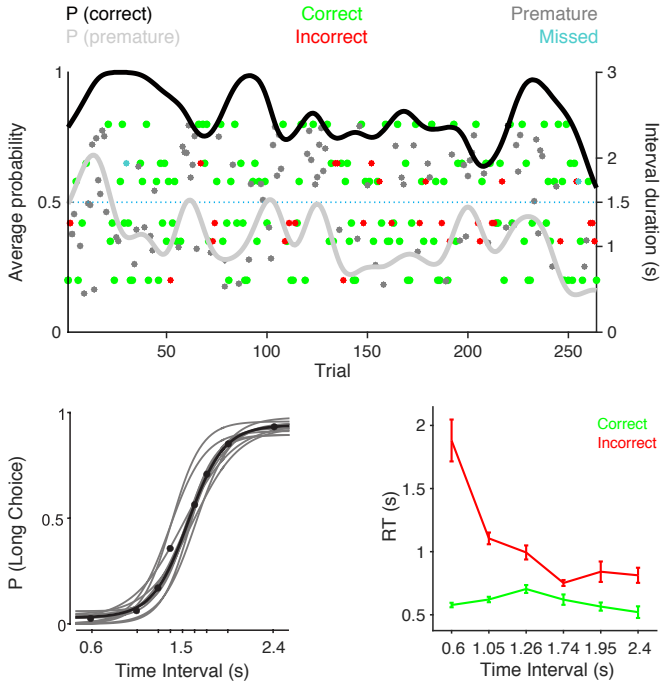


Figure 4.1. **Variable performance in a temporal discrimination task.** *Top* - Single session example. Dots represent single trials and solid lines illustrate the running average of the probability of either a correct choice (black) or a premature choice (gray). *Bottom left* - Same as *middle* panel of Figure 3.1 illustrating a logistic function fit to the daily (gray) and average (black) performance of an example mouse (10 sessions). *Bottom right* - Mean RT across 15 sessions of an example animal, split by correct or incorrect choices (green or red traces, respectively). Error bars, standard error of the mean (SEM).

trial-by-trial basis (Figure 4.1, *top* panel, black curve). An additional layer of variability was observed in cases where animals failed to adhere to task rules. There were three general types of trials in this task: *valid*, *premature* and *missed*. If a mouse waited for the second tone to report its choice, the trial was labeled as *valid* and the choice could be rewarded or not depending on whether it was correct or incorrect (Figure 4.1, *top* panel, green and red dots, respectively). If the animal displayed impulsive-like behavior and did not wait for the second tone before choosing one of the lateral ports, the trial was labeled as *premature* and was not rewarded (Figure 4.1, *top* panel, gray dots and gray solid line). Finally, in a minority of trials, animals initiated a trial but failed to

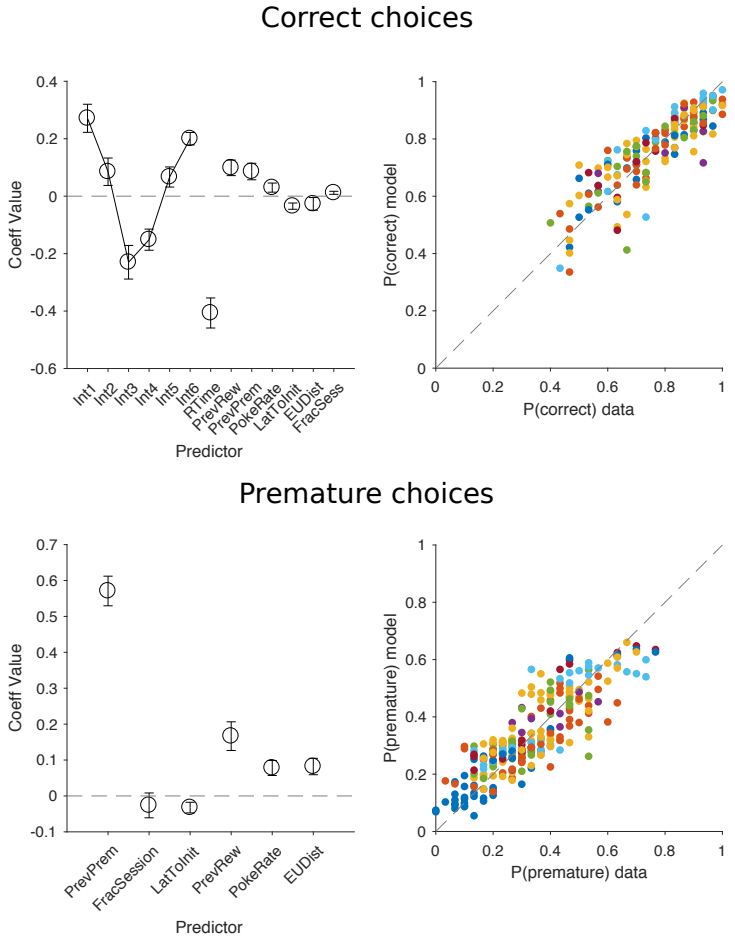


Figure 4.2. **A generalized linear model predicts correct and premature choices**. *Top left* - For each mouse, a GLM was built to predict correct choices. The model was cross-validated (see Section 4.4.5 for details) and the mean of each coefficient value across all cross-validations was calculated per animal. Dots represent the mean across mice of the coefficient values for each predictor variable used in the GLM ($n = 10$ mice). Error bars, SEM. *Top right* - Probability of a correct choice [P(correct)] predicted by the model versus P(correct) in the data. Each color represents a single animal (10 mice). Each dot of the same color represents the mean P(correct) calculated in a bin of 30 trials. For this binning, single trials were sorted in ascending order of model prediction values. *Bottom* - Same as panels on *top* but for GLMs predicting premature choices, using distinct predictors (see x axis in *bottom left* panel).

perform a choice within 5s of the second tone being played (Figure 4.1, *top* panel, blue dots). These trials were labeled as *missed* and were also not rewarded.

In this task, we observed trial-by-trial variability in the probability of correct and premature trials. We sought to identify the factors that contribute to this variability. We started by assessing performance as a function of stimulus identity, since the current stimulus should be the main factor determining choice for optimal performance. As described in Chapter 3, the probability of a long choice was well described by a psychometric function (Figure 4.1, *bottom left* panel): performance was nearly optimal on the shortest and longest intervals, but choices became more variable as the interval durations approached the category boundary. In addition, we found that the RT, or the time between the second tone presentation and choice, varied systematically with stimulus and outcome (Figure 4.1, *bottom right* panel). RTs were shorter for correct trials when compared to incorrect trials.

To identify variables that related to performance on a trial-by-trial basis, and to quantify their relative contributions to performance, we built a GLM for each animal (see Section 4.4.5 for details), in which we estimated animals' performance based on eight predictors (Figure 4.2, *top left* panel, x axis). The first obvious candidate was the stimulus identity on the current trial, since performance depends on stimulus difficulty (Figure 4.1, *bottom left* panel). The second predictor in this model was the RT, a variable that also covaried with performance (Figure 4.1, *bottom right* panel). The third and fourth predictors were related with the history of past rewards: namely the reward outcome and premature behavior on the previous trial. The fifth, sixth and seventh predictors described animals' behavior in a pre-trial period (the 2 seconds leading up to trial initiation): the rate of nose poking in the initiation port, the latency to initiate a trial (the time from trial availability to trial initiation), and the euclidean distance between the animal's trajectory in the pre-trial period of each trial and the mean trajectory in this period across trials. The poke rate and latency to initiate trials can be seen as proxies for an animal's motivation to initiate a trial, and the distance to the mean trajectory is used here as a metric for motor output consistency in this pre-trial period. Lastly, the eighth predictor of the model corresponded to the fraction within a session where each trial occurred, since animals' performance could vary as animals become satiated across a session.

We found that in all animals the current stimulus identity and the respective RT were the best predictors of performance (Figure 4.2, *top left* panel). Other predictors (including past rewards/choices) contributed much less to changes in performance than the evidence presented in the current trial, and in some animals were even excluded

from the model by regularization (see Section 4.4.5). Our GLM predicted performance above chance levels (Figure 4.2, *top right* panel, see Section 4.4.5).

These results indicate that although there was a small influence of trial history and other factors on performance, mice mostly relied on the current stimulus identity to solve the task, which is the strategy they should adopt in order to maximize rewards. In spite of this, animals did not always adhere to task rules, as evidenced by the existence of premature and missed trials. Missed trials were harder to model due to their rare occurrence, but premature trials were more common. Therefore, we built an additional GLM per animal to predict premature choices using the subset of predictors from the model described above that were applicable for a premature choice scenario (Figure 4.2, *bottom left* panel, x axis). We found that the best predictor of a premature response was a premature choice on the previous trial, outweighing the effects of previous rewards, movement or motivation-related predictors (Figure 4.2, *bottom left* panel). As was the case for predictions of performance, this GLM also predicted premature choices above chance levels (Figure 4.2, *bottom right* panel).

4.2.2 Midbrain DA neuron pre-trial activity reflects task engagement

DA neurons have been implicated in a range of internal states, such as arousal (Taylor et al., 2016), motivation (Mogenson et al., 1980; Wise, 2004) and attention (Levy, 1991). We therefore turned to the midbrain DAergic system and began by recording the calcium activity of DA neurons in both the SNc and the VTA during task performance, using fiber photometry as described in Chapter 3. Interestingly, we observed a high level of variability in DA neuron activity during the pre-trial period, and we therefore asked how this variability related to performance in the upcoming trial. To answer this question, we split trials into terciles of pre-trial DA neuron activity in a similar fashion as described in Chapter 3, but now using a 2 second window before trial initiation as our window of interest. This allowed us to identify trials in which pre-trial DA neuron activity was high, medium or low for each session (Figure 4.3). We then selected only valid trials, and plotted the probability of a long choice as a function of interval duration for trials in the high (purple), medium (black) or low (green) tercile (Figure 4.4, *left* panel). Surprisingly, we found that pre-trial DA neuron activity was inversely related with performance: the higher the activity, the poorer the performance. This result was robust across animals and indicates that the activity of DA neurons prior to trial initiation is predictive of animals' performance in the upcoming trial.

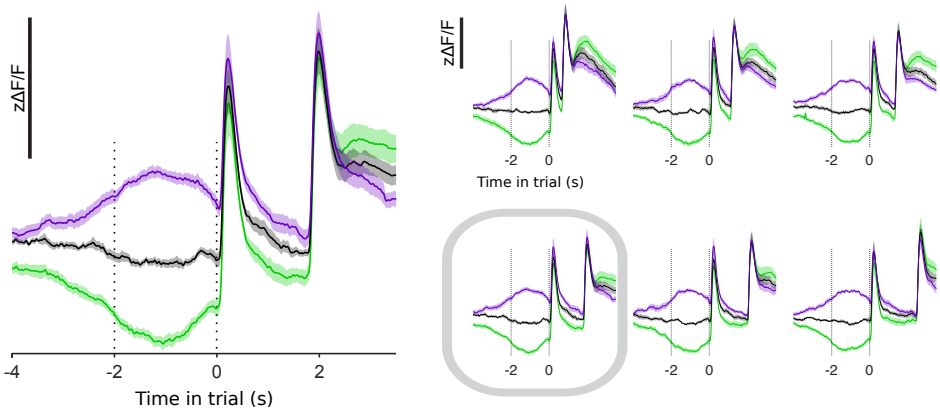


Figure 4.3. **Pre-trial DAergic activity is variable**. *Left* - Average DA neuron response in each tercile of pre-trial activity for the 1.74s interval stimulus (10 mice). Shading, SEM. Low, middle and upper terciles are color coded in green, black and purple, respectively. This color code will be kept throughout this Chapter. *Right* - As in *left* but for all six interval durations; data shown on the *left* are outlined in gray.

We observed the same relationship between DA pre-trial activity and performance independently of whether activity was recorded in the SNc or in the VTA (Figure 4.4, *right* panel), suggesting that the relationship between midbrain DA neuron activity and performance is a global functional signal carried by subpopulations of midbrain DAergic neurons with distinct projection targets. This raised the question of whether this relationship was present in the main output areas of the SNc and VTA, respectively, the DS and the NAc. To answer this question we again expressed GCaMP6f in midbrain DA neurons, but now aimed optical fibers at either DS or NAc (Figure 4.5) and recorded the activity of DAergic projections in these areas. We observed that the relationship between pre-trial DA activity and performance was indeed present at DAergic terminals in both DS and NAc (Figure 4.4, *right* panel), demonstrating that this pre-trial performance-related activity is a global signal that reaches multiple output target areas of midbrain DA neurons.

The data shown so far indicates that pre-trial DA activity is inversely related to animals' performance. However, we wondered whether this relationship might reflect a more general state of task engagement. We operationally define the internal state of task engagement as one where animals initiate and complete trials with high levels of performance and low levels of impulsivity. Conversely, low task engagement is characterized by either poor performance, impulsive behavior (high probability of premature

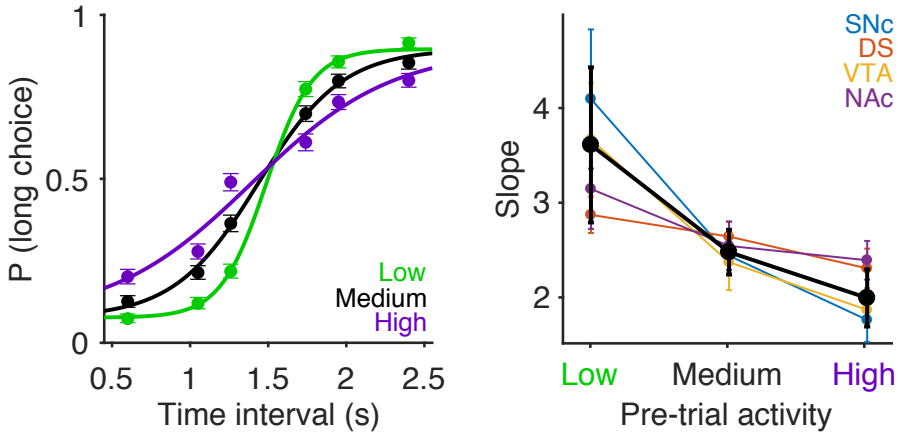


Figure 4.4. **Pre-trial DAergic activity negatively correlates with performance**. *Left* - Psychometric curves constructed using trials from each tercile of DA neuron activity [combining data from animals where optical fibers were aimed at SNc, VTA, dorsal striatum (DS) and nucleus accumbens (NAc)]. Curves are the maximum-likelihood fits of logistic functions with the lowest bayesian information criterion scores ($n = 24769$ trials, 10 mice). Error bars, 95% confidence interval (CI). *Right* - Slope (the $log\frac{1}{s}$, where s is the parameter that controls the slope of the psychometric curve, see Equation 4.1 in Section 4.4.3) is plotted for each tercile of DA activity in each data group of animals where optical fibers were aimed either at the cell bodies in SNc ($n = 7$ mice) or VTA ($n = 4$ mice), or at their main projection sites in DS ($n = 3$ mice) or NAc ($n = 3$ mice).

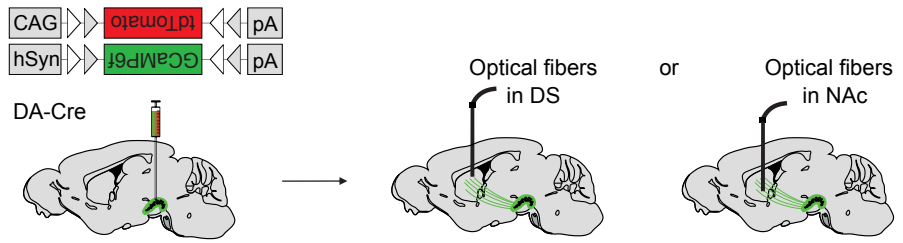


Figure 4.5. **Schematic of the photometry surgical procedure targeting DAergic projection sites**. In a subset of animals, at the same time as SNc and/or VTA adeno-associated virus (AAV) injections and optical fiber implantation, we additionally implanted optical fibers aiming at the dorsal or ventral striatum. After 1-2 weeks of recovery, mice were retrained on the temporal discrimination task and the Ca^{2+} activity of DAergic neurons was recorded during task performance.

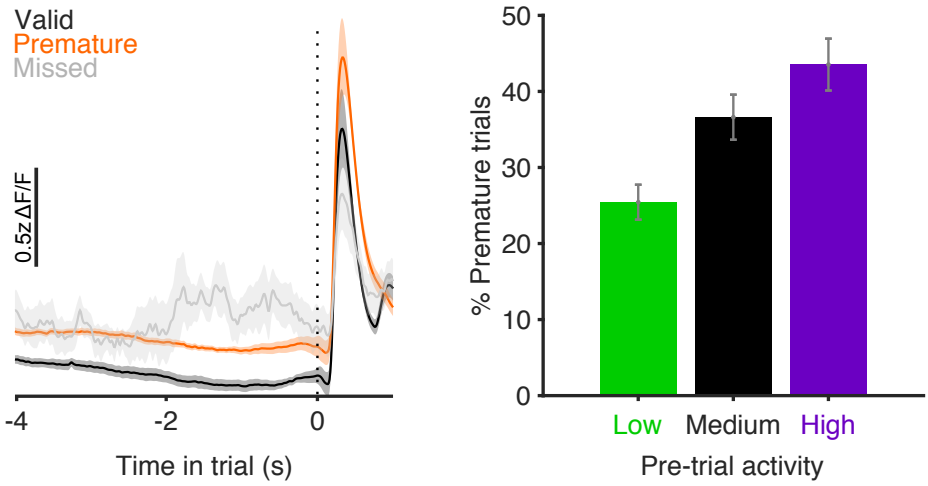


Figure 4.6. **Pre-trial DAergic activity is directly correlated with premature choices and missed trials**. *Left* - Mean DAergic activity aligned on trial initiation, split by the outcome that will follow seconds later (valid trials in black, premature trials in orange and missed trials in gray). Shade, SEM ($n = 10$ mice). *Right* - The mean percentage of premature choices over all trials in each tercile of DA neuron activity. Error bars, SEM ($n = 10$ mice).

trials) or even lack of trial completion (missed trials). If DA neurons reflect this general state, and given the inverse relationship between pre-trial DA neuron activity and performance, we hypothesized that pre-trial DA activity should be higher in premature and missed trials than in valid trials. We tested this hypothesis by plotting DA activity in valid, premature and missed trials, and found that pre-trial DA neuron activity was elevated in premature and missed trials when compared to valid trials (Figure 4.6, *left* panel). Consistent with this result, we found that the higher the pre-trial DA neuron activity, the higher the probability of premature trials (Figure 4.6, *right* panel). These results suggest that pre-trial DA neuron activity relates to more than simply the probability of an animal performing trials correctly, but reflects a broader state of task engagement, predicting the probability of both adherence to task rules and accurate performance.

4.2.3 Pre-trial activity in dMSNs and iMSNs also reflects task engagement

The results presented in the previous Section indicate that pre-trial midbrain DA activity reflects a state of task engagement. Given the reciprocal connectivity patterns between striatal and midbrain DA neurons, we hypothesized that a similar pre-trial signal might be present in striatal neurons. To answer this question, we recorded the calcium activity of medium spiny neurons (MSNs), a γ-aminobutyric acid (GABA)ergic population of projection neurons that represents $> 95\%$ of striatal neurons. These neurons are subdivided into direct pathway MSNs (dMSNs) and indirect pathway MSNs (iMSNs), depending on whether their axonal projections directly target the substantia nigra pars reticulata (SNr) and the internal segment of the globus pallidus (GP), or if they do so indirectly through the subthalamic nucleus (STN) and external segment of the GP (Kreitzer, 2009). These two subpopulations can be genetically targeted since dMSNs mainly express D_1 type DA receptors, while iMSNs mainly express D_2 type DA receptors and the adenosine A_{2A} receptor (Kreitzer, 2009). We used the same Cre-mediated viral expression as described above, but used transgenic mice that express Cre specifically in either dMSNs or iMSNs. This approach allowed us to record the calcium

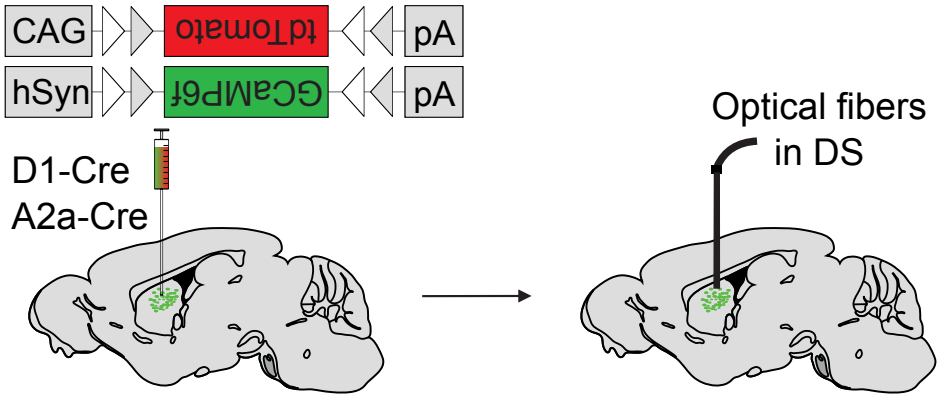


Figure 4.7. **Schematic of the photometry surgical procedure targeting MSNs in either the direct or indirect pathway**. AAV injections and optical fiber implantation in the DS of mice expressing cyclic recombination enzyme (Cre) in either direct pathway MSNs (D1-Cre) or indirect pathway MSNs (A2a-Cre). After 1-2 weeks of recovery, mice were retrained on the temporal discrimination task and the Ca^{2+} activity of DAergic neurons was recorded during task performance.

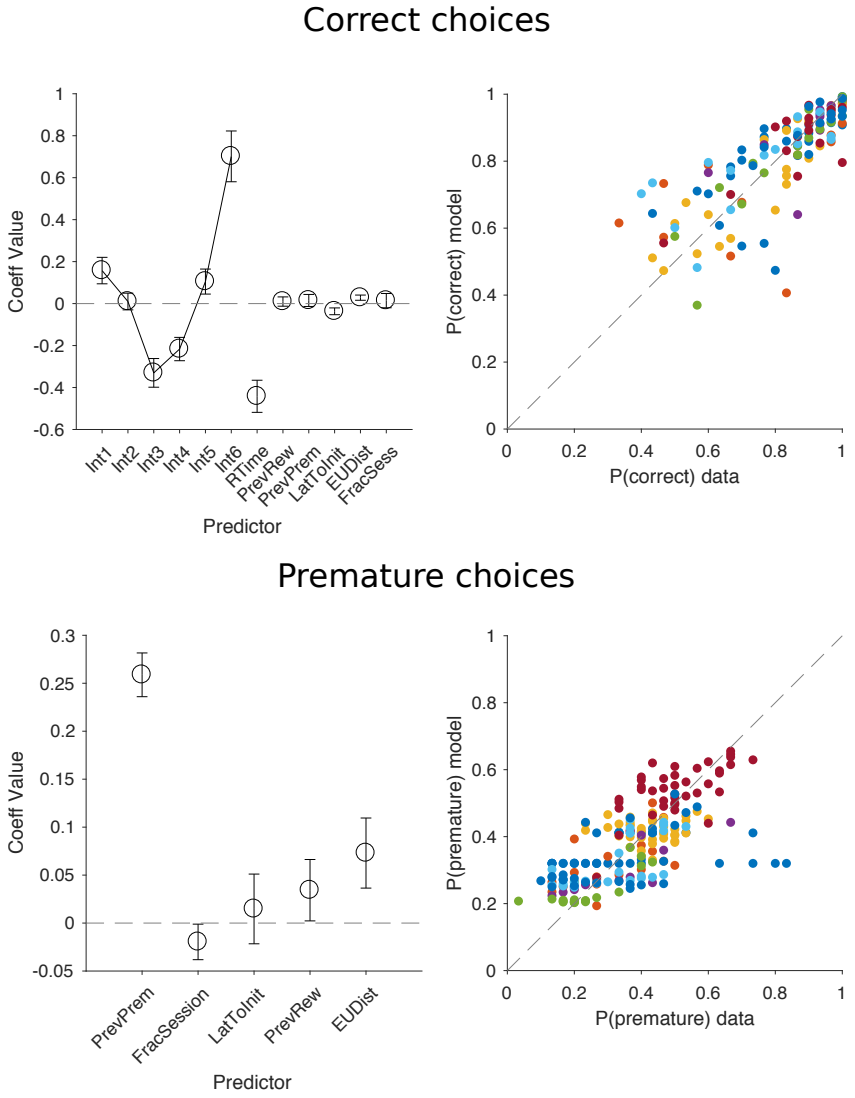


Figure 4.8. **A generalized linear model also predicts correct and premature choices in dMSN-Cre and iMSN-Cre mice**. Same as Figure 4.2 but for animals in which Cre is expressed in either direct pathway MSNs (dMSNs) or indirect pathway MSNs (iMSNs) ($n = 8$ mice).

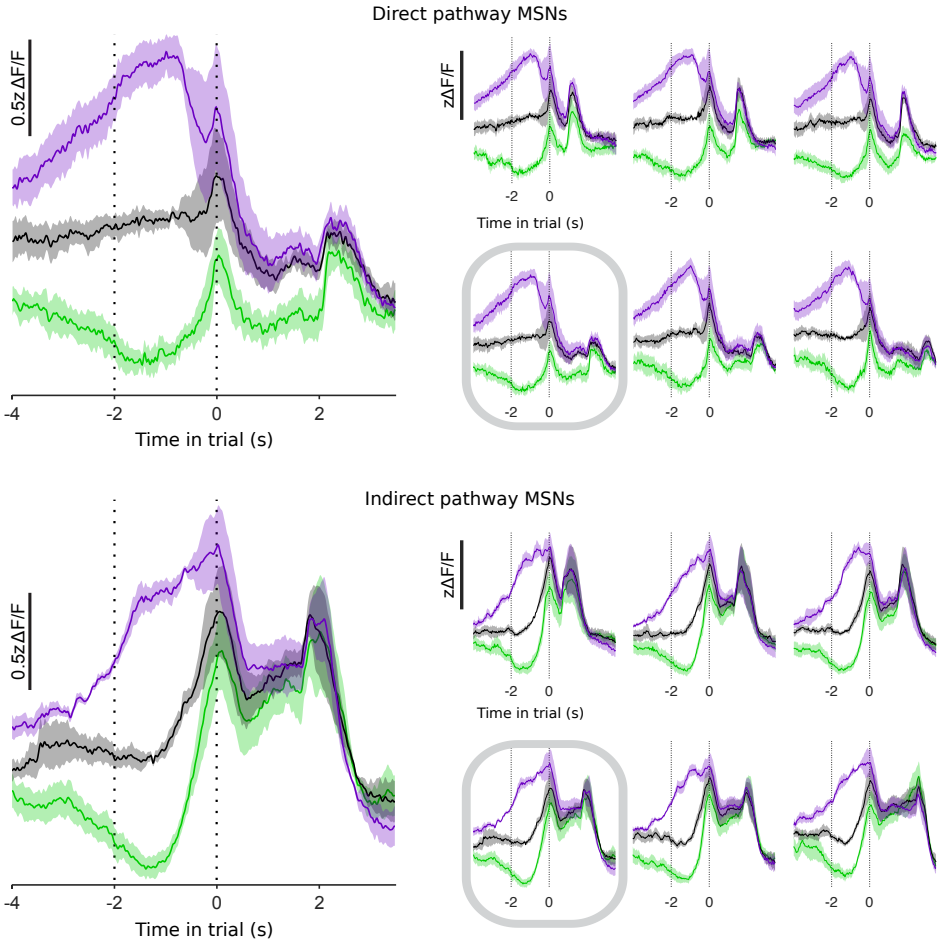


Figure 4.9. **Pre-trial dMSN and iMSN activity is variable**. Same as Figure 4.3 but for activity recorded from dMSNs (*top* panels, $n = 4$ mice) or iMSNs (*bottom* panels, $n = 4$ mice).

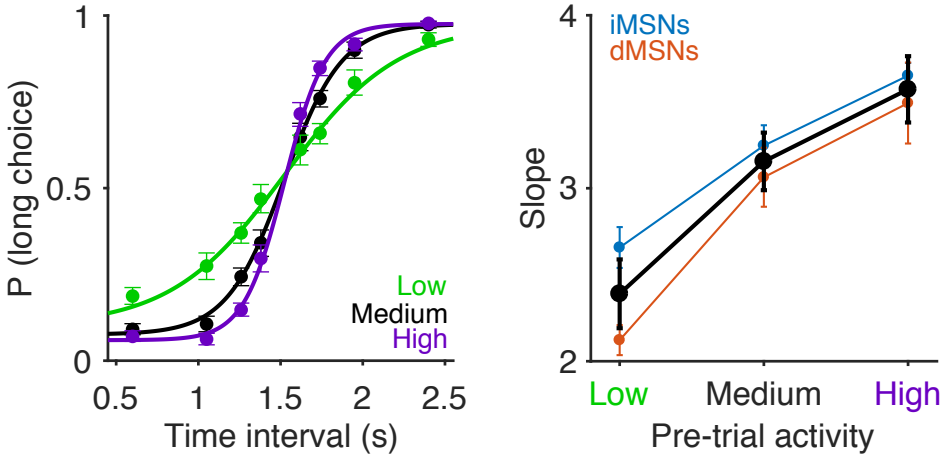


Figure 4.10. **Pre-trial MSN activity is positively correlated with performance**. Same as Figure 4.4 but for recordings performed in the DS from either dMSNs or iMSNs ($n = 8$ mice, 4 in each pathway).

activity of dMSNs and iMSNs while mice performed a similar version of the temporal discrimination task described previously (Figure 4.7, see Section 4.4.3).

We first applied a similar GLM to that in Figure 4.2 to the behavioral data of this different cohort of animals and found identical results to those from DA animals. Specifically, in animals expressing GCaMP6f in MSNs, the best predictors of performance were current stimulus identity and response time (Figure 4.8 *top left* panel), and the best predictor of premature choices was whether the previous choice was also premature (Figure 4.8 *bottom left* panel).

This similarity between DA and MSN animal groups suggests that their behavioral strategies to perform the task are comparable. Similarly to the data in Figure 4.3, both dMSN and iMSN activity displayed variability in the pre-trial period (Figure 4.9). Surprisingly, pre-trial activity in both MSN pathways was positively correlated with performance - the higher the pre-trial activity in MSNs, the better the performance (Figure 4.10). The positive correlation between MSN activity and performance contrasts with the negative correlation between DA neuron activity performance (compare Figures 4.4 and 4.10), and suggests that levels of performance are not encoded in a similar fashion across brain areas. Instead, some transformation must occur between the activity of neurons in these brain areas to explain the opposite sign in their relationship to performance.

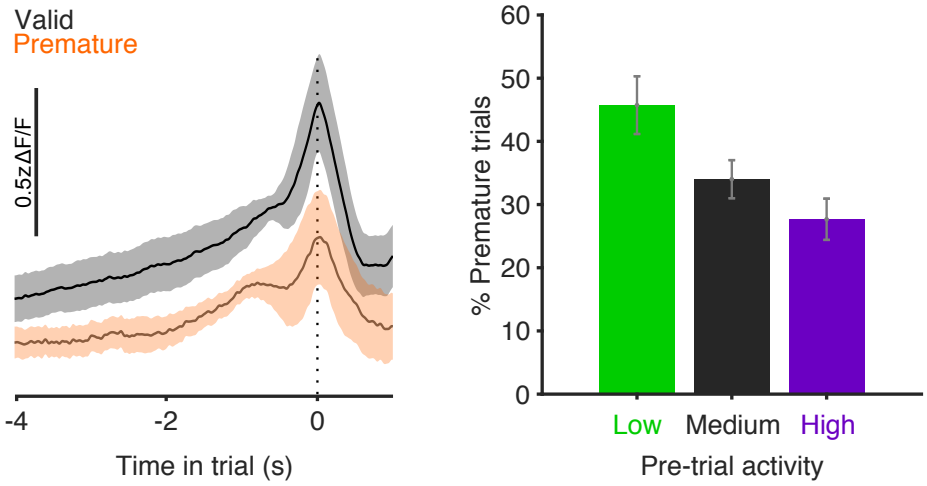


Figure 4.11. **Pre-trial MSN activity is negatively correlated with premature choices**. Same as Figure 4.6 but for recordings performed in the DS from either dMSNs or iMSNs ($n = 8$ mice). Missed trials are not plotted because they were even more rare than in the DA-Cre group of mice.

In addition, we predicted based on the results shown in Figure 4.6 that if MSN activity predicts a general state of engagement, then we should observe lower MSN activity in premature trials than in valid ones (there were not enough missed trials in this data set to be analyzed). Indeed, Figure 4.11 displays this negative relationship between pre-trial MSN activity and premature choices, again consistent with the hypothesis that both midbrain DA neurons and MSNs in the striatum reflect task engagement signals.

4.2.4 Generalized linear model predictions are improved by including pre-trial neural activity

If the pre-trial activity patterns shown in Figures 4.3 and 4.9 are relevant for predicting performance and premature behavior, we should observe an improvement of the models shown in Figures 4.2 and 4.8 if we add pre-trial DA neuron or pre-trial MSN activity as a predictor. To test this hypothesis, we ran four additional GLMs in which we added pre-trial DA neuron or MSN activity as a predictor of performance or of premature choices (Figure 4.12). We compared the performance of these new models that included pre-trial activity with previous models without neural activity as an additional predictor. We found that the coefficient values for the pre-trial DA neuron or MSN activity

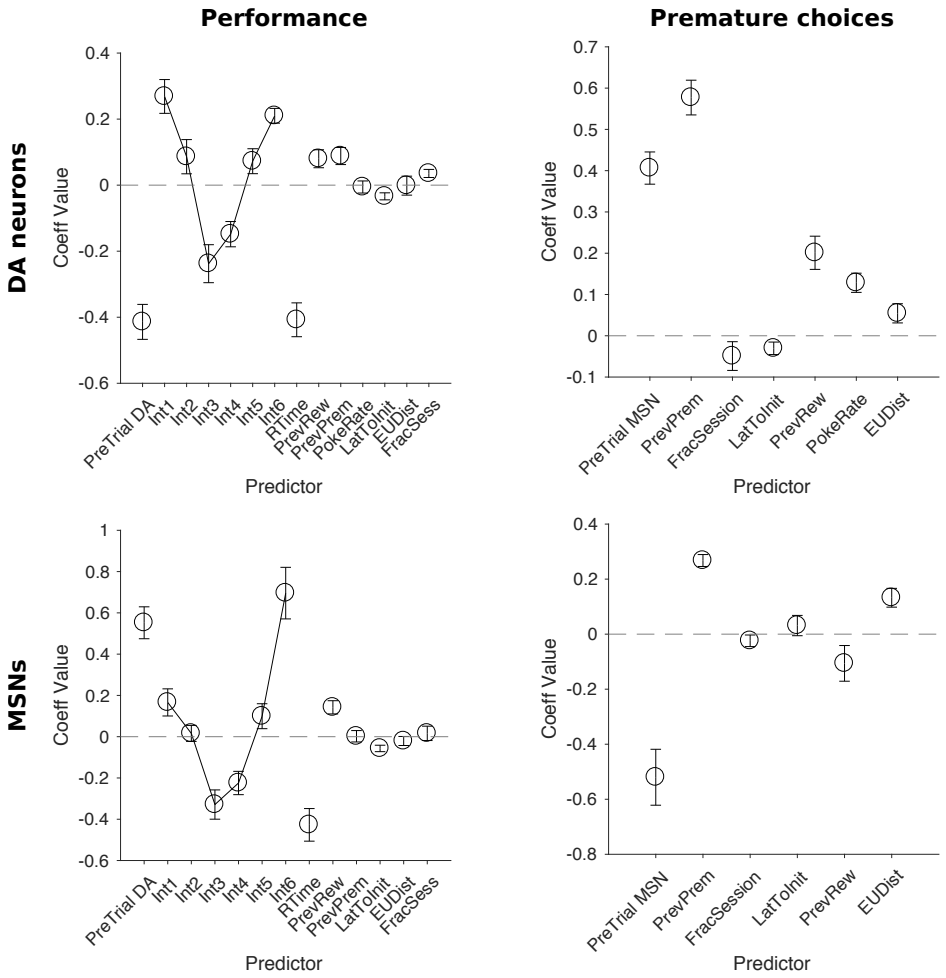


Figure 4.12. **Updated generalized linear models incorporating pre-trial neural activity to predict correct and premature choices**. *Top left* - Same as Figure 4.2 but with the mean DA neuron pre-trial activity in the 2s window leading up to trial initiation as an additional predictor of performance. *Top right* - Same as Figure 4.2 but with the mean pre-trial activity in the 2s window leading up to trial initiation as an additional predictor of premature choices. *Bottom* - Same as Figure 4.8 and *top* panels but for animals where pre-trial activity was recorded from either dMSNs or iMSNs.

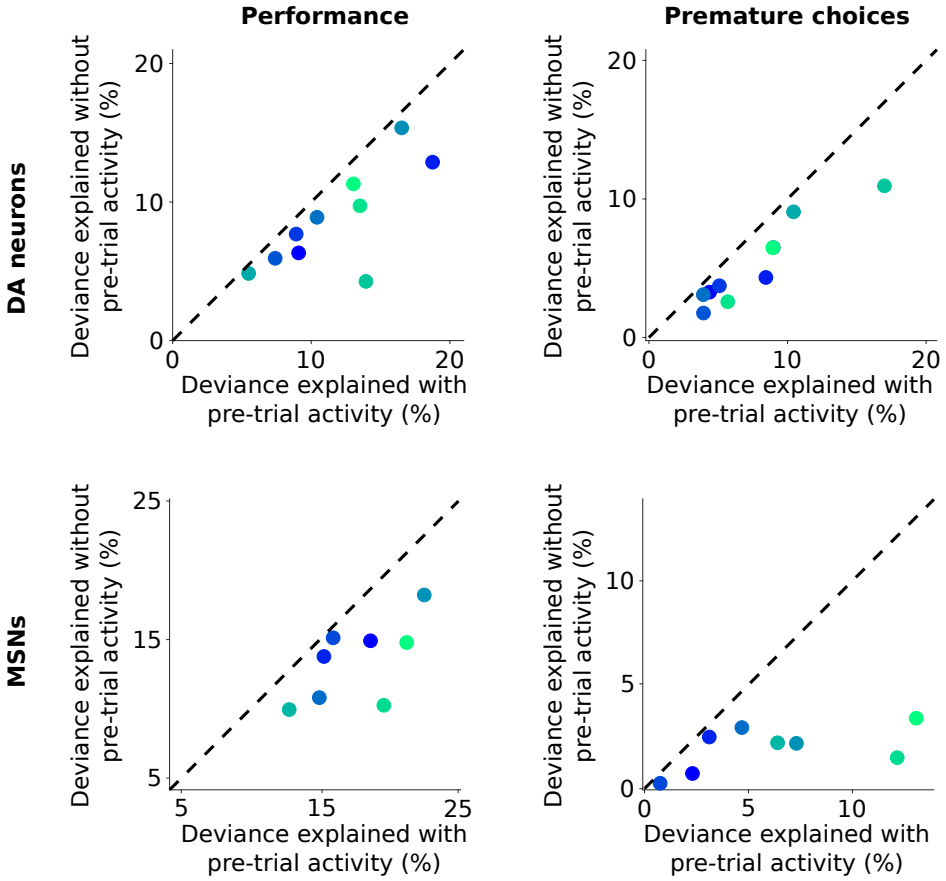


Figure 4.13. **Pre-trial activity of DA neurons and MSNs improves generalized linear model predictions of performance and premature choices.** *Top left* - Deviance explained by the GLM shown in Figure 4.12, which includes pre-trial DA neuron activity as a predictor of performance, versus that explained by the GLM shown in Figure 4.2, which does not include DA neuron activity as a predictor. Each dot represents one mouse ($n = 10$ mice). *Top right* - Same as *top left* panel but for models predicting premature choices. *Bottom* - Same as *top* panels for models including/not including activity of MSNs as a predictor ($n = 8$ mice).

predictors were comparable to other significant predictors such as RT or interval identity (Figure 4.12). As such, the fraction of deviance explained (see Section 4.4.5 for details) by the models with pre-trial neural activity as a predictor was larger than that explained in models without pre-trial neural activity in both the midbrain DAergic system and in the striatum (Figure 4.13, dots lying below the unity line indicate model improvement by adding pre-trial neural activity), indicating that pre-trial activity improves our predictions of both task performance and premature behavior.

4.2.5 Consistency of striatal population dynamics during the delay period reflects task engagement

The results presented so far in this Chapter reveal an interesting correlation between multiple BG neuronal types and behavioral proxies of task engagement, but they simultaneously raise a number of questions.

Firstly, how is the relationship between pre-trial activity and performance inverted between DA neurons and MSNs? There are many possible scenarios consistent with these data (see Section 4.3), one of which involves DA neuron pre-trial activity impacting that of MSNs. We first considered the possibility that DA neurons might act directly through D_1/D_2 type DA receptors mainly expressed in dMSNs and iMSNs, respectively. However, this hypothesis is not parsimonious since DA has opposing effects on these cell types, depolarizing dMSNs and hyperpolarizing iMSNs (Kreitzer, 2009). Therefore, the pre-trial activity of dMSNs should, like that of DA neurons, be inversely related with performance; however, our results indicate a positive relationship between dMSN activity and performance (see Figure 4.10). We then considered the hypothesis that DA neurons might act indirectly on MSNs through striatal interneurons, which could induce the same effect in both dMSNs and iMSNs. As a first step towards testing this hypothesis, we took advantage of an existing dataset in our lab in which the activity of striatal neurons was recorded using electrophysiology in rats performing this temporal discrimination task (Gouvêa et al., 2015). Because distinct neuronal types in the striatum exhibit distinct baseline firing rates and waveform properties (Berke, Okatan, Skurski, & Eichenbaum, 2004; Kreitzer, 2009; Gage et al., 2010), we used these properties to identify putative MSNs (pMSNs) and putative interneurons (pInts) (see Section 4.4.9 for details). We observed that the vast majority of striatal neurons (in particular pMSNs) were silent during the pre-trial period. PInts, however, due to their higher average firing rates, were more active during the pre-trial period. This allowed us to separate trials into octiles based on the pre-trial pInt activity. We compared

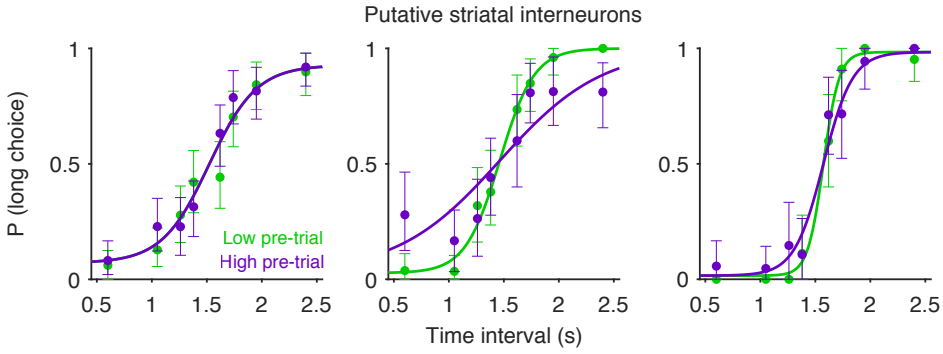


Figure 4.14. **Pre-trial activity of putative striatal interneurons is negatively related with performance**. Psychometric curves constructed using trials from the extreme octiles of putative interneuron pre-trial activity. Curves are the maximum-likelihood fits of logistic functions with the lowest bayesian information criterion scores. Each plot represents one rat. Error bars, 95% CI.

performance between trials in the upper and lower octiles and found that pre-trial pInt activity in 2 out of 3 rats was inversely related with performance (Figure 4.14), the same relationship as that of DA neurons. This observation is suggestive of a scenario where, in the pre-trial period, DA neurons can potentially impact MSN activity indirectly by exciting, for example, parvalbumin positive, fast spiking interneurons (FSIs), which in turn have a hyperpolarizing effect on both dMSNs and iMSNs (Kreitzer, 2009). In fact, given the details of our selection criteria, as well as previous work (e.g., Gage et al., 2010), the population of pInts selected here may indeed be largely comprised of FSIs.

Secondly, how might a signal that occurs in a period before trial initiation influence choices that occur several hundreds of milliseconds to seconds later? From the work of Gouvêa et al. (2015), we know that the population dynamics of striatal neurons in this task resemble the sequential temporal patterns described in Chapter 2. Additionally, Gouvêa et al. (2015) found that subjective time could be decoded from the activity of striatal ensembles during the delay period (the period between the first and second tones). Yet, the photometry technique employed here (see Figure 4.9) sums the activity of many neurons, thereby obscuring these single neuron dynamics during the delay period. To explore these single neuron dynamics and their relationship to task engagement, we re-analyzed the dataset in Gouvêa et al. (2015) during the delay period. For each trial, we calculated how much the striatal population state deviated from the mean population state across all trials (i.e., outlierness, see *left* panel of Figure 4.15 and

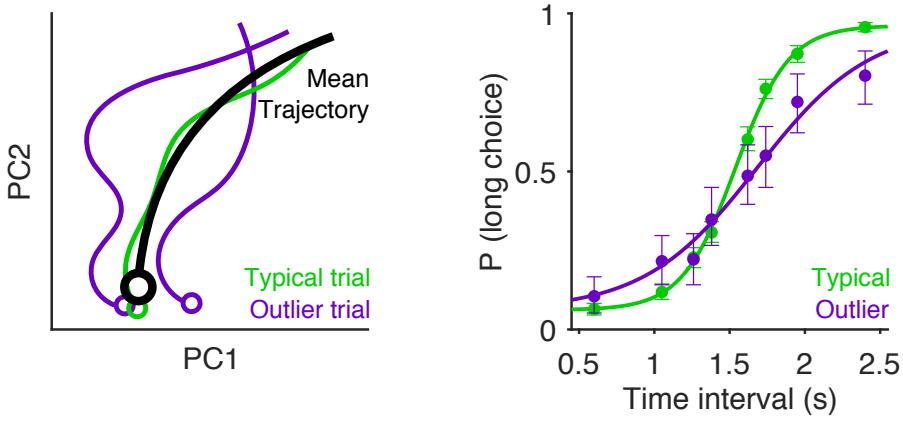


Figure 4.15. **Trial-by-trial consistency in striatal neuron dynamics positively correlates with performance**. *Left* - Cartoon illustrating the selection of outlier trials. Illustration of a low-dimensional representation of population state (i.e., the first two principal components, PCs) during one interval duration. For each interval duration, the euclidean distance between each trial's neural trajectory and the average trajectory across trials in a given interval duration was used as a measure of outlierness (Section 4.4.9 for details). Example outlier trials are color coded in purple, and a typical trial (i.e., similar to the average neural state) is color coded in green. *Right* - The performance of rats on trials falling in the upper octile of the outlierness distribution (purple) are plotted and compared with performance from all other trials (i.e., typical trials, green). Data from the 3 rats in the data set from Gouvêa et al., 2015.

Section 4.4.9 for details). We found that outlier trials (i.e., trials where neural dynamics deviated from the average trajectory) were negatively correlated with performance - the more divergent a trial's trajectory is, the worse the rat's performance (Figure 4.15, *right* panel). These data demonstrate that delay period activity in striatal populations is also predictive of performance, suggesting one possible way in which pre-trial activity may relate to activity during the decision period in striatal neurons.

4.3 Discussion

The data presented in this Chapter highlight behavioral and neural correlates of a state of task engagement. Specifically, animals' performance levels and general adherence to task rules were correlated with the average activity of midbrain DA neurons, striatal MSNs and pInts in a period preceding trial initiation. Moreover, delay period activity in

striatal neurons was also predictive of performance, highlighting a possible link between striatal pre-trial activity and activity during the decision period.

These results add to the known function of DA neurons in regulating states such as arousal (Taylor et al., 2016) and motivation (Mogenson et al., 1980; Wise, 2004), and suggest a basis for the contribution of the BG to a state of task engagement. Here, task engagement is mainly characterized by high levels of performance and low levels of premature behavior. Conversely, lack of task engagement is characterized by poor performance, impulsivity (high probability of premature trials) and even failure to complete trials (missed trials). Interestingly, these traits are hallmarks in the diagnosis of ADHD, a disorder that has been linked to DAergic function (Levy, 1991; Castle et al., 2007). Drugs administered to treat symptoms of ADHD generally act by elevating the levels of DA in the brain. At first glance, this may seem to contradict our observation that low DA neuron activity relates to high task engagement. It is important to note, however, that the relationship between tonic/phasic activity of DA neurons and DA release remains unclear and complex. In fact, due to feedback mechanisms mediated by D2-type auto-receptors, DA agonists have been shown to inhibit the firing of DAergic neurons (Bunney & Aghajanian, 1978), providing a possible explanation for the apparent discrepancy between our results and general ADHD treatments. Additionally, our finding that low MSN activity is associated with decreased performance and increased impulsivity is also consistent with human studies that report low striatal activity in ADHD patients (Lou, Henriksen, Bruhn, Børner, & Nielsen, 1989).

In this study, we monitored DA neuron activity in the context of a temporal discrimination task. It remains unclear, however, whether the relationship between DA neuron activity and task engagement is present in other perceptual decision-making tasks. Other signals carried by DA neurons, such as reward prediction error (RPE) or motivation-related signals, have been described in a wide range of tasks and animal species (Wise, 2004), and have also been observed in both SNc and VTA neurons. Due to the prevalence of this signal across the BG, as well as the potential relevance of a general task engagement signal, we predict that a similar signal might be found in other behavioral contexts.

Although fiber photometry allows one to record the simultaneous activity of many neurons within a target area, it cannot resolve the activity of single neurons. While it remains unclear if this pre-trial activity difference is present in all midbrain DA neurons or only in a subset of them, the fact that we observe this signal in both SNc-DA and VTA-DA neurons suggests that the signal is widespread. In contrast, other response patterns in the same task but in a different trial epoch that reflect and control temporal

judgments are specific to SNc-DA neurons (see Chapter 3). This suggests that, while SNc-DA and VTA-DA neurons may exhibit differences in RPE coding, task engagement signals are similarly represented in these populations.

The presence of neural correlates of task engagement in striatal neurons raises the question of how the signals in distinct BG nuclei relate to each other. Although our results are consistent with a scenario in which DA neurons may indirectly impact MSN activity through striatal interneurons, several other possibilities could also explain our data. For example, through their reciprocal connectivity patterns (Watabe-Uchida et al., 2012), it is possible that MSNs could themselves be impacting DA neuron pre-trial activity. Since MSNs provide inhibitory input onto DA neurons, this could explain the inversion of sign between MSNs and DA neurons in their relationship with task engagement. Further experiments will be required to fully dissect the circuit-level mechanisms underlying the interactions between distinct BG nuclei in the context of task engagement. Along these lines, we performed preliminary experiments in which the activity of both midbrain DA and striatal neurons was simultaneously recorded during task performance (data not shown). Data from such experiments is likely to elucidate some of these mechanisms.

Finally, the observation that animals' performance is diminished when striatal population dynamics are more variable suggests that stereotypic neural dynamics might be associated with states of engagement, while 'wandering' neural dynamics, which are more variable with respect to the average dynamical pattern, can be a sign of task disengagement.

Taken together, the data presented in this Chapter argues for a role of multiple BG components in the process of task engagement. Task engagement, in turn, is combined with current sensory information to guide animals' decisions, adding to the list of internal states that are modulated by these subcortical nuclei.

4.4 Materials & Methods

All experiments were approved by the Champalimaud Centre for the Unknown Bioethics Committee and by the Portuguese Veterinary General Board (Direção Geral Veterinária, project approval 014303 - 0420/000/000/2011) and in accordance with the European Union Directive 86/609/EEC.

4.4.1 Animals

Adult (over two months old) male and female mice of dopamine transporter (DAT):Cre and tyrosine hydroxylase (TH):Cre lines (DA-Cre) were used for the photometry recordings of midbrain DA neurons presented in this Chapter ($n = 10$ mice). Adult (over two months old) male and female mice of D1:Cre (Gensat - EY217, $n = 4$ mice) and A2a:Cre (Gensat - K139, $n = 4$ mice) lines were used for the photometry recordings of dMSN and iMSN activity, respectively. Mice were group housed prior to surgical procedures and single housed following surgery. Mice were maintained under water deprivation for all behavioral experiments (>70% body weight from baseline ad libitum period before deprivation). We also re-analyzed data from an existing data set in the lab (Gouvêa et al., 2015). In this study, three male Long-Evans hooded rats (*Rattus norvegicus*, 6-24 months old) were used for neural recordings using electrophysiology.

4.4.2 Behavioral setup

Mice For the mouse data sets, the behavioral setup was the same as described in Chapter 3, Section 3.4.2.

Rats The behavioral setup for the acquisition of the rat data set in Gouvêa et al., 2015 is described below. Behavioral boxes consisted of a plastic container (IKEA, Alfragide, Portugal) with a speaker (Cover Industrial Co., Guangdong, China) and three noseports (Island Motion), each containing a visible LED and an infra-red emitter/receiver pair for detecting presence at the noseport. Additionally, the choice ports contained a metal tube connected to a solenoid valve for water reward delivery. Valves were calibrated to deliver $25\mu l$ of water per reward event. Interfacing with these sensors and effectors was achieved in the same way as described in Chapter 3, Section 3.4.2.

4.4.3 Behavioral task

Mice and rats were trained to categorize interval durations as either short or long by performing right and left choices as previously described (see Chapter 3, Section 3.4.3 as well as Gouvêa et al., 2014, 2015). The only exception was that animals from the D1:Cre and A2a:Cre genotypes were additionally required to maintain fixation in the

central noseport for the whole delay period between first and second tone presentations. Psychometric functions were fit using a 4-parameter logistic function:

$$P(x) = (u - l)\frac{e^{\frac{x-b}{s}}}{1 + e^{\frac{x-b}{s}}} + l \tag{4.1}$$

Where $P(x)$ is the probability of a long choice as a function of stimulus duration, s controls the slope, b the inflection point and l and u control the minimum and maximum values of the curve respectively. Only sessions where performance was $\geq 70\%$ were used for further analysis.

4.4.4 Video acquisition and tracking

Video data of animals performing the task was acquired using high-speed cameras (Flea3 FL3-U3-13S2C-CS, Point Grey Research Inc., Richmond, Canada) at 120 frames per second. Both the acquisition and offline tracking of each animal were performed using Bonsai (Lopes et al., 2015), which was also used for the acquisition of photometry data as mentioned in Chapter 3, Section 3.4.6. We extracted the position of the center of mass of each animal from the raw videos, by performing background subtraction and thresholding the image so that the animal's body appeared as an isolated element from any other component in the behavioral box. For each frame, we took the centroid of this isolated element and tracked its x and y positions, building a trajectory of the animals' movements in xy coordinates.

4.4.5 Generalized linear models

We fit logistic GLMs to two discrete response variables: i) performance (correct and incorrect trials) or ii) premature behavior (premature trials and valid trials).

Predictor variables to explain correct choices Single-trial measures of behavioral and task-related variables were used. For the model using correct choices as a response variable, the initial GLM included 13 predictor variables. The first 6 variables (Int1, Int2, ..., Int6) corresponded to the six possible stimulus durations in each trial (value of one to the stimulus presented in that trial and zero for all others). The 7^{th} variable (RTime) was the RT of each trial. Because RTs vary with interval duration (see Figure 4.2), we normalized RTs by z-scoring RT values by interval duration. The 8^{th} variable (PrevRew) corresponded to whether the previous trial was rewarded (one) or not (zero). The 9^{th} variable (PrevPrem) corresponded to whether the previous trial

resulted in a premature choice (one) or in a valid one (zero). The 10^{th} variable (Poke rate) corresponded to the number of central nose port interruptions the animal performed in a 4s window leading up to trial initiation. The 11^{th} variable (LatToInit) corresponded to the time from trial availability to trial initiation. The 12^{th} variable (EUDist) corresponded to the eucledian distance taken between each single trial trajectory of the animal in the behavioral box and the mean trajectory of the animal in that session (trajectories were taken in a window of 2s before trial initiation). This variable can be interpreted as a measure of the consistency in xy trajectory traversed by the animal in the behavioral box during the pre-trial period. The 13^{th} variable (FracSess) corresponded to the point in a session at which a given trial occurred (each individual trial number divided by the total number of trials in the session). For the models where pre-trial neural activity was included as a predictor variable, this predictor consisted of the average activity measured in a 2s window before trial initiation. All of these variables were constructed on a session-by-session basis and concatenated per animal.

Predictor variables to explain premature choices Single-trial measures of behavioral and task-relevant variables were used as described above, but some measures, such as interval duration or RT, were not appropriate since the upcoming trial to be predicted would could be premature. Therefore, we used only the applicable subset of the variables described above to predict premature choices: previous premature trials, previous rewarded trials, latency to initiate a trial, fraction of session where a given trial occurred, pre-trial poke rate and consistency in behavioral xy coordinates.

GLM fitting We used the lassoglm package in MATLAB (Mathworks) to fit all the GLMs. For models that estimated performance levels, we used a binary vector of choice outcomes as the response variable ($correct = 1$; $incorrect = 0$). For models that estimated premature choices, we used a similar binary vector ($premature = 1$; $not\_premature = 0$). Each session was divided into two groups: a training group with 80% of the total amount of trials, and a testing group with the remaining 20% of trials. Coefficients were fitted for each animal separately using coordinate descent. We learn the coefficients β by minimizing the cost function:

$$\min_{\beta_0,\beta} \left(\frac{1}{N} Deviance(\beta_0, \beta) + \lambda P_\alpha(\beta) \right) \tag{4.2}$$

where the *Deviance* was calculated using the *log* ratio of the response variable in the training group y_i (e.g., 1 if the trial was correct or premature, 0 otherwise) and the model prediction μ_i for all $i = 1 : N$ trials in the training group. This ratio was

calculated over the positive and null values of the response variable:

$$Deviance = 2\sum_{i=1}^{N}\left(y_i log\Big(\frac{y_i}{\mu_i}\Big) + (1-y_i)log\Big(\frac{1-y_i}{1-\mu_i}\Big)\right) \tag{4.3}$$

The cost function in Equation 4.2 combines the model's *Deviance* with a penalization term $\lambda P_\alpha(\beta)$ that corresponds to elastic net regularization:

$$P_\alpha(\beta) = \sum_{j=1}^{P}\left(\frac{(1-\alpha)}{2}\beta_j^2 + \alpha|\beta_j|\right) \tag{4.4}$$

where, for all $j = 1 : P$ predictor variables, a parameter α interpolates between the L1 (α=1) or the L2 (α=0) norm of β. We selected an α=0.95 which allowed for some predictors to be excluded from the model[1] while preventing degeneracy concerns if predictor variables are strongly correlated. The parameter λ in Equation 4.2, which effectively scales the impact of the penalization $P_\alpha(\beta)$, was chosen by tenfold cross-validation of the training group. After learning the coefficients β using the training set, we evaluated the performance of the model by using the testing group (remaining 20% of trials to validate the model) and calculating the model's *Deviance* (Equation 4.3) but where y_i was now the observed choice in the testing group. This entire procedure was repeated 100 times, generating a distribution of *Deviance* values calculated using only the testing groups of trials.

GLM analysis These *Deviance* values calculated using the test data set were compared to the *Deviance* of a null model (i.e., where the same test data set is used but the model only includes the β_0 bias coefficient parameter). This allowed us to calculate the $Deviance_{explained}$ by the current model:

$$Deviance_{explained} = 1 - \frac{Deviance_{currentmodel}}{Deviance_{nullmodel}} \tag{4.5}$$

To test whether GLM predictions explained the test data sets significantly above chance levels, we compared the models' $Deviance_{explained}$ to a distribution of $Deviance_{explained}$ by models in which the vector of response variables (correct choices or premature choices) were shuffled with respect to the predictor variables. For all animals and all models, the distribution of deviances was significantly different that

[1] Although several predictors were excluded from the models in different runs of the cross-validation procedure, for descriptive purposes we maintained all predictors in the plots shown in this Chapter.

that of the shuffled model ($n = 100$ per distribution, per model, per animal, $p < 0.0001$ in all of them).

4.4.6 Surgery

Mice Surgical procedures were performed as described in Chapter 3, Section 3.4.4 with 2 exceptions. Firstly, when fibers were aimed at the DS or NAc for photometry recordings of DA neuron projections at these sites, the viral injections were performed in the midbrain as described but the optical fibers were aimed at distinct coordinates (for DS - from bregma: 0.74 mm anterior, 1.5 mm lateral, 2.4 mm ventral; for NAc - from bregma: 1.68 mm anterior, 1.25 mm lateral, 4 mm ventral).

Secondly, when fibers were aimed at DS for photometry recordings of dMSN or iMSN activity, we injected 0.3 µL of a mixture of the same AAV1.Syn.Flex.GCaMP6f and AAV1.CAG.Flex.tdTomato virus described in Chapter 3, Section 3.4.4 bilaterally. However, for this experiment, we injected a 5:1 ratio of these viruses in distinct coordinates (from bregma: 0.5 mm anterior, 2.1 mm lateral, 2.3 mm ventral for virus injection, 2.0 mm ventral for optical fiber implantation).

Rats For electrophysiology experiments, we re-analyzed data from (Gouvêa et al., 2015), where rats were implanted with 32-channel tungsten microwire moveable array bundles (Innovative Neurophysiology) under isoflurane anaesthesia. All recordings targeted dorsal striatum (coordinates from bregma: 0.2 mm anterior, 3 mm lateral for rat 1; 0.84 mm anterior, 2.5 mm lateral for rats 2 and 3. Rats were given a week of post-surgical recovery and the placement of the arrays was confirmed with histology.

4.4.7 Fiber photometry setup

For all photometry experiments presented in this Chapter, we used the same photometry setup described in Chapter 3, Section 3.4.6. For the experiments recording from DA neurons, in some animals, optical fibers were simultaneously implanted in multiple areas (DS, NAc, SNc and VTA). In these cases, the location being recorded was alternated daily. For the experiments regarding activity of dMSN and iMSN activity, for each mouse, the hemisphere being recorded was also alternated daily.

4.4.8 Fiber photometry data analysis

Data analysis of photometry experiments performed in all mice of all genotypes presented in this Chapter shared the same basic pre-processing, $\Delta F/F$ calculation, correction using tdTomato signals and z-scoring approaches as described in Chapter 3, Section 3.4.6. Additionally, for splitting trials according to pre-trial activity, for each session, distributions were built using each trial's mean signal in a 2s window before trial initiation. These distributions were then split into terciles, and this separation criterion was used to construct psychometric curves using only the valid trials within each tercile.

4.4.9 Electrophysiology of striatal neurons in the rat

Data acquisition and pre processing For the electrophysiology data analysis described in this Chapter, we re-analyzed data from (Gouvêa et al., 2015). First, neural activity was recorded at a 30kHz sampling frequency, amplified and then filtered between 250-750Hz (Cerebus - Blackrock Microsystems). Each bundle (four bundles per array, 8 electrodes per bundle) was independently moved $50 - 100\mu m$ ventrally after each recording session, hence ensuring that each recording session sampled a distinct neuronal population of striatal neurons. An offline sorting approach (Offline Sorter, Plexon) was used to isolate waveforms from single neurons.

Firing rates calculation and selection of neurons All further analysis were performed using MATLAB (Mathworks). To calculate firing rates from single spike events, we counted spikes in windows of 2ms and convolved these spike counts with Gamma kernel with parameters $\theta = 100$ ms (or 50 bins) and $k = 2$. For further analysis, we selected sessions where ≥ 10 neurons were recorded, and where the minimum population average firing rate across the session was ≥ 2.5Hz.

Identification of pMSN and pInt and pre-trial firing rates calculation We identified as pMSNs all neurons whose average firing rate across all trials was < 5Hz and whose waveform width (i.e., the time from waveform peak to waveform valley) was > 0.4ms. Conversely, pInts were identified as all neurons whose average firing rate across all trials was ≥ 5Hz and whose waveform width was ≤ 0.4ms. Similar criteria have been used previously (Gage et al., 2010), and most likely, the pInts identified here correspond largely to a subpopulation of FSIs (Berke et al., 2004; Gage et al., 2010).

To calculate pre-trial firing rates, for each neuron and each trial, we counted spikes in 2ms bins within a window of 2s prior to trial initiation. Then, for each trial, we averaged the spike rates of specific selected populations (e.g., pInt shown in Figure 4.14). The distributions of these trial-by-trial average pre-trial firing rates in specific subpopulations of striatal neurons were divided into octiles. The performance of each rat for trials that fell in the most extreme octiles of average pInt pre-trial firing rates are plotted in Figure 4.14 (lower octile in green, higher octile in purple).

Analysis of trial-by-trial consistency in striatal neuron dynamics This analysis was performed on individual sessions, splitting trials by stimulus duration. The goal of this analysis was to compare the variability of the neural population state (or trajectory) on single trials to the average state of the population across all trials. We began this analysis by creating a vector of firing rates during the delay period for each trial of a given interval duration in a given session (i.e., population state or trajectory on single trials). We took the median across trials of this population state using only correct trials and refer to it as the average population trajectory. We then calculated the euclidean distance between the population trajectory on a given trial and the average population trajectory, for each time point in the delay period. Next, we took the mean of this euclidean distance across the delay period (i.e., average distance to the mean trajectory on a single trial) and created a distribution of single trial average distances to the mean trajectory. Finally, we selected the trials that fell on the upper octile of this distribution and labeled them as 'outliers', since they were the ones who most deviated from the average population trajectory during the delay period (Figure 4.15, purple trials). The remaining trials were labeled as 'typical' (Figure 4.15, green trials).

4.4.10 Statistics and model comparison

To select the model that best explained differences in the pattern of the subjects' judgments between conditions (i.e. photometry: high, medium, low pre-trial activity; neural activity: outlier or typical delay period trajectory), we fit models as described in Chapter 3, Section 3.4.13. Statistical significance was determined using Wilcoxon sign rank and rank sum tests where appropriate, unless otherwise stated.

Acknowledgments

The mouse photometry of DAergic activity was acquired by Sofia Soares. The mouse photometry of MSNs was acquired by Bruno Cruz. The rat striatal electrophysiology data was acquired by Tiago Monteiro and Thiago Gouvêa. Sofia Soares, Bassam Atallah and Asma Motiwala performed data analysis with contributions from Thiago Gouvêa, Tiago Monteiro and Bruno Cruz. Joseph Paton provided guidance and supervision during all stages of this work. We thank Alessandro Braga for assistance with behavioral training, Gonçalo Lopes for assistance with Bonsai and Margarida Duarte for assistance with mouse colonies. We thank the histopathology and vivarium staff from the Champalimaud Scientific and Technological Platforms for support and animal care.

This work was supported by the following institutions: Champalimaud Foundation; Portuguese FCT (Fundação para a Ciência e a Tecnologia, SFRH/BD/51895/2012 to Sofia Soares); European Molecular Biology Organization (Advanced Long Term Fellowship 983-2012 to Bassam Atallah); Marie Curie Actions (FP7-PEOPLE-2012-IIF 326398 to Bassam Atallah); Bial Foundation (188/12 to Joseph Paton) and Simons Foundation (Simons Collaboration on the Global Brain award 325476 to Joseph Paton).

Chapter 5

Discussion

> "The hours and minutes still seem excruciatingly long when I am bored, and all too short when I am engaged."
>
> Oliver Sacks, 2004

5.1 Brief summary of the main findings

The work presented in this monograph provides strengthening evidence for the role of the basal ganglia (BG) in timing behavior. First, neurons in the dorsal striatum (DS) reflect a scalable population code for time: the responses of single neurons stretched or contracted in time following changes in the duration of the interval being timed. Furthermore, time estimates decoded from the population of striatal neurons were predictive of timing behavior when animals flexibly adjusted to new durations. Midbrain dopamine (DA) neurons in the substantia nigra pars compacta (SNc), which reciprocally connect to the DS, both reflected and controlled temporal judgments, suggesting that these neurons may play a critical role in the moment-by-moment regulation of time judgments. Previous pharmacological studies have demonstrated that many of the effects of DA agonists and antagonists on timing behavior can be explained by changes in attention that accompany DAergic drug administration (Buhusi & Meck, 2002; Ward et al., 2009). Consistent with this observation, we identified signals related with task engagement in midbrain DA and striatal neurons, but importantly, we found that these signals were dissociable from those related to timing behavior.

The following Sections explore open questions that arise from our findings. They also discuss these findings in the context of current theories of interval timing and reinforcement learning (RL), in an attempt at contributing to a unified view of these fields.

5.2 Temporal representations: origins and implications

Our finding that striatal neurons exhibit sequential temporal patterns of activity that support the representation of elapsed time adds to the vast range of studies reporting similar dynamics across the brain in areas such as the cortex (Machens et al., 2010; Harvey et al., 2012; Kim et al., 2013; Schmitt et al., 2017; Wang, Narain, Hosseini, & Jazayeri, 2017), the cerebellum (Buonomano & Mauk, 1994), the hippocampus (Pastalkova et al., 2008; MacDonald et al., 2011), and the thalamus (Schmitt et al., 2017). The widespread nature of these sequential patterns raises the question of where these signals originate.

This Section addresses possible origins and implications of the striatal dynamics we observed, as well as mechanisms that may allow for the flexible control of these dynamics.

5.2.1 Origins and flexible control of striatal dynamics

Due to the long duration of the intervals we probed, and to the existence of sequential activity patterns in striatal afferents, it is unlikely that the dynamics we observed are purely locally generated. The multiplexing of elapsed time with the ongoing sensorimotor state of the animal described in Chapter 2 argues in favor of the hypothesis that striatal dynamics may result from the combination of the sensory, motor and/or associative information of its afferents. Therefore, a key step towards understanding the origin of these striatal dynamics will likely involve studying how they relate to the dynamics present in such input areas. For example, population codes for sensory and behavioral variables have been shown to exhibit distinct timescales across the cortex, with sensory cortices exhibiting shorter timescales than associative ones (Murray et al., 2014; Runyan, Piasini, Panzeri, & Harvey, 2017). At least in part, these distinct timescales are likely to result from stronger and more long-lasting correlations between the activity of neurons in more associative areas when compared to sensory ones (Runyan et al., 2017).

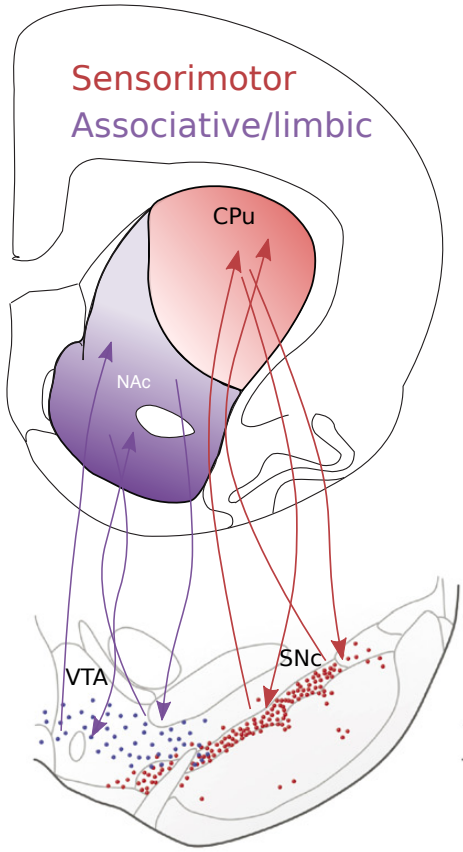


Figure 5.1. **Illustration of the parallel organization of the basal ganglia**. Coronal views of the rat brain at anterior-posterior positions containing either the striatum (top) or the midbrain DAergic system (bottom), illustrating the parallel organization of the BG. In the striatum, dorsal-lateral (DL) areas of the caudate-putamen (CPu) receive predominantly sensory and motor inputs from cortical and thalamic inputs, and tend to reciprocally connect to ventral tier SNc-DA neurons. Conversely, ventral-medial (VM) CPu regions and nucleus accumbens (NAc) neurons preferentially receive more associative/limbic inputs and tend to reciprocally connect to ventral tegmental area (VTA)-DA and dorsal tier SNc-DA neurons. This figure was adapted from Björklund & Dunnett, 2007.

Interestingly, afferent cortical and thalamic projections to the striatum follow a topographical organization, such that sensory and motor inputs mainly target the dorsal-lateral (DL) striatum, while more associative-limbic cortical and thalamic inputs target ventral-medial (VM) striatal regions (see Figure 5.1, Voorn, Vanderschuren, Groenewegen, Robbins, & Pennartz, 2004). This topography may in part explain the distinct functional roles that have been observed across the medial-lateral (ML) axis of the striatum: while the DL striatum seems to reflect stimulus-response associations in habitual behaviors, the dorsal-medial (DM) striatal region seems to reflect stimulus-outcome associations that underlie goal-directed behaviors (Voorn et al., 2004; Gremel & Costa, 2013). Whether this topographical organization can be informative with respect to the representations that we describe in this monograph is not known. One hypothesis is that the long-timescale sequential representations we observed in the striatum might result from the combination of these temporally regular cortical and thalamic inputs, and that the topographical organization of these inputs might be preserved in striatal dynamics. Since the neural recordings described in Chapter 2 did not systematically probe the ML and dorsal-ventral (DV) axes of the striatum, future experiments (e.g., simultaneous recordings of cortical and striatal neuron activity; targeted recordings of striatal neurons on the basis of their afferent inputs) will be able to directly test this hypothesis (see Section 5.3.2 for mentions of such experiments).

Another question that arises from this work regards the observation that striatal neurons flexibly rescale their responses with the interval being timed. Is this property specific to striatal populations? We observed rescaling when the intervals being timed varied from 12s up to 60s, and were presented to the animal in blocks of trials. Additionally, work from Xu, Zhang, Dan, and Poo (2014) showed that cortical neurons were also capable of rescaling their activity when intervals were presented in a block structure, at least on a much shorter timescale (1.5s to 2.5s). Furthermore, our observation that errors in decoded time at block transitions predicted timing behavior suggests that the temporal representation in the striatum may be used to guide behavior. Consistent with this idea, recent work from our group has shown that, when rats are trained in a temporal discrimination task similar to that described in Chapter 3, a low-dimensional representation of the activity of striatal neurons reveals that neural trajectories run ahead or behind when rats judge a near-boundary interval as long or short, respectively (Gouvêa et al., 2015).

Similar trial-by-trial variability has been described in the primate caudate nucleus[1] and dorsomedial frontal cortex (MFC) during a temporal production task (Wang et al., 2017). In both areas, single neuron response profiles were stretched or compressed for longer or shorter production times, respectively, consistent with the the data shown in Chapter 2. Moreover, at the population level, this rescaling was reflected by changes in the speed with which the activity of cortical populations traversed an invariant trajectory, consistent with the results from Gouvêa et al. (2015). Specifically, Wang et al. (2017) found slower speeds for trials where production times were longer and faster speeds for short production times. The dynamics described by Wang et al. (2017) spanned intervals < 1.5s. It remains unknown whether cortical sequential dynamics are able to span tens of seconds, like the striatal dynamics we describe in Chapter 2.

The studies described above indicate that flexible/scalable dynamics, capable of supporting timing behavior, can be found in both cortical and striatal neurons. What mechanisms might underlie this flexibility? Recent work exploring recurrent neural network (RNN) models have made progress towards answering this question. One hypothesis proposed by Wang et al. (2017) is that the scaling and variable speed of cortical dynamics may result from a combination of recurrent dynamics and the input drive to the system. An initial input drive (a cue signal which varies in level providing context regarding the interval duration) may set the initial and final positions of the trajectory along an 'input subspace', while the recurrent dynamics create a 'recurrent subspace' that defines the neural trajectory between these initial and final positions. In their theory, the 'input subspace' is defined by the weights of the inputs to the network, controlling the speed of the trajectory. On the other hand, the 'recurrent subspace' is constrained by the recurrent weights, and is responsible for the rescaling.

While Wang et al. (2017) attempted to model a cortical-like circuit, Murray et al. (2017) have explicitly modeled a striatum-like recurrent inhibitory network with sparse activity to generate sequences of activity with variable duration and speed. Importantly, their network does so without a sequential input, but with a tonic excitatory input that may change in intensity over time. Furthermore, Murray et al. (2017) propose synaptic depression at connections between medium spiny neurons (MSNs) as the mechanism that allows activity to switch from one unit to another. The effect of this synaptic depression is to reduce the amount of inhibition that an active unit applies to inactive units. The basic mechanism that controls the dynamics of how activity switches between

[1] In primates, the DS is composed by the caudate nucleus and the putamen, which are divided by the internal capsule. In rodents, these structures are not distinguishable and are therefore referred to as caudate-putamen (CPu) (Fentress, Stanfield, & Cowan, 1981).

units in this network is the competition between this synaptic depression and the level of excitatory input. In both models, the rescaling of individual units and the speed of population dynamics are determined by a combination of the level of input to the network and the (intrinsic) parameters of recurrent dynamics. Wang et al. (2017) suggest that thalamic inputs may play a role in controlling the input gain and the speed of population dynamics. The authors found that activity in MFC-projecting thalamic neurons was required for performance in their temporal production task; and this activity, unlike that of MFC and caudate neurons, did not scale with production intervals. Instead, thalamic activity was systematically higher for shorter production intervals, providing a possible neural basis for the input drive of their model. Another possibility is that the signals in midbrain DA neurons described in Chapter 3 may play a role in this process. For example, in the model proposed by Murray et al. (2017), DA neuron activity may play a role in the control of temporal rescaling by either modulating the level of input to the recurrent inhibitory network and/or by modifying the levels of recurrent synaptic depression. In support of this hypothesis, work from Tecuapetla, Carrillo-Reid, Bargas, and Galarraga (2007) has shown that DA regulates short-term plasticity at striatal inhibitory synapses.

Interestingly, the concept that the 'intrinsic' properties of neural networks may contribute to their representation of time has been suggested by previous models of timing (Goel & Buonomano, 2014). In fact, recent *in vitro* work has shown that cortical organotypic slices are able to learn the temporal pattern of repeatedly presented temporal intervals, providing evidence in support of such an intrinsic model of timing, at least in the subseconds range (Goel & Buonomano, 2016).

The hypothesis that midbrain DA activity may be involved in regulating the dynamics of striatal population activity during timing behavior may offer an overarching view of the results presented in Chapters 2 and 3. However, other forms of temporal representations are found across the brain. For example, recent work from Emmons et al. (2017) has shown that activity in frontal cortex and striatal neurons varies monotonically during timing behavior (i.e., activity ramps up or down while animals estimate duration). Emmons et al. also found that activity in this frontostriatal circuit scales with the interval being timed, and that inactivating the frontal cortex disrupts ramping activity in striatal neurons. Both linear and nonlinear activity patterns have been previously reported in the the context a temporal reproduction task in the parietal cortex of non-human primates (Jazayeri & Shadlen, 2015). In this study, nonlinear activity profiles were observed while monkeys measured an interval, but activity was reset and proceeded to monotonically increase as monkeys reproduced distinct intervals. Inter-

estingly, the slope of this ramping activity was inversely related to the interval duration that was previously measured, so that activity in the parietal cortex terminated at a similar level at the time of reproduction for all durations. One possibility is that such ramping activity patterns may trigger striatal neurons to become active at different delays. Additionally, the activity of DAergic neurons may play a role in controlling the slope of such ramping activity patterns.

5.2.2 Implications for theoretical timing models

The striatal signals we used to decode time multiplexed information about time and the ongoing sensorimotor state of the animal, therefore implicitly favoring a number of sequential state models of interval timing (Killeen & Fetterman, 1988; Machado et al., 2009; Ahrens & Sahani, 2011). Such models propose that animals may learn the temporal relationship between sensory cues in the environment and their behavior using sensory or behavioral state transitions. Furthermore, our results are consistent with RNN models in which temporal representations are distributed across many neurons and time is therefore encoded by the combined activity of neuronal populations (Buonomano et al., 1995; Wang et al., 2017; Murray et al., 2017).

Additionally, given the ramping activity previously described in cortical and striatal neurons (Leon & Shadlen, 2003; Janssen & Shadlen, 2005; Mita et al., 2009; Jazayeri & Shadlen, 2015; Hanks et al., 2015; Emmons et al., 2017), it is possible that an accumulating process such as in classic pacemaker-accumulator models (Treisman, 1963; Gibbon, 1977) or as in drift-diffusion based timing models (Simen et al., 2011), might serve as a trigger for neurons to become active at different delays, as the accumulator reaches a series of thresholds. Under this framework, DA neuron activity could control temporal judgments by impacting the rate of this accumulation process, but in the opposite direction to that proposed by the DA clock hypothesis (Meck, 1983, 1996). In other words, higher DA neuron activity might slow the rate of an accumulation process rather than accelerating it.

5.2.3 On the distributed nature of temporal representations across the brain

Temporal representations seem to not only be distributed within neurons in a given area, but also across distinct brain areas. How can we make sense of the function of these widespread temporal representations? An attractive hypothesis is that temporal representations can be interpreted within the context of the general functional role of

each brain area. For example, sequential dynamics in the cerebellum may coordinate movements at a fine timescale (Buonomano & Mauk, 1994). Hippocampal temporal representations, conversely, may be combined with spatial representations to organize events in space and time, thereby creating faithful memories of experience (Eichenbaum, 2014). In the cortex, meanwhile, sensory and motor areas may use time-varying signals to represent sensory events (Runyan et al., 2017) and to produce precisely timed movements (Merchant et al., 2013). These data suggest that timing signals are likely distributed across multiple brain areas. Within such a distributed process, what role may striatal temporal representations play when compared to their cortical afferents?

Bakhurin et al. (2017) took steps towards addressing this question by recording the simultaneous activity of striatal and orbitofrontal cortical neurons while mice were trained in a classical conditioning paradigm, in which an odor conditioned stimulus (CS) predicted the delayed delivery of an unconditioned stimulus (US). Mice were able to learn this temporal relationship, as illustrated by their anticipatory licking behavior after the CS presentation. The authors found that, although time from CS presentation could be decoded from both cortical and striatal neurons, time estimates decoded from the striatum were more accurate than those from the cortex. The apparent higher accuracy of striatal than cortical temporal representations might result from the fact that striatal neurons integrate information across a vast range of cortical territories, and may therefore have an advantage in encoding elapsed time over individual cortical areas (Bakhurin et al., 2017).

Additionally, temporal representations in the striatum may be essential for animals to learn how to behave adaptively. The sequential activity patterns we observed in the striatum multiplexed temporal with non-temporal information, such as the ongoing sensorimotor state of the animal. The existence of temporal representations modulated by animals' actions, motor sequences and reinforcement may be critical for the learning and action selection functions attributed to the striatum and to the BG in general (Mink, 1996; Doya, 1999; Niv, 2009).

Finally, the observed rescaling of striatal temporal representations to distinct temporal contexts may provide a neural basis for the observation that learning of associations is timescale-invariant: the number of trials required for learning is constant if the ratio between inter-stimulus-interval and inter-trial-interval is also constant (see Balsam & Gallistel, 2009 for a review on this topic). This timescale invariance can be achieved by the rescaling of temporal representations to the changing temporal context, consistent with our observations and with a general role for striatal dynamics in adaptive behavior.

5.3 Dopamine neurons: interval timing and beyond

The results in Chapter 3 suggest that midbrain DA neurons in the SNc reflect and control temporal judgments. In this Section, I discuss apparent contradictions between our results and previous studies regarding the role of DAergic activity in the control of timekeeping. Additionally, this Section explores other functions that have been attributed to the firing of midbrain DA neurons, and suggests that many of these distinct roles may in fact result from a common computational logic.

5.3.1 Rethinking the DA clock hypothesis

The DA clock hypothesis, a longstanding theory regarding the role of DA in interval timing behavior, posits that high (or low) levels of DA can control the speed of internal timekeeping by increasing (or decreasing) the rate of a pacemaker-like mechanism (Maricq et al., 1981; Meck, 1983; Maricq & Church, 1983; Meck, 1996; Çevik, 2003). The data presented in Chapter 3 adds to a number of studies that seem to contradict the DA clock hypothesis (Holson, Bowyer, Clausing, & Gough, 1996; Buhusi & Meck, 2002; Balcı et al., 2010; Lake & Meck, 2013; Fung, Murawski, & Bode, 2017). One of the main differences between previous literature and the work presented in this monograph is that we recorded and manipulated the activity of midbrain DA neurons on a moment-by-moment basis, while previous work mainly employed long-term lesions (Meck, 2006), pharmacological agents (Maricq & Church, 1983), and genetic (Drew et al., 2007) manipulations of the DAergic system. Such manipulations of tonic DA levels are likely to result in compensatory mechanisms, perhaps offering an explanation for such inconsistencies in the literature. Additionally, *in vivo* microdialysis has shown that DA agonists evoke an initial peak followed by a rapid (< 1 hour) decrease in the levels of striatal DA (Holson et al., 1996), suggesting that DA agonist administration may not always lead to a prolonged increase in DA levels. Furthermore, tonic changes in DA levels across trials have been linked to motivation (Wise, 2004), vigor (Panigrahi et al., 2015) and attention (Levy, 1991), all factors that may confound effects of DA on timekeeping in these long-term manipulations. In fact, superimposition of such effects has been proposed as a possible explanation for how pharmacological manipulations impact timing behavior (for effects on motivation see Odum et al., 2002; Balcı et al., 2010; Balcı, 2014; for effects on attention see Buhusi & Meck, 2002; Ward et al., 2009). We believe that the temporal resolution of the the measurements and manipulations of DA

neuron activity presented in Chapter 3 helped us to circumvent some of these issues. In fact, Chapter 4 illustrates one way in which our approach may have been able to avoid such superimposition effects, since midbrain DA neuron activity in a pre-trial initiation period seems to relate not to timekeeping, but to a general state of task engagement, a distinction that a long-term manipulation would not have identified.

Our data appears consistent with the idea that situations that seem to slow down time estimates may be associated with higher DAergic activity. Our sense of time is known to be influenced by a number of factors. For example, fast internal timekeeping (when external time seems to slow down) has been shown to occur when subjects experience stress (Falk & Bindra, 1954), fear (Watts & Sharrock, 1984), boredom (Watt, 1991) or pain (Ogden, Moore, Redfern, & McGlone, 2015). Conversely, novel music causes consumers to underestimate the time they have spent shopping (Yalch & Spangenberg, 2000), pleasurable events seems to last for a shorter time (Droit-Volet, Ramos, Bueno, & Bigand, 2013), and time estimates are slowed down following the consumption of caloric primary rewards (e.g., fruit juice, Fung et al., 2017). These results are consistent with the notion that novelty, pleasure and primary rewards are associated with slow internal timekeeping (when external time seems to 'fly'). Interestingly, a number of these latter situations are accompanied by changes in midbrain DA activity: for example, DAergic activity is thought to increase following the presentation of novel stimuli (Horvitz, 2000; Menegas, Babayan, Uchida, & Watabe-Uchida, 2017) and the consumption of caloric substances (Frank et al., 2008). These studies are consistent with our observations that increased DAergic activity is related to slower timekeeping.

Still, there might be alternative explanations for our finding that SNc-DA neuron activity relates to temporal judgments. For example, one could argue that the relationship between SNc-DA neuron activity and temporal judgments may be due to a change in the instantaneous value of the option being considered, which could produce results similar to those we describe. Were this the case, we should have observed a similar relationship when recording and manipulating the activity of DA neurons in the ventral tegmental area (VTA), since these neurons are known to be involved in the encoding of value (Hamid et al., 2016). However, no such relationship was present in VTA-DA neurons. Therefore, the most parsimonious explanation of the data shown in Chapter 3 seems to be that SNc-DA neuron activity affects temporal judgments through an impact on timekeeping mechanisms. The idea that DAergic activity might be involved in regulating striatal dynamics, as discussed in Section 5.2.1, suggests one way in which DA neurons may influence temporal judgments. This hypothesis is attractive due to the recurrent connectivity between SNc-DA neurons and neurons in the dorsal striatum

(Björklund & Dunnett, 2007; Watabe-Uchida et al., 2012), which rescale their activity patterns to the interval being timed. Therefore, trials where DA neuron activity is consistently high or low may relate to trials where striatal activity is progressing slowly or quickly. Further data in favor of this hypothesis may be provided from simultaneous recordings of striatal and DA neuron activity (or DA release in the striatum). Alternatively, DA activity may regulate the rate of an accumulation process, as discussed in the previous Section, but in the opposite direction to that suggested by the original DA clock hypothesis (Meck, 1983, 1996).

5.3.2 On the diversity of functions attributed to DA neurons

The hypothesis that midbrain DA neurons in the SNc and VTA encode a reward prediction error (RPE) signal has been supported, for more than two decades, by a number of classical conditioning studies (Schultz et al., 1993, 1997; Cohen et al., 2012; Steinberg et al., 2013). These studies have shown that DA neurons respond phasically to unpredicted rewards, as well as to reward-predicting stimuli after learning, while exhibiting a pause in activity when expected rewards are omitted. Recent work has provided further evidence in favor of this hypothesis, by showing that phasic responses in DA neurons are proportional to the subtraction between expected and experienced outcomes (Eshel et al., 2015, 2016). Additionally, and consistent with the results presented in Chapter 3, DA neurons are modulated by reward expectation, exhibiting larger responses at times when rewards, and reward-predicting cues, are less expected (Fiorillo et al., 2008; Pasquereau & Turner, 2015). Moreover, a number of studies have implicated the firing of DA neurons in model-based learning (Daw, Gershman, Seymour, Dayan, & Dolan, 2011; Bromberg-Martin, Matsumoto, Hong, & Hikosaka, 2010; Sharpe et al., 2017; Langdon, Sharpe, Schoenbaum, & Niv, 2018), expanding the scope of the prediction error signals reflected by these neurons.

Nevertheless, a number of studies have provided data that somewhat challenges this view. For example, some DA neurons seem to respond to novel stimuli that have no predictive power (Horvitz, 2000; Menegas et al., 2017). The projection targets of these neurons are distinct from those of neurons exhibiting more classic RPE responses: DA activity measured at axonal terminals in the nucleus accumbens (NAc) reveals signals following cues that predict rewards, while DA activity at axon terminals in the posterior striatum do not seem to encode an RPE signal, but instead respond to novel stimuli. In addition, it has also been described that DA neurons in the SNc, and

in particular the ones reciprocally connected to relatively DL areas of the striatum, increase their activity in response to both positive and negative USs (juice or air-puffs), as well as to the predicting CS (Matsumoto & Hikosaka, 2009; Lerner et al., 2015). Another class of studies has suggested that SNc-DA neuron activity, as well as axonal DAergic activity in the DS, is modulated by the onset of locomotion, movement direction and adjustments in posture (Howe & Dombeck, 2016; Parker et al., 2016; Dodson et al., 2016). Furthermore, Syed et al. (2016) trained rats to either initiate or suppress movement in response to distinct cues in order to receive reward, and found that DA levels in the NAc are increased when the cue presented triggers movement. In contrast, when the cue presented is associated with movement suppression for reward delivery, DA levels in the NAc are unchanged. Finally, although many of these studies focus on phasic DA responses, and the evoked phasic release of DA, it remains unclear how the sustained (tonic) levels of DA relate to the RPE theory of DA neuron activity. For example, in value-based decision-making paradigms, these tonic modulations of DA neuron activity appear to reflect the average rate of rewards (Hamid et al., 2016). In addition, changes in DA neuron activity on slow timescales have also been linked to vigor (Niv, Daw, Joel, & Dayan, 2007; Panigrahi et al., 2015). Finally, Howard, Li, Geddes, and Jin (2017) have suggested that, in the context of a temporal bisection task, SNc-DA activity biases the selection of actions. Although the authors favor this interpretation, the general pattern of activity they observed seems consistent with our findings: higher DA activity relates to short judgments, but the authors reject the hypothesis that their data might be explained as DA neurons reflecting temporal judgments. This difference in interpretation may, at least in part, result from an attempt of the authors to separate reward expectation and temporal surprise. In our view, an RPE signal can reflect the combination of these two factors, as shown in Chapter 3, and this approach might also explain the signals observed by Howard et al. (2017).

How can we make sense of the vast range of functions that have been attributed to DA neurons, such as those described above? In a recent review, Lau, Monteiro, and Paton (2017) suggested that perhaps DA neurons with apparently distinct functions may, in fact, implement similar algorithms. In their own words:

> "We suggest that the parallel circuit architecture of the BG might be expected to produce variability in the response properties of different dopamine neurons, and that variability in response profile may not reflect variable functions, but rather different arguments that serve as inputs to a common function: the computation of prediction error."

For example, distinct groups of midbrain DA neurons exhibit reciprocal connections with particular striatal regions: DA neurons within the ventral tier of the SNc exhibit preferential reciprocal connections to DL regions of the striatum (Figure 5.1, Björklund & Dunnett, 2007; Watabe-Uchida et al., 2012; Lerner et al., 2015), while those in the VTA and in the dorsal tier of the SNc seem to reciprocally connect to more VM striatal areas, as well as to project to limbic and cortical areas. In turn, the striatum is thought to exhibit a sensory-motor to associative gradient coding in the DL to VM striatal axis, as described in Section 5.2.1 (Figure 5.1, Voorn et al., 2004). Therefore, the apparent localization of motor-related signals in DA neurons that project to dorsal striatal areas may reflect the selective sensorimotor nature of the inputs that these neurons receive from those same areas. Following these ideas proposed by Lau et al. (2017), one way to interpret the differences we observed between SNc-DA and VTA-DA neurons is that a prediction error may be computed in both structures but using distinct inputs, where signals reflecting elapsed time are less predominant and/or precise in VTA-DA neurons than in the ones in the SNc. This might be the case in tasks such as those described in this monograph, where animals likely map the category of interval durations (short vs. long) onto egocentric movements (right vs. left). In such scenarios, combining elapsed time with the ongoing sensorimotor state of an animal may be advantageous. However, unpublished work from our group suggests that, in situations where an allocentric rule might be used to solve the task (e.g., where the animal maps interval duration to cardinal directions such as north and south), ventral striatal neurons seem to encode elapsed time to a higher degree than those in the DS (unpublished work, see conference poster in Monteiro, Rodrigues, Motiwala, Gouvêa, & Paton, 2017). In this context, temporal information might be combined, in the ventral striatum, with more associative signals from areas such as the hippocampus. Therefore, it is possible that DA neurons reciprocally connected to ventral striatal neurons, such as those in the VTA, may reflect and control temporal judgments in this alternative allocentric scenario.

5.4 Towards a unified view of interval timing and RL

A vast range of studies have implicated the BG in interval timing behavior (Maricq et al., 1981; Maricq & Church, 1983; Meck, 1983, 1986; Rammsayer, 1993; Matell et al., 2003, 2004; Xu et al., 2014; Bakhurin et al., 2017; Wang et al., 2017). In what has been a largely separate research field, it has also been highlighted for decades that the BG

are crucial for implementing aspects of RL (Schultz et al., 1993; Mink, 1996; Schultz et al., 1997; Doya, 1999; Waelti et al., 2001; Bayer & Glimcher, 2005; Samejima et al., 2005; Daw et al., 2006; Samejima & Doya, 2007; Lau & Glimcher, 2008; Kim et al., 2009). RL and interval timing might not only share common brain areas, but might share a common theoretical framework, particularly in the form of common temporal representations.

The sequential neural states found in the striatum during timing behavior, as well as the relationship between SNc-DA neuron activity and temporal judgments, may provide evidence towards a unified view of the role of the BG in interval timing and in RL. First, the signals we observed in the striatum resemble temporal basis functions, which have been proposed by a number of learning models as efficient and neurally plausible temporal representations (Grossberg & Schmajuk, 1989; Suri & Schultz, 1999; Ludvig et al., 2008). In fact, similar temporal representations have also been applied in the context of interval timing models (see Section 1.4.4; Grossberg & Schmajuk, 1989; Machado, 1997; Staddon & Higa, 1999).

In an attempt at reconciling these two fields, Gershman et al. (2014) proposed that:

> "(...) by incorporating a time-sensitive action selection mechanism into RL models, a single computational system can support both RL and interval timing."

One way to achieve such a time-sensitive action selection mechanism is to implement an Actor-Critic architecture combined with temporal basis representations. According to Gershman et al., such a combination reproduces the basic behavioral patterns observed in probe trials during a peak-interval (PI) procedure (i.e., increase in response rate, peaking at the estimated reward time, followed by a decrease in response rate). Most importantly, the authors points out how basis functions that peak early in time are more precise than those that peak later, and as a consequence, the credit for rewards that occur at longer delays is assigned to a larger number of basis functions. Due to this dispersed credit assignment for later rewards, the timing of responses that occur for those later rewards is more variable than that of those that occur for earlier ones. This feature allows this RL architecture to display the scalar property (Gibbon, 1977). Furthermore, DA neurons may play a crucial role in this process, since there is evidence that prediction errors might be computed based on a value signal that decreases in precision over time. This idea is based on data showing that the response of DA neurons to reward delivery is positively related to the logarithm of the delay between stimulus and reward (Fiorillo et al., 2008). The results presented in this monograph support

the view discussed above, where RL and interval timing may be implemented under similar theoretical and biological bases. Yet, it remains unclear whether there might be an advantage for a system so intrinsically linked to reward processing to also control subjective time estimates. This question is discussed in the next and final Section.

5.5 Combining time and reward may be adaptive

Having a system that combines the encoding of rewards with that of variable time estimates might be particularly relevant in ethological contexts such as foraging. Optimal foraging theory addresses the strategies underlying animals' foraging behavior in the face of dynamic environments, where they must balance the relative costs of staying in a particular food patch and switching to a new one (Stephens & Krebs, 1986; Kolling & Akam, 2017). A central problem that foraging animals must solve is to estimate how long they should stay in the current food patch. On one hand, food patches become less valuable the longer an animal engages in depleting them, but on the other, switching to another food patch takes time and is therefore costly. Under such constraints, the optimal (reward-maximizing) strategy has been described by marginal value theory (MVT), which posits that animals should stay in a given patch for as long as the reward rate at that patch is greater than or equal to the average reward rate in the environment (Charnov, 1976). In most cases, foraging behavior can be well explained by an MVT comparison rule (Constantino & Daw, 2015; Kolling & Akam, 2017). To implement an MVT-like model, the brain must compute both the average reward rate in the environment and the reward rate at the current patch, and must compare these values in order to make a decision. To calculate such reward rates, an estimate of elapsed time is crucial. Combining the encoding of reward and time may be adaptive because it could potentially allow estimates of reward rate to be adjusted when patches are better (or worse) than expected, such that animals engage with such patches for longer (or shorter) times. Under this hypothesis, one might expect that better-than-expected situations should be associated with shorter time estimates, as we propose in this monograph. Therefore, the current reward rate at better-than-expected patches would be calculated as a higher value than if time estimates were precise.

However, MVT models, as well as model-free RL algorithms, fail to explain foraging behavior in situations where the reward rate at a given patch may either increase or decrease as time elapses (Wittmann et al., 2016). Under such circumstances, subjects are more likely to leave the current patch if reward rates are decreasing. These data suggest that, when deciding when to leave a patch, humans might use recent changes

in reward rate to extrapolate the future reward rate, and use this estimate to guide behavior (Kolling & Akam, 2017). Kolling and Akam suggest that model-based average reward RL may explain these apparently distinct foraging strategies. Additionally, it is possible that RL models that combine more flexible and variable temporal representations may explain behavior under complex foraging scenarios, offering a possible adaptive reason for timing control and reward processing to be combined in one system.

5.6 Concluding remarks

Representing elapsed time is crucial for animals to behave adaptively. In this monograph, we trained rodents to estimate duration and probed the roles of striatal and midbrain DA neurons in timing behavior. We propose that the striatum encodes a scalable population code for time, and that midbrain DA neuron activity reflects and controls temporal judgments.

Our work offers a possible biological basis for temporal representations (and their variability) within interval timing models. Additionally, because we found data indicating that higher DAergic activity relates to slow timekeeping, our results challenge the DA clock hypothesis, a major theory in the field. Furthermore, our moment-by-moment activity measurements and manipulations allowed us to dissociate signals related to timekeeping from those related to animals' task engagement levels.

The results presented in this monograph provide strengthening evidence for the role of the BG in interval timing. Furthermore, our results may also be used as a basis for a continued effort in unifying interval timing and learning theories such as RL.

Finally (and thank you for reading this far), our results also raise a number of questions regarding, for example, the origin of these BG dynamics, how they might be used to produce behavior and what types of behaviors they may influence. Exciting times are ahead, as researchers continue to combine multidisciplinary approaches to advance our understanding of these questions, and ultimately to explain how organisms behave adaptively.

Appendix A

Articles published in peer-reviewed journals

A.1 A scalable population code for time in the striatum

The following Section is a reprint of the article published as in Mello et al., 2015. Copyright information and permissions can be found at http://www.cell.com/permissions.

Current Biology
Article

A Scalable Population Code for Time in the Striatum

Gustavo B.M. Mello,[1,2,3] Sofia Soares,[1,2,3] and Joseph J. Paton[1,2,*]
[1]Champalimaud Neuroscience Programme, Champalimaud Centre for the Unknown, Lisbon 1400-038, Portugal
[2]Instituto Gulbenkian de Ciência, Oeiras 2780-156, Portugal
[3]Co-first author
*Correspondence: joe.paton@neuro.fchampalimaud.org
http://dx.doi.org/10.1016/j.cub.2015.02.036

SUMMARY

To guide behavior and learn from its consequences, the brain must represent time over many scales. Yet, the neural signals used to encode time in the seconds-to-minute range are not known. The striatum is a major input area of the basal ganglia associated with learning and motor function. Previous studies have also shown that the striatum is necessary for normal timing behavior. To address how striatal signals might be involved in timing, we recorded from striatal neurons in rats performing an interval timing task. We found that neurons fired at delays spanning tens of seconds and that this pattern of responding reflected the interaction between time and the animals' ongoing sensorimotor state. Surprisingly, cells rescaled responses in time when intervals changed, indicating that striatal populations encoded relative time. Moreover, time estimates decoded from activity predicted timing behavior as animals adjusted to new intervals, and disrupting striatal function led to a decrease in timing performance. These results suggest that striatal activity forms a scalable population code for time, providing timing signals that animals use to guide their actions.

INTRODUCTION

To behave adaptively in complex, ever-changing environments, animals must learn which actions to take in a particular context based on their past experience. However, to learn about the sometimes-delayed consequences of actions and to guide future behavior, it is absolutely necessary that the brain represent not only actions and consequences but also temporal information about when those actions and consequences occur [1].

Multiple lines of evidence implicate the basal ganglia (BG) as a locus for the representation of such temporal information. Lesions of the striatum in rats [2], disease states that affect the BG such as Parkinson's [3] and Huntington's disease [4], drugs that affect dopamine (DA) signaling [5], and genetic manipulations that affect the DA system in the BG [6] all result in interval timing dysfunction. Furthermore, human fMRI studies have found that the striatum, a main input area of the BG, is activated by tasks that involve the processing of interval information [7, 8].

In addition, many theoretical models have been proposed to explain timing behavior. These models can be grouped into at least three categories. Pacemaker-accumulator models integrate pulses emitted from a central pacemaker to measure elapsed time [9, 10]. Beat frequency models detect patterns of activation across resettable oscillatory processes at different frequencies to encode time delays from a resetting event [11]. Sequential state models contain orderly transitions between different activity states that can be used to encode time [12–14]. These theories reproduce various aspects of timing behavior in many interval timing tasks. However, neural data in conflict or in support of the various theories are lacking.

To understand how time is encoded in neural circuits, we recorded the spiking activity of neurons as rats performed an interval timing task. Specifically, given the apparent localization of timing function in striatal tissue, we asked whether striatal neural activity could encode elapsed time over durations of tens of seconds to 1 min while we measured behavior that reflected animals' estimates of time.

We found that different striatal neurons fired maximally at different delays from reward receipt and that information about animals' time estimates could be extracted from striatal populations by simply treating neurons as tuned for time. Importantly, this tuning for time, while affected by sensorimotor event-related neural responses, could not be fully explained by ongoing behavior, as even cells that displayed responses locked to a specific behavior varied their responses depending on when that behavior was executed within a given interval. Strikingly, we found that temporal tuning stretched or contracted, rescaling with the interval being timed. Thus, striatal populations encoded relative time, flexibly adapting to the immediate demands of the environment. Finally, we ran a simple simulation of the task and show that neural responses resembling those we observe in the striatum are suitable as a basis for timing behavior. These results provide important biological insight into how a major brain system encodes time during behavior.

RESULTS

Lever Pressing Start Time under Fixed Interval Reinforcement Schedules Is a Behavioral Measure of Rats' Expectation of Time until Reward

To elicit robust time-dependent behavior over a broad range of timescales, we employed operant conditioning procedures under fixed interval (FI) schedules of reinforcement (Figure 1A).

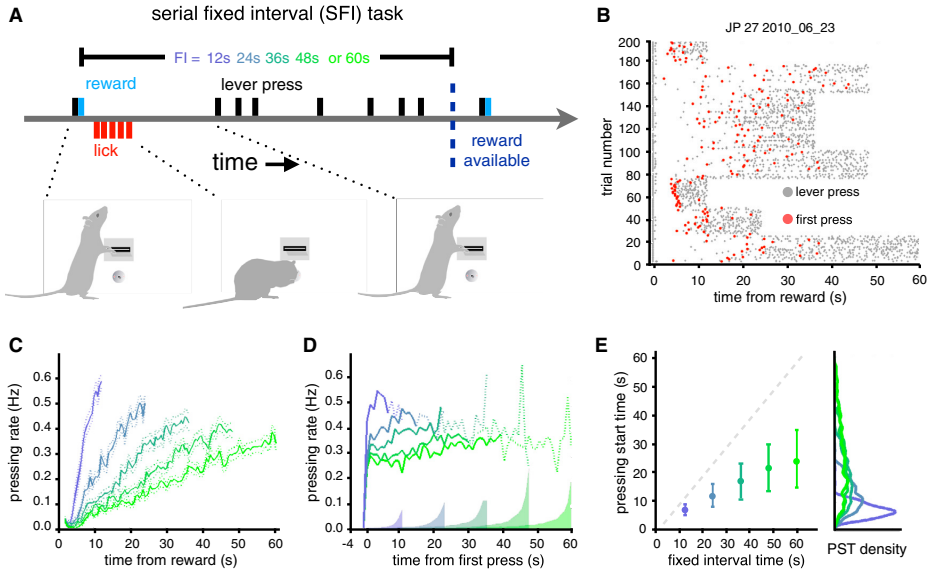


Figure 1. The SFI Task Produces Systematic Changes in Lever PST

(A) Task structure. The following color code will be commonly used: blue represents short FIs, and green represents longer FIs.

(B) Example of lever pressing behavior in a single session of the SFI task. Gray markers indicate a lever press; red markers indicate the PST.

(C) Average lever pressing rate in each of the five FIs, aligned on preceding reward. Dashed lines represent SEM.

(D) Average rate of lever pressing in each block, aligned on PST. Traces are plotted on a solid line for the period for which more than half of the trials contribute data and on a dotted line after that point. Shaded patches along the horizontal axis represent SEM. 1 s bins are indicated in (C) and (D).

(E) Median and interquartile range of PST for each of the five FIs. Smoothed density functions depicting the full distributions of PST are shown on the right. See also Figure S1.

Briefly, rats were placed in a behavioral box containing a lever positioned over a liquid delivery port and were trained to press the lever to receive water reward. Reward delivery triggered a timer, and reward became available again only after the timer exceeded a FI ranging from 12 s to 60 s in multiples of 12 s. Lever presses occurring after reward delivery but before the FI had elapsed were not reinforced. A FI was maintained for between 18 and 40 rewards before changing to another FI, randomly chosen from the interval set.

In single sessions, rats tended to distribute lever pressing toward the latter portion of the FI, shifting when they responded as FI changes occurred (Figures 1B and S1A). This pattern of responding produced ramps in block-wise averaged pressing as a function of time that varied in slope in relation to FI (Figures 1C and S1B). However, this did not reflect the pattern of responding in single trials. We asked how pressing evolved after pressing onset (pressing start times, PSTs) in each trial by aligning on the PST and averaging lever press rates across trials and within blocks of the same FI (Figure 1D). Rats pressed at a relatively constant rate after the first press in each trial, with a rate determined by the experienced reward rate (Figure S1C). The ramps in the reward-aligned pressing as a function of time largely result from changing distributions of PSTs (Figure 1E), as these vary systematically with FI, and averaging a group of step functions with onset times drawn from these distributions will produce ramps of varying slope.

This serial fixed interval (SFI) lever pressing task produced systematic variation in the distributions of PSTs of bouts of anticipatory pressing, consistent with previous timing studies employing FI schedules of reinforcement [9]. These bouts were of a relatively constant rate that varied with reward rate over time (Figures 1D and S1C). The PST thus provided a behavioral metric that covaried with the animals' changing expectation about time until the next available reward, which we compared to the activity of neurons recorded in the striatum during performance of the task as described below.

Striatal Neurons Display Temporal Tuning

In the SFI task, reward delivery is both the timing cue and the reinforcer. Since animals reported knowledge of time between reward availability by when they began to press a lever, we asked whether neuronal responses in the striatum aligned on reward might reveal a signal that animals could use to guide the decision of when to begin pressing. We recorded broadly in the dorsal striatum so as to sample neurons from regions previously shown to be important for interval timing behavior [2] (inset in Figure S1D), and the vast majority of units we recorded exhibited average firing rates of less than five spikes per second, consistent with a population made up of mostly medium spiny projection neurons [15] (Figure S1D).

Aligned on reward delivery, the population of recorded cells exhibited a broad distribution of activity patterns, as reflected

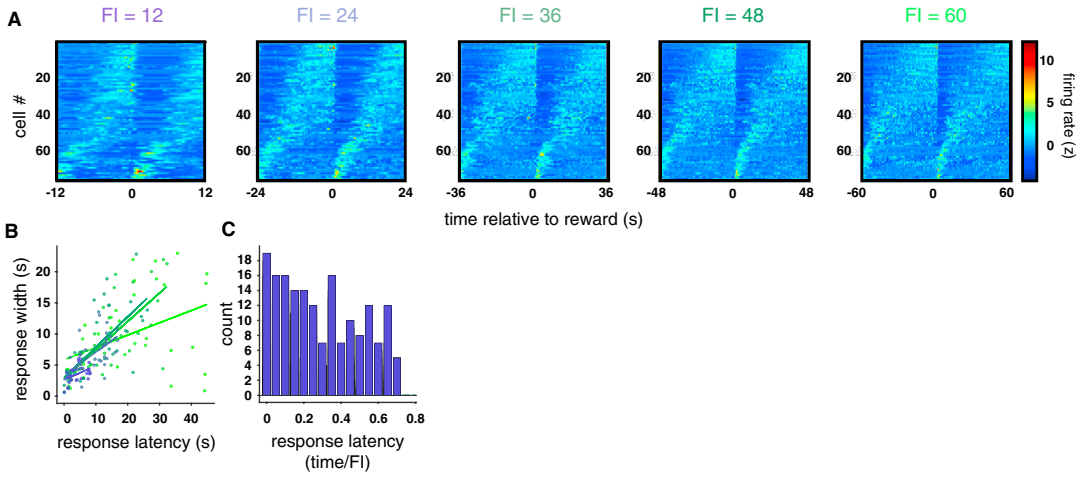


Figure 2. Striatal Neurons Display Variable Responses that Tile Tens of Seconds to 1 Min

(A) SDFs of neurons that maintained their relative ordinal position in time within the population across all five FIs, aligned on reward.

(B) Width of each cell's response within each FI as a function of latency to peak firing. Colored lines represent the best linear fit to the data.

(C) Histogram of relative peak latencies pooled over all FIs, using data shown in (B).

See also Figure S2.

in the normalized spike density functions (SDFs; see Supplemental Experimental Procedures for details) shown in Figure 2A. Some cells fired just after reward delivery, others fired in the middle of the delay, and others fired leading up to the next reward (Figures 2A, S2, and S3). This produced a slow-moving "bump" of activity that traversed the population during each FI. In theory, reading out the location of this bump in the population could provide an estimate of time within the FI. However, a core feature of interval timing behavior is that timing accuracy decreases with the magnitude of the interval being timed [9]. Two features of the neural data could potentially contribute to this phenomenon: an increased spread of each neuron's responses as a function of their peak latency and a decreasing density of neurons displaying peak firing rates as time progresses. We found that the widths of responses were indeed correlated with their latencies to peak firing within each FI (Figure 2B, linear regression, FI 12 s, $R = 0.4443$, $p < 0.001$; FI 24 s, $R = 0.7563$, $p < 0.001$; FI 36 s $R = 0.7188$, $p < 0.001$; FI 48 s, $R = 0.5910$, $p < 0.001$; FI 60 s $R = 0.4733$, $p < 0.001$; see Supplemental Experimental Procedures for details). In addition, the density of peak firing rate latencies in our population decreased over time within the FI (Figure 2C). Thus, the bump in activity within the striatum population moved progressively slower as the FI wore on. Strikingly, the overall time taken by this bump to traverse the population appeared to scale with the FI (Figures 2A and S4A). To begin to assess apparent scaling of response times, we first selected cells that we had recorded in all five FIs and that maintained their ordinal position within the population when responses within each FI were ordered by firing dynamics [16]. Of the 112 neurons recorded in all FIs, we found that 76 neurons (68%) maintained their ordinal position in time across the population (see Supplemental Experimental Procedures for details). The responses of these neurons can be observed in Figure 2A, wherein the position of cells along the y axis is the same across the panels displaying average responses in each of the FIs (for all recorded cells, see Figure S4A).

To quantify to what degree responses rescaled, we computed a scale factor for each neuron as the ratio of the center of mass (COM) of the SDF in the 12-s FI over the COM of the SDF in each of the other four FIs (Figure 3A). The distributions of these scale factors were sharper than and significantly different from null distributions generated by shuffling cell identity across FIs and recomputing the scale factors (red distributions in Figure 3A, Kolmogorov-Smirnov test, $p < 0.001$ for all pairwise comparisons). The medians of these distributions, were the population to have rescaled its responses in direct proportion to the FI, should lie at 1/2, 1/3, 1/4, and 1/5 for the scale factors corresponding to 12/24 s, 12/36 s, 12/48 s, and 12/60 s FIs, respectively. We observed median values of 0.59, 0.39, 0.30, 0.24 for the corresponding distributions, indicating near-proportional rescaling of response times across the recorded striatal population. A more-complete description of the relative scale of responses can be seen in Figures 3B–3E, where the COM of each cell's SDF in the 12-s FI against each of the other FIs are displayed. These data demonstrate a strong tendency for rescaling of neural responses across the population, suggesting that the state of striatal populations may convey relative elapsed time information scaled to the animal's estimate of the current behaviorally relevant timescale in the environment. We explore this hypothesis in greater detail below.

Striatal Populations Encode Information about Timing Behavior

The above analyses of striatal neural responses indicate a gross correspondence between striatal activity and timing behavior across blocks of trials, suggesting that striatal activity patterns

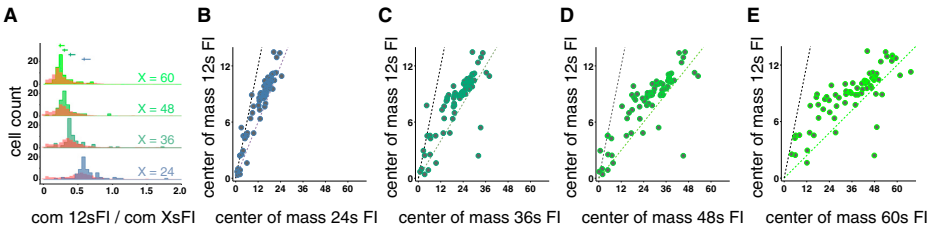


Figure 3. Striatal Neurons Rescale Their Response Time with FI

(A) Distributions of scale factors obtained by calculating the ratio of the center of mass (COM) of the SDF between the 12-s FI and X-s FI (24 s, 36 s, 48 s, and 60 s, respectively, from blue to green) for each cell. For each distribution of scale factors, a null distribution was generated by shuffling cell identity across FIs and recomputing the scale factors (red).

(B–E) COM of each cell's SDF in the 12-s FI against each of the other FIs. The black dotted line signifies no change in COM from block to block. The colored dotted line signifies a change in COM that is proportional to the change in FI relative to the 12-s FI.

See also Figure S3.

might guide decisions about when to begin pressing the lever during each FI. To test this hypothesis, we applied a decoding approach to data collected from single trials near block transitions, wherein animals systematically changed the time that they began to press the lever. Specifically, we asked three questions. First, we asked whether decoded time estimates covaried with true time. Second, we asked whether systematic errors in estimated time as compared to true time occurred at these block transitions. Lastly, we asked whether any observed errors in time encoding correlated with timing behavior.

We first built a probabilistic decoder to derive an estimate of elapsed time from reward in single trials given the observed spiking response of the population. We focused on the first trials of the 12-s and 60-s FI blocks because these blocks were the shortest and longest FIs employed, respectively. Thus, animals consistently overestimated and underestimated the amount of time remaining until reward as they entered 12-s and 60-s blocks. Briefly, our decoder was constructed as follows. In each of the first seven trials of a block, we counted spikes within defined time bins and asked how likely we were to have observed that number of spikes at each time given the observed distributions of spike counts in trials 8 onward of the corresponding block. This generated a likelihood function for current time, given an observed spike count in each bin, for each individual cell. To derive a measure of the population's estimate of the likelihood for current time, we multiplied together the individual cells' likelihood functions. We then took the mean of this likelihood function as our estimate for current time [17].

In Figures 4A and 4C, we display decoded estimates as a function of time for the first seven trials of 12-s and 60-s FI blocks. We found that decoded estimates tracked true time but that systematic errors between estimates and true time were present in the first few trials of the 12-s and 60-s FI blocks. This feature can be observed more readily when estimates derived from multiple trials are plotted on the same axes (Figures 4B and 4D, quadratic fits). Initial estimates were relatively slow and fast in the first trials of the 12-s and 60-s FI blocks, respectively, and became more accurate after the first few trials.

Next we asked whether such timing signals may be used by animals to guide timing behavior. We first asked whether errors in decoded time estimates over the first trials of blocks were correlated with timing behavior. We found that the mean PST was significantly correlated with the errors in time estimates derived from the population over the first seven trials of 12-s and 60-s FI blocks (Figure 5; FI = 12, $R^2 = 0.63$, $p = 0.03$; FI = 60, $R^2 = 0.64$, $p = 0.03$). In the initial trials of the 12-s FI block, rats began pressing late relative to subsequent trials, and likewise, the decoded estimate of time relative to reward ran slow (Figures 4B and 5). The first trials of the 60-s FI block showed a similar relationship, yet opposite in direction: the decoded estimate ran quickly in early trials, and rats were early to press (Figures 4D and 5). We then tested in two control animals whether manipulating striatal circuitry via bilateral infusions of the GABAa agonist muscimol produced deficits in timing behavior (Figure S5). Indeed, at a dose that rendered rats able to perform the task, muscimol reversibly and significantly diminished the relationship between PST and FI (linear regression, likelihood ratio test, significant effect of treatment, $p < 0.001$), showing that a normally functioning striatum is critical for normal timing behavior. The consistency between time estimates decoded from striatal populations and trial-by-trial variations in timing behavior at block transitions, together with observed dependence of a normally functioning striatum for normal timing behavior, suggests that the brain uses a population code for time that samples broadly from striatal neurons to guide decisions about when to act.

Striatal Neurons Multiplexed Information about Action and Time

Based on previous studies [18–20], we expected that striatal neurons would display significant modulation by behaviors during the FI. Could behaviors that accompany task performance fully explain the sequential neural responses we observed? Several features of the data argue against this possibility. Rats consistently licked at the reward port from 0.5 s to 5.5 s after reward delivery (Figure S4B), and yet, our ability to decode time was unaffected by the animal being engaged in a fixed behavior over this time (see initial ~5 s of decoded time estimates in Figures 4A and 4C). After departing from the reward port, however, it is possible that observed dynamics in neural responses are accounted for by ongoing behaviors. Were this the case, responses related to a particular behavior should not vary

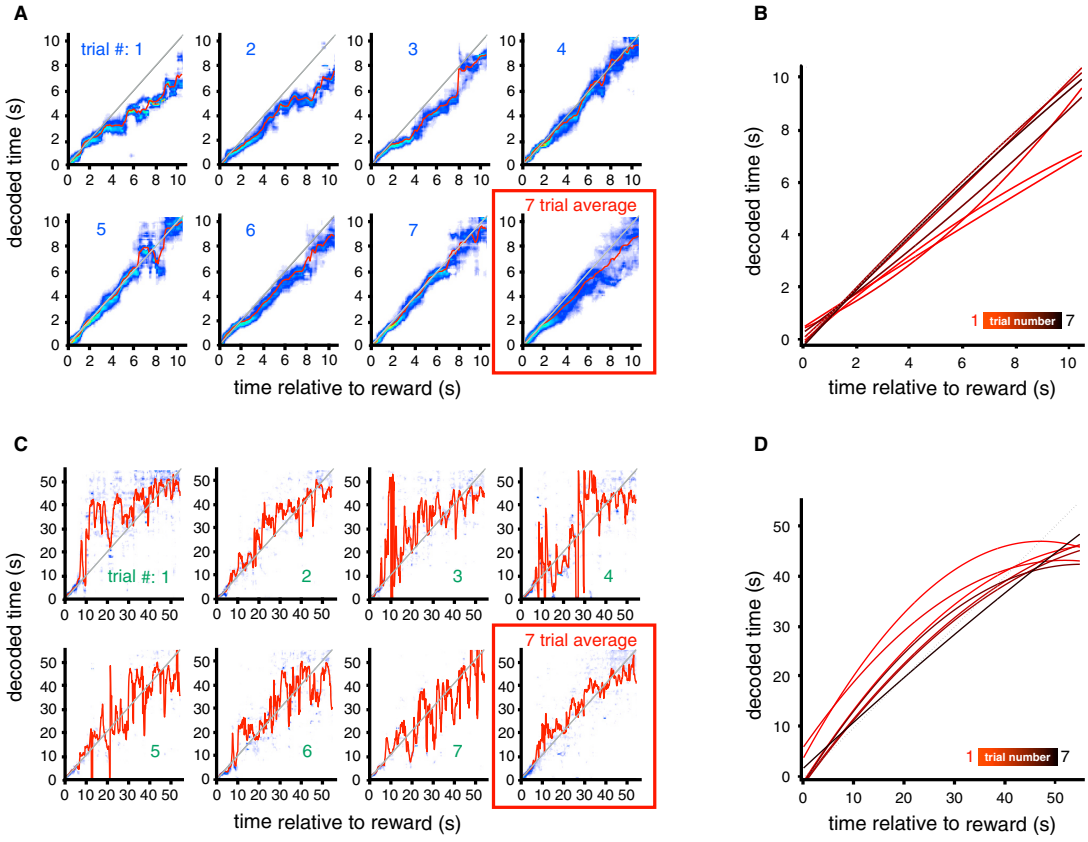


Figure 4. Single-Trial Estimates of Elapsed Time Decoded from the Population Response Correlate with True Time during Initial Trials of 12-s and 60-s FI Blocks

(A) Decoded population estimates of elapsed time from reward in single trials, for the first seven trials of the 12-s FI block plotted against true time. Red traces indicate the mean of the population likelihood function, and the underlying heatmap indicates the population likelihood function. The last panel shows a seven-trial average likelihood function using the first seven trials of the 12-s block.

(B) Decoded estimates of elapsed time for the first seven trials of the 12-s FI block plotted on the same axis. Curves are quadratic fits to the mean likelihood function of each individual trial (red lines in first seven panels). Red curves represent early trials, and black curves represent later trials.

(C) Same description as in (A), but for the 60-s FI.

(D) Same description as in (B), but for the 60-s FI.

See also Figure S4.

depending on when in a trial the rat engaged in that behavior. To identify neurons that were significantly modulated by a measured behavior in our task, we focused on a 2.5-s epoch centered on the PST in each trial. We found that of the 76 neurons displayed in Figure 2A, 31 exhibited significant modulations around the onset of lever pressing. Next, we asked whether spiking observed in time bins aligned on the PST was additionally correlated with the time, relative to the FI, that pressing onset occurred. More than half of pressing onset-modulated neurons (16/31, 52%) displayed a significant correlation between spiking around each press initiation and the relative time that press onset occurred within the FI (Pearson's linear regression, $p < 0.01$). Figures 6A–6D show examples of four such neurons from three different animals, all of which vary in their responses around the PST, from none at all to robust firing.

The regression approach described above is only expected to identify neurons that display a monotonic relationship between pressing onset response and the relative time of pressing onset. Other cells may have displayed significant time-dependent modulations in pressing onset response that were not monotonic (for example, see Figures S2B and S3A). To identify such cells, we asked whether the median of distributions of spikes counts, collected around pressing onsets and falling into each of five quintiles of relative PSTs, differed from each other. We found that 53 out of 76 neurons (70%) displayed in Figure 2A exhibited significantly different median spike counts across relative time

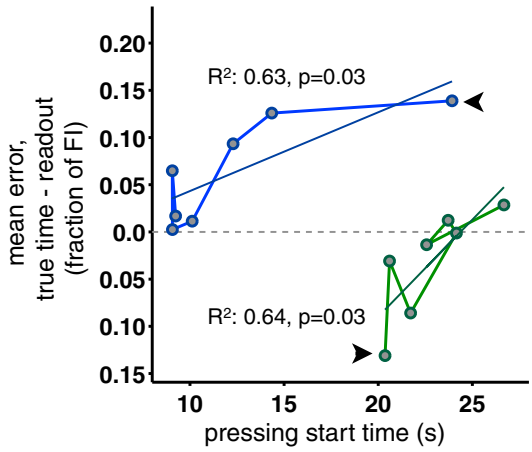


Figure 5. Errors in Decoded Time Predicted Timing Behavior
Mean error between true time and the decoded population estimate in the first seven trials of the 12-s (blue) and 60-s (green) FI blocks. Contiguous trials are connected by solid lines to display the trajectory of the data over trials, and the first trial on each block is indicated by the black arrow. Dashed horizontal gray line represents zero error average decoding as compared to true time. See also Figure S5.

within the FI ($p < 0.01$, Kruskal-Wallis). Of these, nine cells were significantly modulated by the onset of lever pressing and were not identified in the linear regression analysis. Overall, only six cells that displayed response modulation around PST did not exhibit additional modulation by relative time in the FI as assessed by linear regression and/or nonparametric testing for median difference in spike count. These results suggest that striatal neurons multiplex information about time and immediate sensorimotor state of the animal and argue strongly against the possibility that the striatal population responses we observed can be explained by purely non-time-related responses to specific sensory or motor components of ongoing behavior.

A Simple Simulation of Timing Behavior

In order to understand the relationship between the recorded striatal signals and rats' behavior, we ran a simple simulation that performed the SFI task (Figure 7A). The core of this simulation is comprised of a set of temporal basis functions that were inspired by the diverse single-neuron responses observed in our striatal dataset as well as existing timing and learning models [21–24]. We used the method described in [23] to generate temporal bases. Each function was used as a rate function for generating inhomogeneous Poisson spike trains from which time was read out during task performance. Whenever this time readout passed a threshold, presses were produced at a fixed rate. In order to adapt to the changing FIs, we implemented a simple learning rule to update a temporal scale factor for the basis functions depending on the difference between expected time of reward and encoded time at the time of reward delivery. Lastly, to account for our observation that many striatal neurons multiplexed information about action and time, each press produced a response in the temporal bases that was proportional to the product of the original time-dependent rate function at the time of the press and a rate function generated by the press itself. With these elements, we ran the simulation under the conditions contained in the SFI task.

The simulation produced qualitatively similar behavior to that of rats (Figures 7B and S6) and reproduced the three main features that we observed in striatal neurons: temporal tuning, rescaling of neural responses (Figure 7C), and multiplexing of information about action and time (Figure 7D). Although simple, the simulation serves as proof of principle that neural activity with the properties that we observe in this study can serve as a basis for timing behavior and suggests candidate computational elements such as a scale factor and temporal error signal for which there might exist functional analogs in the brain.

DISCUSSION

Time is a fundamental dimension of animals' experience in the world. As such, it plays an integral role in many brain processes, from perception to motor control to learning and memory formation. What is the role of temporal representation within the BG? A dominant view supported by a wide range of neurobiological data posits that the BG implements aspects of reinforcement learning (RL) [1, 20, 25–28], learning how an organism ought to act in order to maximize reward. However, to learn about the sometimes-delayed consequences of actions and to guide future behavior toward rewarding outcomes, it is absolutely necessary that the brain represent situations and actions through time [1, 29]. Indeed, temporal relations among actions and events contain the causal information that learning systems have evolved to detect through a process sometimes referred to as credit assignment [30]. Once credit for the occurrence of predictable events has been assigned, this information must be used to profitably guide the course and timing of action as situations arise. This continuous learning-behaving cycle is what RL algorithms naturally account for [29]. Yet, it is not known how the BG, the brain system most often associated with RL, represents temporal relationships over the durations necessary to explain its purported role in animal learning and behavior.

The sequential neural states that we describe in the striatum during timing behavior can provide a unifying view of the BG's role in timing and RL. These signals are strikingly similar to temporal basis functions proposed in existing learning models as more neurally plausible and efficient representations of time [21–23], which we show can be used to generate timing behavior similar to what we observed experimentally. Such models operate by learning a set of weights used in a weighted sum of the temporal bases to construct a moment-by-moment prediction about future events such as expected reward. In theory, a weighted combination of activity patterns in the cortical or thalamic inputs to the striatum could act as such temporal bases and modulate the responses of striatal neurons that we observed.

An important question for future studies concerns the mechanism that generated the striatal dynamics we observed. We find it unlikely, given the duration of the intervals we examined, that striatal dynamics were purely locally generated, although several modeling studies suggest mechanisms for generating sequential activity states using striatum-like circuitry over

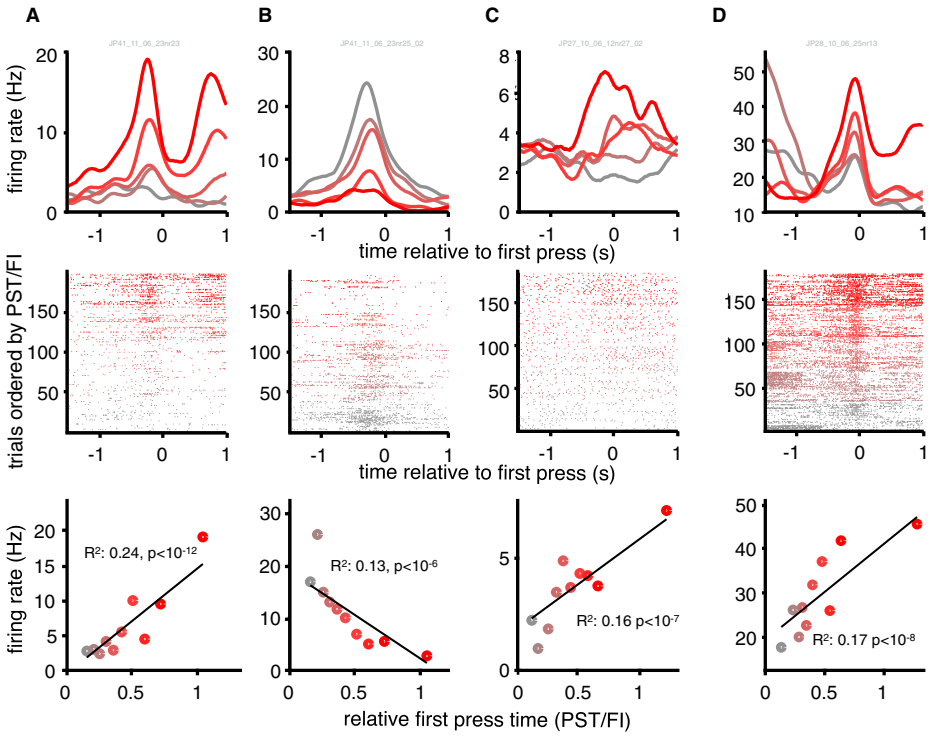


Figure 6. Pressing Onset Responsive Neurons Display Sensitivity to the Time Relative to the FIs

(A–D) Four single-neuron peri-stimulus time histograms (top) and raster plots (middle) of 2.5-s epochs aligned on pressing onset event (from three animals; the two first columns display data from two neurons recorded in the same animal and same session). Trials were sorted in ascendant fashion from bottom to top on the vertical axis by the pressing onset time relative to the FI (middle) and grouped into quintiles. Here, the colors from gray to red represent the first to the fifth quintile, respectively (middle and top panels). Bottom panels: correlation between the firing rate of the respective neuron and the time of the pressing onset relative to FI. Each data point is color coded from gray to red for the first to the tenth decile of the relative pressing onset time. Firing rates were extracted from the most modulated 500 ms bin of the four bins surrounding the pressing onset event.

shorter timescales [31, 32]. Indeed, the signals we use to decode time were affected, but not fully explained by, the ongoing sensorimotor state of the animal. Thus, our decoding approach implicitly endorses a number of prominent interval timing theories, positing that animals may use behavioral [12, 14] or sensory state [33] transitions to learn to time events in the environment and their own behavior.

Our data appear most consistent with theoretical models that suggest distributed representations of time encoded by the joint activity of populations of neurons [13]. Indeed, the decoder used in the current study assumes that time information may be present in many different neurons. However, we cannot rule out that upstream of the population we recorded in the striatum, other forms of temporal representations may exist. For instance, an accumulating process such as that contained within pacemaker accumulator models [9] might act to trigger neurons to become active at different delays as the accumulator passes a series of thresholds.

We show that sequential neural activation in the striatum can be used to encode time on a scale of tens of seconds up to 1 min. These results add to a growing list of studies that demonstrate sequential activation of neurons over multi-second timescales in other brain areas, such as the hippocampus [34, 35], the cerebellum [36], the parietal cortex [37], and the prefrontal cortex [38–40]. Unlike previous studies, we found that many individual striatal neurons exhibited responses that dynamically rescaled with the timing of events in the environment and that this scaling of responses produced changes in time encoding by the population that correlated with timing behavior. Combined with previous studies highlighting the importance of a normally functioning striatum for timing behavior [2–4, 6], the effect of striatal inactivation in the current study, and other work that demonstrated time encoding by striatal populations over shorter timescales [41], our results suggest that information about where in time a subject finds itself relative to anticipated events in the environment is present in populations of striatal neurons and is used to guide behavior.

Similar timing signals observed in areas other than the striatum are viewed within the larger context of the functional role of those areas where they were recorded. Timing signals in the

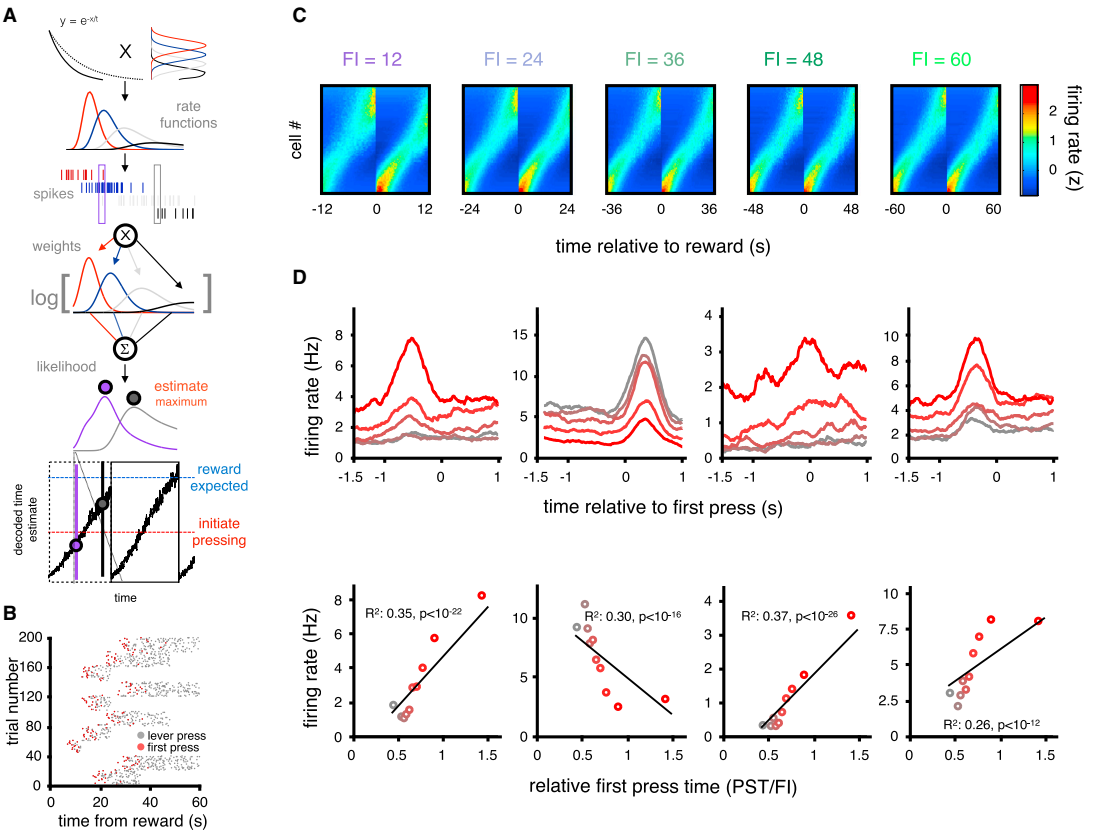


Figure 7. A Simple Simulation of Timing Behavior

(A) Firing of striatal neurons was modeled based on receptive fields for the height of a decaying trace that is reset in each trial by reward delivery (top left). This trace can decay faster (solid line) or slower (dotted line) by adjusting the parameter γ. The Gaussian functions (top right) represent receptive fields evenly spaced along the height of the trace function. The trace function was multiplied by the receptive fields to generate rate functions, the levels of which vary across time as the memory trace decays. Spike counts observed within defined time bins were then multiplied by the logarithm of their respective rate functions and summed to compute the population log likelihood function for current time given the population response, from t = 0 to t = FI. The maximum of this likelihood function was used to derive our estimate for current time relative to reward, for each time bin. Decoded time estimates can run faster or slower depending on whether the trace function decays quickly or slowly. For each trial, when the decoded time estimate reached a given threshold (red dotted line), we simulated a probabilistic pressing process. If the decoded estimate runs too slowly, it fails to reach the threshold value for expected reward (blue dotted line) before the current FI elapses, and the reward happens before it was expected (dotted black box), generating a large prediction error that drives appropriate updating of γ in the next trial. If the decoded estimate runs more accurately (solid black box), a small prediction error is generated, and γ is minimally adjusted in the next trial.

(B) Example of simulated lever pressing behavior on the SFI task. Gray markers indicate a lever press; red markers indicate the PST.

(C) SDFs of simulated units ordered by response profile. Each panel corresponds to one FI.

(D) Four single-unit peri-stimulus time histograms of 2.5-s epochs aligned on pressing onset event (top). Trials were grouped in quintiles of the relative PST. The colors from gray to red represent the average firing in the first to the fifth quintile, respectively. The bottom panel shows the correlation between the firing rate of the correspondent unit on the top panel and the PST relative to FI. Each data point is color coded from gray to red for the first to the tenth decile of the relative PST. See also Figure S6.

hippocampus might endow explicit memories with accurate information about the order and temporal context of events [24], and timing signals in the cerebellum might coordinate learned actions at a fine timescale [36], while timing signals in premotor cortex might enable accurate timing of movement in general [42]. The striatal neurons we observed appear to multiplex temporal information with other, non-temporal types of information, such as signals related to the ongoing sensorimotor state of the animal and likely other previously identified striatal signals related to actions, motor sequences, or reinforcement [19, 26–28]. Such multiplexing of temporal and other information in populations of striatal neurons as observed in the current study is likely to be critical to the previously ascribed and often-studied function of the BG in learning and action selection.

EXPERIMENTAL PROCEDURES

All experiments were in accordance with the European Union Directive 86/609/EEC and approved by the Portuguese Veterinary General Board (Direcção-Geral de Veterinária, project approval 014303 - 0420/000/000/2011). Five male Long-Evans hooded rats were used in the neurophysiological experiments, and two male Long-Evans rats were used for the muscimol experiments. All isolated units (179 total from 5 rats, 25 R1, 9 R2, 21 R3, 28 R4, 96 R5) recorded for at least three blocks in sessions in which PSTs correlated significantly with FI ($p < 0.05$) were included in subsequent analyses. All analyses and simulations were performed using custom software in MATLAB (MathWorks). See Supplemental Experimental Procedures for a detailed description of methods and procedures.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures and six figures and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2015.02.036.

AUTHOR CONTRIBUTIONS

G.B.M.M. and J.J.P. designed the experiments. G.B.M.M. and S.S. carried out the experiments. G.B.M.M., S.S., and J.J.P. analyzed the data and wrote the manuscript.

ACKNOWLEDGMENTS

We thank Bassam Atallah, Brian Lau, Kenway Louie, Christian Machens, Zachary Mainen, Thiago Gouvêa, Eric DeWitt, Alfonso Renart, and Masayoshi Murakami for critical comments on versions of the manuscript and discussions. We thank the histopathology and vivarium staff from the Champalimaud Scientific and Technological Platforms for support. This work was supported by Champalimaud and Gulbenkian Foundations and fellowships to G.B.M.M. and S.S from the Portuguese Foundation for Science and Technology.

Received: December 15, 2014
Revised: January 23, 2015
Accepted: February 11, 2015
Published: April 23, 2015

REFERENCES

1. Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. Science *275*, 1593–1599.
2. Meck, W.H. (2006). Neuroanatomical localization of an internal clock: a functional link between mesolimbic, nigrostriatal, and mesocortical dopaminergic systems. Brain Res. *1109*, 93–107.
3. Malapani, C., Rakitin, B., Levy, R., Meck, W.H., Deweer, B., Dubois, B., and Gibbon, J. (1998). Coupled temporal memories in Parkinson's disease: a dopamine-related dysfunction. J. Cogn. Neurosci. *10*, 316–331.
4. Rowe, K.C., Paulsen, J.S., Langbehn, D.R., Duff, K., Beglinger, L.J., Wang, C., O'Rourke, J.J., Stout, J.C., and Moser, D.J. (2010). Self-paced timing detects and tracks change in prodromal Huntington disease. Neuropsychology *24*, 435–442.
5. Maricq, A.V., and Church, R.M. (1983). The differential effects of haloperidol and methamphetamine on time estimation in the rat. Psychopharmacology (Berl.) *79*, 10–15.
6. Ward, R.D., Kellendonk, C., Simpson, E.H., Lipatova, O., Drew, M.R., Fairhurst, S., Kandel, E.R., and Balsam, P.D. (2009). Impaired timing precision produced by striatal D2 receptor overexpression is mediated by cognitive and motivational deficits. Behav. Neurosci. *123*, 720–730.
7. Hinton, S.C., and Meck, W.H. (2004). Frontal-striatal circuitry activated by human peak-interval timing in the supra-seconds range. Brain Res. Cogn. Brain Res. *21*, 171–182.
8. Tanaka, S.C., Doya, K., Okada, G., Ueda, K., Okamoto, Y., and Yamawaki, S. (2004). Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. Nat. Neurosci. *7*, 887–893.
9. Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. Psychol. Rev. *84*, 279–325.
10. Simen, P., Balci, F., de Souza, L., Cohen, J.D., and Holmes, P. (2011). A model of interval timing by neural integration. J. Neurosci. *31*, 9238–9253.
11. Meck, W.H., Penney, T.B., and Pouthas, V. (2008). Cortico-striatal representation of time in animals and humans. Curr. Opin. Neurobiol. *18*, 145–152.
12. Killeen, P.R., and Fetterman, J.G. (1988). A behavioral theory of timing. Psychol. Rev. *95*, 274–295.
13. Buonomano, D.V., and Merzenich, M.M. (1995). Temporal information transformed into a spatial code by a neural network with realistic properties. Science *267*, 1028–1030.
14. Machado, A., Malheiro, M.T., and Erlhagen, W. (2009). Learning to time: a perspective. J. Exp. Anal. Behav. *92*, 423–458.
15. Gage, G.J., Stoetzner, C.R., Wiltschko, A.B., and Berke, J.D. (2010). Selective activation of striatal fast-spiking interneurons during choice execution. Neuron *67*, 466–479.
16. Geffen, M.N., Broome, B.M., Laurent, G., and Meister, M. (2009). Neural encoding of rapidly fluctuating odors. Neuron *61*, 570–586.
17. Dayan, P., and Abbott, L.F. (2005). Theoretical Neuroscience, Second Edition. (Cambridge: MIT Press).
18. Mink, J.W. (1996). The basal ganglia: focused selection and inhibition of competing motor programs. Prog. Neurobiol. *50*, 381–425.
19. Jin, X., and Costa, R.M. (2010). Start/stop signals emerge in nigrostriatal circuits during sequence learning. Nature *466*, 457–462.
20. Kim, H., Sul, J.H., Huh, N., Lee, D., and Jung, M.W. (2009). Role of striatum in updating values of chosen actions. J. Neurosci. *29*, 14701–14712.
21. Grossberg, S., and Schmajuk, N.A. (1989). Neural dynamics of adaptive timing and temporal discrimination during associative learning. Neural Netw. *2*, 79–102.
22. Suri, R.E., and Schultz, W. (1999). A neural network model with dopamine-like reinforcement signal that learns a spatial delayed response task. Neuroscience *91*, 871–890.
23. Ludvig, E.A., Sutton, R.S., and Kehoe, E.J. (2008). Stimulus representation and the timing of reward-prediction errors in models of the dopamine system. Neural Comput. *20*, 3034–3054.
24. Howard, M.W., MacDonald, C.J., Tiganj, Z., Shankar, K.H., Du, Q., Hasselmo, M.E., and Eichenbaum, H. (2014). A unified mathematical framework for coding time, space, and sequences in the hippocampal region. J. Neurosci. *34*, 4692–4707.
25. Doya, K. (1999). What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? Neural Netw. *12*, 961–974.
26. Lauwereyns, J., Watanabe, K., Coe, B., and Hikosaka, O. (2002). A neural correlate of response bias in monkey caudate nucleus. Nature *418*, 413–417.
27. Samejima, K., Ueda, Y., Doya, K., and Kimura, M. (2005). Representation of action-specific reward values in the striatum. Science *310*, 1337–1340.
28. Lau, B., and Glimcher, P.W. (2008). Value representations in the primate striatum during matching behavior. Neuron *58*, 451–463.
29. Sutton, R.S., and Barto, A.G. (1998). Reinforcement Learning. (Cambridge: MIT Press).
30. Balsam, P.D., and Gallistel, C.R. (2009). Temporal maps and informativeness in associative learning. Trends Neurosci. *32*, 73–78.
31. Ponzi, A., and Wickens, J. (2010). Sequentially switching cell assemblies in random inhibitory networks of spiking neurons in the striatum. J. Neurosci. *30*, 5894–5911.
32. Berns, G.S., and Sejnowski, T.J. (1998). A computational model of how the basal ganglia produce sequences. J. Cogn. Neurosci. *10*, 108–121.

33. Ahrens, M.B., and Sahani, M. (2011). Observers exploit stochastic models of sensory change to help judge the passage of time. Curr. Biol. *21*, 200–206.

34. Pastalkova, E., Itskov, V., Amarasingham, A., and Buzsáki, G. (2008). Internally generated cell assembly sequences in the rat hippocampus. Science *321*, 1322–1327.

35. MacDonald, C.J., Lepage, K.Q., Eden, U.T., and Eichenbaum, H. (2011). Hippocampal "time cells" bridge the gap in memory for discontiguous events. Neuron *71*, 737–749.

36. Buonomano, D.V., and Mauk, M.D. (1994). Neural network model of the cerebellum: temporal discrimination and the timing of motor responses. Neural Comput. *6*, 38–55.

37. Harvey, C.D., Coen, P., and Tank, D.W. (2012). Choice-specific sequences in parietal cortex during a virtual-navigation decision task. Nature *484*, 62–68.

38. Machens, C.K., Romo, R., and Brody, C.D. (2010). Functional, but not anatomical, separation of "what" and "when" in prefrontal cortex. J. Neurosci. *30*, 350–360.

39. Shinomoto, S., Omi, T., Mita, A., Mushiake, H., Shima, K., Matsuzaka, Y., and Tanji, J. (2011). Deciphering elapsed time and predicting action timing from neuronal population signals. Front. Comput. Neurosci. *5*, 29.

40. Kim, J., Ghim, J.W., Lee, J.H., and Jung, M.W. (2013). Neural correlates of interval timing in rodent prefrontal cortex. J. Neurosci. *33*, 13834–13847.

41. Jin, D.Z., Fujii, N., and Graybiel, A.M. (2009). Neural representation of time in cortico-basal ganglia circuits. Proc. Natl. Acad. Sci. USA *106*, 19156–19161.

42. Merchant, H., Pérez, O., Zarco, W., and Gámez, J. (2013). Interval tuning in the primate medial premotor cortex as a general timing mechanism. J. Neurosci. *33*, 9082–9096.

A.2 Midbrain dopamine neurons control judgment of time

The following Section is a reprint of the article published as in Soares et al., 2016. Copyright information and permissions can be found at http://www.sciencemag.org/help/reprints-and-permissions.

and eventually merge close to the bottom of the image. As qualitatively evidenced by the dI/dU map of Fig. 4B and quantitatively supported by the line sections plotted at the bottom of this panel, the edge state disappears as soon as the step-step separation decreases below the spatial extent of the edge state (*25*)—i.e., about 10 nm. We further analyze the response of these edge states to high magnetic fields B. Figure 4C reports a dI/dU map (top) and STS data acquired on an odd step edge (bottom) at B = 11 T; contrary to the quantum spin Hall state found in HgTe, the 1D TCI state investigated here is robust against time-reversal symmetry breaking perturbations. Finally, Fig. 4D shows that the edge state also persists at elevated temperatures (T = 80 K). Despite the reduced intensity evidenced by the STS spectrum, a well-defined 1D channel is still clearly present.

The observation of a distinct type of one-dimensional states at odd step edges of topological crystalline insulators with relatively wide bulk band gaps opens up opportunities for the use of topological materials for sensing and information processing purposes well beyond existing materials (*4*, *10*, *11*). Furthermore, the absence of scattering and the high degree of spin polarization observed in tight-binding calculations indicate that the 1D midgap state might be useful for spintronics applications. By patterning the step-and-terrace structure of TCI surfaces, this may allow for the creation of well-separated conductive channels with a width of only about 10 nm. This may lead to interconnections between functional units at ultrahigh packing densities. To fully explore whether the 1D midgap state found at odd TCI step edges display quantum conductance effects, further investigations by, for example, four-probe transport measurements, will be needed.

REFERENCES AND NOTES

1. M. Z. Hasan, C. L. Kane, *Rev. Mod. Phys.* **82**, 3045–3067 (2010).
2. X.-L. Qi, S.-C. Zhang, *Rev. Mod. Phys.* **83**, 1057–1110 (2011).
3. M. König *et al.*, *Science* **318**, 766–770 (2007).
4. D. Hsieh *et al.*, *Nature* **452**, 970–974 (2008).
5. Y. L. Chen *et al.*, *Science* **325**, 178–181 (2009).
6. P. Roushan *et al.*, *Nature* **460**, 1106–1109 (2009).
7. T. Zhang *et al.*, *Phys. Rev. Lett.* **103**, 266803 (2009).
8. B. A. Bernevig, T. L. Hughes, S.-C. Zhang, *Science* **314**, 1757–1761 (2006).
9. L. Fu, C. L. Kane, E. J. Mele, *Phys. Rev. Lett.* **98**, 106803 (2007).
10. I. Knez, R.-R. Du, G. Sullivan, *Phys. Rev. Lett.* **107**, 136603 (2011).
11. T. Hirahara *et al.*, *Phys. Rev. Lett.* **107**, 166801 (2011).
12. J. Liu *et al.*, *Nat. Mater.* **13**, 178–183 (2014).
13. T. H. Hsieh *et al.*, *Nat. Commun.* **3**, 982 (2012).
14. P. Dziawa *et al.*, *Nat. Mater.* **11**, 1023–1027 (2012).
15. B. M. Wojek *et al.*, *Phys. Rev. B* **90**, 161202 (2014).
16. Y. Tanaka *et al.*, *Nat. Phys.* **8**, 800–803 (2012).
17. Y. Okada *et al.*, *Science* **341**, 1496–1499 (2013).
18. A. Gyenis *et al.*, *Phys. Rev. B* **88**, 125414 (2013).
19. D. Zhang *et al.*, *Phys. Rev. B* **89**, 245445 (2014).
20. See supplementary materials on *Science* Online.
21. I. Zeljkovic *et al.*, *Nat. Mater.* **14**, 318–324 (2015).
22. I. Zeljkovic *et al.*, *Nat. Phys.* **10**, 572–577 (2014).
23. C. Pauly *et al.*, *Nat. Phys.* **11**, 338–343 (2015).
24. W. P. Su, J. R. Schrieffer, A. J. Heeger, *Phys. Rev. Lett.* **42**, 1698–1701 (1979).
25. Y. Zhang *et al.*, *Nat. Phys.* **6**, 584–588 (2010).

ACKNOWLEDGMENTS

This research was supported by DFG (through SFB 1170 "ToCoTronics"; projects A02, B04, and C05) and by the Polish National Science Centre NCN grants 2014/15/B/ST3/03833 and 2012/07/B/ST3/03607. We further acknowledge support by the European Research Council (ERC) through ERC-StG-TOPOLECTRICS-Thomale-336012. D.D.S. and G.S. gratefully acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) for funding this project by providing computing time on the GCS Supercomputer SuperMUC at Leibniz Supercomputing Centre (LRZ, www.lrz.de).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/354/6317/1269/suppl/DC1
Materials and Methods
SupplementaryText
Figs. S1 to S7
References (*26*, *27*)

23 July 2016; accepted 11 November 2016
10.1126/science.aah6233

BRAIN RESEARCH

Midbrain dopamine neurons control judgment of time

Sofia Soares,* Bassam V. Atallah,*† Joseph J. Paton†

Our sense of time is far from constant. For instance, time flies when we are having fun, and it slows to a trickle when we are bored. Midbrain dopamine neurons have been implicated in variable time estimation. However, a direct link between signals carried by dopamine neurons and temporal judgments is lacking. We measured and manipulated the activity of dopamine neurons as mice judged the duration of time intervals. We found that pharmacogenetic suppression of dopamine neurons decreased behavioral sensitivity to time and that dopamine neurons encoded information about trial-to-trial variability in time estimates. Last, we found that transient activation or inhibition of dopamine neurons was sufficient to slow down or speed up time estimation, respectively. Dopamine neuron activity thus reflects and can directly control the judgment of time.

Our ability to accurately estimate and reproduce time intervals is variable and depends on many factors, including motivation (*1*), attention (*2*), sensory change (*3*), novelty (*4*), and emotions (*5*). In addition, several neurological and neuropsychiatric disorders (*6*–*9*) are accompanied by changes in timing behavior. Midbrain dopamine (DA) neurons are implicated in many of the psychological factors (*10*) and disorders (*6*, *8*, *11*) associated with changes in time estimation.

Midbrain DA neurons also encode reward prediction errors (RPEs) (*12*–*15*), an important teaching signal in reinforcement learning (*16*). Phasic DA responses to reward-predicting cues reflect the magnitude of (*17*, *18*), probability of (*19*), and expected time delay until the reward (*20*, *21*). When expectation varies over time, DA neuron responses are smaller at times when rewards and reward-predicting cues are more expected (*21*, *22*), indicating that DA neurons receive temporal information. Manipulations of the DAergic system by pharmacological (*23*) or genetic (*24*) approaches disrupt timing behavior, suggesting that DA neurons may directly modulate timing. However, the data from pharmacological and genetic manipulations are inconsistent: In some cases, DA seems to speed up timekeeping (*23*, *25*), and in others, DA seems to slow down or not affect timekeeping (*26*, *27*).

Champalimaud Research, Champalimaud Centre for the Unknown, Lisbon, Portugal.
*These authors contributed equally to this work. **†Corresponding author. Email: bassam.atallah@gmail.com (B.V.A.); joe.paton@neuro.fchampalimaud.org (J.J.P.)**

To determine (i) what signals are encoded by midbrain DA neurons during timing behavior and (ii) how DA neurons contribute to variability in temporal judgments, we measured and manipulated the activity of DA neurons in mice as they performed categorical decisions about duration (*28*). We first trained mice to perform a temporal discrimination task (Fig. 1A, left). Mice initiated trials at a central nose port, immediately triggering the delivery of two identical tones separated by a variable delay. Mice reported the delay between tones as shorter or longer than 1.5 s at one of two lateral nose ports for water reward. Incorrect choices were not rewarded. Performance was nearly perfect for the easiest intervals but more variable for intervals near 1.5 s (the boundary between the "short" and "long" categories) and was well described by a sigmoid psychometric function (Fig. 1A, middle).

We then pharmacogenetically suppressed DAergic neuronal activity and observed impaired temporal judgments on treatment days as compared with adjacent nontreatment days (P < 0.004, n = 3 mice; Fig. 1A, right). We also observed a tendency to perform fewer trials [control group, 177 ± 15 trials; clozapine N-oxide (CNO)–treated group, 115 ± 54 trials; mean ± SD; P = 0.05], suggesting that the animals' motivation was affected by DAergic suppression. To test whether fluctuations in endogenous DA neuron activity predicted systematic changes in temporal judgments, we used fiber photometry (*29*) to measure Ca^{2+} activity in DAergic neurons, targeting the substantia nigra pars compacta (SNc) (Fig. 1, B and C, and figs. S1 and S2).

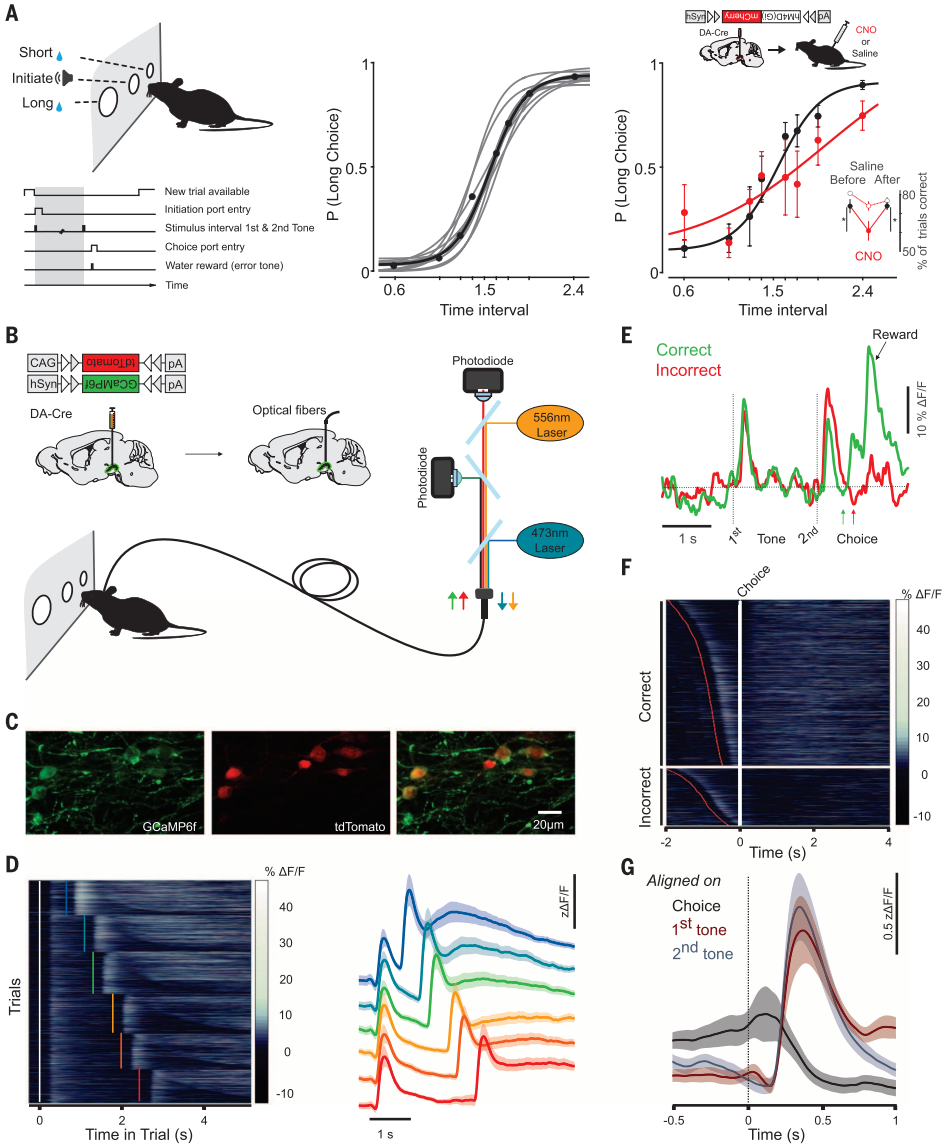


Fig. 1. Dopaminergic (DAergic) signaling is required and precisely aligned to temporal cues, not movement, during performance of a temporal categorization task. (**A**) Shown on the left is the task schematic and order of events (circles in the upper panel, nose-ports; gray shading in the lower panel, interval period). A logistic function fit to the daily (gray) and average (black) performance of an example mouse (10 sessions) is shown in the middle. Pharmacogenetic suppression (hM4D) was targeted to midbrain DAergic neurons, and mice were injected with either CNO or saline on adjacent days; shown on the right is mean psychometric performance on days with saline or CNO treatment (black or red, respectively; $n = 3$ mice). Error bars, SEM. The inset shows the percent of correct trials on days before and after CNO treatment in mice expressing hM4D (filled circles, $n = 3$; $^*P < 0.005$) or non–hM4D-expressing controls (open circles, $n = 4$). Error bars, SEM. (**B**) Schematic of the photometry apparatus and viral and surgical procedure. (**C**) Image of the substantia nigra pars compacta (SNc) histology. (**D**) On the left, all trials of DA neuronal activity recorded from a single subject are shown, split by interval duration and aligned on trial initiation (first tone delivery; white vertical line). Each row represents a trial, and within each interval, trials are sorted from fast (top) to slow (bottom) response time (RT, time from the second tone to choice; 3759 trials). Shown on the right are mean DAergic neuron responses, split by interval duration ($n = 5$ mice; intervals are color-coded as throughout). Shading, SEM across mice. *z*, *z*-score, $\Delta F/F$, see the methods. (**E**) Example photometric traces recorded during a single correct and incorrect trial of the 1.74-s interval. (**F**) Photometric recordings of DA neuronal activity from a single subject, split by outcome (correct choices, top; incorrect choices, bottom) and aligned on choice (white). Within each outcome, trials were sorted by RTs [slow (top) to fast (bottom)]. Red dots mark the time of second-tone presentation (2426 trials). (**G**) Mean DAergic responses of incorrect trials aligned on the three main task events (first tone, second tone, and choice; $n = 5$ mice). Shading, SEM across mice.

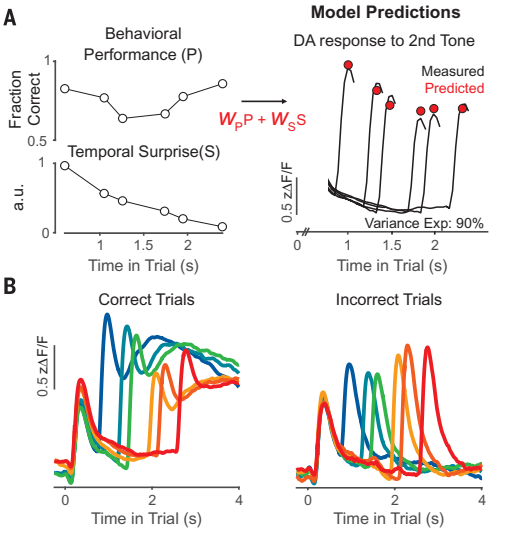


Fig. 2. DAergic responses correlate with temporal judgments and are explained by a simple model of reward prediction error (RPE). (**A**) Linear model (left) including RPE components: expectation of reward *P* (subject performance, top left) and temporal expectation *S* (surprise, the inverse of the subjective hazard function; bottom left). *w*, weight; a.u., arbitrary units. In the middle panel, measured second-tone DAergic response for six time intervals (black traces; $n = 5$ mice) are compared to predicted DA response (red dots). The graph on the right shows model predictions versus measured DAergic activity (gray symbols, individual mice; mean responses across mice, black filled circles). (**B**) Average measured DA response for all intervals during correct and incorrect trials. (**C**) Mean DA response to the second tone when an interval was judged as long versus short. Each shape represents a different mouse. Black symbols represent responses averaged across all interval stimuli.

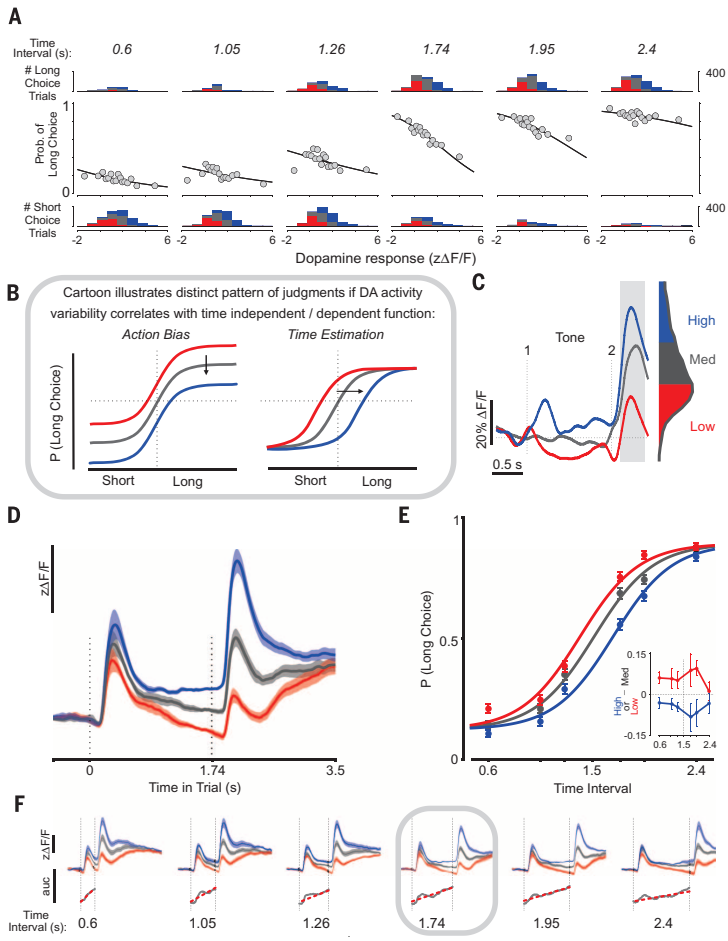


Fig. 3. Changes in a time-dependent component of choice behavior are predicted by DAergic activity. (**A**) Trial-by-trial logistic regression (black) that predicts choice from the amplitude of the second-tone DA response (gray), for each of the six time intervals (left to right). The top and bottom histograms illustrate the number of trials, as a function of DA response, in which the subject made long and short choices, respectively ($n = 8533$ trials, 5 mice). For each session and interval, DA responses are grouped into terciles—high (blue), medium (gray), and low (red)—throughout the figure. (**B**) Distinct patterns of temporal judgments are expected depending on the nature of the relationship between DA response and choice. (**C**) Three individual trials illustrating low, medium, and high second-tone DA responses (quantified as the mean response in the gray-shaded box) and grouped by tercile within the entire second-tone response distribution, depicted at right. (**D**) Average DA response in each tercile for the 1.74-s interval stimulus ($n = 1868$ trials, 5 mice). Shading, SEM. (**E**) Psychometric curves constructed using trials from each tercile of DA response. Curves are the maximum-likelihood fits of logistic functions with the lowest Bayesian information criterion scores ($n = 8533$ trials, 5 mice). Error bars, 95% confidence interval (CI). The inset shows the difference in the probability of making a long choice between medium and low or high (red or blue) DA response trials. Error bars, SEM. (**F**) The top row is as in (D) but for all six interval durations; data shown in (D) are outlined in gray. The bottom row shows the area under the curve (auc), distinguishing high- and low-tercile DA responses. This difference in DA response increased during the course of the trial (red linear regression; coefficient of determination r^2 ranging from 0.72 to 0.98; $P < 0.0001$).

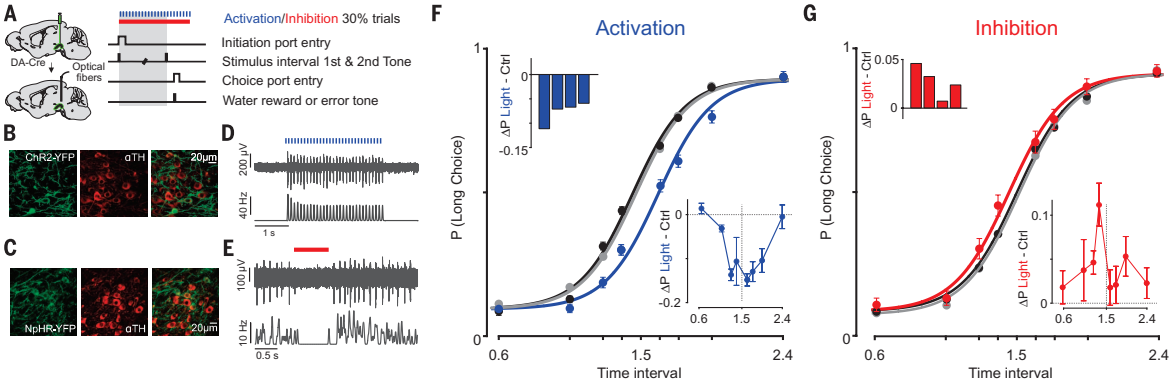


Fig. 4. Optogenetic manipulation of dopamine neurons is sufficient to change judgment of time. (**A**) Schematics illustrating viral strategy and subsequent fiber implantation (left) and stimulation protocol (right). (**B** and **C**) Histology confirming membrane expression of ChR2-YFP or NpHR-YFP (both green) in neurons of the SNc expressing tyrosine hydroxylase (TH, red). (**D** and **E**) Single-trial (top panels) and peri-stimulus time histogram (bottom panels) of in vivo electrophysiological measurement of two DA neurons reliably activated and inactivated by light (n = 53 and 8 trials, respectively). (**F**) Choice behavior and psychometric curves during control trials (black), photoactivated trials (blue), and unstimulated trials immediately after photoactivation (gray) (n = 4 mice). Error bars, 95% CI. Insets show the mean difference in the probability of a long choice between photoactivated and control trials (top, one bar per animal; bottom, one data point per stimulus). Error bars, SEM. (**G**) Same as (F) but for animals whose DA neurons were inhibited (n = 4).

We observed DAergic responses locked to the three main task events on single trials: the first tone, the second tone, and reward delivery (or omission thereof) (Fig. 1E). Activity increased after reward delivery and decreased when the reward was omitted in the case of incorrect choices (Fig. 1F) (*30*). DAergic signaling has also been implicated in movement; however, DA neuron activity in this task did not reflect movement per se (Fig. 1, F and G, and fig. S3).

In this task, the second tone marks the end of the interval to be discriminated and is a sensory cue that predicts reward. The amplitude of a RPE at the time of the second tone should be modulated by two factors: the subject's expectation of reward at tone delivery and their temporal expectation of the second tone itself. First, expectation of reward varies as a function of stimulus difficulty, where the more difficult the interval to be discriminated, the lower the probability of reward (Fig. 2A). Second, because delay intervals were randomly selected from the stimulus set on each trial, occurrence of the second tone becomes less surprising with time (Fig. 2A). Indeed, animals were sensitive to changing temporal expectation, as indicated by a systematic decrease in response time (RT, the delay between second-tone delivery and choice execution) with increasing interval duration (RT for the shortest interval greater than RT for the longest interval; $P < 0.005$ in each of five mice). To test whether second-tone responses reflected a RPE that integrated information about temporal expectation and expected reward, we asked how well the pattern of average responses to all six second tones could be explained by a linear combination of temporal expectation (i.e., surprise, the inverse of the subjective hazard function; fig. S4) and performance (the probability of reward for each stimulus). On average, 90% of variance in mean responses could be explained by a relatively equal contribution of these two factors (range, 58 to 99%; n = 5 mice; Fig. 2A). Reward responses were also consistent with RPE coding: Within a given choice category, they tended to be larger for intervals that animals miscategorized more often (fig. S5).

On average, DA neuron responses to the second tone contained information about elapsed time through their encoding of temporal expectation. Do these responses relate to variations in judgments of time? When animals correctly judged intervals, the response to the second tone was, on average, larger for intervals in the short category (Fig. 2B). However, on incorrect trials, the pattern was reversed: The response to the second tone was larger for intervals in the long category. Thus, DA response magnitude reflected the animals' assessment of the interval duration, not the actual interval duration. Over all intervals, the second-tone response for a given interval was significantly larger when that interval was judged as short ($P < 0.001$; Fig. 2, B and C). How do these results relate to the underlying decision and motor processes that guide choice during the task?

In principle, the trial-to-trial variations in DA neuron activity could be related to a time-dependent component of the decision, such as the speed of internal timekeeping or the location of the decision boundary in time. Alternatively, variations in DA activity might reflect a time-independent component of the behavior, such as a constant action bias. To quantitatively evaluate these two possibilities, we performed a logistic regression to assess the degree to which the magnitude of the DA neuron response to the second tone predicted animals' choices on single trials. We found that activity predicted choice to a lesser extent in the case of easy stimuli than in the case of difficult stimuli (Fig. 3A). These data suggest that the DA neuron response was systematically related to the horizontal position of the psychometric curve along the time axis and not the vertical position along the choice axis (Fig. 3B). To test this, we split trials into high, medium, and low terciles of the distribution of responses to the second tone [Fig. 3, A (histograms) and C]. While the second-tone response amplitude was used to group trials, the systematic ordering of DA neuron responses emerged toward the beginning of the trial and persisted throughout an interval (Fig. 3, D and F). We next constructed psychometric curves for trials in each tercile and compared a range of models for the psychometric curve. The model that best explained the behavioral data collected from high-, medium-, and low-tercile trials consisted of three sigmoid curves that differed only in their horizontal location along the time axis (Fig. 3E). We observed a shift toward long choices when DAergic activity was low, and the opposite shift when activity was high. Specifically, as DA activity varied from the lower to the upper tercile, the psychometric threshold shifted by ~340 ms (i.e., ~20% of the 1.5-s category boundary; range, 90 to 620 ms; 6 to 42%; n = 5 mice). The relationship between DAergic response and psychometric shift was observed for recordings in either hemisphere (fig. S6), thus ruling out an explanation based on the laterality of short versus long choices. Instead, these results indicate that higher or lower midbrain DAergic activity is correlated with a change in a time-dependent component of the decision.

How might this correlation between DA neuron activity and the location of the psychometric curve along the time axis relate to our initial finding that temporal expectation contributed to the average second-tone response? The theory of DAergic RPE coding predicts that slower (faster)

timekeeping, by stretching (contracting) temporal surprise along the time axis, should increase (decrease) DAergic responses to the second tone (fig. S7). We observed a pattern of DAergic response to the second tone that was consistent with this (Fig. 2, B and C, and fig. S7). Furthermore, if DAergic activity reflects RPE continuously throughout a trial, differences in activity associated with slower or faster timekeeping (i.e., the separation between low- and high-activity terciles) should also grow continuously over time, and indeed, this is the case in our data (Fig. 3F and fig. S7). In contrast to the expected impact of variability in the speed of timekeeping on RPE coding, it is not apparent to us how changes in the location of the decision boundary along an animal's internal notion of time should change RPEs arising at the presentation of the second tone. The most parsimonious explanation of the data is that DA neuron activity reflects variability in the speed of internal timekeeping.

These results demonstrate a correlation between temporal judgments and DA neuron activity. However, it is unclear whether DA neuron activity simply reflects, or whether it is sufficient to cause changes in, time judgments. We mimicked the observed variability in DAergic responses by optogenetically activating or inhibiting DA neurons (Fig. 4, A to E) on a minority of randomly chosen trials. Notably, we found that increasing or decreasing DA activity resulted in a horizontal shift in the psychometric curve in the directions predicted by the photometry data, albeit more modestly in the case of photoinhibition (excitation, 140 ± 20 ms, $n = 4$ mice; inhibition, −68 ± 23 ms, $n = 4$ mice; Fig. 4, F and G, and fig. S8). These effects were transient, occurring only on stimulated trials, and thus could not be explained as resulting from learning (Fig. 4, F and G), nor were they observed in control animals (fig. S9). In addition, as was the case when sorting trials on the basis of DA response to the second tone, we observed no systematic effect on RTs, arguing against DAergic neuron activity affecting the subjects' movement toward or incentive salience of choice options during the task (fig. S10).

Here we demonstrate a direct link between signals carried by midbrain DA neurons and judgments of elapsed time. Higher or lower levels of DAergic activity not only correlated with but could directly control timekeeping. These data are in agreement with some results of pharmacological manipulations of the DAergic system during timing tasks (*26*), but appear at odds with some others that showed accelerated timekeeping with increased DAergic tone (*23*, *25*). However, recent studies demonstrate that many of the pharmacological effects on timing behavior can be explained by the changes in motivation (*27*, *31*) that accompany DAergic drug administration (*32*). Indeed, pharmacogenetic DAergic manipulation in our task affected motivated behavior. Variability in the effects of pharmacology on timing may result from its relatively slow time course, which allows for compensation and/or the superposition of multiple distinct behavioral effects. Our approach circumvents these issues with genetically targeted, transient manipulations of DA neuron activity. Additionally, we focused on DA neurons in the SNc because many project to a dorsocentral region of the striatum where removal of DA input can cause a selective deficit in timing (*33*); however, whether DA neurons in other regions, such as the ventral tegmental area, contribute to timing variability is unknown. Last, we monitored and manipulated the activity of midbrain DA neurons, and not the levels of released DA. The relationship between tonic and phasic firing of DA neurons and DA release is not entirely clear, and it is complicated by feedback mechanisms by which released DA can affect the firing of DA neurons (*34*).

Although unexpected, the data presented here may explain existing behavioral data. Situations in which DAergic activity is elevated naturally, such as states of high approach motivation (*35*), response uncertainty (*36*), or cognitive engagement (*37*), are associated with underestimation of time (*1*, *2*, *38*). Conversely, situations that decrease DAergic activity, such as when fearful or aversive stimuli are presented (*39*), are associated with overestimation of time (*40*). These observations, together with our data, suggest that flexibility in time estimation may confer an adaptive advantage on the individual. For example, underestimating duration in better-than-expected situations may lead to longer engagement in those situations, resulting in even greater reward than if time estimation were not flexible. In other words, there may be a normative explanation for why "time flies when we are having fun" underlying our observation that DA neurons, which are so central to reward processing, exert control over time estimation.

REFERENCES AND NOTES

1. P. A. Gable, B. D. Poole, *Psychol. Sci.* **23**, 879–886 (2012).
2. J. T. Coull, F. Vidal, B. Nazarian, F. Macar, *Science* **303**, 1506–1508 (2004).
3. M. B. Ahrens, M. Sahani, *Curr. Biol.* **21**, 200–206 (2011).
4. V. Pariyadath, D. Eagleman, *PLOS ONE* **2**, e1264 (2007).
5. S. Droit-Volet, W. H. Meck, *Trends Cogn. Sci.* **11**, 504–513 (2007).
6. M. A. Pastor, J. Artieda, M. Jahanshahi, J. A. Obeso, *Brain* **115**, 211–225 (1992).
7. M. Wittmann, D. S. Leland, J. Churan, M. P. Paulus, *Drug Alcohol Depend.* **90**, 183–192 (2007).
8. V. Noreika, C. M. Falter, K. Rubia, *Neuropsychologia* **51**, 235–266 (2013).
9. O. F. Wahl, D. Sieg, *Percept. Mot. Skills* **50**, 535–541 (1980).
10. R. Cools, *Neuroscientist* **14**, 381–395 (2008).
11. A. Lüthi, C. Lüscher, *Nat. Neurosci.* **17**, 1635–1643 (2014).
12. W. Schultz, P. Dayan, P. R. Montague, *Science* **275**, 1593–1599 (1997).
13. H. M. Bayer, P. W. Glimcher, *Neuron* **47**, 129–141 (2005).
14. N. Eshel, J. Tian, M. Bukwich, N. Uchida, *Nat. Neurosci.* **19**, 479–486 (2016).
15. E. E. Steinberg *et al.*, *Nat. Neurosci.* **16**, 966–973 (2013).
16. R. S. Sutton, A. G. Barto, *Introduction to Reinforcement Learning*, vol. 135 (MIT Press, 1998).
17. P. N. Tobler, C. D. Fiorillo, W. Schultz, *Science* **307**, 1642–1645 (2005).
18. J. Y. Cohen, S. Haesler, L. Vong, B. B. Lowell, N. Uchida, *Nature* **482**, 85–88 (2012).
19. C. D. Fiorillo, P. N. Tobler, W. Schultz, *Science* **299**, 1898–1902 (2003).
20. S. Kobayashi, W. Schultz, *J. Neurosci.* **28**, 7837–7846 (2008).
21. C. D. Fiorillo, W. T. Newsome, W. Schultz, *Nat. Neurosci.* **11**, 966–973 (2008).
22. B. Pasquereau, R. S. Turner, *J. Neurophysiol.* **113**, 1110–1123 (2015).
23. A. V. Maricq, R. M. Church, *Psychopharmacology* **79**, 10–15 (1983).
24. M. R. Drew *et al.*, *J. Neurosci.* **27**, 7731–7739 (2007).
25. C. V. Buhusi, W. H. Meck, *Behav. Neurosci.* **116**, 291–297 (2002).
26. J. I. Lake, W. H. Meck, *Neuropsychologia* **51**, 284–292 (2013).
27. F. Balci *et al.*, *Brain Res.* **1325**, 89–99 (2010).
28. T. S. Gouvêa *et al.*, *eLife* **4**, e11386 (2015).
29. S. P. dos Santos Matias, E. Lottem, G. P. Dugue, Z. F. Mainen, http://biorxiv.org/content/early/2016/06/18/059758 (2016).
30. W. Schultz, P. Apicella, T. Ljungberg, *J. Neurosci.* **13**, 900–913 (1993).
31. A. L. Odum, L. M. Lieving, D. W. Schaai, *J. Exp. Anal. Behav.* **78**, 195–214 (2002).
32. B. Panigrahi *et al.*, *Cell* **162**, 1418–1430 (2015).
33. W. H. Meck, *Brain Res.* **1109**, 93–107 (2006).
34. B. S. Bunney, G. K. Aghajanian, *Naunyn Schmiedeberg's Arch. Pharmacol.* **304**, 255–261 (1978).
35. E. S. Bromberg-Martin, M. Matsumoto, O. Hikosaka, *Neuron* **67**, 144–155 (2010).
36. V. de Lafuente, R. Romo, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 19767–19771 (2011).
37. I. Fried *et al.*, *Nat. Neurosci.* **4**, 201–206 (2001).
38. R. E. Hicks, G. W. Miller, M. Kinsbourne, *Am. J. Psychol.* **89**, 719–730 (1976).
39. E. B. Oleson, R. N. Gentry, V. C. Chioma, J. F. Cheer, *J. Neurosci.* **32**, 14804–14808 (2012).
40. F. N. Watts, R. Sharrock, *Percept. Mot. Skills* **59**, 597–598 (1984).

ACKNOWLEDGMENTS

We thank A. Braga for assistance with behavioral training; M. Duarte for assistance with mouse colonies; G. Lopes for assistance with Bonsai; T. Monteiro, T. Gouvêa, other members of the Paton laboratory, B. Lau, E. Lottem, M. Murakami, C. Poo, A. Renart, and T. Akam for discussions and/or comments on the manuscript; Z. Mainen for support; platforms at the Champalimaud Centre for histology support and animal care; and V. Jayaraman, R. A. Kerr, D. S. Kim, L. L. Looger, and K. Svoboda from the GENIE (Genetically-Encoded Neuronal Indicator and Effector) Project at the Howard Hughes Medical Institute's Janelia Farm Research Campus for providing the AAV-GCaMP6f through the University of Pennsylvania Vector Core. Viruses for expression of NpHR3.0 and EYFP are available from the University of North Carolina Vector Core under a material transfer agreement with K. Deisseroth. Viruses for expression of GCaMP6f and TdTomato are available from the University of Pennsylvania Vector Core under a material transfer agreement with the trustees of the University of Pennsylvania on behalf of J. Wilson. The work was funded by the Bial Foundation (188/12 to J.J.P.), the Simons Foundation (Simons Collaboration on the Global Brain award 325476 to J.J.P.), Fundação para Ciência e Tecnologia (SFRH/BD/51895/2012 to S.S.), the European Molecular Biology Organization (Advanced Long Term Fellowship 983-2012 to B.V.A.), Marie Curie Actions (FP7-PEOPLE-2012-IIF 326398 to B.V.A.), and the Champalimaud Foundation (internal funding to J.J.P.). Data presented in this paper can be found at www.dropbox.com/sh/ip6forddl84028j/AAAsa3ry41bu4acYk1Bl3KDra?dl=0.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/354/6317/1273/suppl/DC1
Materials and Methods
Figs. S1 to S10
References (*41*–*43*)

8 July 2016; accepted 4 November 2016
10.1126/science.aah5234

References

Adler, A., Katabi, S., Finkes, I., Israel, Z., Prut, Y., & Bergman, H. (2012). Temporal convergence of dynamic cell assemblies in the striato-pallidal network. *Journal of Neuroscience*, *32*(7), 2473–2484.

Ahrens, M. B., & Sahani, M. (2011). Observers exploit stochastic models of sensory change to help judge the passage of time. *Current Biology*, *21*(3), 200–206.

Anderson, D. J. (2016). Circuit modules linking internal states and social behaviour in flies and mice. *Nature Reviews Neuroscience*, *17*(11), 692–704.

Artieda, J., Pastor, M. A., Lacruz, F., & Obeso, J. A. (1992). Temporal discrimination is abnormal in parkinson's disease. *Brain*, *115*(1), 199–210.

Bakhurin, K. I., Goudar, V., Shobe, J. L., Claar, L. D., Buonomano, D. V., & Masmanidis, S. C. (2017). Differential encoding of time by prefrontal and striatal network dynamics. *Journal of Neuroscience*, *37*(4), 854–870.

Balcı, F. (2014). Interval timing, dopamine, and motivation. *Timing & Time Perception*, *2*(3), 379–410.

Balcı, F., Ludvig, E. A., Abner, R., Zhuang, X., Poon, P., & Brunner, D. (2010). Motivational effects on interval timing in dopamine transporter (dat) knockdown mice. *Brain Research*, *1325*, 89–99.

Balcı, F., Papachristos, E., Gallistel, C., Brunner, D., Gibson, J., & Shumyatsky, G. (2008). Interval timing in genetically modified mice: a simple paradigm. *Genes, Brain and Behavior*, *7*(3), 373–384.

Balleine, B. W., & Dickinson, A. (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology*, *37*(4), 407–419.

Balsam, P. D., & Gallistel, C. R. (2009). Temporal maps and informativeness in associative learning. *Trends in neurosciences*, *32*(2), 73–78.

Barto, A. G., Sutton, R. S., & Anderson, C. W. (1983). Neuronlike adaptive elements that can solve difficult learning control problems. *IEEE transactions on systems, man, and cybernetics*(5), 834–846.

Bayer, H. M., & Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron*, *47*(1), 129–141.

Berditchevskaia, A., Cazé, R., & Schultz, S. R. (2016). Performance in a go/nogo perceptual task reflects a balance between impulsive and instrumental components of behaviour. *Scientific reports*, *6*.

Berke, J. D., Okatan, M., Skurski, J., & Eichenbaum, H. B. (2004). Oscillatory entrainment of striatal neurons in freely moving rats. *Neuron*, *43*(6), 883–896.

Berns, G. S., & Sejnowski, T. J. (1998). A computational model of how the basal ganglia produce sequences. *Journal of cognitive neuroscience*, *10*(1), 108–121.

Berridge, K. C. (2004). Motivation concepts in behavioral neuroscience. *Physiology & behavior*, *81*(2), 179–209.

Bertler, Å., & Rosengren, E. (1959). Occurrence and distribution of dopamine in brain and other tissues. *Cellular and Molecular Life Sciences*, *15*(1), 10–11.

Björklund, A., & Dunnett, S. B. (2007). Dopamine neuron systems in the brain: an update. *Trends in neurosciences*, *30*(5), 194–202.

Braitenberg, V. (1967). Is the cerebellar cortex a biological clock in the millisecond range? *Progress in brain research*, *25*, 334–346.

Bromberg-Martin, E. S., Matsumoto, M., & Hikosaka, O. (2010). Distinct tonic and phasic anticipatory activity in lateral habenula and dopamine neurons. *Neuron*, *67*(1), 144–155.

Bromberg-Martin, E. S., Matsumoto, M., Hong, S., & Hikosaka, O. (2010). A pallidus-habenula-dopamine pathway signals inferred stimulus values. *Journal of neurophysiology*, *104*(2), 1068–1076.

Brunton, B. W., Botvinick, M. M., & Brody, C. D. (2013). Rats and humans can optimally accumulate evidence for decision-making. *Science*, *340*(6128), 95–98.

Buhusi, C. V., & Meck, W. H. (2002). Differential effects of methamphetamine and haloperidol on the control of an internal clock. *Behavioral neuroscience*, *116*(2), 291.

Buhusi, C. V., & Meck, W. H. (2005). What makes us tick? functional and neural mechanisms of interval timing. *Nature reviews. Neuroscience*, *6*(10), 755.

Bunney, B., & Aghajanian, G. (1978). d-amphetamine-induced depression of central dopamine neurons: evidence for mediation by both autoreceptors and a striato-nigral feedback pathway. *Naunyn-Schmiedeberg's archives of pharmacology*, *304*(3), 255–261.

Buonomano, D. V. (2014). Neural dynamics based timing in the subsecond to seconds range. In *Neurobiology of interval timing* (pp. 101–117). Springer.

Buonomano, D. V., & Laje, R. (2010). Population clocks: motor timing with

neural dynamics. *Trends in cognitive sciences*, *14*(12), 520–527.

Buonomano, D. V., & Maass, W. (2009). State-dependent computations: spatiotemporal processing in cortical networks. *Nature reviews. Neuroscience*, *10*(2), 113.

Buonomano, D. V., & Mauk, M. D. (1994). Neural network model of the cerebellum: temporal discrimination and the timing of motor responses. *Neural computation*, *6*(1), 38–55.

Buonomano, D. V., Merzenich, M. M., et al. (1995). Temporal information transformed into a spatial code by a neural network with realistic properties. *Science*, *267*(5200), 1028–1030.

Carandini, M., & Churchland, A. K. (2013). Probing perceptual decisions in rodents. *Nature neuroscience*, *16*(7), 824–831.

Carlsson, A. (1959). The occurrence, distribution and physiological role of catecholamines in the nervous system. *Pharmacological reviews*, *11*(2), 490–493.

Carlsson, A., Lindqvist, M., & Magnusson, T. (1957). 3, 4-dihydroxyphenylalanine and 5-hydroxytryptophan as reserpine antagonists. *Nature*, *180*(4596), 1200–1200.

Carlsson, A., Lindqvist, M., Magnusson, T., & Waldeck, B. (1958). On the presence of 3-hydroxytyramine in brain. *Science*, *127*(3296), 471–471.

Castle, L., Aubert, R. E., Verbrugge, R. R., Khalid, M., & Epstein, R. S. (2007). Trends in medication treatment for adhd. *Journal of attention disorders*, *10*(4), 335–342.

Catania, A. (1970). Reinforcement schedules and psychophysical judgment: A study of some temporal properties of behavior. *The theory of reinforcement schedules*.

Catania, A. C., & Reynolds, G. (1968). A quantitative analysis of the responding maintained by interval schedules of reinforcement. *Journal of the Experimental analysis of behavior*, *11*(3S2), 327–383.

Çevik, M. Ö. (2003). Effects of methamphetanine on duration discrimination. *Behavioral neuroscience*, *117*(4), 774.

Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theoretical population biology*, *9*(2), 129–136.

Cheyne, J. A., Solman, G. J., Carriere, J. S., & Smilek, D. (2009). Anatomy of an error: A bidirectional state model of task engagement/disengagement and attention-related errors. *Cognition*, *111*(1), 98–113.

Church, R. M., & Deluty, M. Z. (1977). Bisection of temporal intervals. *Journal of Experimental Psychology: Animal Behavior Processes*, *3*(3), 216.

Church, R. M., Meck, W. H., & Gibbon, J. (1994). Application of scalar timing theory to individual trials. *Journal of Experimental Psychology: Animal Behavior Processes*, *20*(2), 135.

Clarke, S., Ivry, R., Grinband, J., Roberts, S., & Shimizu, N. (1996). Exploring the domain of the cerebellar timing system. *Advances in psychology*, *115*, 257–280.

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*(7383), 85.

Constantino, S. M., & Daw, N. D. (2015). Learning the opportunity cost of time in a patch-foraging task. *Cognitive, Affective, & Behavioral Neuroscience*, *15*(4), 837–853.

Cools, R. (2008). Role of dopamine in the motivational and cognitive control of behavior. *The Neuroscientist*, *14*(4), 381–395.

Cotzias, G. C., Papavasiliou, P. S., & Gellene, R. (1969). Modification of parkinsonism - chronic treatment with l-dopa. *New England Journal of Medicine*, *280*(7), 337–345.

Cotzias, G. C., Van Woert, M. H., & Schiffer, L. M. (1967). Aromatic amino acids and modification of parkinsonism. *New England Journal of Medicine*, *276*(7), 374–379.

Coull, J. T., Vidal, F., Nazarian, B., & Macar, F. (2004). Functional anatomy of the attentional modulation of time estimation. *Science*, *303*(5663), 1506–1508.

Cowles, J. T., & Finan, J. L. (1941). An improved method for establishing temporal discrimination in white rats. *The Journal of Psychology*, *11*(2), 335–342.

Creelman, C. D. (1962). Human discrimination of auditory duration. *The Journal of the Acoustical Society of America*, *34*(5), 582–593.

Cumming, W., & Schoenfeld, W. (1958). Behavior under extended exposure to a high-value fixed interval reinforcement schedule. *Journal of the Experimental Analysis of Behavior*, *1*(3), 245–263.

Daw, N. D., Courville, A. C., & Touretzky, D. S. (2006). Representation and timing in theories of the dopamine system. *Neural computation*, *18*(7), 1637–1677.

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, *69*(6), 1204–1215.

Dayan, P., & Abbott, L. F. (2005). *Theoretical neuroscience* (Vol. 806). Cambridge, MA: MIT Press.

de Lafuente, V., & Romo, R. (2011). Dopamine neurons code subjective sensory experience and uncertainty of perceptual decisions. *Proceedings of the National Academy of Sciences*, *108*(49), 19767–19771.

Dews, P. (1978). Studies on responding under fixed-interval schedules of reinforcement: Ii. the scalloped pattern of the cumulative record. *Journal of*

the Experimental Analysis of Behavior, *29*(1), 67–75.

Dias-Ferreira, E., Sousa, J. C., Melo, I., Morgado, P., Mesquita, A. R., Cerqueira, J. J., ... Sousa, N. (2009). Chronic stress causes frontostriatal reorganization and affects decision-making. *Science*, *325*(5940), 621–625.

Dodson, P. D., Dreyer, J. K., Jennings, K. A., Syed, E. C., Wade-Martins, R., Cragg, S. J., ... Magill, P. J. (2016). Representation of spontaneous movement by dopaminergic neurons is cell-type selective and disrupted in parkinsonism. *Proceedings of the National Academy of Sciences*, *113*(15), E2180–E2188.

Doya, K. (1999). What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural networks*, *12*(7), 961–974.

Drew, M. R., Simpson, E. H., Kellendonk, C., Herzberg, W. G., Lipatova, O., Fairhurst, S., ... Balsam, P. D. (2007). Transient overexpression of striatal d2 receptors impairs operant motivation and interval timing. *Journal of Neuroscience*, *27*(29), 7731–7739.

Droit-Volet, S., & Meck, W. H. (2007). How emotions colour our perception of time. *Trends in Cognitive Sciences*, *11*(12), 504–513.

Droit-Volet, S., Ramos, D., Bueno, J. L., & Bigand, E. (2013). Music, emotion, and time perception: the influence of subjective emotional valence and arousal? *Frontiers in psychology*, *4*.

Ehringer, H., & Hornykiewicz, O. (1960). Distribution of noradrenaline and dopamine (3-hydroxytyramine) in the human brain and their behavior in diseases of the extrapyramidal system. *Klinische Wochenschrift*, *38*, 1236–1239.

Eichenbaum, H. (2014). Time cells in the hippocampus: a new dimension for mapping memories. *Nature Reviews. Neuroscience*, *15*(11), 732.

Emmons, E. B., De Corte, B. J., Kim, Y., Parker, K. L., Matell, M. S., & Narayanan, N. S. (2017). Rodent medial frontal control of temporal processing in the dorsomedial striatum. *Journal of Neuroscience*, *37*(36), 8718–8733.

Eshel, N., Bukwich, M., Rao, V., Hemmelder, V., Tian, J., & Uchida, N. (2015). Arithmetic and local circuitry underlying dopamine prediction errors. *Nature*, *525*(7568), 243.

Eshel, N., Tian, J., Bukwich, M., & Uchida, N. (2016). Dopamine neurons share common response function for reward prediction error. *Nature Neuroscience*, *19*(3), 479–486.

Falk, J. L., & Bindra, D. (1954). Judgment of time as a function of serial position and stress. *Journal of Experimental Psychology*, *47*(4), 279.

Fechner, G. (1966). Elements of psychophysics. vol. i.

Fentress, J., Stanfield, B., & Cowan, W. (1981). Observations on the development of the striatum in mice and rats. *Anatomy and embryology*, *163*(3), 275–

298.
Ferster, C. B., & Skinner, B. F. (1957). Schedules of reinforcement.
Fiorillo, C. D., Newsome, W. T., & Schultz, W. (2008). The temporal precision of reward prediction in dopamine neurons. *Nature neuroscience*, *11*(8), 966–973.
Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*, *299*(5614), 1898–1902.
Frank, G. K., Oberndorfer, T. A., Simmons, A. N., Paulus, M. P., Fudge, J. L., Yang, T. T., & Kaye, W. H. (2008). Sucrose activates human taste pathways differently from artificial sweetener. *Neuroimage*, *39*(4), 1559–1569.
Freeman, W. J., & Skarda, C. A. (1985). Spatial eeg patterns, non-linear dynamics and perception: the neo-sherringtonian view. *Brain Research Reviews*, *10*(3), 147–175.
Freeman, W. J., & van Dijk, B. W. (1987). Spatial patterns of visual cortical fast eeg during conditioned reflex in a rhesus monkey. *Brain research*, *422*(2), 267–276.
Fried, I., Wilson, C. L., Morrow, J. W., Cameron, K. A., Behnke, E. D., Ackerson, L. C., & Maidment, N. T. (2001). Increased dopamine release in the human amygdala during performance of cognitive tasks. *Nature neuroscience*, *4*(2), 201.
Fung, B., Murawski, C., & Bode, S. (2017). Caloric primary rewards systematically alter time perception.
Gable, P. A., & Poole, B. D. (2012). Time flies when you're having approach-motivated fun: Effects of motivational intensity on time perception. *Psychological science*, *23*(8), 879–886.
Gage, G. J., Stoetzner, C. R., Wiltschko, A. B., & Berke, J. D. (2010). Selective activation of striatal fast-spiking interneurons during choice execution. *Neuron*, *67*(3), 466–479.
Gallistel, C. R., & Gibbon, J. (2000). Time, rate, and conditioning. *Psychological review*, *107*, 289.
Geffen, M. N., Broome, B. M., Laurent, G., & Meister, M. (2009). Neural encoding of rapidly fluctuating odors. *Neuron*, *61*(4), 570–586.
Gerfen, C. R., & Bolam, J. P. (2010). The neuroanatomical organization of the basal banglia. In *Hanbook of basal ganglia structure and function* (pp. 3–28).
Gershman, S. J., Moustafa, A. A., & Ludvig, E. A. (2014). Time representation in reinforcement learning models of the basal ganglia. *Frontiers in computational neuroscience*, *7*.
Gibbon, J. (1977). Scalar expectancy theory and weber's law in animal timing. *Psychological review*, *84*(3), 279.

Gibbon, J., & Church, R. M. (1984). Sources of variance in an information processing theory of timing. In H. L. Roitblat, H. S. Terrace, & T. G. Bever (Eds.), *Animal cognition* (Vol. 1, p. 465). Psychology Press.

Gibbon, J., & Church, R. M. (1992). Comparison of variance and covariance patterns in parallel and serial theories of timing. *Journal of the experimental analysis of behavior*, *57*(3), 393–406.

Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar timing in memory. *Annals of the New York Academy of sciences*, *423*(1), 52–77.

Gibbon, J., & Malapani, C. (2006). Time perception and timing, neural basis of. In *Encyclopedia of cognitive science.* John Wiley & Sons, Ltd.

Gmehlin, D., Fuermaier, A. B., Walther, S., Debelak, R., Rentrop, M., Westermann, C., ... others (2014). Intraindividual variability in inhibitory function in adults with adhd - an ex-gaussian approach. *PloS one*, *9*(12), e112298.

Goel, A., & Buonomano, D. V. (2014). Timing as an intrinsic property of neural networks: evidence.

Goel, A., & Buonomano, D. V. (2016). Temporal interval learning in cortical cultures is encoded in intrinsic network dynamics. *Neuron*, *91*(2), 320–327.

Gooch, C. M., Wiener, M., Wencil, E. B., & Coslett, H. B. (2010). Interval timing disruptions in subjects with cerebellar lesions. *Neuropsychologia*, *48*(4), 1022–1031.

Gouvêa, T. S., Monteiro, T., Motiwala, A., Soares, S., Machens, C., & Paton, J. J. (2015). Striatal dynamics explain duration judgments. *Elife*, *4*, e11386.

Gouvêa, T. S., Monteiro, T., Soares, S., Atallah, B. V., & Paton, J. J. (2014). Ongoing behavior predicts perceptual report of interval duration. *Frontiers in neurorobotics*, *8*.

Graybiel, A. M., Aosaki, T., Flaherty, A. W., Kimura, M., et al. (1994). The basal ganglia and adaptive motor control. *Science*, 1826–1826.

Gremel, C. M., & Costa, R. M. (2013). Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. *Nature communications*, *4*, 2264.

Grondin, S. (2010). Timing and time perception: a review of recent behavioral and neuroscience findings and theoretical directions. *Attention, Perception, & Psychophysics*, *72*(3), 561–582.

Grossberg, S., & Merrill, J. W. (1992). A neural network model of adaptively timed reinforcement learning and hippocampal dynamics. *Cognitive brain research*, *1*(1), 3–38.

Grossberg, S., & Schmajuk, N. A. (1989). Neural dynamics of adaptive timing and temporal discrimination during associative learning. *Neural Networks*, *2*(2), 79–102.

Grossmann, K. E. (1973). Continuous, fixed-ratio, and fixed-interval reinforce-

ment in honey bees. *Journal of the Experimental Analysis of Behavior*, *20*(1), 105–109.

Hamid, A. A., Pettibone, J. R., Mabrouk, O. S., Hetrick, V. L., Schmidt, R., Vander Weele, C. M., ... Berke, J. D. (2016). Mesolimbic dopamine signals the value of work. *Nature neuroscience*, *19*(1), 117–126.

Hanks, T. D., Kopec, C. D., Brunton, B. W., Duan, C. A., Erlich, J. C., & Brody, C. D. (2015). Distinct relationships of parietal and prefrontal cortices to evidence accumulation. *Nature*, *520*(7546), 220–223.

Harrington, D. L., Lee, R. R., Boyd, L. A., Rapcsak, S. Z., & Knight, R. T. (2004). Does the representation of time depend on the cerebellum? effect of cerebellar stroke. *Brain*, *127*(3), 561–574.

Harvey, C. D., Coen, P., & Tank, D. W. (2012). Choice-specific sequences in parietal cortex during a virtual-navigation decision task. *Nature*, *484*(7392), 62.

Heim, A. (1892). Notizen über den tod durch absturz. *Jahrbuch des Schweizer Alpenclub*, *27*, 327–37.

Hicks, R. E., Miller, G. W., & Kinsbourne, M. (1976). Prospective and retrospective judgments of time as a function of amount of information processed. *The American journal of psychology*, 719–730.

Hinton, S. C., & Meck, W. H. (2004). Frontal-striatal circuitry activated by human peak-interval timing in the supra-seconds range. *Cognitive Brain Research*, *21*(2), 171–182.

Hoagland, H. (1933). The physiological control of judgments of duration: Evidence for a chemical clock. *The Journal of General Psychology*, *9*(2), 267–287.

Hollerman, J. R., & Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nature neuroscience*, *1*(4).

Holson, R. R., Bowyer, J. F., Clausing, P., & Gough, B. (1996). Methamphetamine-stimulated striatal dopamine release declines rapidly over time following microdialysis probe insertion. *Brain research*, *739*(1), 301–307.

Hornykiewicz, O. (1963). The tropical localization and content of noradrenalin and dopamine (3-hydroxytyramine) in the substantia nigra of normal persons and patients with parkinson's disease. *Wiener klinische Wochenschrift*, *75*, 309.

Horvitz, J. C. (2000). Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. *Neuroscience*, *96*(4), 651–656.

Howard, C. D., Li, H., Geddes, C. E., & Jin, X. (2017). Dynamic nigrostriatal dopamine biases action selection. *Neuron*, *93*(6), 1436–1450.

Howard, M. W., MacDonald, C. J., Tiganj, Z., Shankar, K. H., Du, Q., Hasselmo,

M. E., & Eichenbaum, H. (2014). A unified mathematical framework for coding time, space, and sequences in the hippocampal region. *Journal of Neuroscience*, *34*(13), 4692–4707.

Howe, M., & Dombeck, D. (2016). Rapid signalling in distinct dopaminergic axons during locomotion and reward. *Nature*, *535*(7613), 505–510.

Ivry, R. B., & Keele, S. W. (1989). Timing functions of the cerebellum. *Journal of cognitive neuroscience*, *1*(2), 136–152.

Ivry, R. B., & Schlerf, J. E. (2008). Dedicated and intrinsic models of time perception. *Trends in cognitive sciences*, *12*(7), 273–280.

Jaldow, E. J., Oakley, D. A., & Davey, G. C. (1989). Performance of decorticated rats on fixed interval and fixed time schedules. *European Journal of Neuroscience*, *1*(5), 461–470.

James, W. (1886). The perception of time. *The Journal of speculative philosophy*, *20*(4), 374–407.

Janssen, P., & Shadlen, M. N. (2005). A representation of the hazard rate of elapsed time in macaque area lip. *Nature neuroscience*, *8*(2), 234–241.

Jazayeri, M., & Movshon, J. A. (2006). Optimal representation of sensory information by neural populations. *Nature neuroscience*, *9*(5), 690.

Jazayeri, M., & Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nature neuroscience*, *13*(8), 1020–1026.

Jazayeri, M., & Shadlen, M. N. (2015). A neural mechanism for sensing and reproducing a time interval. *Current Biology*, *25*(20), 2599–2609.

Jin, D. Z., Fujii, N., & Graybiel, A. M. (2009). Neural representation of time in cortico-basal ganglia circuits. *Proceedings of the National Academy of Sciences*, *106*(45), 19156–19161.

Jin, X., & Costa, R. M. (2010). Start/stop signals emerge in nigrostriatal circuits during sequence learning. *Nature*, *466*(7305), 457.

Jordan, R. (1955). Time and contingency in st. augustine. *The Review of Metaphysics*, 394–417.

Judd, C., et al. (1899). Time-sense.

Kamin, L. J. (1969). Predictability, surprise, attention, and conditioning. *Punishment and aversive behavior*.

Kant, I. (1770). Inaugural dissertation: On the form and principles of the sensible and the intelligible world. *Kant. Selected Pre-Critical Writings*, 45–92.

Karmarkar, U. R., & Buonomano, D. V. (2007). Timing in the absence of clocks: encoding time in neural network states. *Neuron*, *53*(3), 427–438.

Killeen, P. R., & Fetterman, J. G. (1988). A behavioral theory of timing. *Psychological review*, *95*(2), 274.

Kim, H., Sul, J. H., Huh, N., Lee, D., & Jung, M. W. (2009). Role of striatum in updating values of chosen actions. *Journal of neuroscience*, *29*(47), 14701–14712.

Kim, J., Ghim, J.-W., Lee, J. H., & Jung, M. W. (2013). Neural correlates of interval timing in rodent prefrontal cortex. *Journal of Neuroscience*, *33*(34), 13834–13847.

Kim, J., Jung, A. H., Byun, J., Jo, S., & Jung, M. W. (2009). Inactivation of medial prefrontal cortex impairs time interval discrimination in rats. *Frontiers in behavioral neuroscience*, *3*.

Kobayashi, S., & Schultz, W. (2008). Influence of reward delays on responses of dopamine neurons. *Journal of neuroscience*, *28*(31), 7837–7846.

Kolling, N., & Akam, T. (2017). (reinforcement?) learning to forage optimally. *Current Opinion in Neurobiology*, *46*, 162–169.

Kreitzer, A. C. (2009). Physiology and pharmacology of striatal neurons. *Annual review of neuroscience*, *32*, 127–147.

Lake, J. I., & Meck, W. H. (2013). Differential effects of amphetamine and haloperidol on temporal reproduction: dopaminergic regulation of attention and clock speed. *Neuropsychologia*, *51*(2), 284–292.

Lammel, S., Ion, D. I., Roeper, J., & Malenka, R. C. (2011). Projection-specific modulation of dopamine neuron synapses by aversive and rewarding stimuli. *Neuron*, *70*(5), 855–862.

Lammel, S., Lim, B. K., Ran, C., Huang, K. W., Betley, M. J., Tye, K., ... Malenka, R. C. (2012). Input-specific control of reward and aversion in the ventral tegmental area. *Nature*, *491*(7423), 212.

Langdon, A. J., Sharpe, M. J., Schoenbaum, G., & Niv, Y. (2018). Model-based predictions for dopamine. *Current Opinion in Neurobiology*, *49*, 1–7.

Lau, B., & Glimcher, P. W. (2008). Value representations in the primate striatum during matching behavior. *Neuron*, *58*(3), 451–463.

Lau, B., Monteiro, T., & Paton, J. J. (2017). The many worlds hypothesis of dopamine prediction error: implications of a parallel circuit architecture in the basal ganglia. *Current Opinion in Neurobiology*, *46*, 241–247.

Lauwereyns, J., Watanabe, K., Coe, B., & Hikosaka, O. (2002). A neural correlate of response bias in monkey caudate nucleus. *Nature*, *418*(6896), 413.

Lejeune, H., & Wearden, J. (1991). The comparative psychology of fixed-interval responding: Some quantitative analyses. *Learning and Motivation*, *22*(1), 84–111.

Leon, M. I., & Shadlen, M. N. (2003). Representation of time by neurons in the posterior parietal cortex of the macaque. *Neuron*, *38*(2), 317–327.

Lerner, T. N., Shilyansky, C., Davidson, T. J., Evans, K. E., Beier, K. T., Zalocusky, K. A., ... others (2015). Intact-brain analyses reveal distinct information carried by snc dopamine subcircuits. *Cell*, *162*(3), 635–647.

Levy, F. (1991). The dopamine theory of attention deficit hyperactivity disorder (adhd). *Australian and New Zealand Journal of Psychiatry*, *25*(2), 277–283.

Liljeholm, M., & O'Doherty, J. P. (2012). Contributions of the striatum to learning, motivation, and performance: an associative account. *Trends in cognitive sciences*, *16*(9), 467–475.

Lopes, G., Bonacchi, N., Frazão, J., Neto, J. P., Atallah, B. V., Soares, S., ... others (2015). Bonsai: an event-based framework for processing and controlling data streams. *Frontiers in neuroinformatics*, *9*.

Lou, H. C., Henriksen, L., Bruhn, P., Børner, H., & Nielsen, J. B. (1989). Striatal dysfunction in attention deficit and hyperkinetic disorder. *Archives of Neurology*, *46*(1), 48–52.

Ludvig, E. A., Bellemare, M. G., & Pearson, K. G. (2011). A primer on reinforcement learning in the brain: Psychological, computational, and neural perspectives. In *Computational neuroscience for advancing artificial intelligence: Models, methods and applications* (pp. 111–144). IGI Global.

Ludvig, E. A., Sutton, R. S., & Kehoe, E. J. (2008). Stimulus representation and the timing of reward-prediction errors in models of the dopamine system. *Neural computation*, *20*(12), 3034–3054.

Ludvig, E. A., Sutton, R. S., & Kehoe, E. J. (2012). Evaluating the td model of classical conditioning. *Learning & behavior*, *40*(3), 305–319.

Lüthi, A., & Lüscher, C. (2014). Pathological circuit function underlying addiction and anxiety disorders. *Nature Neuroscience*, *17*(12), 1635–1643.

Maass, W., Natschläger, T., & Markram, H. (2002). Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural computation*, *14*(11), 2531–2560.

MacDonald, C. J., Lepage, K. Q., Eden, U. T., & Eichenbaum, H. (2011). Hippocampal "time cells" bridge the gap in memory for discontiguous events. *Neuron*, *71*(4), 737–749.

MacDonald, C. J., & Meck, W. H. (2005). Differential effects of clozapine and haloperidol on interval timing in the supraseconds range. *Psychopharmacology*, *182*(2), 232–244.

Machado, A. (1997). Learning the temporal dynamics of behavior. *Psychological review*, *104*(2), 241.

Machado, A., Malheiro, M. T., & Erlhagen, W. (2009). Learning to time: A perspective. *Journal of the Experimental Analysis of Behavior*, *92*(3), 423–458.

Machens, C. K., Romo, R., & Brody, C. D. (2010). Functional, but not anatomical, separation of "what" and "when" in prefrontal cortex. *Journal of Neuroscience*, *30*(1), 350–360.

Malapani, C., Dubois, B., Rancurel, G., & Gibbon, J. (1998). Cerebellar dysfunctions of temporal processing in the seconds range in humans. *Neuroreport*, *9*(17), 3907–3912.

Malapani, C., Rakitin, B., Levy, R., Meck, W. H., Deweer, B., Dubois, B., &

Gibbon, J. (1998). Coupled temporal memories in parkinson's disease: a dopamine-related dysfunction. *Journal of Cognitive Neuroscience*, *10*(3), 316–331.

Malenka, R., Nestler, E., & Hyman, S. (2009). Chapter 6: widely projecting systems: monoamines, acetylcholine, and orexin. *Sydor A, Brown RY. Molecular Neuropharmacology: A Foundation for Clinical Neuroscience (2nd ed.). New York: McGraw-Hill Medical*, 147–148.

Manly, T., Robertson, I. H., Galloway, M., & Hawkins, K. (1999). The absent mind: further investigations of sustained attention to response. *Neuropsychologia*, *37*(6), 661–670.

Manns, J. R., Howard, M. W., & Eichenbaum, H. (2007). Gradual changes in hippocampal activity support remembering the order of events. *Neuron*, *56*(3), 530–540.

Maricq, A. V., & Church, R. M. (1983). The differential effects of haloperidol and methamphetamine on time estimation in the rat. *Psychopharmacology*, *79*(1), 10–15.

Maricq, A. V., Roberts, S., & Church, R. M. (1981). Methamphetamine and time estimation. *Journal of Experimental Psychology: Animal Behavior Processes*, *7*(1), 18.

Matell, M. S., King, G. R., & Meck, W. H. (2004). Differential modulation of clock speed by the administration of intermittent versus continuous cocaine. *Behavioral neuroscience*, *118*(1), 150.

Matell, M. S., & Meck, W. H. (2004). Cortico-striatal circuits and interval timing: coincidence detection of oscillatory processes. *Cognitive brain research*, *21*(2), 139–170.

Matell, M. S., Meck, W. H., & Nicolelis, M. A. (2003). Interval timing and the encoding of signal duration by ensembles of cortical and striatal neurons. *Behavioral neuroscience*, *117*(4), 760.

Matell, M. S., Meck, W. H., et al. (2000). Neuropsychological mechanisms of interval timing behavior. *Bioessays*, *22*(1), 94–103.

Matias, S., Lottem, E., Dugue, G. P., & Mainen, Z. F. (2017). Activity patterns of serotonin neurons underlying cognitive flexibility. *Elife*, *6*, e20552.

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

Meck, W. H. (1983). Selective adjustment of the speed of internal clock and memory processes. *Journal of Experimental Psychology: Animal Behavior Processes*, *9*(2), 171.

Meck, W. H. (1986). Affinity for the dopamine d2 receptor predicts neuroleptic potency in decreasing the speed of an internal clock. *Pharmacology Biochemistry and Behavior*, *25*(6), 1185–1189.

Meck, W. H. (1996). Neuropharmacology of timing and time perception. *Cognitive brain research*, *3*(3), 227–242.

Meck, W. H. (2006). Neuroanatomical localization of an internal clock: A functional link between mesolimbic, nigrostriatal, and mesocortical dopaminergic systems. *Brain Research*, *1109*(1), 93–107.

Meck, W. H., Penney, T. B., & Pouthas, V. (2008). Cortico-striatal representation of time in animals and humans. *Current opinion in neurobiology*, *18*(2), 145–152.

Mello, G. B., Soares, S., & Paton, J. J. (2015). A scalable population code for time in the striatum. *Current Biology*, *25*(9), 1113–1122.

Menegas, W., Babayan, B. M., Uchida, N., & Watabe-Uchida, M. (2017). Opposite initialization to novel cues in dopamine signaling in ventral and posterior striatum in mice. *elife*, *6*, e21886.

Merchant, H., Pérez, O., Zarco, W., & Gámez, J. (2013). Interval tuning in the primate medial premotor cortex as a general timing mechanism. *Journal of Neuroscience*, *33*(21), 9082–9096.

Miall, C. (1989). The storage of time intervals using oscillating neurons. *Neural Computation*, *1*(3), 359–371.

Mink, J. W. (1996). The basal ganglia: focused selection and inhibition of competing motor programs. *Progress in neurobiology*, *50*(4), 381–425.

Minsky, M. (1961). Steps toward artificial intelligence. *Proceedings of the IRE*, *49*(1), 8–30.

Mita, A., Mushiake, H., Shima, K., Matsuzaka, Y., & Tanji, J. (2009). Interval time coding by neurons in the presupplementary and supplementary motor areas. *Nature neuroscience*, *12*(4), 502–507.

Mogenson, G. J., Jones, D. L., & Yim, C. Y. (1980). From motivation to action: functional interface between the limbic system and the motor system. *Progress in neurobiology*, *14*(2), 69–97.

Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive hebbian learning. *Journal of neuroscience*, *16*(5), 1936–1947.

Monteiro, T., Rodrigues, F., Motiwala, A., Gouvêa, T. S., & Paton, J. J. (2017). Decoupling categorical choice from action reveals pure timing signals in striatal populations. *Cosyne Abstract*.

Mountcastle, V. B., Talbot, W. H., Sakata, H., & Hyvarinen, J. (1969). Cortical neuronal mechanisms in flutter-vibration studied in unanesthetized monkeys: Neuronal periodicity and frequency discrimination. *Journal of neurophysiology*.

Murray, J. D., Bernacchia, A., Freedman, D. J., Romo, R., Wallis, J. D., Cai, X., ... others (2014). A hierarchy of intrinsic timescales across primate cortex. *Nature neuroscience*, *17*(12), 1661–1663.

Murray, J. M., et al. (2017). Learning multiple variable-speed sequences in striatum via cortical tutoring. *eLife*, *6*, e26084.

Newsome, W. T., & Pare, E. B. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area (mt). *Journal of Neuroscience*, *8*(6), 2201–2211.

Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology*, *53*(3), 139–154.

Niv, Y., Daw, N. D., Joel, D., & Dayan, P. (2007). Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology*, *191*(3), 507–520.

Noreika, V., Falter, C. M., & Rubia, K. (2013). Timing deficits in attention-deficit/hyperactivity disorder (adhd): Evidence from neurocognitive and neuroimaging studies. *Neuropsychologia*, *51*(2), 235–266.

Noyes Jr, R., & Kletti, R. (1972). The experience of dying from falls. *Omega-Journal of Death and Dying*, *3*(1), 45–52.

Odum, A. L., Lieving, L. M., & Schaal, D. W. (2002). Effects of d-amphetamine in a temporal discrimination procedure: Selective changes in timing or rate dependency? *Journal of the experimental analysis of behavior*, *78*(2), 195–214.

Ogden, R. S., Moore, D., Redfern, L., & McGlone, F. (2015). The effect of pain and the anticipation of pain on temporal perception: A role for attention and arousal. *Cognition and Emotion*, *29*(5), 910–922.

O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map. preliminary evidence from unit activity in the freely-moving rat. *Brain research*, *34*(1), 171–175.

O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford: Clarendon Press.

Oleson, E. B., Gentry, R. N., Chioma, V. C., & Cheer, J. F. (2012). Subsecond dopamine release in the nucleus accumbens predicts conditioned punishment and its successful avoidance. *Journal of Neuroscience*, *32*(42), 14804–14808.

Panigrahi, B., Martin, K. A., Li, Y., Graves, A. R., Vollmer, A., Olson, L., ... Dudman, J. T. (2015). Dopamine is required for the neural representation and control of movement vigor. *Cell*, *162*(6), 1418–1430.

Pariyadath, V., & Eagleman, D. (2007). The effect of predictability on subjective duration. *PLoS ONE*, *2*(11), e1264.

Parker, A. J., & Newsome, W. T. (1998). Sense and the single neuron: probing the physiology of perception. *Annual review of neuroscience*, *21*(1), 227–277.

Parker, N. F., Cameron, C. M., Taliaferro, J. P., Lee, J., Choi, J. Y., Davidson, T. J., ... Witten, I. B. (2016). Reward and choice encoding in terminals

of midbrain dopamine neurons depends on striatal target. *Nature neuroscience*, *19*(6), 845.

Pasquereau, B., & Turner, R. S. (2015). Dopamine neurons encode errors in predicting movement trigger occurrence. *Journal of neurophysiology*, *113*(4), 1110–1123.

Pastalkova, E., Itskov, V., Amarasingham, A., & Buzsáki, G. (2008). Internally generated cell assembly sequences in the rat hippocampus. *Science*, *321*(5894), 1322–1327.

Pastor, M., Artieda, J., Jahanshahi, M., & Obeso, J. (1992). Time estimation and reproduction is abnormal in parkinson's disease. *Brain*, *115*(1), 211–225.

Pavlov, I. P. (1927). *Conditional reflexes: An investigation of the physiological activity of the cerebral cortex*. H. Milford.

Pavlov, I. P., & Anrep, G. V. (2003). *Conditioned reflexes*. Courier Corporation.

Platt, J. R., & Davis, E. R. (1983). Bisection of temporal intervals by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, *9*(2), 160.

Ponzi, A., & Wickens, J. (2010). Sequentially switching cell assemblies in random inhibitory networks of spiking neurons in the striatum. *Journal of Neuroscience*, *30*(17), 5894–5911.

Rakitin, B. C., Gibbon, J., Penney, T. B., Malapani, C., Hinton, S. C., & Meck, W. H. (1998). Scalar expectancy theory and peak-interval timing in humans. *Journal of Experimental Psychology: Animal Behavior Processes*, *24*(1), 15.

Rammsayer, T. (1990). Temporal discrimination in schizophrenic and affective disorders: evidence for a dopamine-dependent internal clock. *International Journal of Neuroscience*, *53*(2-4), 111–120.

Rammsayer, T. H. (1993). On dopaminergic modulation of temporal information processing. *Biological psychology*, *36*(3), 209–222.

Reppert, S. M., & Weaver, D. R. (2002). Coordination of circadian timing in mammals. *Nature*, *418*(6901), 935.

Rescorla, R. A. (1966). Predictability and number of pairings in pavlovian fear conditioning. *Psychonomic Science*, *4*(11), 383–384.

Rescorla, R. A. (1967). Pavlovian conditioning and its proper control procedures. *Psychological review*, *74*(1), 71.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of pavlovian conditioning: The effectiveness of reinforcement and non-reinforcement. *Classical conditioning II: Current research and theory*.

Reynolds, G., & Catania, A. C. (1962). Temporal discrimination in pigeons. *Science*, *135*(3500), 314–315.

Reynolds, J. N., Hyland, B. I., & Wickens, J. R. (2001). A cellular mechanism of reward-related learning. *Nature*, *413*(6851), 67.

Reynolds, J. N., & Wickens, J. R. (2002). Dopamine-dependent plasticity of corticostriatal synapses. *Neural Networks*, *15*(4), 507–521.

Roberts, S. (1981). Isolation of an internal clock. *Journal of Experimental Psychology: Animal Behavior Processes*, *7*(3), 242.

Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997). Oops!': performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia*, *35*(6), 747–758.

Rowe, K. C., Paulsen, J. S., Langbehn, D. R., Duff, K., Beglinger, L. J., Wang, C., ... Moser, D. J. (2010). Self-paced timing detects and tracks change in prodromal huntington disease. *Neuropsychology*, *24*(4), 435.

Runyan, C. A., Piasini, E., Panzeri, S., & Harvey, C. D. (2017). Distinct timescales of population coding across cortex. *Nature*, *548*(7665), 92–96.

Sacks, O. (2004). Speed. *New Yorker*, *80*, 48–59.

Sakurai, T. (2007). The neural circuit of orexin (hypocretin): maintaining sleep and wakefulness. *Nature Reviews Neuroscience*, *8*(3), 171–181.

Samejima, K., & Doya, K. (2007). Multiple representations of belief states and action values in corticobasal ganglia loops. *Annals of the New York Academy of Sciences*, *1104*(1), 213–228.

Samejima, K., Ueda, Y., Doya, K., & Kimura, M. (2005). Representation of action-specific reward values in the striatum. *Science*, *310*(5752), 1337–1340.

Schmitt, L. I., Wimmer, R. D., Nakajima, M., Happ, M., Mofakham, S., & Halassa, M. M. (2017). Thalamic amplification of cortical connectivity sustains attentional control. *Nature*, *545*(7653), 219–223.

Schultz, W. (1986). Responses of midbrain dopamine neurons to behavioral trigger stimuli in the monkey. *Journal of neurophysiology*, *56*(5), 1439–1461.

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*(3), 900–913.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593–1599.

Schultz, W., & Romo, R. (1987). Responses of nigrostriatal dopamine neurons to high-intensity somatosensory stimulation in the anesthetized monkey. *Journal of neurophysiology*, *57*(1), 201–217.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of neurology, neurosurgery, and psychiatry*, *20*(1), 11.

Sharpe, M. J., Chang, C. Y., Liu, M. A., Batchelor, H. M., Mueller, L. E., Jones, J. L., ... Schoenbaum, G. (2017). Dopamine transients are sufficient and necessary for acquisition of model-based associations. *Nature Neuroscience*.

Shinomoto, S., Omi, T., Mita, A., Mushiake, H., Shima, K., Matsuzaka, Y., & Tanji, J. (2011). Deciphering elapsed time and predicting action timing from neuronal population signals. *Frontiers in computational neuroscience*, *5*.

Silva, L. R., Amitai, Y., & Connors, B. W. (1991). Intrinsic oscillations of neocortex generated by layer 5 pyramidal neurons. *Science*, *251*(4992), 432.

Simen, P., Balci, F., Cohen, J. D., Holmes, P., et al. (2011). A model of interval timing by neural integration. *Journal of Neuroscience*, *31*(25), 9238–9253.

Simen, P., Rivest, F., Ludvig, E. A., Balci, F., & Killeen, P. (2013). Timescale invariance in the pacemaker-accumulator family of timing models. *Timing & Time Perception*, *1*(2), 159–188.

Skinner, B. F. (1938). *The behaviour of organisms: An experimental analysis*. D. Appleton-Century Company Incorporated.

Smallwood, J., Davies, J. B., Heim, D., Finnigan, F., Sudberry, M., O'Connor, R., & Obonsawin, M. (2004). Subjective experience and the attentional lapse: Task engagement and disengagement during sustained attention. *Consciousness and cognition*, *13*(4), 657–690.

Smallwood, J., McSpadden, M., & Schooler, J. W. (2007). The lights are on but no one's home: Meta-awareness and the decoupling of attention when the mind wanders. *Psychonomic Bulletin & Review*, *14*(3), 527–533.

Smallwood, J. M., Baracaia, S. F., Lowe, M., & Obonsawin, M. (2003). Task unrelated thought whilst encoding information. *Consciousness and cognition*, *12*(3), 452–484.

Soares, S., Atallah, B., & Paton, J. (2016). Midbrain dopamine neurons control judgment of time. *Science*, *354*(6317), 1273–1277.

Spencer, R. M., & Ivry, R. B. (2005). Comparison of patients with parkinson's disease or cerebellar lesions in the production of periodic movements involving event-based or emergent timing. *Brain and cognition*, *58*(1), 84–93.

Staddon, J., & Higa, J. (1999). Time and memory: towards a pacemaker-free theory of interval timing. *Journal of the experimental analysis of behavior*, *71*(2), 215–251.

Stamatakis, A. M., Jennings, J. H., Ung, R. L., Blair, G. A., Weinberg, R. J., Neve, R. L., ... others (2013). A unique population of ventral tegmental area neurons inhibits the lateral habenula to promote reward. *Neuron*, *80*(4), 1039–1053.

Starkweather, C. K., Babayan, B. M., Uchida, N., & Gershman, S. J. (2017). Dopamine reward prediction errors reflect hidden-state inference across time. *Nature Neuroscience*, *20*(4), 581–589.

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. *Nature neuroscience*, *16*(7), 966–973.
Stephens, D. W., & Krebs, J. R. (1986). *Foraging theory*. Princeton University Press.
Stubbs, A. (1968). The discrimination of stimulus duration by pigeons. *Journal of the experimental analysis of behavior*, *11*(3), 223–238.
Sulzer, D., Zhang, H., Benoit-Marand, M., & Gonon, F. (2010). Regulation of extracellular dopamine: release and reuptake. In (pp. 297–319).
Suri, R. E., & Schultz, W. (1999). A neural network model with dopamine-like reinforcement signal that learns a spatial delayed response task. *Neuroscience*, *91*(3), 871–890.
Sutton, R. S., & Barto, A. G. (1990). Time-derivative models of pavlovian reinforcement. In *Learning and computational neuroscience: Foundations of adaptive networks.*
Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction* (Vol. 1) (No. 1). MIT press Cambridge.
Syed, E. C., Grima, L. L., Magill, P. J., Bogacz, R., Brown, P., & Walton, M. E. (2016). Action initiation shapes mesolimbic dopamine encoding of future rewards. *Nature neuroscience*, *19*(1), 34–36.
Tanaka, S. C., Doya, K., Okada, G., Ueda, K., Okamoto, Y., & Yamawaki, S. (2004). Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nature neuroscience*, *7*(8), 887.
Taylor, N. E., Van Dort, C. J., Kenny, J. D., Pei, J., Guidera, J. A., Vlasov, K. Y., ... Solt, K. (2016). Optogenetic activation of dopamine neurons in the ventral tegmental area induces reanimation from general anesthesia. *Proceedings of the National Academy of Sciences*, *113*(45), 12826–12831.
Tecuapetla, F., Carrillo-Reid, L., Bargas, J., & Galarraga, E. (2007). Dopaminergic modulation of short-term synaptic plasticity at striatal inhibitory synapses. *Proceedings of the National Academy of Sciences*, *104*(24), 10258–10263.
Thorndike, E. L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *Psychological Monographs*, *2*.
Thorndike, E. L. (1911). *Animal intelligence: Experimental studies*. Macmillan.
Tobler, P. N., Fiorillo, C. D., & Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science*, *307*(5715), 1642–1645.
Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological review*, *55*(4), 189.
Treisman, M. (1963). Temporal discrimination and the indifference interval: Implications for a model of the" internal clock". *Psychological Monographs: General and Applied*, *77*(13), 1.
Treisman, M. (1966). A statistical decision model for sensory discrimination which predicts weber's law and other sensory laws: Some results of a com-

puter simulation. *Perception & Psychophysics*, *1*(4), 203–230.
Treisman, M., Cook, N., Naish, P. L., & MacCrone, J. K. (1994). The internal clock: electroencephalographic evidence for oscillatory processes underlying time perception. *The Quarterly Journal of Experimental Psychology*, *47*(2), 241–289.
Treisman, M., Faulkner, A., Naish, P. L., & Brogan, D. (1990). The internal clock: Evidence for a temporal oscillator underlying time perception with some estimates of its characteristic frequency. *Perception*, *19*(6), 705–742.
Voorn, P., Vanderschuren, L. J., Groenewegen, H. J., Robbins, T. W., & Pennartz, C. M. (2004). Putting a spin on the dorsal - ventral divide of the striatum. *Trends in neurosciences*, *27*(8), 468–474.
Waelti, P., Dickinson, A., & Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. *Nature*, *412*(6842), 43–48.
Wahl, O. F., & Sieg, D. (1980). Time estimation among schizophrenics. *Perceptual and motor skills*.
Wang, J., Narain, D., Hosseini, E., & Jazayeri, M. (2017). Flexible control of speed of cortical dynamics. *bioRxiv*, 155390.
Wang, X.-J. (2001). Synaptic reverberation underlying mnemonic persistent activity. *Trends in neurosciences*, *24*(8), 455–463.
Ward, R. D., Kellendonk, C., Simpson, E. H., Lipatova, O., Drew, M. R., Fairhurst, S., ... Balsam, P. D. (2009). Impaired timing precision produced by striatal d2 receptor overexpression is mediated by cognitive and motivational deficits. *Behavioral neuroscience*, *123*(4), 720.
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*(5), 858–873.
Watt, J. D. (1991). Effect of boredom proneness on time perception. *Psychological Reports*, *69*(1), 323–327.
Watts, F. N., & Sharrock, R. (1984). Fear and time estimation. *Perceptual and motor skills*, *59*(2), 597–598.
Wearden, J. H., Smith-Spark, J., Cousins, R., Edelstyn, N., Cody, F., & O'Boyle, D. (2008). Stimulus timing by people with parkinson's disease. *Brain and cognition*, *67*(3), 264–279.
Wiener, M., Turkeltaub, P., & Coslett, H. B. (2010). The image of time: a voxel-wise meta-analysis. *Neuroimage*, *49*(2), 1728–1740.
Wise, R. A. (2004). Dopamine, learning and motivation. *Nature reviews. Neuroscience*, *5*(6), 483.
Witten, I. B., Steinberg, E. E., Lee, S. Y., Davidson, T. J., Zalocusky, K. A., Brodsky, M., ... others (2011). Recombinase-driver rat lines: tools, techniques, and optogenetic application to dopamine-mediated reinforcement. *Neuron*, *72*(5), 721–733.

Wittmann, M., Leland, D. S., Churan, J., & Paulus, M. P. (2007). Impaired time perception and motor timing in stimulant-dependent subjects. *Drug and alcohol dependence*, *90*(2), 183–192.

Wittmann, M. K., Kolling, N., Akaishi, R., Chau, B. K., Brown, J. W., Nelissen, N., & Rushworth, M. F. (2016). Predictive decision making driven by multiple time-linked reward representations in the anterior cingulate cortex. *Nature communications*, *7*.

Wodka, E. L., Mark Mahone, E., Blankner, J. G., Gidley Larson, J. C., Fotedar, S., Denckla, M. B., & Mostofsky, S. H. (2007). Evidence that response inhibition is a primary deficit in adhd. *Journal of clinical and experimental neuropsychology*, *29*(4), 345–356.

Woodrow, H. (1928). Temporal discrimination in the monkey. *Journal of Comparative Psychology*, *8*(5), 395.

Woodrow, H. (1930). The reproduction of temporal intervals. *Journal of Experimental Psychology*, *13*(6), 473.

Xu, M., Zhang, S.-y., Dan, Y., & Poo, M.-m. (2014). Representation of interval timing by temporally scalable firing patterns in rat prefrontal cortex. *Proceedings of the National Academy of Sciences*, *111*(1), 480–485.

Yalch, R. F., & Spangenberg, E. R. (2000). The effects of music in a retail setting on real and perceived shopping times. *Journal of business Research*, *49*(2), 139–147.

Znamenskiy, P., & Zador, A. M. (2013). Corticostriatal neurones in auditory cortex drive decisions during auditory discrimination. *Nature*, *497*(7450), 482.

ITQB-UNL | Av. da República, 2780-157 Oeiras, Portugal
Tel (+351) 214 469 100 | Fax (+351) 214 411 277

www.itqb.unl.pt