



HAL
open science

Activity of dopaminergic neurons during skilled reaching in head-fixed rats

Clara Besserer

► **To cite this version:**

Clara Besserer. Activity of dopaminergic neurons during skilled reaching in head-fixed rats. Neuro-
science. Université de Bordeaux, 2023. English. NNT : 2023BORD0210 . tel-04699847

HAL Id: tel-04699847

<https://theses.hal.science/tel-04699847v1>

Submitted on 17 Sep 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITE DE BORDEAUX

Doctoral School ED Sciences de la Vie et de la Santé 154

University Department Institut des Maladies Neurodégénératives, CNRS UMR 5293

Thesis defended by Clara BESSERER

Defended on September 15, 2023

In order to become Doctor from UNIVERSITE DE BORDEAUX

Academic Field **Neurosciences**

Activity of dopaminergic neurons during skilled reaching in head-fixed rats

Thesis supervised by Nicolas MALLET

Committee members

<i>Referees</i>	Corinne BEURRIER	Institut de Neurosciences de la Timone UMR7289
	David ROBBE	Inmed UMR1249
<i>Examiners</i>	Catherine LEMOINE	INCIA – UMR 5287
	Robert SCHMIDT	Ruhr-Universität Bochum
<i>Supervisor</i>	Nicolas MALLET	IMN CNRS UMR5293

COLOPHON

Doctoral dissertation entitled “Activity of dopaminergic neurons during skilled reaching in head-fixed rats”, written by Clara BESSERER, completed on October 25, 2023, typeset with the document preparation system \LaTeX and the `yathesis` class dedicated to theses prepared in France.

UNIVERSITE DE BORDEAUX

Doctoral School ED Sciences de la Vie et de la Santé 154

University Department Institut des Maladies Neurodégénératives, CNRS UMR 5293

Thesis defended by Clara BESSERER

Defended on September 15, 2023

In order to become Doctor from UNIVERSITE DE BORDEAUX

Academic Field **Neurosciences**

Activity of dopaminergic neurons during skilled reaching in head-fixed rats

Thesis supervised by Nicolas MALLET

Committee members

<i>Referees</i>	Corinne BEURRIER	Institut de Neurosciences de la Timone UMR7289
	David ROBBE	Inmed UMR1249
<i>Examiners</i>	Catherine LEMOINE	INCIA – UMR 5287
	Robert SCHMIDT	Ruhr-Universität Bochum
<i>Supervisor</i>	Nicolas MALLET	IMN CNRS UMR5293

UNIVERSITE DE BORDEAUX

École doctorale **ED Sciences de la Vie et de la Santé 154**

Unité de recherche **Institut des Maladies Neurodégénératives, CNRS UMR 5293**

Thèse présentée par **Clara BESSERER**

Soutenue le **15 septembre 2023**

En vue de l'obtention du grade de docteur de l'UNIVERSITE DE BORDEAUX

Discipline **Neurosciences**

Activité des neurones dopaminergiques pendant l'exécution d'une tâche comportementale de préhension chez des rats tête-restreinte

Thèse dirigée par Nicolas MALLET

Composition du jury

<i>Rapporteurs</i>	Corinne BEURRIER	Institut de Neurosciences de la Timone UMR7289
	David ROBBE	Inmed UMR1249
<i>Examineurs</i>	Catherine LEMOINE	INCIA – UMR 5287
	Robert SCHMIDT	Ruhr-Universität Bochum
<i>Directeur de thèse</i>	Nicolas MALLET	IMN CNRS UMR5293

Activity of dopaminergic neurons during skilled reaching in head-fixed rats**Abstract**

Midbrain dopaminergic neurons (DANs) are essential for associative learning and goal-directed movements. Several studies have recorded their activity during motor tasks and most have observed an increase of dopaminergic activity during movement or around its initiation. However, the functional role played by DANs during instrumental tasks is still unclear. It has been hypothesized that dopamine could modulate movement vigor in an online manner instead of only encoding reward prediction error, but causal experiments have yielded conflicting results. Other theories argue for a role of movement-related dopamine in habit formation or retrospective causal learning, but the extent of dopaminergic contribution to sensory-motor behaviors in physiological conditions is unknown. In this study, we sought to investigate the functional heterogeneity of DANs in the SNc and characterize their activity during a reach-and-grab motor task in TH-Cre rats. We recorded DANs electrical activity and GCaMP calcium signal in head-fixed rats performing the operant task using high-density silicon probes and fiber photometry, respectively. We observed that DANs increased their activity, as measured through global calcium signal, both after the auditory cue and during movement execution. Interestingly, the movement-evoked responses were highly lateralized depending on the limb used. We further dissected the neuronal mechanism causing these increases in calcium transients using in vivo electrophysiological recordings. We found that the activity of DANs neurons recorded in the SNc was functionally heterogeneous during the task with individual neurons responding to task events in the specialized-manner. Finally, we observed a correlation between dopaminergic activity and movement vigor, as well as an inhibition of dopaminergic activity during movement repression. Our results reveal the functional diversity present in DANs during a fine operant motor task, and highlight the close relationship between dopamine and motor execution.

Keywords: dopamine, snc, dorsolateral striatum, intrumental skilled-reaching behavior, in vivo electrophysiology & calcium imaging, vigor

Institut des Maladies Neurodégénératives, CNRS UMR 5293

Bâtiment Neurocampus, Université de Bordeaux, 146 Rue Léo Saignat, 33000 Bordeaux, France

Activité des neurones dopaminergiques pendant l'exécution d'une tâche comportementale de préhension chez des rats tête-restreinte**Résumé**

Les neurones dopaminergiques sont essentiels à l'apprentissage associatif et aux mouvements volontaires. De multiples études ont étudié leur activité lors de tâches motrices et la plupart ont observé une augmentation de l'activité dopaminergique pendant le mouvement ou à son initiation. Cependant, le rôle fonctionnel joué par les neurones dopaminergiques pendant les tâches instrumentales est toujours inconnu. Si on leur attribue classiquement un rôle dans l'apprentissage ainsi que l'encodage de la valeur des récompenses attendues, une hypothèse dominante de cette dernière décennie suggère que les neurones dopaminergiques pourraient également moduler la vigueur des mouvements à venir, mais les preuves causales sont encore limitées. La contribution de la dopamine aux comportements instrumentaux complexes reste à élucider. Dans cette étude, nous avons enregistré des neurones dopaminergiques de la substance noire compacte afin de caractériser leur hétérogénéité fonctionnelle et leur activité lors d'une tâche motrice à choix. Pour cela nous avons mis en place une méthode de tête restreinte chez le rat qui nous permet d'effectuer des enregistrements électrophysiologiques aigus simultanément à des enregistrements de l'activité calcique des terminaisons dopaminergiques dans le striatum via fiber photometry. À l'échelle de la population, nous avons observé une activité dopaminergique en réponse au signal sonore ainsi que durant le mouvement et à la consommation de la récompense. À l'échelle des neurones individuels, nous avons observé une grande spécialisation de leur activité. Nous avons montré que l'activité dopaminergique était latéralisée spécifiquement pendant le mouvement, et que cette activité ainsi que celle qui précède le mouvement était corrélée à sa vigueur. Nos résultats révèlent la diversité fonctionnelle des neurones dopaminergiques pendant une tâche comportementale complexe, et mettent en valeur les liens entre la dopamine et l'exécution motrice.

Mots clés : dopamine, snc, striatum dorsolatéral, comportement instrumental de préhension, électrophysiologie & imagerie calcique in vivo, vigueur

Remerciements

I would first like to thank Dr. Corrine Beurrier, Dr. David Robbe, Pr. Dr. Robert Schmidt and Pr. Dr. Catherine Le Moine for accepting to review my work and be part of my jury.

Je remercie Thomas Boraud, directeur de l'IMN, de m'avoir accueillie dans son laboratoire et dans son équipe.

Je remercie tout particulièrement Nicolas Mallet, sans conteste le meilleur directeur de thèse qu'il est possible d'avoir. Merci de m'avoir accepté comme thésarde alors que nous ne nous connaissions pas, que tu avais déjà 2 autres étudiantes à encadrer, et que mon projet à l'origine n'était pas du tout ton domaine. Tu m'as tout appris et même si j'aurais encore du mal à refaire correctement tous les branchements du CED from scratch, j'ai l'impression d'être un peu plus autonome en fin de thèse qu'au début. Ma confiance en moi à toujours été limitante mais tu ne m'as jamais fait sentir que j'étais trop dépendante; merci de m'avoir aidé pour toutes ces manip.

Merci Sophie d'avoir été un exemple à suivre, de m'avoir appris comment entraîner les rats, comment faire les chirurgies et l'histologie. Merci Lise, partenaire de bureau depuis le tout début, et à Victor et Julien qui nous y ont rejoints en cours de route. Grace à vous venir travailler à toujours été agréable.

Merci à tout le reste de l'équipe 4; Arthur, Marc, Alexis et Aude, ainsi que les anciens thésards; Roman et Remya, et ceux qui sont en cours de thèse; Eduarda, Fjola, Milesa. Ce qui m'a poussé à faire ma thèse dans cette équipe est sa bonne ambiance et sa pluridisciplinarité, et ce sont des qualités que j'apprécie d'autant plus maintenant.

Merci Alejandro et Morgane, qui ont effectué un stage auprès de moi cette année, et désolée de ne pas avoir pu être une meilleure encadrante!

Je remercie Alexandra Elbakyan pour sa contribution sans prix à la communauté scientifique. L'introduction de cette thèse serait 3 fois plus courte sans Sci-Hub.

Enfin, merci à l'équipe du PIV-EXPE pour avoir pris soin de nos animaux et pour nous permettre de faire nos expériences dans les meilleures conditions possibles. Même si ils ne pourront pas le lire, je souhaite également remercier les rats avec lesquels j'ai travaillé. Leur contribution à ce travail est plus importante

que la mienne. J'ai entraîné chaque rat individuellement pendant de nombreux mois et pourrais encore vous donner les caractéristiques distinctifs de chacun d'entre eux. Ils m'ont apporté les plus grand plaisirs pendant ma thèse mais également les plus grands stress et plus grandes tristesses. Merci les minous et désolée pour tout.

Double cursus oblige, après les remerciements du coté des scientifiques viennent les remerciements coté médecine.

Merci à Thomas Bienvenu, sans qui je n'aurais jamais fait ma thèse avec Nico. Tu m'as dit que c'était l'un des meilleur scientifique que tu connaises et tu ne l'a pas survendu. En revanche, tu as aussi dit que tout le monde finissait par detester son directeur de thèse et j'attend toujours. Merci surtout pour l'ESS, pour m'avoir accepté dans ta toute premiere promotion de bébés médecins-chercheurs, et pour avoir toujours été un exemple vers lequel tendre. Je remercie également Isabelle Dupin et Marie-Edith Lafon, les deux autres piliers de l'ESS. Vous avez été des modèles de bienveillance et je vous serai eternellement reconnaissante pour tout ce que vous avez fait pour nous.

Merci à Boris Barbour et Eric Clauser de m'avoir acceptée à l'Ecole de l'Inserm Lilianne Bettencourt, ainsi qu'à la Fondation Bettencourt Schueller de rendre tout cela possible. L'Ecole de l'Inserm est l'une des chose qui me motive le plus dans ce difficile parcours de médecin-chercheur. Merci à tous les étudiants du réseau double cursus que j'ai rencontré alors que je commençais médecine et qui m'ont inspiré à poursuivre dans cette voie.

Merci à ma promo de l'ESS, Jérémy, Louis, Pierre, Ugo, Yoan, ainsi que nos confères simple cursus Achille, Marcus et Etienne. Vous avez été ma seconde famille depuis de nombreuses années maintenant et avez été d'un support sans faille. Merci Aliona, Théo et Jérémy, la petite team thèse précoce bordelaise, que j'abandonne à regret.

Merci à ma famille, à mes parents qui m'ont transmis le gout de la science à la fois sur le plan de l'innée et de l'acquis, et pour leur support pendant cette première partie de mes études, ainsi qu'à mes petites soeurs Alice et Solène.

Last but not least, merci à mon autre petite famille, Louis et Sweden. Merci Louis pour ton support pendant toute cette thèse, rien de ce que je ne pourrais écrire ici serait suffisant. Merci Sweden, pour tout le recomfort, c'est toi le meilleur minou de tous.

Acronyms

A | B | C | D | G | L | M | N | P | R | S | T | V

A

AMPA α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid. 16

B

BG Basal ganglia. 8

BLA Baso-lateral amygdala. 10, 11

C

cAMP-PKA Cyclic adenosine monophosphate - Protein kinase A. 17

CS conditioned stimulus. 23

D

D1-MSNs Direct medium spiny neurons. 8, 14

D1R Dopamine receptors of class 1. 14–16

D2-MSNs Indirect medium spiny neurons. 8

D2R Dopamine receptors of class 2. 14–16

DANs Dopaminergic neurons. 4

DAT Dopaminergic transporter. 4

DLS Dorsolateral striatum. 8, 13

DMS Dorsomedial striatum. 8, 10, 13

G

GABA γ aminobutyric acid. 8

GP Globus pallidus. 8

L

L-DOPA L-3,4-dihydroxy phenylalanine. 3

LTD Long term depression. 18
LTP Long term potentiation. 18

M

MAO Monoamine oxydase. 4
mRNA Messenger RNA. 4
MSNs Medium spiny neurons. 8, 11, 14, 16

N

NAc Nucleus accumbens. 7, 10
NMDA N-methyl-D-aspartate. 15, 16

P

PET Positron emission tomography. 18
PETH Peri event time histogram. 48
PFC Prefrontal cortex. 8
PKA Protein kinase A. 19
PPN Pediculo-pontine nucleus. 8

R

R-O Response-Outcome. 12
RPE Reward Prediction Error. 26

S

S-R Stimulus-Response. 13
SNc Substantia Nigra pars compacta. 4, 7
STDP Spike-timing-dependent plasticity. 19
STN Subthalamic Nucleus. 8

T

TDRL Temporal difference reinforcement learning. 25
TH Tyrosine hydroxylase. 3
TS Tail of the striatum. 12

V

VMAT2 vesicular monoamine transporter 2. 3
VTA Ventral tegmental area. 7

Avant-propos

If dopamine (DA) neurotransmission has been historically associated with the processing of hedonic or reward information, work performed in the past 10 years has highlighted an amazing variety of DA functions in basal ganglia circuits. Seminal work from the 1990's has first linked the activity of DA neurons with the prediction of upcoming rewards in associative learning. DA has since been considered as a key signal for reinforcement learning to enable the associations between sensory events or actions and rewards. Since most of our daily movements are executed to achieve specific goals or reward maximization, it is thus not surprising that DA transmission has also been tightly involved in sensorimotor processing and movement execution. Such implication of DA into motor behaviors is perhaps best illustrated by the devastating motor deficits associated with the death of these neural population in Parkinson's disease which manifest as akinesia (absence of voluntary movements) and bradykinesia (slowness of movements). That being said, the precise characterization and the causal contribution of midbrain DA dynamics during the execution of dexterous movements in physiological conditions remains to be clearly demonstrated.

In the context of associative learning, dopamine has been shown to encode reward prediction error (RPE), that is the discrepancy between the 'expected reward' vs. the amount of reward effectively 'received' by the animal. Upon presentation of a cue that has been associated with a delayed reward, dopaminergic neurons exhibit a phasic activity scaling with the value of the expected reward, but also with the probability to receive it, and the delay between the cue and the reward (Bayer and Glimcher, 2005; Bromberg-Martin and Hikosaka, 2009; Eshel et al., 2016; Schultz et al., 1997). However, less is known about the activity of DANs during instrumental learning. Indeed, despite early recordings of DA activity being described as purely encoding sensory stimuli (Schultz, 2016), with no contribution to movement generation (Schultz and Romo, 1990), multiple studies from the past decade have brought attention to motor-related dopaminergic activity. Spontaneous rewarded actions are often associated with an increase of dopamine activity around movement initiation (Da Silva et al., 2018; Howe MW, 2016), and this dopaminergic signal only appears after learning the association between the

action and the reward (Coddington and Dudman, 2018).

In more complex, naturalistic situations such as when a specific action must be executed based on external cues in order to obtain a reward, the contribution of dopamine is still unclear. Do DANs drive learning of a cue-action or an action-reward association? Do DANs modulate motor execution in cases where a specific speed or strength is required? Do DANs signal an evolving RPE at each moment of the task? Are DANs involved in sensory-motor processing? Recent studies investigating these questions have reported a phasic activation of DANs after cue presentation and/or movement initiation (Coddington and Dudman, 2018). Some studies have also observed a ramping signal leading to reward (Jin and Costa, 2010; Syed et al., 2016) while others have hypothesized that dopamine could invigorate movement in an online manner (Da Silva et al., 2018; Niv et al., 2007). However, the causal evidences between dopamine and movement vigor are still conflicting (Hamid et al., 2015; Lee et al., 2020; H. Liu et al., 2022).

Considering the key contribution of DANs in many different brain functions, it is not surprising to find anatomo-functional specialization in DAergic systems. DANs are located in two main brain regions: the ventral tegmental area (VTA), specialized in encoding motivational and reward aspect (Steinberg et al., 2013), and the substantia nigra pars compacta (SNc), more related to motor activity (Da Silva et al., 2018; Dodson et al., 2016a; Jin and Costa, 2010; Mendonça et al., 2021). Interestingly, the topographical organization of these DAergic projections are coherent with the different behavioral contribution of the striatal territories their preferentially innervate. Indeed, the dorsolateral part of the striatum (DLS), that receives DA inputs from the SNc, is involved in the execution of learned motor sequences (Dhawale et al., 2021) whereas the dorsomedial striatum (DMS), which receive DA inputs from the VTA, is involved in the learning of this sequences (Yin et al., 2005; Yin and Knowlton, 2006). This dichotomic view of DLS and DMS function is however challenged by recent findings showing that the DLS, but not the DMS, is also required for the acquisition of complex movement pattern (Wolff et al., 2022).

Overall, the activity and function of DA neurons in dexterous instrumental tasks still requires further investigations and many questions remain to be addressed: What is the activity of SNc DA neurons during such complex and lateralized motor task? Is there a functional heterogeneity of response in SNc DA neurons population? Could a rapid change in DAergic signaling alter upcoming/on-going movement execution (for example by encoding vigor), or does it guide future changes of motor behavior through learning?

Addressing all these questions was the objective of my PhD project. To do so, we have designed a sensory-discrimination reach-and-grasp motor task in head-

fixed rats that gave us access to fine movement parameters such as reaction time, movement time and velocity, as well as a diversity of motor behaviors (i.e. missed trials, anticipations, wrong lever choice, etc.). We have combined this complex motor task with in vivo electrophysiological and calcium activity recordings of DA neurons to better understand the spatio-temporal dynamics and contribution of dopamine signal to action execution. Our in vivo electrophysiological recordings were performed using high-density silicon probes (64 channels) coupled with local optogenetic stimulation through an optic fiber glued to the recording electrode. DA neurons of the SNc were identified based on their electrophysiological properties as well as their excitatory responses to laser pulse using a ChR2 opto-tagging strategy opsins). To this end, we used transgenic TH-Cre Long Evans rats that allow us to specifically access dopaminergic neurons. These in vivo single-cell electrophysiological recordings of DANs were combined with fiber photometric measurements of DA signals in the DLS using the GCaMP8f calcium sensor that gave us access to population-level monitoring of DA dynamics during different motor outputs and across training days. So far, our dataset includes more than a hundred identified DANs, most of them recorded during successful Go trials.

Our results demonstrate that DANs responses are heterogeneous during this motor task with a clear behavioral specialization across single-cell neurons. Indeed, a subset of DA neurons was mostly responsive to the ‘Go’ Cue, as evinced by the presence of a phasic activation time-locked at 60/80ms after the auditory signal. Another subset of DA neuron, was mostly active during the reaching phase of the movement, although this population of DANs were less stereotyped than the ‘cue-responding’ neurons. We also demonstrated that these DANs responses effectively translate into dopamine axonal changes in the DLS as measured through global calcium signal dynamics present at the cue and during movement. Interestingly, the motor-related activity was lateralized both at the global and single-cell levels, with a higher dopaminergic response in the hemisphere contralateral to the movement.

Both the cue-evoked and movement-evoked signals were correlated to the duration of reaction time and movement time of the animal, indicating an eventual link between dopaminergic activity and movement vigor. Furthermore, in cases where the animal decides not to press any lever following the cue, a subtype of error trials, we observed a much lower dopaminergic response to the cue, that was immediately followed by a brief inhibition of dopaminergic activity in the DLS. This type of activity was also observed in No-Go trials, in which a particular cue instructed the animal to not move but instead stay immobile for a predefined duration of time in order to get the reward.

In conclusion, our work reveals the highly dynamic nature of dopamine signaling in the DLS during fine movement execution. In particular, behavioral events such as the sensory cue and the movement initiation appears to encoded by DANs activity

and could be related to multiple aspects of the upcoming motor behavior, such the direction, the laterality, its vigor or, on the opposite, its immediate suppression.

Sommaire

Abstract	vii
Remerciements	ix
Acronyms	xi
Avant-propos	xiii
Sommaire	xvii
List of Tables	xix
List of Figures	xxi
I Introduction	1
1 Dopamine; from molecule to network	3
2 Dopamine and associative learning	23
3 Dopamine and instrumental learning	41
II Methods	61
4 Data acquisition	63
5 Data analysis	71
III Results	75

6 Recordings in head-fixed rats during the instrumental task	77
7 Dopaminergic activity during an instrumental skilled-reaching task	85
IV Discussion	105
8 Discussion	107
References	115
Contents	133

List of Tables

- 4.1 List of viruses used and their injection sites 66
- 4.2 Stereotaxic coordinates 66

List of Figures

1.1	Molecular heterogeneity of dopaminergic neurons.	5
1.2	Firing pattern of dopaminergic neurons.	6
1.3	Functional organisation of the striatum	9
1.4	D1-like receptors and D2-like receptors	14
1.5	Molecular mechanisms of action of dopamine	16
1.6	Model for reinforcement plasticity	21
2.1	Dopamine encodes Reward Prediction Error	29
2.2	Dopamine RPE signal scales with reward value	30
2.3	Dopamine signals time intervals.	33
2.4	ANCCR model of retrospective learning	35
3.1	Dopaminergic activity during spontaneous movements	45
3.2	Dopaminergic activity during cue-triggered movements	48
3.3	Lateralization of dopaminergic activity	50
3.4	Dopamine is involved in the control of vigor	54
3.5	The Action Prediction Error model	56
4.1	Behavioral training setup and head-fixation implantation.	64
4.2	Video of a rat performing the task	68
6.1	Characteristics of the reaction time and the movement time.	78
6.2	GCaMP8f signal and kinetics.	80
6.3	Electrophysiological identification of dopaminergic neurons.	83
7.1	Dopaminergic activity during contralateral trials.	86
7.2	Activity of putative SNc dopaminergic neurons during contralateral trials.	87
7.3	Lateralisation of dopaminergic signal	89
7.4	Lateralisation of cue units	90
7.5	Lateralisation of movement units	91
7.6	Movement-evoked dopaminergic activity is dependant on context.	92

7.7	Correlation between dopaminergic activity in the DLS and the vigor of the movement.	95
7.8	Correlation between dopaminergic activity in the DLS and the vigor of the movement.	96
7.9	Correlation between movement speed and reward-evoked dopaminergic activity.	98
7.10	Dopaminergic activity during trials when the animal did not press any lever.	101
7.11	Dopaminergic activity during NoGo trials.	103

Part I

Introduction

Dopamine; from molecule to network

1.1 Dopamine as a neurotransmitter

Dopamine is a catecholamine produced in both the central nervous system where it acts as a neurotransmitter and in the peripheral tissues where it regulates pancreatic endocrine function including insulin release.

In the brain, dopamine is produced by specific dopaminergic neurons and is derived from the amino acid tyrosine which is hydroxylated into L-3,4-dihydroxy phenylalanine (L-DOPA) and finally decarboxylated into dopamine. The two key enzymes to this synthesis are the tyrosine hydroxylase (Tyrosine hydroxylase (TH)) and the DOPA decarboxylase. As they are specific to this pathway we can use the TH to specifically target dopamine producing neurons for histological purposes or gene therapy.

After its synthesis in the cell body, dopamine is stored in synaptic vesicles by the vesicular monoamine transporter 2 (vesicular monoamine transporter 2 (VMAT2)). An increase in axonal calcium concentration (due to either an action potential or an axo-axonic stimulation) causes the vesicles to release the dopamine into the synaptic cleft. Dopamine can then diffuse up to 7 μm away from its release site. In the synaptic space, dopamine can diffuse up to 12 μm away from the release site (Gonon, 1997). However at this distance it can only be detected by high sensitivity receptors, whereas low affinity receptors are only efficient within a maximum radius of 2 μm from the quantal release site.

After release, the synaptic dopamine is cleared by the dopaminergic transporter Dopaminergic transporter (DAT). The concentration of dopamine in the synapse returns to baseline after less than a few hundreds of a millisecond (Phillips and Stamford, 2000). Once its back in the cell, dopamine is catabolised by the monoamine oxidase enzyme (Monoamine oxydase (MAO)) and recycled.

1.2 Dopaminergic neurons

Dopaminergic neurons (Dopaminergic neurons (DANs)) are the main source of dopamine in the brain as peripheral dopamine cannot cross the blood-brain barrier. These neurons are concentrated in mesolimbic nuclei that we will describe in the following section. The entire population amounts to 45 000 neurons in rats, 165 000 in monkeys and 600 000 in humans (German and Manaye, 1993).

Dopaminergic neurons share morphological traits. They all participate in dense dendritic networks with widespread axonal projections. Furthermore, it was widely theorized that these neurons would also share common molecular phenotypes. Namely through the ubiquitous expression of TH and DOPA decarboxylase enzymes, as well as the VMAT2 transporter. However, a recent review by Morales and Margolis, 2017 put paid to this notion. Indeed they showed that the molecular phenotype of dopaminergic neurons was heterogeneous. Some subsets of TH-expressing VTA neurons lacked detectable levels of transcripts encoding the VMAT2 or DAT proteins. Whereas others expressed TH Messenger RNA (mRNA) but lacked measurable concentrations of TH proteins.

In addition, other genetic heterogeneity has been described with Substantia Nigra pars compacta (SNc) dopaminergic neurons subtypes that have topographically biased projections along the rostrocaudal, mediolateral, and dorsoventral axes of the striatum Poulin et al., 2018 1.1.

1.3 DANs electrophysiological properties

DANs are tonically active neurons. They usually fire at low frequencies (1-5Hz) but can also exhibit phasic activity with bursts of action potentials (above 15Hz) (Grace

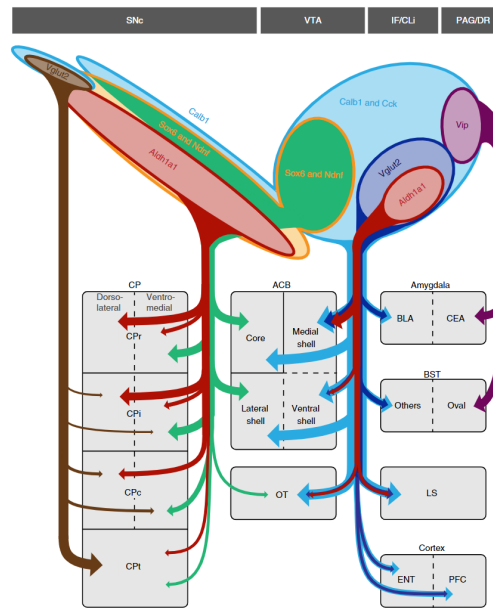


Figure 1.1: **Molecular heterogeneity of dopaminergic neurons** (From Poulin et al., 2018)

and Bunney, 1980) (fig. 1.2). They are characterized by long action potential duration and large hyperpolarization. This broad spike shape has mainly been observed with glass electrode recordings and is often absent if silicone probes are used.

The two modes of DAN's activities, tonic and phasic, are thought to serve different functions. Tonic dopamine firing provides a background dopamine tone that is supposed to be important for numerous striatal-dependent functions such as motivation (Goto and Grace, 2007; Schultz, 2007). By contrast, phasic activity plays a crucial role in reward processing as well as during associative learning (Schultz et al., 1997; Wise, 2004). These aspects will be further developed in the second part of this introduction.

Tonic and phasic firing differentially affect the release of dopamine at the synapse. There is a very low probability of release for dopamine-containing vesicles. This means that neurotransmitter release will usually vary as a function of firing frequency (Lebowitz and Khoshbouei, 2020). DANs always have a reserve pool of waiting dopamine vesicles that are relatively insensitive to stimuli (Martel 2020).

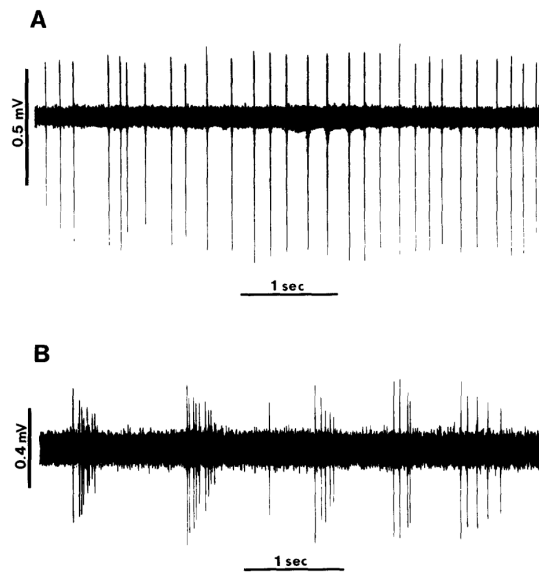


Figure 1.2: **Firing pattern of dopaminergic neurons.** **A:** Tonic activity around 5Hz. **B:** Phasic activity with a characteristic decrease in the amplitude of the action potential during the burst. (From Freeman et al., 1985)

The phasic firing or bursting mode of DANs is characterized by a minimum of 2 action potentials with an inter-spike interval ranging from 80 to 160ms (Grace and Bunney, 1984) (see fig. 1.2 B). Some observers have shown that the inter-spike interval tends to lengthen as the burst goes on and that the amplitude of the spikes decreases. Bursts are then always followed by a period of silence.

1.4 Dopamine release

DAN terminals do not form classical synapses but have varicosities that contain the release sites. Dopamine is stored in vesicles from which they are then released at the varicosities. Exocytosis is triggered by an increase in intracellular calcium. The release sites respond robustly to initial stimulation. However, their vesicles are rapidly depleted, even under a tonic firing mode. During burst firing, only the first few action potentials lead to a significant dopamine release. Thus, synchronization of the neuronal population is a far more important factor than individual activity in order to trigger the mass release of dopamine during behaviorally significant events (for review, see P. G. Liu and Kaeser, 2021). Phasic release depends more

on the simultaneous recruitment of a dopamine neuron population than actual burst firing.

In contrast to classical neurotransmitters, dopamine is not necessarily released directly into synapses. Instead, it is often released into the extracellular space, where it diffuses and reaches the cell membrane, creating a gradient of neurotransmitter concentration. The probability of receptor activation thus depends on its distance from the release site (Cragg et al., 2001). This mechanism is called volume transmission.

1.5 Dopaminergic nuclei

Dopaminergic neurons are mainly located in two midbrain nuclei; the Ventral Tegmental Area (Ventral tegmental area (VTA)) and the Substantia Nigra pars compacta (SNc). In primates, these are called the A10 and A9 regions respectively (Dahlström and Fuxe, 1964). These two nuclei are close to each other but are anatomically distinct. Their axons also project to different target structures (Menegas et al., 2015).

1.5.1 VTA

The VTA is the most medial nucleus of the two. VTA dopaminergic neurons mainly project onto the Nucleus Accumbens (Nucleus accumbens (NAC)), the ventral part of the striatum, but also onto the lateral habenula, the dorsal raphe nucleus, the lateral hypothalamus, the amygdala, and the cortex. These last two connections are part of the mesolimbic and mesocortical pathways, respectively. The mesolimbic pathway influences motivation and reward characterization, whilst the mesocortical pathway is implicated in associative memory and motor functions (Hosp et al., 2011; Puig et al., 2014).

1.5.2 SNc

The SNc is found more laterally and caudally than the VTA. It forms a thin layer of neurons above the SNr. The highest density of neurons is located at its most

cranial part. Sparse dopaminergic neurons are sometimes found in the SNr.

SNc neurons receive collateral inputs from SNr neurons, as well as striatal and Globus pallidus (GP) inputs, all of which are γ aminobutyric acid (GABA)ergic. These inhibitory inputs negatively modulate the tonic activity of DANs. On the other hand, the SNc receives strong excitatory inputs from the Prefrontal cortex (PFC), Subthalamic Nucleus (STN), and Pediculo-pontine nucleus (PPN), which serve to drive phasic dopaminergic activity.

SNc neurons send axonal projections principally to the dorsal striatum (Dorsomedial striatum (DMS) and Dorsolateral striatum (DLS)). There is a topographical organization of these projections, with sub-circuits distributed on the mediolateral axis (Haber et al., 2000). This means that lateral SNc neurons, projecting to the DLS will mainly receive inputs from the DLS (Beier et al., 2015; Menegas et al., 2015).

1.6 Dopaminergic targets

DANs send densely innervate the cortex (mostly the frontal cortex) as well as the striatum. SNc neurons mostly target the dorsal striatum.

1.6.1 Striatum

The striatum is the largest nucleus of the Basal ganglia (BG) and one of its main input structures with the STN. It receives sensory, motor, and cognitive information from the cortex and the thalamus, and then transmits them to downstream BG structures.

The striatal output neurons are called medium spiny neurons (Medium spiny neurons (MSNs)). These neurons project onto the SNr directly or indirectly through one of two main projection pathways: the direct and the indirect pathway. The direct pathway has no intermediate neurons. Whereas the indirect pathway has synapses in the globus pallidus (GP) and subthalamic nucleus (STN) before reaching the SNr. MSNs belonging to the direct pathway (Direct medium spiny neurons (D1-MSNs)) express D1-type dopamine receptors whilst indirect pathway MSNs (Indirect medium spiny neurons (D2-MSNs)) express D2-type dopamine

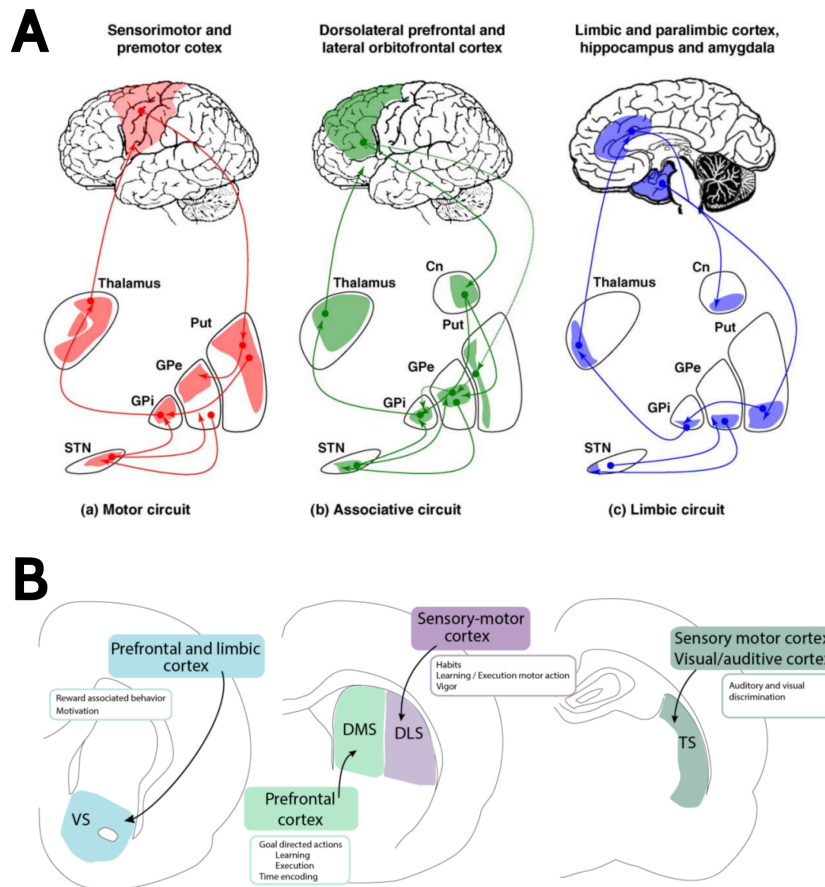


Figure 1.3: **Functional organisation of the striatum.** **A:** Topographical organization of the basal ganglia loops (from Krack et al., 2010) **B:** Schematic coronal sections of the striatal sub-regions and their main input structures.

receptors (Gerfen et al., 1990). This difference in terms of receptors affects the way these neurons respond to dopamine (see 1.7.1). Indeed, dopamine influx increases D1-MSNs excitability but decrease that of D2-MSNs (Planert et al., 2013).

1.6.2 Direct vs indirect pathway

It has long been believed that direct and indirect pathways played an opposing role in motor function, the D1-MSNs being active and facilitating movements while the D2-MSNs would inhibit movements (Kravitz et al., 2010). However, electrophysiological recordings have since shown both pathways' MSNs to be active

during movement initiation. Furthermore, they both seem to be inactive during immobility. This suggests that these two pathways have a synergistic relationship (Cui et al., 2013; Jin et al., 2014; Tecuapetla et al., 2014).

1.6.3 Striatal sub-regions

The striatum can be anatomically separated into four main territories with differing functions (Hunnicutt et al., 2016). The dorsolateral striatum (DLS) is the sensorimotor region. The dorsomedial striatum (DMS) is the associative region. The last two regions, the ventral and caudal parts of the striatum (namely the NAC and the tail of the striatum) are both limbic structures. These sub-regions receive different inputs and form distinct sub-circuits inside the basal ganglia loop, both anatomically and functionally (fig. 1.3). In addition to this organization in sub-nuclei, the striatum is also composed of two components; the striosome and the matrix. Their respective roles are still unclear but they exhibit very different inputs and projections (Miyamoto et al., 2018).

DMS

The DMS is classically described as the associative striatum. It receives inputs from the prefrontal cortex as well as the basolateral amygdala (Baso-lateral amygdala (BLA)) and the visual cortex (Hintiryan et al., 2016; Hunnicutt et al., 2016). The DMS has been much studied in the context of goal-directed actions (actions performed with the goal of obtaining a reward of known value) because this type of behavior relies heavily on environmental inputs and the animal's internal state. Lesion studies have provided evidence for the role of the DMS in the action-outcome association, including the learning and the execution phases (Yin and Knowlton, 2006; Yin et al., 2005). Electrophysiological recording of DMS neurons has shown that they were active during learning and action selection (Coynel et al., 2010; Kupferschmidt et al., 2017). After extensive training, when the action becomes habitual instead of goal-directed, the activity in the DMS decreases in favor of the DLS (Yin et al., 2009).

DLS

The DLS is commonly described as the sensorimotor region of the striatum and is associated with habitual behavior. The DLS receives most of its inputs from the somatosensory and motor cortices. Its role in habit formation has been demonstrated by lesion studies (Balleine and Dickinson, 1998; Dickinson and Weiskrantz, 1985; Tricomi et al., 2009; Yin and Knowlton, 2006; Yin et al., 2004; Yin et al., 2005). The predominating theory regarding goal-directed and habitual behaviors is the existence of a switch from DMS to DLS: initially a new behavior will be coordinated by the DMS. Once it becomes habitual this role transfers from the DMS to the DLS (Barnes et al., 2005; Yin et al., 2010). The DLS has also been identified as an important region for the learning and execution of motor actions (Wolff et al., 2022), as well as the control of movement vigor (Dudman and Krakauer, 2016; Smith and Graybiel, 2013).

Ventral striatum

The ventral striatum is mainly composed of the nucleus accumbens (NAc), which is further subdivided into a shell and a core (the medial and lateral portions of the NAc, respectively). It receives inputs from the pre-frontal cortex, the hippocampus, and the BLA that relays limbic information such as outcome values or emotional valence. The NAc is known to participate in motivated and reward-related behaviors. Indeed, it receives large dopaminergic inputs from the VTA and has been identified as the main target of the dopaminergic reward prediction error signals (which will be described in 2.2).

The NAc also contributes to social behavior (Le Merrer et al., 2023) and motivation. Induced lesions of the NAc in rats cause a decrease in movement vigor, attributed to a decrease in motivation. These lesions however do not impact instrumental learning. Electrophysiological recordings of NAc MSNs show that they encode the predictive value of stimuli as well as the associated motor behavior (Nicola et al., 2004).

Tail of the striatum

The tail of the striatum is the most caudal part of the striatum. It has recently garnered interest as a functionally important area of the striatum (Valjent and Gangarossa, 2021). It receives inputs from the sensorimotor cortex as well as the auditory and visual cortices (Deniau et al., 1996; Gangarossa et al., 2013; Hunnicutt et al., 2016; Jiang and Kim, 2018). Because of the strong inputs coming from the auditory cortex, the Tail of the striatum (TS) has been shown to encode auditory stimuli and is necessary for auditory discrimination (Gao et al., 2018; Guo et al., 2018; Xiong et al., 2015). As with the other striatal regions, the TS is densely innervated by dopaminergic projections coming mainly from the SNc. This nigrostriatal pathway has recently been shown to carry an action prediction error (see 3.3.2) signal during an auditory discrimination task (Greenstreet et al., 2022)

1.6.4 Contribution of the dorsal striatum to goal-directed and habitual behaviors

The distinction between goal-directed and habitual behavior originated in psychology. According to psychologists, goal-directed behavior is purposeful because the individual has foreknowledge both of the reward value as well as of the causal relationship between actions and their outcomes. In nature, this behavior has the advantage of being quickly adjustable to environmental changes (Dayan, 2009). Because it depends on knowing the consequences of a given action it can be called a response-outcome (Response-Outcome (R-O)) control. In experimental settings, behavior is usually considered goal-directed if it is sensitive to reinforcer devaluation (giving unlimited access to the reward before the training session) and contingency degradation (granting randomly distributed rewards dissociated from the animal's behavior) (Adams and Dickinson, 1981; Dickinson and Weiskrantz, 1985). Rodent studies have repeatedly highlighted the contribution of the DMS to goal-directed behavior (Balleine, 2005; Corbit and Balleine, 2005; Yin et al., 2005).

Habitual behaviors, on the other hand, are automatic and inflexible. They develop over time after repeated performance. Whilst goal-directed behaviors are controlled by internal knowledge of the outcome value, habitual behaviors

are triggered by external stimuli. Habits have the advantage of requiring little cognitive effort, but their inflexibility makes them unable to adjust to changes in the environment (Dayan, 2009).

Habits, in essence, are an association between an environmental cue and an action. Thus they are sometimes called stimulus-response (Stimulus-Response (S-R)) associations. In an experimental setting, habitual behaviors can be developed in rodents by over-training and employing variable-interval reinforcement that leads to reward-timing uncertainty (Adams and Dickinson, 1981; Colwill and Rescorla, 1985). The experimental definition of habitual behavior is negative in nature: habits are insensitive to devaluation and contingency degradation. Indeed, habits are not affected by (short-term) changes in the value of the reward (Wood and Runger, 2016).

Many studies have shown that DLS supports habit learning (Barnes et al., 2005; Graybiel, 2008; Smith and Graybiel, 2013; Thorn et al., 2010; Yin and Knowlton, 2006; Yin et al., 2004). Lesions to the DLS result in the maintenance of goal-directed behavior even with over-training. This contrasts with the effect of lesions to the DMS. Indeed these lesions result in the early emergence of habitual behavior (Yin and Knowlton, 2006). There is also a large body of evidence suggesting that training and habit formation induces the transfer of motor control from the DMS to DLS (Belin et al., 2009; Graybiel, 2008; Thorn et al., 2010; Yin et al., 2004; Yin et al., 2009; Yin et al., 2005).

This shift from DMS to DLS is thought to depend on dopaminergic inputs. Indeed, lesions to nigrostriatal pathways impair habitual learning (Faure et al., 2005), whereas pharmacological increase of dopamine concentration biases behavior towards habit formation (Nelson and Killcross, 2006). The main hypothesis underlying a switch of the DA signaling is the ascending spiral hypothesis (Haber et al., 2000). This hypothesis postulates that in a tri-synaptic circuit, DMS neurons could modulate the activity of dopaminergic neurons projecting to the DLS. However, it has only been shown that dopamine in the DMS, and not the DLS, is necessary for habit formation. DMS dopamine reinforces habitual behaviors while DLS dopamine is stable in rats developing a habit strategy in a lever preference test (van Elzelingen et al., 2022).

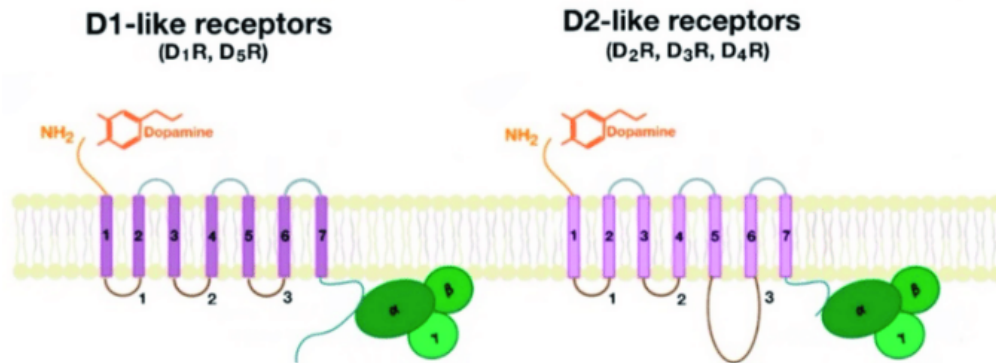


Figure 1.4: **Representation of D1-like receptors and D2-like receptors** in the plasma membrane, coupled with their G proteins (From Martinez et al., 2020)

1.7 Dopamine receptors

Dopaminergic receptors are G protein-coupled receptors and are classified into two main families depending on the type of G protein. D1-class receptors (Dopamine receptors of class 1 (D1R)) are associated with excitatory G proteins whilst D2-class receptors (Dopamine receptors of class 2 (D2R)) are associated with an inhibitory G protein. The two types of dopaminergic receptors, therefore, modulate the activity of the adenylate cyclase in opposite ways (Kebabian and Calne, 1979). If the D1R is found almost exclusively at post-synaptic sites, the D2R is expressed both pre- and post-synaptically on dopaminergic neurons, acting as auto-receptors and providing feedback modulation on dopaminergic activity.

1.7.1 The different effects of D1R and D2R activation on MSN activity

As described above, the D1 and D2 dopamine receptors are expressed in direct and indirect MSNs respectively. In 1987, Li et al. showed that dopamine exerts an opposite effect on these two neuronal populations; it increases the activity of D1-MSNs but decreases that of D2-MSNs (Li et al., 1987). Conversely, lesions to the dopaminergic nuclei cause a decrease of D1-MSNs but an increase of D2-MSNs activity (Young et al., 1986). Surprisingly, the precise effect of dopamine on the

MSNs electrophysiological activity and plasticity is still unclear.

Dopamine, being a neuromodulator, does not have a direct excitatory or inhibitory effect on the postsynaptic neurons. Rather it modulates the response of neurons to other neurotransmitters. This modulation is in turn greatly dependent on the region and target neuron. For example, differences have been observed between the ventral and dorsal striatum. Here, we will focus on the effect observed in the dorsal striatum.

Dopaminergic modulation of presynaptic neurotransmitter release

Dopaminergic receptors D2R can be found on presynaptic sites of the corticostriatal synapses (D1R can also sometimes be found pre-synaptically : Dumartin et al., 2007). Measuring the release of glutamatergic vesicles shows that dopamine seems to act as a high-pass filter on cortical projections. It suppresses glutamate release sites that are less active but has no effect on fast release sites (Bamford et al., 2004). This effect scaled with the frequency of cortical stimulation. Optical recordings of presynaptic release concomitant to whole-cell electrophysiology have shown that D1- and D2-like receptor agonists can have either a slight presynaptic excitatory or inhibitory effect on glutamate release respectively (W. Wang et al., 2012). It seems that dopamine acts as a frequency filter of cortical inputs by inhibiting the weaker signals. If true this suggests a potential learning mechanism present when both cortical and dopaminergic inputs are highly active at the same time. For example, if an action and a reward occur simultaneously (for a detailed review of dopamine's effect on presynaptic sites, see Bamford et al., 2018; Tritsch and Sabatini, 2012).

Dopaminergic modulation of postsynaptic receptors

Dopamine's roles extend beyond its presynaptic effects. It also modulates the presence of postsynaptic neurotransmitter receptors. D1R stimulation can potentiate extra synaptic N-methyl-D-aspartate (NMDA) receptor function by increasing currents, membrane depolarisation, and cytosolic Ca²⁺ levels evoked by application of NMDA receptor agonists (André et al., 2010; Jocoy et al., 2011). The mechanisms underlying this effect are still unclear but likely include NMDA receptors

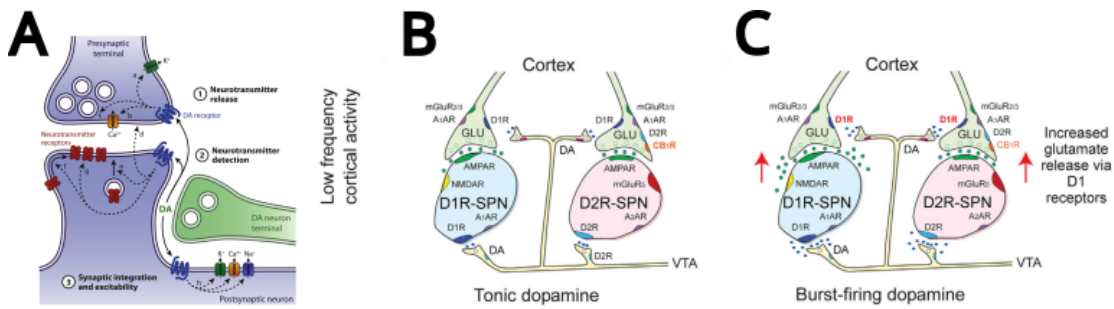


Figure 1.5: **Molecular mechanisms of action of dopamine on cortico-striatal synapses.** **A:** From Tritsch and Sabatini, 2012. **B-C:** From Bamford et al., 2018

phosphorylation and membrane trafficking. On the other hand, D2R activation does not directly modulate NMDA receptors (Higley and Sabatini, 2010). In addition, dopamine can bidirectionally modulate α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptors. D1R agonists and D2R antagonists promote their phosphorylation while D2R agonists decrease it (Håkansson et al., 2006). This impacts the membrane currents caused by local application of AMPA receptor agonists. Membrane currents are decreased by D2R activation and increased by D1R activation (André et al., 2010).

Dopaminergic modulation of MSNs intrinsic excitability

Besides its synaptic effects, dopamine also modulates MSNs intrinsic excitability. Early studies surprisingly observed that the D1Rs tended to decrease MSNs excitability but that the D2Rs increased it (Nicola et al., 2000). However further investigations showed that the opposite effect was more likely, with a dopamine-induced increase of D1-MSN excitability but a decrease of D2-MSN excitability (Gerfen and Surmeier, 2011). Dopamine's effects on excitability occur through modulation of the MSNs voltage-gated K^+ , Na^+ , and Ca^{2+} channels. Thus, the activation of D1Rs drives a switch from a hyper-polarized state to a more depolarized state that is closer to the spike threshold, promoting synaptic integration and spiking rate (Wickens and Arbuthnott, 2005). By contrast, D2R activation increases membrane repolarisation and decreases calcium influx (Higley and Sabatini, 2010).

This effect on MSN excitability has been confirmed *in vivo* by Kirkpatrick et al., 2016 and Owesson-White et al., (2016). To characterize the effect of dopamine on

the activity of D1 and D2-MSNs in an awake, freely moving animal, Kirkpatrick et al. implanted an electrode combined with a microiontophoresis probe, allowing simultaneous fast-scan cyclic voltammetry and drug injection. The effects observed in vivo were similar to those recorded in slices, despite a higher baseline activity of the MSNs: activation of D1 receptors enhanced the excitation of MSNs, whereas activation of D2R's reduced their excitation (Kirkpatrick et al., 2016, Owesson-White et al., 2016).

1.7.2 Different sensitivity of the D1R and D2R to dopamine levels

Another critical difference between D1R and D2R is their affinity for dopamine. Early pharmacological experiments showed that the affinity for dopamine and dopaminergic agonists was higher for D2R than for D1R (Richfield et al., 1989). In addition, the D1R has a higher dissociation constant. This means that the D1R is saturated by greater concentrations of dopamine than the D2R. This difference in affinity led to the emergence of the theory that the D1 and D2 receptors respond differently to phasic and tonic dopamine. Phasic activity causes a high concentration of dopamine and is thought to activate the predominantly low-affinity D1R, whilst the much lower level of tonic dopamine would be sensed selectively by the higher-affinity D2 receptors. The nanomolar background dopamine due to the tonic activity of dopaminergic neurons has been hypothesized to maintain constant activation of striatal D2R in vivo

However, the very existence of this high-low affinity difference is still debated (Skinbjerg et al., 2012). Using optogenetic stimulations, Marcott et al. showed that the D2R responded to both tonic and phasic dopamine release. This indicates that both D2 MSNs and D1 MSNs are potentially sensitive to phasic dopaminergic activity (Marcott et al., 2014). This effect has also been observed in a computational model incorporating binding and unbinding kinetics of both the D1R and D2R (Hunger et al., 2020). Because receptor activation and functional effect are not necessarily identical, Yapo et al. investigated the effect of dopamine on Cyclic adenosine monophosphate - Protein kinase A (cAMP-PKA) signaling in D1 and D2 MSNs. They showed that phasic dopamine transients could activate both the D1R

and D2R within the same range of concentration, D1R being only two times more sensitive than D2R (Yapo et al., 2017). In addition, Positron emission tomography (PET) imaging studies have shown that only 10 to 20% of striatal D2Rs were actually occupied by dopamine at baseline, contradicting the hypothesis that the D2Rs are constantly activated in vivo due to their high affinity. Thus it seems that contrary to the prevailing belief, the difference of affinity initially observed between D1Rs and D2Rs might not have important functional implications, and that both D1 and D2 MSNs could be sensitive to phasic dopaminergic transients.

1.7.3 Functional effect of dopamine on MSNs

It is not possible to explain dopamine-driven behavioral learning simply via its modulating effect on MSNs excitability. If dopamine is mainly released at the time of reward consumption, it is likely that the motor program that needs to be reinforced must have happened up to a few seconds prior. Changing the excitability of the cortico-striatal synapses would only make sense if the dopaminergic input occurred before the movement, not after.

Hebbian plasticity

A first possibility to explain dopaminergic-driven learning would be the long-term modification of specific cortico-striatal synapses. Indeed dopaminergic receptors appear to modulate long-term changes in glutamatergic synaptic plasticity in MSNs. D1R signaling could promote long-term potentiation (Long term potentiation (LTP)) (Calabresi et al., 2007; Reynolds et al., 2001). Whereas D2R receptor signaling would promote long-term depression (Long term depression (LTD)) (Gerdeman et al., 2002; Kreitzer and Malenka, 2007). In addition, the loss of dopamine in Parkinson's disease models leads to a decrease of both forms of synaptic plasticity in MSNs. Shen et al., 2008 showed that dopamine was necessary for the induction of bidirectional, Hebbian plasticity at glutamatergic synapses of MSNs, in concert with glutamate and adenosine. They hypothesized that under basal condition, the low levels of dopamine in the striatum were enough to enable bidirectional plasticity in D2-MSN (due to their supposedly higher affinity) but not in D1 MSNs where the lack of D1R activation would only permit LTD and not

LTP. Stronger dopamine transients caused by a behaviorally significant stimulus would then activate the D1R as well and enable the induction of Hebbian LTP in D1-MSNs (Shen et al., 2008).

Structural plasticity

A second possibility, in addition to synaptic plasticity, is that dopamine induces structural plasticity changes in MSNs. In 2014, Yagishita et al. observed that spike-timing dependant plasticity (Spike-timing-dependent plasticity (STDP)) stimulation protocol (consisting of glutamate release concomitant to MSN action potential) induced a robust dendritic spine enlargement if dopamine was present but only a weak one if dopamine was absent (Yagishita et al., 2014). Very interestingly, this effect was still present if the dopaminergic input was delayed by a second after the STDP protocol. The mechanism underlying this temporal contingency depends on the temporary activation of the Protein kinase A (PKA) in distal dendrites after a calcium influx. This could provide an explanation to the reinforcement of past actions by dopamine: a reward-induced dopaminergic activation closely following an action would enable structural and Hebbian plasticity only in the dendrites that have recently received glutamatergic inputs.

More recently, Iino et al., 2020) paired STDP protocols with optogenetic stimulations of dopaminergic neurons. They observed that under tonic levels of dopamine, neither the D1-MSNs nor the D2-MSNs showed spine enlargement after STDP stimulation. On the other hand, a transient decrease in DANs firing rate (caused, for example, by a reward omission) was able to induce structural plasticity in the D2-MSNs. This structural plasticity was concomitant to an increase in glutamate sensitivity (Iino et al., 2020).

1.8 From cell plasticity to behavioral learning

How does dopamine drive learning? In the previous section, we saw that dopamine could induce synaptic plasticity and intrinsic changes in striatal MSNs. Dopamine transients closely following a sensory-motor event can induce LTP and dendritic spine enlargements as well as modify the MSN excitability for a few seconds.

Dopamine thus causes both retroactive and proactive changes. We can relate these mechanisms to different aspects of behavior:

- A reward obtained immediately after action will reinforce this action. This can be linked to the plasticity mechanisms described in 1.7.3, where a STDP stimulation could induce plasticity if dopamine was released within a 1-2s time window. This time window is explained by the presence of a synaptic 'eligibility trace' that is maximal during the cortical stimulation and then progressively decreases (see fig. 1.6). This kind of time window was observed in behavioral settings where increasing the delay between the cue or action and the reward decreased the learning rate.
- Once an association is learned (see 2.2) and dopaminergic neurons are physically active at the time of the cue presentation and not the reward delivery, dopamine would have a proactive effect on MSN activity and corticostriatal plasticity. By increasing MSN's excitability, dopamine would potentiate cortical inputs and facilitate movement execution.

1.9 Basal Ganglia and motor control

Coordinated movements rely on intact motor centers in the brainstem and spinal cord. For example, specific brainstem circuits can generate coordinated limb movements or locomotor behaviors (Esposito et al. 2014, Caggiano et al. 2018). However, those circuits need input from other brain areas to evaluate when and how these behaviors should be performed (Arber and Costa, 2018). The basal ganglia are at the core of motor functions and dysfunctions. They integrate information about the context and motor plan and then select the most appropriate behavior. The striatum is the input nucleus of the basal ganglia. It receives inputs from the cortex and allows the initiation of a specific action by removing the inhibition on the BG output nucleus, the SNr while preventing the occurrence of involuntary movements (Hikosaka et al., 2000). Such key inputs that form cortical synapses in the striatum are heavily modulated by dopamine transients. Dopamine is mostly produced by midbrain nuclei, the VTA and SNc, whose primary target is the

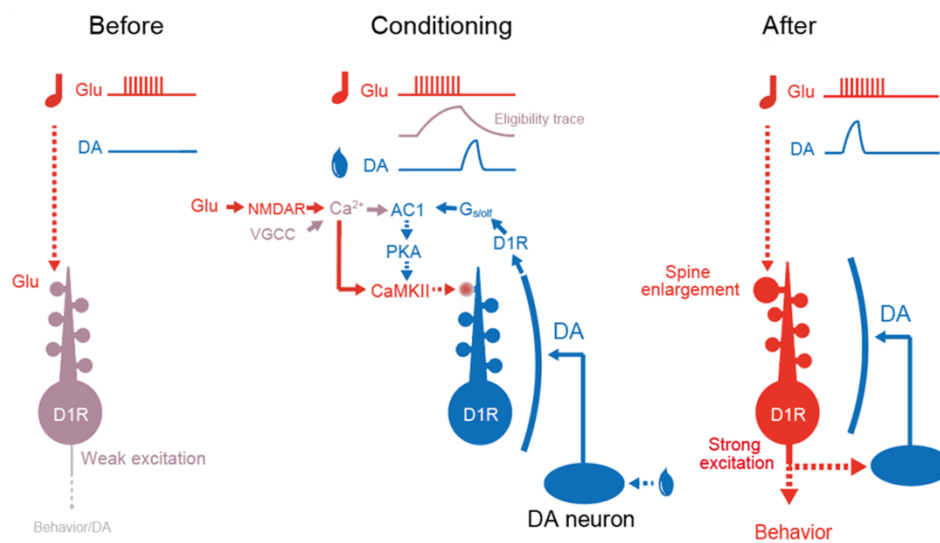


Figure 1.6: **Proposed model for reinforcement plasticity.** Before conditioning (left), cortical inputs could not excite D1-MSNs strongly enough to induce a response. During conditioning (middle), when dopamine is released within the synaptic eligibility trace after the synaptic activity, it activates the process of spine potentiation. After conditioning (right), with the potentiated dendritic spine, the same inputs can now activate D1-MSNs to induce behavioral and dopaminergic activities. (From Yagishita, 2023)

striatum. Depending on the type of dopamine receptors present at the surface of the striatal medium spiny neurons (MSNs), this dopamine neuromodulation in the striatum either invigorates cortico-striatal inputs or opposes them. In the following chapters, we will investigate how dopamine affects behavioral learning and motor execution.

Dopamine and associative learning

In this chapter we will present the theories and models of associative learning and review the contribution of dopaminergic neurons to this process.

2.1 Associative learning

The first theory of associative learning in animals was proposed more than a century ago by Thorndike, 1898. He defined learning as the association of a stimulus with an individual's response. These connections are reinforced when the response is followed by a reward.

This type of learning is called "instrumental conditioning" as it involves the animal producing a specific behavior in response to a stimulus. It can be contrasted with "classical conditioning" or "Pavlovian conditioning" which is characterized by a stimulus-stimulus association. In this kind of associative learning, a neutral stimulus (the conditioned stimulus, conditioned stimulus (CS)), such as a tone, signals the delivery of a rewarding or aversive event (the unconditioned stimulus, US), such as a food reward or an electric shock. As the animal learns the association between the two stimuli, it adapts its behavior (the conditioned response, CR) in consequence, either by salivating or licking in anticipation of a reward or by freezing in expectation of a foot shock.

Pavlovian conditioning, as a behavioral mechanism, plays an important role in animal behavior. It allows us to recognize and adapt to significant cues in our

environment. For example, the sound of a car alerts us to an imminent danger whilst the sight of a cookie box signals the presence of a tasty snack. As an experimental paradigm, it is no less interesting because it provides an easy and controllable way to study how animals learn to associate events. It requires less training than more complex instrumental conditioning tasks and is highly customizable.

We will now present the different models that have been developed to describe how associative learning works in animals.

2.1.1 Reinforcement learning

Following the discovery of Pavlovian conditioning, the theory of reinforcement learning was mathematically described by Bush and Mosteller, 1951 with the following equation:

$$A_{next_trial} = A_{current_trial} + \alpha(R_{current_trial} - A_{last_trial}) \quad (2.1)$$

In this equation, the probability that the animal will exhibit the conditioned behavior on the next trial A_{next_trial} is dependent on the rewards obtained during previous trials and the difference with how much reward was expected. α is the learning rate. This was the first model to suggest a sequential error-based rule for reinforcement learning.

2.1.2 Rescorla–Wagner model

The most influential theory of associative learning was proposed by Rescorla & Wagner (Rescorla, 1972). The Rescorla-Wagner model posits that the strength of the behavioral response (CR) from the animal depends on the strength of the association between the CS and the US. During each trial or co-occurrence of two stimuli, the associative strength of the CS is updated according to the following equation :

$$\Delta V_S = \alpha\beta(\lambda - V_T) \quad (2.2)$$

λ being the magnitude of the US, V_T the sum of the associative strengths of all the stimuli present during the trial, and $\alpha\beta$ two learning-rate parameters whose

values vary between 0 and 1. The value of α is determined by the salience of the CS, and that of β by the characteristics of the reinforcer. This model was especially interesting as it explained behavioral phenomena not yet explained by previous models such as blocking and inhibitory conditioning.

Although the Rescorla-Wagner model seems to describe well the growth of associations between the CS and US, it does not explain how the internal representations of these stimuli are formed, and how they influence behavioral performance.

In 1975, Mackintosh proposed that the associability α of a stimulus depends on how accurately it predicts reinforcement (Mackintosh, 1975). If the conditioned stimulus is the best available predictor of a reward, its associative strength will be high, and conversely, if the CS is a poor predictor of reinforcement, its associative strength will be low. A stimulus S is considered a good predictor of the US if the difference between its current associative strength and the asymptote for conditioning ($\lambda - V_S$), is less than ($\lambda - V_X$), where V_X is given by the sum of the associative strengths of all the other stimuli present in the environment at the same time as S, and inversely. If S is a poor predictor, its associability α_S will decrease. Equation 2 shows how the associability of S α_S influences the rate of conditioning to S:

$$\Delta V_S = \alpha_S \beta (\lambda - V_S) \quad (2.3)$$

This theory predicts that attention to a stimulus will increase if it is the best available predictor of reinforcement.

2.1.3 Temporal difference learning model

The need for the temporal difference model (Temporal difference reinforcement learning (TDRL)) arose from two main limitations of the Rescorla-Wagner model:

- Time is treated as a fixed period ("trials") that always follows one another. In reality, time is continuous and trials can be irregularly spaced. Intuitively, one assumes that learning is influenced by the salience of the conditioned stimulus as well as by the time interval between the CS and the US. A very rare event is less likely to form a strong association than a frequent one, and

an event immediately followed by a reward is more likely to be considered biologically relevant than if the reward is delayed a few minutes afterward (this last phenomenon is also partially explained by temporal discounting which will be treated in section 2.2.3).

- This model only handles simple associations between two events but could not deal with sequential cues (an event A predicting an event B that predicts an event C).

In the Bush and Mosteller reinforcement learning theory, reward prediction error (Reward Prediction Error (RPE)) is the difference between the weighted average of past rewards and the reward received during the previous trial (Bush and Mosteller, 1951). On the other hand, Sutton and Barto's RPE is the difference between the animal's general expectation of future rewards and the clues that comfort or contradict this expectation (Sutton and Barto, 1981).

To achieve a more continuous representation of time, the TDRL model uses "states" that represent the environment at different time points of the task (it could be one state per second for example). Each state is associated with a value (learned and estimated) that represents the reward expected at this state and the sum of the rewards (adjusted on time) expected in the subsequent states (equation 2.4).

$$V(S_t) = \sum_{\tau=1}^{\infty} \gamma^{\tau-t} r(S_{\tau}) \quad (2.4)$$

Where S_t is the state S that the animal is in at the current time t, $r(S_{\tau})$ is the reward at time τ , and γ is a discount factor that decrements future rewards.

The value attributed to each state is then updated once the animal has progressed to another state. This is done by adding the prediction error to the previous state's value. If the prediction error is positive (after an unexpected reward) then the value will increase, but if it is negative (i.e. a predicted reward doesn't materialize) the value will decrease.

To generate a positive reward prediction error, one would have to give an unexpected reward to the animal or give him a clue that he'll receive such an initially unexpected reward in the future. This last point in particular is im-

portant as it allows the learning of sequential cues. If the reward is predicted by cue B but cue B is predicted by cue A, then cue A conveys the entirety of the information and no prediction error will be observed at the presentation of cue B.

But how does this model allow for learning? When an animal is first presented with the CS, it conveys no information (a dog will not salivate when it first hears a bell). With this model, the learning process only happens at the time of reward delivery, this generates a prediction error. So how do we go from this situation to a situation in which the CS generates the prediction error and not the reward? To this end, the TDRL model modifies the values of several states preceding the reward delivery based on the reward itself. This implies that the values predicting the reward progressively shift from the moment of reward delivery back to the first point in time from when it is possible to predict this reward, i.e. the CS. This phenomenon is called back-propagation and it is one of the main limitations of the TDRL model since it has never been observed in biological conditions. Alternative hypotheses will be discussed in section 2.3.

2.2 Dopamine and Reward Prediction Error

These models for associative and reinforcement learning were made primarily for computer science and machine learning. It has since turned out that they can adequately describe biological observations, both on the behavioral and cellular scale. The first parallels between reward prediction error and dopaminergic neurons were made in the 1990's.

Dopamine was already known to be released at reward delivery and was hypothesized to play a role in motivation Satoh et al., 2003.

What Schultz et al. observed in 1993 is that, in the context of associative conditioning, dopaminergic neurons only fired at reward delivery during the early phase of learning, when the association is not yet formed. Afterward, in experienced animals that have learned the association between the CS and the US, dopaminergic neurons were no longer activated anymore by reward delivery (Schultz et al., 1997). In addition, they observed that in expert animals, dopaminergic neurons fire

bursts of action potentials during CS presentation. At first, these responses were interpreted as transient changes of attentional and motivational processes, but then another group noticed that the activity of DANs very closely resembled the reward prediction error as it is defined in the TDRL model (Montague et al., 1996). Subsequent recordings in monkeys subjected to a very simple Pavlovian conditioning task confirmed that the increase of dopaminergic activity shifts from reward delivery to CS presentation throughout training. Another aspect that really confirmed that DANs were encoding reward prediction errors was their reaction to reward omission. If the CS is presented to an animal having learned the CS-US association, but then the US is then omitted, the reward prediction error at the time of expected reward delivery should be negative, because the predicted value associated with this state was high but the actual value was very low. This negative RPE was observed in DANs activity in the form of a pause in their firing at the time of reward omission. As such, the theory that DANs encode RPE during associative tasks was formulated and is still the dominant paradigm today. Many experiments performed since 1997 corroborate the RPE theory and allow for a better understanding of associative learning. In the following sections, we will describe their main findings.

2.2.1 Value

The dopaminergic response to either an unexpected reward or to a cue predicting a future reward reflects the value of the said reward. This effect has principally been observed in the VTA. In a task where different cues predict different volumes of the same reward, the activity of dopaminergic neurons scales linearly with the amount of future reward (Fiorillo et al., 2003; Roesch et al., 2007; Tobler et al., 2005). This scaling is observed both at the population level (Cohen et al., 2012; Roesch et al., 2007) and at the single neuron level (Tobler et al., 2005, 2.2 A and B).

Some sub-populations of dopaminergic neurons in the VTA have also shown decreased activity when the animal receives a punishment, or when a cue associated with punishment is presented (Masayuki Matsumoto and Okihide Hikosaka, 2014; Matsumoto and Hikosaka, 2009; Ungless et al., 2004).

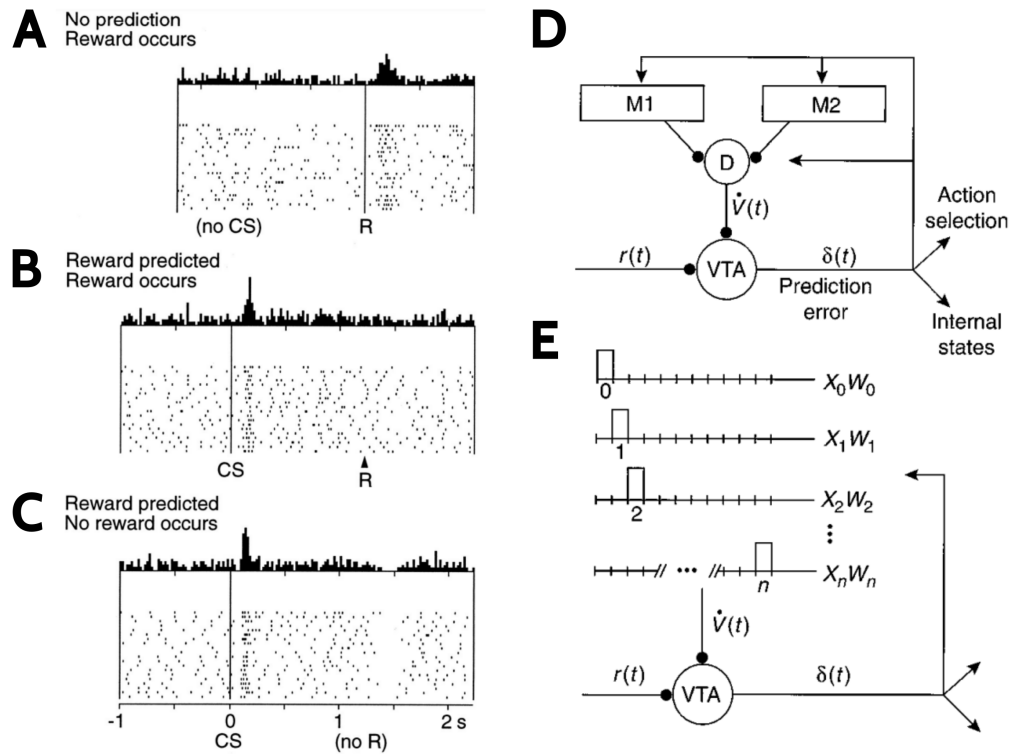


Figure 2.1: **Dopamine encodes Reward Prediction Error.** **A:** When a reward is given unexpectedly, dopaminergic neurons respond to the reward. **B:** When the reward is associated with a conditioned stimulus (CS) such as an auditory cue, the dopaminergic neurons respond to the CS and are unaffected by the reward delivery. **C:** When the reward is associated with a CS and the reward is omitted after the CS, the dopaminergic neurons respond to the reward omission by a brief inhibition. **D:** Interpretation of the anatomical arrangement of inputs and outputs of the ventral tegmental area (VTA) in the context of a TDRL model. $V(t)$ represents the degree to which the current sensory state differs from the previous sensory state. Information about reward $r(t)$ also converges on the VTA. The VTA output is taken as a simple linear sum $\delta(t) = r(t) + \dot{V}$. **E** Temporal representation of a sensory cue, at multiple delays x_n from its initial time of onset, and each delay is associated with a separately adjustable weight w_n (From Schultz et al., 1997).

Recently, Y. Wang et al., 2021 argued that, in addition to phasic activity, the tonic activity of dopaminergic neurons also encodes the reward value. They used a task in which a changing visual cue indicated the amount of future reward. The amount of reward increased gradually before settling at a final value. Dopaminergic neurons recorded during this task increased their tonic activity as the visual cue increased in value. At the population level, a 'ramping' of dopaminergic activity was observed, increasing until the visual cue stopped changing. The opposite effect was also observed if the changing visual cue indicated a decrease in the amount of reward. In this case, the firing rate of dopaminergic neurons progressively decreased as the predicted value decreased (2.2 C and D).

Thus, dopaminergic neurons of the VTA encode the perceived value of upcoming rewards.

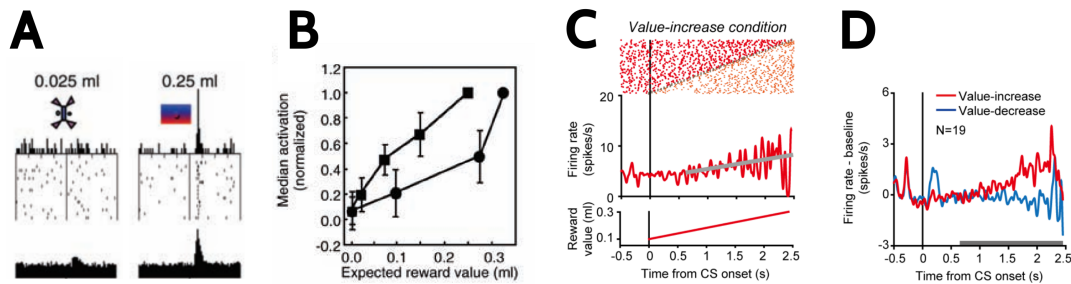


Figure 2.2: **Dopamine RPE signal scales with reward value.** **A:** Single-neuron (top) and population responses (bottom) (=7 neurons) to visual conditioned stimuli. The expected magnitudes of the rewards are shown above the rasters. **B:** Average response of dopaminergic neurons as a function of expected liquid volume after the onset of visual stimuli for two monkeys (**A** and **B** adapted from Tobler et al., 2005). **C:** Activity of an example dopamine neurons in a "progressive value-increase" task. Rasters and spike density functions are aligned by the conditioned stimulus (CS) onset. Red and orange rasters indicate spikes occurring before and after, respectively, the reward value has stopped increasing. The spikes occurring before the stop were used to calculate the SDF. The thick gray line indicates the regression line between firing rate and time. The bottom graph indicates the reward value as a function of time. **D:** Averaged spike density functions of 19 dopamine neurons with a significantly positive regression slope in the value-increase condition (**C** and **D** from Y. Wang et al., 2021).

2.2.2 Uncertainty

Not only can dopaminergic neurons encode the expected intrinsic value of the future reward, they can also encode the probability of getting a reward. A cue associated with a low probability of reward will elicit a lower dopaminergic response than a cue associated with a high probability of reward. This scales in a linear fashion and was first observed by Fiorillo et al., 2003. Interestingly, as the dopaminergic response to the cue increased with reward probability, the dopaminergic response to the reward delivery followed an opposite trend. When the cue signaled a low reward probability, then the surprise occasioned by the delivery of the reward generates a reward prediction error.

More surprisingly, Fiorillo et al also observed a 'ramping' activity of dopaminergic neurons from the onset of the cue to the expected time of the reward, and this sustained increase in activity was at its highest when the reward probability was 0.5. This suggests that dopaminergic activity correlates with uncertainty and that a higher uncertainty is associated with a stronger increase in dopaminergic tonic activity (Fiorillo et al., 2003). Although this particular result has not been replicated since, De Lafuente and Romo, 2011 also reported an increase in dopaminergic activity in situations with higher perceptual uncertainty.

In 2021, Wei et al., showed that this dopaminergic scaling to reward probability translated to different dopaminergic release profiles depending on striatal target regions, with the highest scaling in the DMS. This scaling effect is also present in the DLS and the VS but there is less difference between a high- and low reward probability cue (Wei et al., 2021).

2.2.3 Timing

The potential implication of dopaminergic activity in timing is very interesting. The most studied and better-understood mechanism is the temporal discounting of reward, but other aspects of time management could involve dopaminergic neurons. This includes but is not limited to the measure of interval duration or the timing of movement initiation.

The basal ganglia, and more precisely the striatum and the dopaminergic nuclei, are thought to manage time intervals ranging from seconds to minutes (hours being

tracked by the circadian rhythm mechanism and structures like suprachiasmatic nuclei, and sub-seconds intervals by the cerebellum) (Buhusi and Meck, 2005). Dopaminergic neurons appear to be well situated for such function, and some models suggest that their pacemaking property is responsible for time perception. Since dopaminergic neurons have very regular activity, the number of action potentials between event A and event B could be used to measure the A-B interval duration, and is analogous to memorizing the length of A-B interval duration.

Temporal discounting

In the TDRL model, the reward expectation is calculated using the discounted sum of future rewards. What this means is that a reward expected within an hour has less value than a reward expected in 10 seconds. It has been widely observed across species, humans included, that animals tend to value short-term rather than long-term rewards (Richards et al., 1997). Behavioral experiments have shown that the value of a future reward decreases exponentially (or hyperbolically; Kobayashi and Schultz, 2008) with the time delay to reward delivery.

Temporal delay also affects associative learning, as longer delays between the CS and the US decrease the learning efficacy (Holland, 1980).

In accordance with the RPE model, it has been observed that the dopaminergic neurons indeed respond differently depending on reward delay. The phasic response to the CS, supposed to convey the expectation for future reward, decreases as the CS-reward interval increases (Fiorillo et al., 2008; Kobayashi and Schultz, 2008; Roesch et al., 2007, 2.3 A). Conversely, dopaminergic activity at the time of reward delivery, which disappears when the reward is completely predicted by the CS, is in fact modulated in a non-binary way depending on the CS-reward interval. Indeed the longer the interval, the bigger the dopaminergic response. (Fiorillo et al., 2008). This last mechanism is due to temporal uncertainty, a lack of precision in the reward prediction, and is going to be explained in more detail in the next subsection.

It, therefore, seems that the dopaminergic response is affected by temporal delays in the same way as the behavioral response is. Dopaminergic response to a delayed reward is similar to the response to an immediate smaller reward, indicating

that DANs encode the subjective rather than the objective value of future rewards.

Interval duration

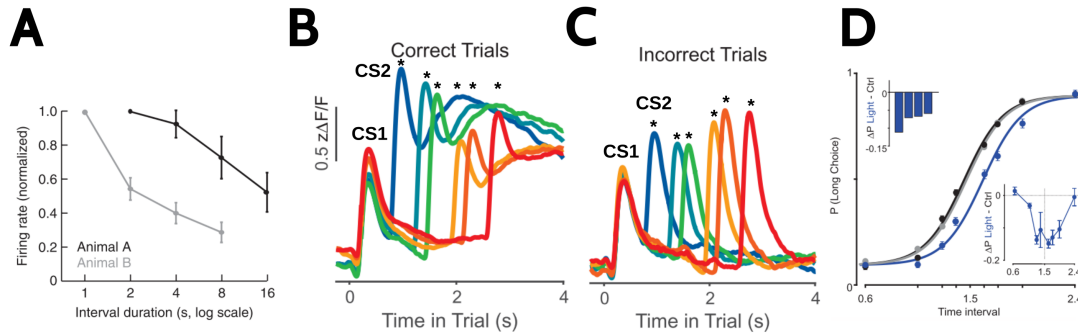


Figure 2.3: **Dopamine signals time intervals.** **A:** Mean population responses (\pm s.e.m., 11 neurons in animal A, 30 in animal B) to each conditioned stimulus as a function of stimulus-reward interval (from Fiorillo et al., 2008). **B-C:** Average measured dopaminergic responses for all intervals during correct (**B**) and incorrect trials (**C**). The black stars represent the times at which the CS2 was played. Blue and green lines represent 'short interval' trials while yellow and red lines represent 'long interval' trials. **D:** Choice behavior and psychometric curves during control trials (black), photoactivated trials (blue), and unstimulated trials immediately after photoactivation (gray) ($n = 4$ mice). 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). (B-D from Soares et al., 2016)

In the TDLR model, dopaminergic neurons are supposed to know the interval between the CS and the US. This can be observed if an anticipated reward is omitted. Indeed reward omission causes dopaminergic neurons to pause their activity at the expected time of reward delivery.

In 2008, Fiorillo et al. attempted to determine the precision with which dopaminergic neurons predict the timing of reward delivery (Fiorillo et al., 2008). Monkeys were trained to do a classical conditioning task. Visual stimuli were associated with a fixed delay to reward, ranging from 1 to 16s depending on the CS. Anticipatory licking was used as a readout of learning and time prediction. They observed that the animals started licking about halfway through the stimulus-

reward interval, but that the precision decreased with longer intervals. The intensity of licking gradually increased until it peaked around the time of reward delivery. If the reward was omitted then the licking would continue for a few seconds before stopping. This low precision (in the seconds range) shows that there is a certain temporal uncertainty, that scales linearly with interval duration.

More recent work by Soares et al. sought to investigate the contribution of DANs to the timing of behavior (Soares et al., 2016). They trained mice to perform a temporal discrimination task in which the animals had to measure the interval between two stimuli and determine if it was shorter or longer than 1.5s. The mice's performance decreased when the interval duration was close to 1.5s (the decision boundary), following a standard sigmoid psychometric function. Interestingly, pharmacological inhibition of dopaminergic activity decreased overall performance. Calcium recordings of dopaminergic activity showed that DANs response to the second stimuli (marking the end of the interval) encoded information about elapsed time. The response was on average larger for shorter intervals. However, when the animal incorrectly classified the interval, the pattern of dopaminergic activity was reversed, and the response was larger for intervals in the long category (2.3 B and C : in C the dopaminergic response during short intervals is lower as they were incorrectly classified as long intervals, and vice et versa). This suggests that the dopaminergic signal reflects the subjective assessment of the interval duration and not its actual duration. This activity was correlated with the behavioral output as they observed that when dopaminergic activity was low the animal tended to classify the interval as "long", versus "short" when dopaminergic activity was higher. Manipulating dopaminergic activity by optogenetic activation or inhibition successfully shifted the behavioral output, indicating that dopaminergic neurons might play a causal role in timekeeping (2.3 D).

2.3 Alternative hypothesis to RPE

2.3.1 'Adjusted net contingency for causal relations' model

Although the RPE paradigm is well established and has been backed by many studies in the past few decades, a recent paper by Jeong et al., 2022 called for the

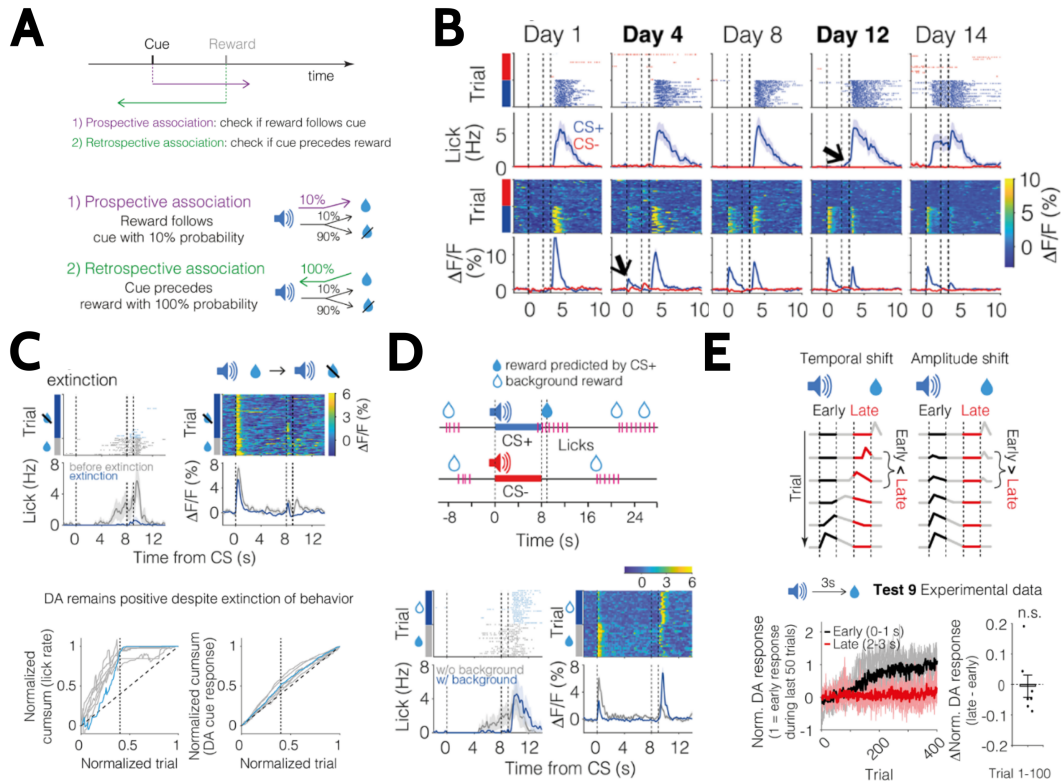


Figure 2.4: **The ANCCR model suggests that dopamine might drive retrospective learning.** **A:** Animals can learn cue-reward associations either prospectively (“does reward follow cue?”) or retrospectively (“does cue precede reward?”). **B:** Licking and dopamine measurements from an example animal showing that the dopamine response to CS+ significantly precedes the emergence of anticipatory licking (Days 4 vs 12 respectively, shown by the arrows). **C:** Top: Anticipatory licking and dopamine release in an example animal before and after extinction. Trials are shown in chronological order from bottom to top. The three vertical dashed lines indicate cue onset, cue offset, and reward delivery. The dopaminergic cue response of an example animal remains positive well after it learns the extinction of the cue-reward association. Bottom: Across all animals, the dopaminergic cue response remains significantly positive despite abrupt behavioral learning of extinction. Each line is one animal, with the blue line corresponding to the example. The dashed vertical line is the trial during which the experimental condition transitions. **D:** Top: Experimental paradigm to reduce retrospective association while maintaining prospective association, by degrading the cue-outcome contingency. Bottom: Licking and dopamine release from an example animal. **E:** Top: Schematic of learning dynamics for pre-reward dopamine dynamics based on RPE or ANCCR signaling. If there is a temporal shift, the difference in dopamine response between early and late phases of a trial will be negative in the initial trials. Bottom: Dynamics of dopamine response during early and late periods within a trial over training (left), and their difference during the first 100 trials. (A to E adapted from Jeong et al., 2022)

reevaluation of dopaminergic signaling meaning.

The Rescorla-Wagner or TDRL models use prospective predictions to associate a cue with a reward. In this framework, animals learn how often a reward follows a given cue; the cue predicts future rewards. Reward prediction error, conveyed by dopaminergic neurons, updates this prediction when the outcome does not match. The prospective nature of this type of learning model is susceptible to criticism. In a natural context, the number of cues far outweighs the number of rewards. How would dopaminergic neurons signal reward prediction for every potentially relevant cue in a complex environment?

An alternative way to learn cue-reward associations would be to search for the cause of the reward because rewarding events are relatively rare. If there is a causal relationship between a cue and a reward, then the cue always precedes the reward, even if the reward probability associated with this cue is not 100%. In this framework, animals learn, retrospectively, whether the cue consistently precedes the reward (fig. 2.4 A).

Jeong et al., hypothesized that dopamine might play a role in such retrospective learning instead of encoding RPE. They designed a retrospective causal learning algorithm that infers whether a cue is the cause of a reward by measuring whether it precedes the reward more than that expected by chance, and builds a cognitive map of causal associations. With this algorithm, potentially relevant cues are signaled by an 'adjusted net contingency for causal relations' (ANCCR), and the authors propose that dopaminergic neurons convey ANCCR instead of RPE. Simulations of this ANCCR-based causal learning model are consistent with traditional features such as scaling with reward magnitude and probability, blocking or conditioned inhibition. However, in some experimental settings, the two models make different predictions. The authors tested a few of them to identify which model fitted the best with experimental data. The most interesting of these discrepant predictions are discussed below.

Dynamics of dopaminergic and behavioral learning

In the TDRL model, RPE signal has a direct effect on behavioral learning. As we saw in 1.7.3, dopamine can strengthen cortico-striatal synapses retroactively

and increase striatal MSNs excitability, thereby reinforcing rewarded behavior. In Pavlovian experiments, dopamine has been shown to be necessary for the learning of anticipatory licking (Morrens et al., 2020; Yamaguchi et al., 2022). The TDLR model predicts that during learning, the dopaminergic shift from the reward to the cue happens concomitantly to the appearance of anticipatory behaviors. By contrast, in the ANCCR model, associations are learned first and timing is learned second. This means that a cue will become relevant (and therefore will have an ANCCR signal) before the structure of the task can be learned, leading in fine to an updated and adapted behavior. The process associating the cue with its rewards is distinct/separate from the one that encodes the temporal delay between cue and reward. Thus the ANCCR model predicts that the dopaminergic response to the cue will appear long before anticipatory behaviors. They used the dopamine sensor dLight 1.3b, expressed in the NAc, to investigate the evolution of dopaminergic signaling during learning, and observed that the dopamine response to the cue became significant several days before the animals showed anticipatory licking, meaning that the dopaminergic activity and the behavioral output are indeed dissociated (fig. 2.4 B).

Extinction

In the same paper, they then tested the extinction of a learned cue-reward association. If the reward abruptly ceased to be delivered after the cue, the TDLR model predicts that the RPE signal associated with the cue will quickly be updated to reflect the current reward prediction, and reduced to zero. The ANCCR model however predicts that the dopaminergic response to the cue will not decrease because the retrospective association between the reward and the cue has not changed. In the ANCCR framework, only a reward can update a cue-reward contingency, and a once-rewarded cue remains relevant even in the absence of reward. Because ANCCR signal and behavior are dissociated, the animal is able to adjust its behavior while the dopaminergic activity remains the same. Their experiment showed an abrupt decrease of licks following the extinction, but only a slight decrease of the dopamine signal over dozens of trials (fig. 2.4 C).

Contingency degradation

This extinction experiment can be described as a change in the prospective association between the cue and the reward, as the probability of receiving a reward after the cue decreases whilst the probability of a cue preceding the reward remains high. An opposite experiment, with a decrease of the retrospective but not of the prospective association, would be similar to a contingency degradation: by giving random unpredicted reward in addition to predicted reward, the probability that a reward is preceded by a cue decreases. The TDLR model predicts no change in the dopaminergic response to the cue in this case, as the reward prediction remains the same. On the other hand, the ANCCR model predicts a rapid drop in dopamine cue response. The experimental data showed a rapid decrease in the dopamine cue response, as well as a disappearance of anticipatory licking and an increase in dopamine reward response (fig. 2.4 D).

Back-propagation

In the TDLR model, a value is attributed to each 'state' or time point of the trial. Because the value of each state is updated when the animal moves to the next state, the model predicts that the reward will influence the prediction error of the state immediately preceding it at first and that this reward prediction error signal will back-propagate until it reaches the first cue that predicts this reward. However, this phenomenon has very rarely been observed *in vivo*.

One such rare observation was made by Amo et al., 2022 in the VTA and VS. They showed that mice learning the association between an odor (lasting 1s) and water reward (after a 2s delay), exhibited a temporal shift of dopaminergic activation after the odor presentation. During the first few trials, dopamine (measured with a GRAB-DA sensor) was released in the VS just before reward delivery. However, after a few dozen trials it was being released earlier and earlier before reaching a plateau. Similar experiments in the VS using odors as CS had not shown any back-propagation phenomenon (Flagel et al., 2011; Menegas et al., 2017). This apparent discrepancy concerning back-propagation in Amo et al., 2022 could be due to the unreliable nature of odor cues relative to auditory ones. Indeed olfaction is an active process, that benefits from training. Animals are able to detect smells earlier

and earlier (at lower concentrations) with training. This progressively shorting delay between odor release during the experiment and its detection by the animal can give the false impression of a back-propagating signal.

Unlike the TDRL model, the ANCCR model does not predict a back-propagation of the ANCCR dopaminergic signal, because in this model time is not divided in states. Jeong et al., 2022 did not observe a back-propagating dopaminergic activity in their experimental task (fig. 2.4 E). Their experiment also relied on the measurement of dopamine through a fluorescent dopamine sensor in the VS (like Amo et al., 2022), but used an auditory cue instead of an olfactory one.

This new model presented by Jeong et al., 2022 could therefore completely change the way we perceive the role of dopamine in associative learning. Whilst staying consistent with past findings, this ANCCR model could resolve some inconsistencies that exist between TDRL model predictions and experimental data.

Dopamine and instrumental learning

In this chapter, we will expand on associative learning to include situations and tasks in which an action is executed by the animal in order to obtain rewards. We will describe the pattern of dopaminergic activity during such instrumental tasks and discuss the potential functional roles of dopamine in instrumental behaviors.

3.1 What is instrumental learning?

Operant conditioning, or instrumental conditioning, is a form of associative learning but differs from classical conditioning in that the reinforcer is contingent on the performance of behavior by the animal. The animal thus learns the consequences of its own actions and alters its behavior as a result of training. As in classical conditioning, the learned behavior is extinguished when repeating it is no longer followed by food reinforcement.

Experimental paradigms of instrumental conditioning can reinforce either an action that is spontaneously initiated by the animal or trigger the desired behavior with a CS. In both instrumental and classical conditioning, a close temporal association between two stimuli is necessary. However, whereas classical conditioning consistently provides the US/reward upon presentation of the CS, instrumental conditioning demands a correct behavioral response from the animal to receive the reward.

As we saw in the previous chapter, associative learning is fairly well replicated by

the temporal difference reinforcement learning model. In such models, the actions of the subject have no impact on reward delivery. The TDRL model computes a predictive signal that is neither dependent on nor influences the subject behavior. Considering that animals are able to optimize their behavior to maximize rewards, models of reinforcement learning should take actions into consideration. This is done by introducing a component called the Actor, which is modulated by the RPE signal, sometimes called the Critic (Sutton and Barto, 1998) This approach has been shown to be efficient in machine learning studies. In a biological context, the Actor would correspond to the cortico-striatal pathway while the Critic would correspond to the nigro-striatal pathway.

3.2 Goal-directed vs habitual behavior

Behavioral responses, in the context of instrumental tasks, are generally divided into two distinct mechanisms: habitual actions and goal-directed actions. As described in 1.6.4, habits are stable, automatic responses to a specific environment, while goal-directed actions are flexible strategies to obtain a reward. It is usually considered that, early on during training, instrumental behavior is flexible and sensitive to changes in the outcome, and that as training progresses behavioral responses become more dependent on the environmental cues and less on the outcome. Such change in response mode is thought to be mediated by an association between cues that preceded the action through selective reinforcement. In a way, we could compare this process to second-order associative learning (or sequential conditioning): when cue A predicts a cue B that predicts a reward, an association is first formed between the reward and cue B, and then between cue B and cue A. In a simple instrumental task, such as a lever press, the association between the lever press (action) and the reward would appear first, after which the association between the environment (appearance of the lever, a cue, etc) and the action would be formed in a second phase. If we consider the RPE signal encoded by dopamine, when the first association is formed (between the lever press and the reward), we could assume that the RPE signal switches from reward delivery to the moment of the lever press, as is observed during Pavlovian conditioning. This dopamine influx could then serve as a reinforcing signal for a second-order association (between a

cue and the action).

In practice, distinguishing habits from goal-directed actions is difficult and relies mostly on evaluating the sensitivity of the performance to outcome value (Balleine and Dickinson, 1998). For example one of the hallmarks of habitual behaviors is their insensitivity to contingency degradation and outcome devaluation.

Moreover, the opposition and mutual exclusion of these two types of behavioral response have been debated, with recent models hypothesizing that they could actually function in parallel (Bogacz, 2020), using distinct prediction error signals (Miller, 2019).

3.3 Dopaminergic contribution to instrumental behaviors

It is well established, by decades worth of evidence, that dopamine acts as a learning signal both in associative and instrumental learning. However, it is unclear if dopamine is involved in other parts of behavioral control, such as decision-making, motivation, or movement vigor modulation. Furthermore, it is unknown if dopamine affects behavior proactively (on a trial-to-trial basis, as opposed to a progressive modification of behavior through learning). Indeed, the functional role of the dopamine RPE signal after it switches from reward delivery to cue presentation is still unknown. During instrumental tasks, a dopamine influx just before movement execution could potentially affect said movement. Going back to the cellular effect of dopamine on striatal neurons, it is not possible to exclude an 'online' modulation of behavior. Indeed it has been shown that dopamine can temporarily modify the excitability or gain of the target neurons (Lahiri and Bevan, 2020; Thurley et al., 2008).

In the following sections, we will investigate the activity pattern and potential functions of dopamine during instrumental tasks. We will consider the dopaminergic activity during two periods of the task, firstly during the animal's actual movements and then secondly an earlier transient caused by the cue but preceding movement onset.

3.3.1 Dopaminergic activity during movement

As we have seen in the previous chapter, the dopaminergic activity in response to a cue or conditioned stimulus has been well described. By contrast, dopaminergic activity during movement is less well-known. This is due in part to the diversity of possible movements and instrumental tasks. In this section we will describe the patterns of dopamine activity observed during instrumental tasks, starting with tasks that do not include an explicit cue before the animal's action, and then by going over tasks in which the action is triggered by a specific cue (for review, see Coddington and Dudman, 2019).

Spontaneous movements

Spontaneous movements are internally generated by the animal without any explicit environmental cue. The most commonly studied type of spontaneous movement is locomotion but more complex behaviors, such as a lever press task in which the lever is always present in the area and where there is no temporally delimited trial, can also be considered spontaneous. Reinforcing such behaviors with rewards increases the probability that the animal will perform the reinforced action.

As seen in 1.7.3, if a spontaneous behavior is rewarded, the dopamine release at reward delivery will act on cortico-striatal synapses to increase the gain of previously activated synapses. However, once the behavior is reinforced and the animal performs it regularly, the activity pattern of dopaminergic neurons is unclear. Is there a shift of the RPE signal similar to that observed during Pavlovian conditioning?

Dodson et al., 2016a recorded the electrophysiological activity of SNc and VTA neurons during spontaneous and unrewarded locomotion. They observed a tendency of SNc DANs to decrease their firing rate at the onset of the locomotion bout (fig. 3.1 A, top). Da Silva et al., 2018 also recorded SNc DANs during spontaneous unrewarded locomotion. They reported that half of their recorded units (13/25) increased their activity around movement initiation, while 7/25 units decreased their activity at movement onset (fig. 3.1 A, middle). Howe MW, 2016 used the calcium sensor GCaMP6f to record dopaminergic activity in the dorsal striatum during spontaneous locomotion. They observed frequent transients of activity

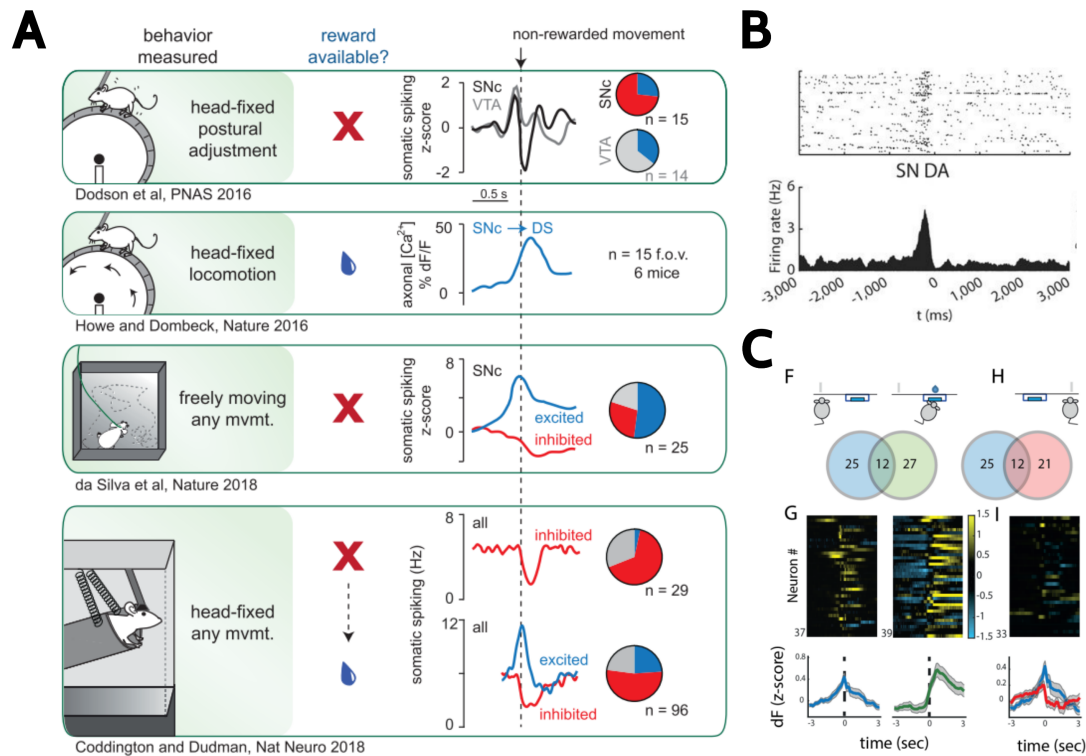


Figure 3.1: Dopaminergic activity during spontaneous movements. A: A comparison of results showing movement-related dopaminergic signals in rodents during spontaneous unskilled tasks (figures extracted from the review Coddington and Dudman, 2019). **B:** Example of positive modulation of firing rate in a SNc DA neuron before lever pressing. Top: raster plot where each dot indicates a spike; bottom: PETH for the neuron; time zero is the time of lever pressing (from Jin and Costa, 2010). **C:** Activity of all recorded ROIs of dopaminergic neurons from one session aligned to the first contralateral lever press (left), to the beginning of reward consumption (middle), and first ipsilateral press (right). PETH of positively modulated neurons for first contralateral press, reward, and ipsilateral press (bottom) and corresponding heat maps (top) (from Mendonça et al., 2021).

during locomotion, that on average occurred just prior to acceleration onsets (fig. 3.1 A, middle). The amplitude of those calcium transients correlated with the amplitude of the immediately following acceleration bursts. As the spontaneous movements studied in these three examples were unrewarded, we cannot really speak of an instrumental task per se. Moreover, locomotion is an innate and unskilled behavior, that is less likely to rely on basal ganglia control than more complex actions.

Since there was no learning required, these recordings show the pure motor-related activity of dopamine neurons, without contamination by environmental cues or reward prediction. We will now review studies that have looked at spontaneous and rewarded movements, both with untrained behaviors and with trained, self-paced instrumental behaviors.

Coddington and Dudman, 2018 first recorded SNc neurons during locomotion in head-fixed mice that were never exposed to sucrose reward. Most of the recorded neurons were inhibited at movement onset prior to any training. They then trained the mice with a Pavlovian task and looked at the dopamine activity at the onset of spontaneous movements occurring during inter-trial interval. Even though more than half of the recorded neurons still showed inhibition, almost a quarter of them were physically excited at movement onset (fig. 3.1 A, bottom). Jin and Costa, 2010 recorded DANs activity during a spontaneous lever press task. They trained mice to perform a FR8 (fixed-ratio 8) task, in which the animal spontaneously initiated the sequence of lever press. More than 50% of their recorded SNc DANs showed a phasic burst of activity right before the first lever press of the sequence, possibly starting at movement initiation (fig. 3.1 B). Mendonça et al., 2021 recorded single cell SNc neurons using the GCaMP6f during a lever press task. All the neurons that were modulated by the motor activity showed an increase of fluorescence around the time of the lever press (with some units being active before and some after the press). Only half of those neurons were also excited by the reward consumption (fig. 3.1 C).

To summarize, except for Dodson et al., 2016a and Coddington and Dudman, 2018, recordings of dopaminergic neurons activity during spontaneous movement show that there is an increase of their activity around movement onset. It would also seem that there is a heterogeneity of responses, with a subpopulation of

dopaminergic neurons being inhibited during movement and another one being excited, and that the proportion of excited neurons increased with training and reward expectation.

Triggered movements

By opposition to spontaneous movements, most instrumental motor tasks use a 'Go' cue to trigger the behavioral response from the animal. If there is only one possible behavioral response, the cue can be part of a habitual association, but if there are multiple cues and different possible actions, a step of decision-making can be required between the cue and action initiation, and these two conditions might involve distinct dopaminergic activity.

Early recordings of dopaminergic neurons in the SNc of monkeys during a Go/NoGo reaching task showed that they increased their activity during the whole movement without correlation with movement parameters and no responses to the sensory cue (Schultz, W. and A. Ruffieux and Aebischer, 1983). However, subsequent similar experiments also using monkeys observed more dopamine response to the Go cue presentation than movement onset (Romo and Schultz, 1990; Schultz, 1986, fig. 3.2 A).

Most studies have focused on movements that are part of an instrumental task (in which the reward is contingent on correct behavioral response). A few however have looked at cue-triggered movements that are not required by the task. Those movements can generally be called anticipatory behaviors, such as anticipatory licking in a Pavlovian task. They arise after the cue-reward association has been learned and reflect the reward expectation of the animal. Coddington and Dudman, 2018 observed that the excitation of DANs at movement onset was higher for anticipatory movements than spontaneous movements (fig. 3.2 B). Barter et al., 2015 found that the firing rate of dopaminergic neurons between the cue presentation and the reward delivery correlated to the velocity and the acceleration of the mice's anticipatory head movements.

Amongst studies in which the reward is contingent on the instrumental behavior, some have observed an increase of dopamine activity at movement onset similar to spontaneous movements (Howe et al., 2014; Nomoto et al., 2010; Roitman et al.,

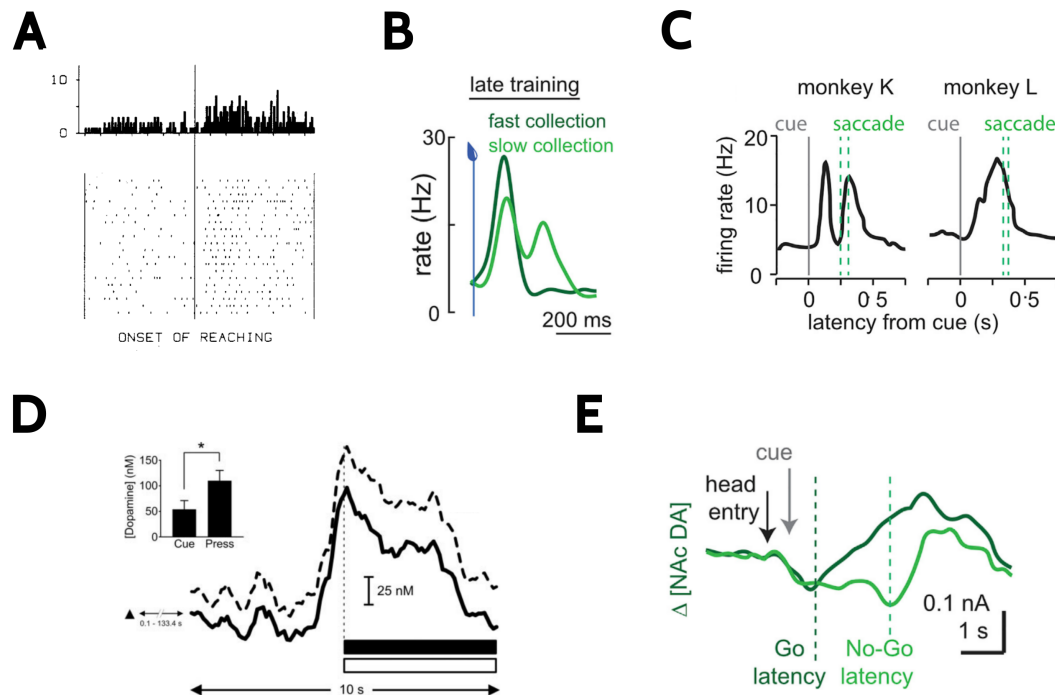


Figure 3.2: Dopaminergic activity during cue-triggered movements. **A:** The activity of a monkey’s SNc neurons during a Go / NoGo reaching task. ‘Onset of reaching’ denotes the release of the holding key. Only trials in the Go-situation are shown. The Peri event time histogram (PETH) shows the sum of impulses that are depicted as dots for individual trials ($N = 20$) below (from Schultz, 1986). **B:** Mice were head fixed on a movable platform and DA neuron activity was measured with juxtacellular recordings during Pavlovian conditioning. In late training, bimodal DA responses were seen to depend on the timing of reward-related action (first lick after reward delivery), with a delayed response component during trials with relatively slow reward collection (light green trace) (from Coddington and Dudman, 2019, adapted from Coddington and Dudman, 2018). **C:** Head-fixed monkeys performed a perceptual discrimination task: a reward for correctly indicating the motion of a random dot stimulus by saccading to one of two targets. Bimodal responses were seen in SNc-DA neurons from two monkeys, with a delayed component aligned equally as well to cue presentation as to the initiation of reward-related action (saccades) (from Coddington and Dudman, 2019, adapted from Nomoto et al., 2010). **D:** FSCV recordings in the VTA during a cued lever press task. Even when the reaction time was superior to 5s, the lever press still coincides with dopamine peak. A dramatic increase in signal (mean+SEM represented by solid and dashed black lines, respectively) was observed immediately before and peaked at the operant response (denoted by a vertical dashed line). The cue presentation (black triangle) occurred (0.1–133.4 sec) before the start of the trace. The cue presentation elicited a rise in dopamine, but rats did not press the lever until a second larger rise in dopamine happened ($p < 0.05$; left inset) (from Roitman et al., 2004). **E:** Rats performed a food-rewarded go/no-go task where following a nose poke, different cues instructed a ‘go’ to press a lever, or a ‘no-go’ to remain at the nose port for a delay before going to press the lever. Mean NAc DA concentration changes as determined by FSCV, aligned to cue delivery, with the mean action timings (‘go latency’ and ‘no-go latency’) indicated by dotted lines. Transient increases were time-locked to the moment of reward-related action in both trial types, rather than to cue delivery (from Coddington and Dudman, 2019, adapted from Syed et al., 2016).

2004; Syed et al., 2016). When the instrumental behavior was locomotion based, the dopamine started increasing at the beginning of locomotion and kept increasing until the animal reached the target (Howe et al., 2014 with mice; Syed et al., 2016 with rats). This was sometimes called 'ramping' of dopamine activity and has also been observed in a more recent study (Krausz et al., 2023). With a shorter time-scale, in a lever press task, Roitman et al., 2004 observed that dopamine's increase started at movement onset and peaked around the time of the lever press, but it is difficult to be certain that the co-occurrence of the peak and the press is not simply due to the temporal kinematics of the FSCV and the relatively uniform movement time of the animal. Finally, in one of the two monkeys of Nomoto et al., 2010, a clear dopamine burst can be observed at the onset of the eye saccade required by the task (fig. 3.2 C).

Two of the main difficulties when studying such instrumental tasks are to identify with precision the movement onset and to separate the response to the go cue from the response to the movement. Indeed, animals usually react very quickly to the go cue, and depending on the recording method used it can be difficult to isolate a movement-specific activity. Two examples of papers that have succeeded in separating them are Roitman et al., 2004 and Syed et al., 2016. The former used FSCV to record dopaminergic activity during a lever press task. As the FSCV lacked the temporal precision to separate the cue and movement responses, they used trials with a longer reaction time ($>5s$) and observed a distinct increase in dopaminergic activity at movement onset, that peaked at the time of the lever press (fig. 3.2 D). The latter designed a task that was particularly suited to counter this problem, with go trials in which the rat had to immediately Go and press a lever, and NoGo trials in which the rat had to wait before going to press the lever. In the NoGo trials, the dopamine increase in the NAc only occurred when the rat initiated movement (fig. 3.2 E).

Some studies, such as Parker et al., 2016 and Hart et al., 2022 report an increase of dopaminergic activity at the time of the lever press and not at movement onset. And finally, some studies have not observed any kind of movement-related dopaminergic activity (Mohebi et al., 2019; Owesson-White et al., 2016; Wei et al., 2021)

Lateralization

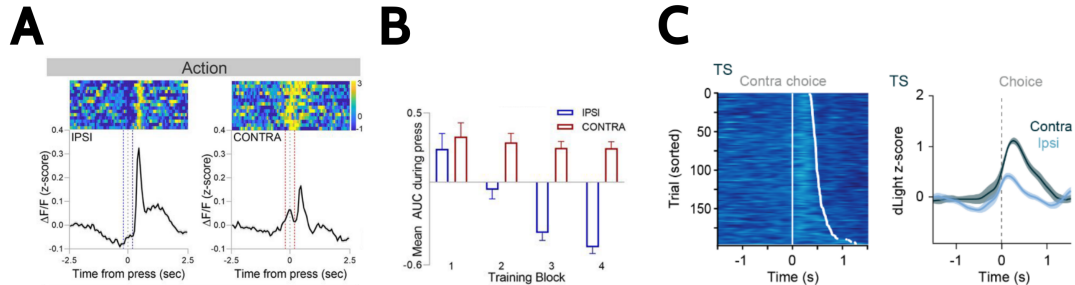


Figure 3.3: **Lateralization of dopaminergic activity.** **A:** $\Delta F/F$ dLight1.1 signals over time aligned to ipsilateral (left) and contralateral (right) lever presses. Top: Example trials showing color-coded $\Delta F/F$ signals. Bottom: the mean $\Delta F/F$ averaged across all trials during instrumental training, Action Window indicated by blue (ipsi) and red (contra) lines. **B:** The mean (+ SEM) area under the curve (AUC) of the $\Delta F/F$ signal during the Action Window for all ipsilateral and contralateral lever presses, averaged within each training block (A and B from Hart et al., 2022). **C:** Left: Example of a dLight1.1 recording session in the tail of striatum (TS), aligned to time of leaving the center port to make a contralateral choice. White dot shows the time of entering contralateral or ipsilateral choice port. Right: Average photometry traces in TS ($n=9$ mice) aligned to choice, time leaving the center port, for contralateral and ipsilateral choices. Lines represent the average response profile across mice (from Greenstreet et al., 2022).

Lateralization of motor control is the ability to produce unilateral or asymmetric movements. Pure unilateral movements are difficult to observe in rodents; indeed, even a one-arm lever press requires the activation of contralateral muscles to stabilize the body. Motor control circuitry is lateralized. Each cerebral hemisphere controls movement on the contralateral side of the body. Unilateral lesions of the basal ganglia nuclei cause contralateral symptoms such as locomotion impairment and difficulty using the contralateral forelimb.

Although the effect of a unilateral lesion of the dopaminergic neurons is well described, it remains unclear if the physiological activity of DANs is lateralized in any way during behavior. In the context of reward prediction error signaling, it would seem logical for dopaminergic activity to be identical in the two hemispheres,

as rewards are not lateralized. However, SNc DANs mainly receive inputs from the ipsilateral hemisphere, so it would be possible for the two SNc to show activity discrepancies (Watabe-Uchida et al., 2012).

Barter et al., 2015 observed that most DANs correlated with a rightward movement of the head were located in the left SNc, and vice et versa.

Greenstreet et al., 2022 recorded dopamine release in the tail of the striatum using the dopamine sensor dLight1.1 during a sensory discrimination task where the mice had to choose and move to a port (ipsi or contralateral to the recording site). They observed an increase of dopamine activity in the TS contralateral to the chosen port, which was absent in the ipsilateral TS (fig. 3.3 C). This lateralization was only present during choice and movement, and not during cue presentation or reward delivery.

Hart et al., 2022 recorded dopamine release in the DMS using the dopamine sensor dLight1.1 during ipsi and contralateral lever presses. During the movement, dopamine release was only increased in the contralateral striatum (fig. 3.3 A). In the ipsilateral striatum, although there was some dopaminergic activity observed during the early phase of training, in late training they observed an inhibition of dopaminergic activity (fig. 3.3 B). In addition, dopamine release during reward delivery and consumption was lateralized as well, with a bigger dopamine response in the ipsilateral DMS. After degrading the action-outcome contingency of one of the two levers, they observed a diminution of dopamine lateralization, with dopamine release in both striatum regions reaching almost similar levels.

If dopamine indeed is involved in motor execution or motor learning, we would expect its release to be lateralized during unilateral movements. However, it is unclear whether the lateralization occurs only for movement-related dopamine signals or for cue-generated RPE as well. If the dopamine response to cue merely encodes RPE and has no impact on motor behavior, it would probably not be lateralized; but if dopamine modulates motor parameters such as vigor, or if it encodes stimulus-response association, lateralization is to be expected.

3.3.2 Potential functions of the dopamine activity during movement

Thus we have seen that in most reported cases, dopaminergic neurons are active during instrumental movements, either at movement onset or as a ramping activity culminating at the end of the movement. There is currently no consensus as to what this movement-related activity means nor what its purpose is. Here we explore some possible hypotheses:

- Movement dopamine is a RPE signal similar to a cue-evoked RPE signal: If the cue RPE is a response to sensory input, then the movement RPE could, in a similar way, be a response to the sensory-motor inputs generated by the movement, or else a response to a copy of the motor signal sent by the cortex. This movement RPE could be used to predict upcoming rewards; if the animal knows that it needs to press a lever in order to receive a reward, the sensory-motor feedback from the lever press signals that the animal completed the task successfully and that this will lead to a reward. In a maze, walking towards the reward gives continuous updates on the (spatial and temporal) proximity of the reward (Krausz et al., 2023).
- Dopamine activity modulates the vigor of the current movement. As it has been observed that dopamine can change the excitability of striatal MSNs, we could assume that dopamine released at the onset of the movement facilitates cortico-striatal communication during the movement and thus increases its speed or strength (Da Silva et al., 2018; Mendonça et al., 2021). We will describe this hypothesis in further detail in 3.3.2.
- Dopamine during movements drives learning across trials to strengthen the stimulus-action association. This hypothesis has been investigated by (Greenstreet et al., 2022). They recorded dopamine release in the tail of the striatum during a decision-making task and found that the dopamine release during the movement could be modeled as an 'Action Prediction Error' (APE) signal, that encodes how close the current movement is to the movement that is usually performed in response to a given cue. They hypothesized that this

signal contributes to habit formation. We will describe this hypothesis in further detail in 3.3.2.

Vigor

Vigor is a term that is generally used to describe the speed of a movement or an association of the movement speed and the reaction time between a stimulus and the movement. Our vigor can be influenced by our energy level as well as our motivation and the expected reward of an action. In humans, reward expectation has been shown to influence both motor speed (Summerside et al., 2018) and reaction time (Manohar et al., 2015). In a decision-making task, Thura et al., 2014 observed that the urgency of an ongoing decision strongly influenced the speed and duration of the movement.

Due to the bradykinesia symptom observed in patients with Parkinson's disease, dopamine was very quickly associated with vigor. However, the actual contribution of dopamine to the control of vigor in physiological conditions is still heavily debated. Dopaminergic activity has been shown to encode kinematic parameters of movements (Barter et al., 2015; Da Silva et al., 2018; Engelhard et al., 2019; Howe MW, 2016; Saunders et al., 2018, see fig. 3.4 A-C) and it has thus been hypothesized that dopamine could control the gain of cortico-striatal inputs during motor execution and hence modulate vigor.

Bova et al., 2020 found that consistent dopaminergic stimulations changed the reaching kinematics of rats over dozens of trials. But once this change was established, the kinematics for a given trial depended strongly on the dopamine levels during that trial.

This potential role of dopamine in controlling vigor was modeled by Niv et al., 2007. The authors argued that it is the tonic levels of dopamine that influence the average movement vigor. This is also the conclusion of a more recent paper by Liu et al., who measured the speed and amplitude of movement of mice in a lever push task. They observed that dopamine depletion impaired the movement vigor and that levodopa injections were sufficient to restore it (H. Liu et al., 2022, fig. 3.4 D). In addition, optogenetic stimulation of dopaminergic neurons in healthy mice failed to increase vigor.

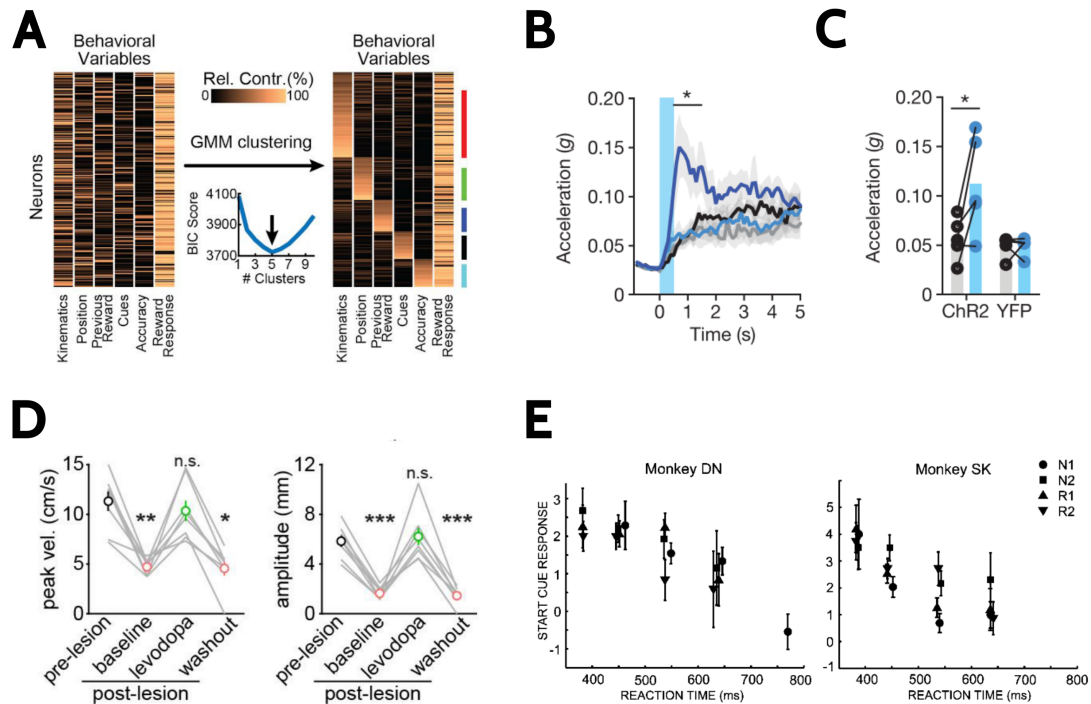


Figure 3.4: **Dopamine is involved in the control of vigor.** **A:** Functional heterogeneity of VTA DA neurons. Left: Relative contribution of each behavioral variable to explain variance of neural activity for each neuron before clustering (all neurons and variables are shown). Right: Same data grouped based on GMM clustering (ordered within each cluster by each neuron’s probability to belong to the cluster) (Adapted from Engelhard et al., 2019). **B:** Transient SNc dopamine neuron activation promotes movement initiation. Mean acceleration after 20Hz ChR2 stimulation. $n = 5$ mice per group. Dark blue, ChR2 on; light blue, ChR2 off; black, YFP on; grey, YFP off. **C:** Mean acceleration from 0 to 1 s. $n = 5$ mice per group (B and C adapted from Da Silva et al., 2018). **D:** Left: Median peak velocity of lever pushes across conditions. $p < 0.001$, Friedman; Dunn–Šidák post-hoc comparisons between pre-lesion vs. post-6OH-DA-lesion baseline ($p = 0.002$), vs. levodopa ($p = 0.99$), vs. washout ($p = 0.012$). $N = 8$ mice. Right: Peak amplitude of lever pushes. $p = 0.002$, Friedman; Dunn–Šidák: pre-lesion vs. post-lesion ($p = 0.022$), vs. levodopa ($p = 1.000$), vs. washout ($p = 0.006$) (From H. Liu et al., 2022). **E:** Correlation of the magnitude of neural responses to the CS in 52 DA neurons in monkey DN and 54 DA neurons in monkey SK to reaction time (RT) to depress the start button after the CS. The trials were classified into five groups on the basis of the RTs, and the mean and SEM of DA neuron responses in these groups of trials are plotted (From Satoh et al., 2003).

All of these studies assumed that it was the dopaminergic activity during (Barter et al., 2015; Bova et al., 2020) or immediately before the movement (Da Silva et al., 2018; Mendonça et al., 2021) that might be responsible for vigor control (or at least vigor encoding). But it would seem unlikely that dopamine could modulate in real time a movement already in execution. However, when the movement is triggered by a go cue, the cue-evoked dopaminergic response could change the striatal excitability before the cortex sends inputs to initiate the movement. Indeed, if the cue-evoked response conveys information about the value of the eventual reward, it could very well be responsible for vigor control. Satoh et al., 2003 investigated this question, and found that the magnitude of the cue-evoked responses was negatively correlated with the reaction time of the animal (fig. 3.4 E). They hypothesized that this signal represented motivational properties, and it could also be defined as vigor.

Action Prediction Error

APE was defined as the prediction error of the current action as compared to the same action as habitually performed in the same environment. This theory has roots in the "DopAct" framework proposed by Bogacz in 2020, stating that parallel systems in the basal ganglia were in charge of value, goal, habit, and that over time a goal-directed action stopped being overseen only by the goal system and was managed by the habit system as well (Bogacz, 2020). This conjoined control by a "habitual controller" and a "goal-directed controller" had also been hypothesized by Miller, 2019, who showed that a model combining habit and goal-directed control more closely fitted behavioral data than traditional models. Both of these theoretical works show that basal ganglia can implement a value-free learning system in addition to the value-based learning system that is usually described (in which the RPE is involved). The two learning systems could function in synergy and be more or less important at different stages of learning or during different tasks. If the RPE is the teaching signal used by the value-based system, then another type of dopaminergic signal could be the teaching signal for the value-free system. RPE teaches the animal to repeat an action that lead to a reward in the past, and APE would teach the animal to repeat an action that has been taken in

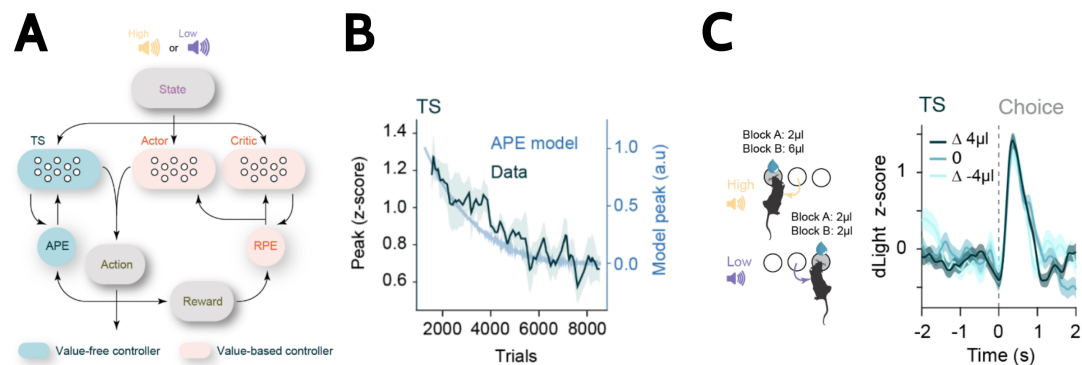


Figure 3.5: **The Action Prediction Error model.** **A:** Schematic of the BG network model including APE. The value-based control system where learning is driven by RPE, consists of an actor and a critic striatal region that learn to control actions and generate reward prediction, respectively. The value-free control system that mimics the tail of striatum (TS), is updated by APE and learns to control action and generate action predictions. **B:** Movement-evoked dopamine release in the TS decreases with learning, which is inconsistent with the RPE model but in line with the APE model. **C:** Left: Schematic of the predicted value manipulation task design. For each mouse, the value of the left and right port was altered in separate sessions. The RPE model predicts an increased response for $\Delta 4$ blocks whereas the APE model predicts no change. Right: Example of movement-aligned TS dopamine response when the relative value of the cues changed. The solid line represents the mean responses for one mouse across trials for each trial type (A to C from Greenstreet et al., 2022).

the same conditions in the past (fig. 3.5 A). The hypothesis of Greenstreet et al., 2022, is that APE is the difference between the action performed by the animal and the extent to which this action was predicted given the context.

The notion of context or environment is key here and stems back to the definition of habits. In opposition to goal-directed behavior, that is driven by the expectation of the outcome, habits are driven by associations between cues present in the environment and the action that often follows those cues. It is described as "stimulus-response" association. In the same environment, an animal is likely to repeat the same actions. It is thus hypothesized that dopamine could form a prediction of which action is going to be taken given a set of cues (light, sound, objects). If the action ultimately performed differs from this prediction, an APE would occur.

In their paper, Greenstreet et al., 2022, studied the dopaminergic activity in the tail of the striatum (TS) during an auditory discrimination task. They observed that dopaminergic neurons were active during contralateral movement (locomotion toward a lever), and that this activity tended to decrease during learning (fig. 3.5 B). Their interpretation was that as a behavior becomes more and more habitual, the basal ganglia forms better predictions of expected actions and thus the action prediction error encoded by dopaminergic neurons decreases. The dopaminergic signal recorded in the TS during this task was insensitive to reward or value changes, in accordance with habits theory (fig. 3.5 C). Finally, they observed that, on a trial-to-trial basis, the amplitude of the dopaminergic signal during an action biases the subsequent trial, with larger dopaminergic responses increasing the probability of repeating the same action.

3.3.3 Cue-evoked dopaminergic response in instrumental tasks

In this section, we will describe the dopaminergic response to 'Go' cues in instrumental tasks. Indeed, it is unclear if the dopaminergic response to the cue in such a task would be identical and have the same properties as the cue-evoked RPE signal that is observed during Pavlovian conditioning. One could expect, considering the frequent need to choose an appropriate action in response to environmental stimuli,

and to sometimes perform the selected movement with a precise speed, that the cue-evoked dopamine release plays a role in these processes. Dopamine could also help the animal decide whether or not it wants to act based on the expected reward associated with the stimulus. In the next paragraphs, we will investigate these hypotheses.

Action selection

In 2006, Morris et al., recorded monkeys during decision-making tasks and showed that the activity of dopamine neurons in response to the cue was modulated according to the value of the upcoming action. In addition, they observed that the action chosen by the animal was determined more by the dopaminergic activity than by the actual reward of the chosen action. This shows that dopamine may play a role in immediate decision making, but that it could also drive long-term changes in decision policy (the probability to choose a certain action in response to a given cue) (Morris et al., 2006).

Intuitively, we would assume that in order to select the action associated with the largest reward, or the largest probability of reward, the animal needs to estimate the expected rewards associated with each choice. However, Roesch et al., 2007 showed that in fact, the cue-evoked response of dopaminergic neurons in rats reflected the value of the most valuable option. This was true both for the difference in the size of the reward as well as the delay before the reward, showing similarities with the RPE signal of Pavlovian tasks. Moreover, and contrarily to the previously mentioned paper by Morris et al., the cue-evoked dopaminergic activity always reflected the highest value option, even when this option was not selected by the animal. Only after the option was selected, did dopamine signal the actual value of the chosen action.

Parker et al., 2016 tried manipulating action selection by inhibiting dopaminergic neurons in mice, and found that optogenetic inhibition during the entire duration of the trial (randomly selected trial) alters subsequent choices. However, if inhibition was limited to a subset of time within the trial (before the action for example) it did not have any effect. Similarly, Lee et al., 2020 did not observe any effect of optogenetic inhibition before and during the movement.

Evaluation of cost and effort

Decision-making, either to choose between two actions or to decide whether or not to execute an action, requires the consideration of the cost-to-benefit ratio of the action (Westbrook et al., 2021). Indeed, it is not economically profitable for animals to execute a movement that requires more energy than the reward will provide. Pharmacological inhibition or lesion of dopamine inputs to NAc biases animals towards low-cost low-reward actions (Salamone et al., 2018). Although dopamine is more often implicated in the evaluation of benefits (reward value), it has also been hypothesized that dopamine could signal cost or, more specifically, a cost to benefits ratio. A recent preprint by Eshel et al., investigated this hypothesis and observed that striatal dopamine encoded both the reward magnitude and the sunk cost of action. The term sunk cost is used because it refers to an action that was already executed (Eshel et al., 2022). Moreover, the authors found that high motivation was associated with lower dopamine levels in the striatum and a weaker relationship between dopamine and cost, a mechanism that could explain why motivated animals are more likely to choose high-effort/high-reward actions.

3.3.4 Short vs long-term causal effects of dopamine on motor behavior

We have already described the molecular timescale during which dopamine might act on cortico-striatal synapses in 1.7.1 However, it is important to keep in mind the behavioral timescale of the dopamine signal we are examining during motor execution. There is no doubt that long-term modification of dopamine levels has an impact on motor behavior. Dopamine release at reward delivery drives learning over dozens of trials or even days. In addition, chronic diminution of dopamine by pharmacological manipulation or neuronal ablation prevents learning and reduces the vigor of movement (Baraduc et al., 2013; Mazzoni et al., 2007; Palmiter, 2008; Panigrahi et al., 2015; Salamone and Correa, 2012). But does dopamine act on shorter timescales as well, ranging from single trial to a few subsequent trials? The existence of a clear distinction between 'online control' and a role solely in reinforcement learning is highly uncertain and controversial. Considering our

understanding of the molecular effects of dopamine, it is possible that it exerts an influence on forthcoming cortico-striatal transmissions that is separate from its effects on synaptic plasticity. On short timescales, a modulation of MSN excitability could influence the receptiveness of certain striatal neurons to cortical input, and on longer timescales, the variation in basal dopamine concentration in the striatum could drive a global change in striatal response (as is observed during change in motivation or during L-DOPA treatment for example).

Some studies have tried manipulating dopaminergic activity to observe behavioral changes on short timescales. An optogenetic inhibition of SNc dopaminergic neurons over several seconds tends to inhibit movement initiation (locomotion, Bova et al., 2020; Da Silva et al., 2018; Hunter et al., 2022). Even though relatively short optogenetic stimulations can affect immediate behavior by triggering locomotion (Da Silva et al., 2018), increasing speed (Barter et al., 2015), or decreasing reaction time (Howard et al., 2017; Soares et al., 2016), dopaminergic stimulation will most often modify behavior only during subsequent trials and not during the immediate trial following stimulation (Hamid et al., 2015; Saunders et al., 2018).

In this introduction, we have seen that dopamine can change the excitability of striatal neurons but mainly acts by modulating the strength of cortico-striatal synapses. This mechanism allows dopamine to retrospectively attribute a value to environmental cues and to reinforce certain actions performed by the animal. However, not much is known about the role of dopamine once the associations and behaviors have been learned. Several hypotheses suggest that dopamine could play a role in time measurement, action selection, or movement vigor control. Most of the accounts on dopaminergic control of vigor have relied on optogenetic stimulations during locomotion or simple tasks, but we lack a precise characterization of dopaminergic activity during a skilled motor task and its link to vigor in physiological conditions.

Part II

Methods

Data acquisition

4.1 Animals

All experiments were performed on Long Evans rats. The animals were housed under artificial conditions of light (following a 12h light-dark reverse cycle), temperature (24°C) and humidity (45%) with food available ad libitum. Behavioral training and experiments were performed during the dark phase cycle and water intake was restricted to 15mL of water per day per animal. Transgenic TH-cre animals were acquired from the Rat Resource & Research Center (line LE-Tg(TH-Cre)^{3.1Deis}, RRRC #00659). These TH-Cre rats offer a powerful means to selectively target dopaminergic neurons as shown by previous work (Witten et al., 2011). Rats were usually housed 2 or 3 per cage, but remained isolated if their cage-mate started to damage their implant. Both male and female rats were used in our study.

4.2 Behavioral training prior to implantation

Rats commenced training for the behavioral task at least 3 weeks after their arrival at the animal facility. They were first familiarized with human handling and with the experimental training setup.

Each training operant box (IMetronic, France) is composed of two levers placed on either side of a reward delivery port. Behind the levers are 3 LEDs (central, left, and right) as well as 2 speakers. On the floor in front of the levers are two

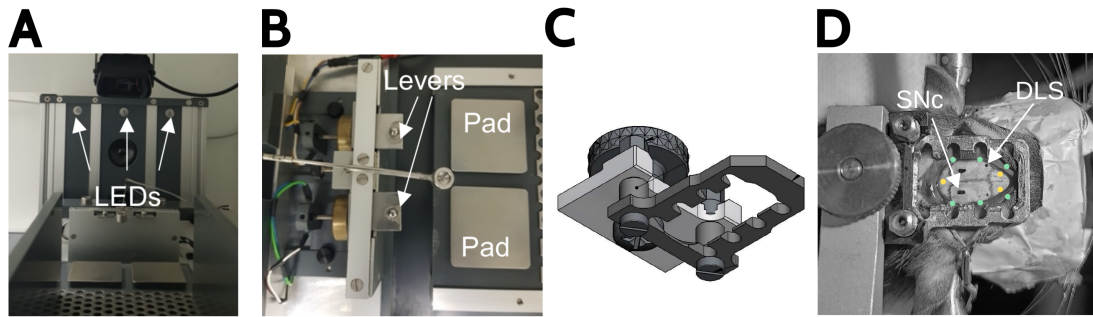


Figure 4.1: **Behavioral training setup and head-fixation implantation.** **A:** Front view of the behavioral training setup without the head-fixation support. In front of the animal are three LED lights, the levers, and the pads. **B:** View from above, with the pads, levers, and reward port. **C:** 3D model of the head-fixation implant, viewed from below, with the intermediate parts that can be removed after the training. **D:** Implantation surgery, viewed from above. The approximate locations of the SNc and DLS were drawn onto the skull with black ink. Green dots represent the locations where the anchor screws will be placed, and yellow dots mark the spots where the ECoG, reference, and ground screws will be placed. The implant is then cemented in place.

metallic pads that detect contact with the rat's forepaws (fig. 4.1 A and B).

During the motor training and experiments, brown noise was always played in the background to avoid the rats from being distracted by surrounding experimental noise.

The initial training is conducted on freely moving rats inside these operant boxes. The animals first learn to consume a sucrose solution (40 μ l of sucrose 5%) delivered randomly at the reward spout. They are then trained to press the left and right levers in order to receive a reward. This behavior is either spontaneously learned by the animal or shaped by giving rewards for imperfect movements until it can complete the task on its own. If a rat had a preference for one lever over the other, the reward associated with the weak lever was temporarily reinforced until both levers were equally pressed. Rats were trained on a daily basis usually during 1 hour-long sessions.

4.3 Surgery

4.3.1 Head-fixation implantation surgery

Custom-made head-fixation implants (fig. 4.1 C) were implanted on animals that were at least 8 weeks old and weighed more than 200g for females or 250g for males. The implants consist of a flat rectangular structure with a window in the middle. Surrounding this window are dents where the anchor screws are placed. On the caudal side of the implant, two large thread screws oriented upwards are used to connect it to the head-fixation device. The weight of the implant is 7g. After being cemented in place, only two fixation screws are still visible.

Implantation was performed under stereotaxic conditions and isoflurane anesthesia (induction 4%, maintenance 1.5%, Iso-vet). Animals were placed on a heating blanket and fixed with ear bars in a flat skull position. They received 6mg/kg of local analgesic (Lidocaine, Lurocaine) and 0.5mg/kg of opioid analgesic (Buprenorphine, Buprecare). An ophthalmic ointment was used throughout the surgery to prevent eye dehydration as well as opaque screens to protect from the intense surgical lighting. The surface of the skull was exposed, and cleaned and then 6 screw holes were drilled in the lateral ridges. An additional 3 to 5 holes were drilled for reference and ECoG screws. The implant was then positioned on the skull and cemented in place using bone cement (Zimmer Biomet). Importantly, the skull located above the stereotaxic coordinates of the DLS and the SNc (fig. 4.1 D) was not covered by the bone cement to enable the virus injection performed later on. Once the cement was set the exposed skull was covered with a removable silicon biopolymer (Silbione, Elkem). At the end of the surgery, animals received 5mg/kg of anti-inflammatory (Carprofen, Rimadyl or Metacam), and were allowed to recover on a heated blanket for 1-2 hours.

4.3.2 Viral injections and optic fiber implantation

In order to manipulate the activity of dopaminergic neurons, we injected viruses expressing an excitatory opsin in the SNc of TH-cre rats. Alternatively, to record their calcium activity, we used the calcium sensor GCaMP8f (Zhang et al., 2021).

The table 4.1 lists the viral constructs used and their injection sites and the table 4.2 specifies the stereotaxic coordinates for each injection site.

Virus List						
Virus	Target site	Volume	Stock concentration	Injected concentration	Provider	Reference
AAV5-EF1 α -DIO-hChR2(H134R)-eYFP	SNc 1,2	2*500nl	$4 * 10^{12}$	$4 * 10^{12}$	UNC	AV4313-2A
AAV9-hSyn-Flex-jGCaMP8f-WPRE.SV40	SNc 1,2	2*500nl	$2.4 * 10^{13}$	$1.2 * 10^{13}$	Addgene	#162379

Table 4.1: List of viruses used and their injection sites

Injection sites			
Target nucleus	AP(mm)	ML(mm)	DV(mm)
SNc 1	-5.0	2.2	-7.4
SNc 2	-5.5	2.4	-7.4
SNc 3	-5.2	2.2	-6.5
DLS	0.9	3.8	-4.4

Table 4.2: Stereotaxic coordinates

Optic fibers were implanted in the DLS or the SNc (at site SNc3) to record fiber photometry signals or to carry out optogenetic stimulations. For recordings we used 400 μ m diameter fibers but for stimulations we used 200 μ m diameter fibers. In the DLS the fibers were implanted at the same location as the viral injection. However, in the SNc we implanted the fibers slightly above the viral injection site.

We alternated the hemisphere chosen for the electrophysiological or calcium recordings but often tried to choose the hemisphere contralateral to the preferred paw for electrophysiological recordings.

The virus surgeries were performed after head implantation surgeries but under the same conditions. They were placed on the stereotaxic frame using the head-fixation system, and the skull was accessed by removing the protective silicone from the implant.

4.4 Head-fixed behavioral training

After implantation, the animals were given at least one week to recover before starting the training. For the first few days, water-deprived animals were accustomed to being restrained by the handler and getting the intermediate head-fixation parts screwed onto their implant. They received sucrose at each step of this training. Around day 3, animals were placed in the behavioral chamber without being head-fixed. On day 4 we started to restrain them via their head implant. Initially, the restraintment was limited to a few seconds and rewarded with sucrose. Once they were comfortable with the setup and handling, they were fixed on the head-fixation system for a few seconds and given access to unlimited sucrose by the reward delivery port. We then gradually increased their fixation time, making sure that the animals remained as stress-free as possible and that they received enough reward to form a positive association with the procedure. A full week was usually sufficient for the rats to be comfortable on the setup for an unlimited amount of time.

4.4.1 Early training: spontaneous lever press

When the animals became comfortable enough with head fixation, we started the behavioral task training by rewarding spontaneous lever presses. Because these lever-press movements remained relatively rare in most head-fixed rats, we also rewarded any movement occurring in proximity to the lever in order to progressively shape the desired motor output and encourage the rats to press the levers. Since rats often have a preferred side of lever pressing, we also trained them to press the other lever by degrading the first one until both sides were pressed in equal amounts.

4.4.2 Increasing task difficulty

Behavioral task difficulty gradually increased. First, the animal learned to initiate a trial by putting both of its paws on the metal pads and staying in this position for an increasing amount of time. Then, we introduced side cues that indicated which lever the rat had to press. These cues consisted of a LED light that turns

on above the side of the lever to be pressed, and a 50ms cue tone at frequency 1 kHz or 5 kHz for the left lever and right lever press, respectively. As the animals learned to associate the cues to each lever, we introduced punishments for mistakes. Mistakes included pressing the wrong lever (bad hit), or pressing a lever before the Go cue tone (anticipation). These errors lead to a 6s time off before the next trial could be initiated. When the rat's performance was good enough on these Go trials, we introduced a third type of cue for 'NoGo' trials. In this task (occurring 5 to 10% of the trials), white noise was played for 300ms, signaling the rat that it had to stay still and not press any lever in order to obtain the reward. Surprisingly, this type of trial had a relatively low success rate compared to Go trials and rats frequently pressed a lever during NoGo trials.

You can find a video of a rat performing the task by scanning the QR code below.



Figure 4.2: **Video of a rat performing the task.** The first trial of the video is annotated with the significance of the LED lights below the rat. It is also paused at key moments such as cue, movement onset, lever press, and reward delivery. All trials are at 0.125 speed (each trial is video recorded for the first 1s). The video can also be found by following this link: [Link to a Google Drive](#)).

4.4.3 From training to recordings

Fiber photometry required few preparatory steps before starting the recording sessions, however acute head-fixed electrophysiological data acquisition required additional preparatory manipulations that needed to be performed on the animal before the recordings could begin (i.e inserting the probe, lowering it into the SNc, etc). This additional handling could be stressful if rats were not used to it through training. Thus, for a few weeks preceding the start of the recording period, the rats were habituated to these pre-recording steps. This habituation includes noises,

light illumination directed on the implant, handling of tools above the head of the animals, and a 15-minute waiting period before starting the behavioral session (corresponding to the time it takes to lower the electrode into the SNc recording site).

4.5 Recordings

4.5.1 Electrophysiological recordings

Acute head-fixed electrophysiological recordings were performed using 64 channels silicon probes (ASSY 77-M1, Cambridge Neurotech). The electrode was connected to the eCube Server (White Matter) acquisition system and the signal was recorded using the Open Ephys software.

Rats were placed on the head-fixed setup and the silicon protecting the recording area was removed. Lurocain was applied to the craniotomy and the electrode was inserted above the SNc. Fiber photometry recordings were sometimes performed simultaneously. The electrode was lowered into the SNc so that a few fast-firing putative SNr neurons would be visible on the lower recording sites of the probe then, the behavioral task and the recording were started. Dopaminergic neurons were opto-tagged with single pulses of laser (5ms at 15mW, 470nm) at random moments during the inter-trial interval. The recordings sessions lasted until the rat displayed signs of fatigue or stopped performing the motor task. The recording electrode was then removed from the brain and left to clean overnight in a trypsin solution (1x, Sigma). The craniotomy was cleaned and covered with duraGel (Cambridge Neurotech) and silicon.

4.5.2 Fiber photometry recordings

Fiber photometry recordings started one week after viral injection and optic fiber implantation. Signal intensity was measured from day 7 to day 20 by looking at the dopamine response to an unpredicted reward. The first discernible responses were usually observed around day 10-15, responses with good signal-to-noise ratio followed this period and the signal usually reached its plateau by day 20. Fiber

photometry recordings cause minimal stress to the animal; the only preparation necessary is a brief cleaning of the implanted fiber with ethanol and connection to the recording cable. We used a Dual Color Multichannel Fiber Photometry System (R810, RWD) and the accompanying acquisition software for the recordings. Laser excitation at 410nm was used for the isosbestic signal and stimulation at 470nm was used for the GCaMP global calcium signal. All the behavioral events (cue, pads off, lever press, etc) were copied as TTL inputs in the fiber photometry system to allow for synchronization of the different data sources. Recording sessions were limited to less than an hour per day to avoid bleaching the signal.

Data analysis

All data analysis was performed using python.

5.1 Behavior

Behavioral events and parameters were recorded with Spike2 (CED) as TTL digital signals. We had access to the onset and offset of the sound cue, lever presses, left and right pad contacts, central LED and reward delivery. Spike2 was also used to control optogenetic stimulations. These signals were analyzed to extract the timestamps of every trial start, classify the trials based on their outcomes, verify that the correct paw was used to press the lever and measure the reaction time(RT)/movement time (MT) associated with each trial.

5.2 Fiber photometry

Fiber photometry data was preprocessed by fitting the isobestic signal (410nm) to the 470nm signal. The isobestic signal includes artifacts, but not calcium-dependent events, and so can be subtracted from the signal and used to normalize the data. The $\Delta F/F$ signal thus obtained was filtered to remove very slow signal variations ($<0.01\text{Hz}$) caused by bleaching, and lightly smoothed.

5.3 Electrophysiology

Electrophysiological data were recorded with the Open Ephys software, and synchronized with other data sources using the optogenetic stimulation pulses that were copied from Spike2 to Open Ephys. We performed spike sorting on a selected period of interest of the recording: the first 150s were usually excluded due to excessive drift of the units. We used SpikeInterface (Buccino et al., 2020) to perform the preprocessing and spike sorting, followed by manual curation of the extracted units using the open-source Python library Phy (see <https://phy.readthedocs.io/en/latest/>). As our recordings were performed on head-fixed animals, some drift in the spike recordings was observed. The most abrupt drift classically occurred at the beginning of the session following the first few trials of the reach-and-grasp motor task. A slower drift could also cause the units to move upwards and change channels. This phenomenon was relatively easy to identify during the manual curation of the spike sorting data, and we could merge clusters of spikes that belonged to a unit that drifted across channels. During the manual curation, we also removed units that were not present during the majority of the recording, units that had a bad signal-to-noise ratio, and units that showed an excessive violation of the refractory period.

All electrophysiological data were then analyzed using the Neo python package (see <https://pypi.org/project/neo/>).

All Z-scores (both for fiber photometry and electrophysiology) were calculated by resampling the data and computing the standard deviation and mean of the bootstrap distribution of the mean across samples.

5.4 Histology

On the last day of recording, the shank of the silicon probe was painted with red fluorescent dye (DilC18, Thermofisher) and lowered into the SNc. The animal was then sacrificed with an overdose of pentobarbital (Exagon) and intracardiac perfusion was performed with PBS 0.01mM followed by paraformaldehyde 4%. The brain was kept for 48h in a solution of PBS 0.01 mM - formaldehyde 4% and then cut into 50 μ m slices with a vibratome (VT1000, Leica Microsystems). The slices

were stored at 4°C in PBS 0.01 mM - Azide 0.1%.

For rats having been injected with the GCaMP8f, GRAB-DA or dLight virus, a GFP amplification was performed. Briefly, brain sections were rinsed 3 times in PBS 0.01 mM, and incubated overnight with an anti-GFP chicken primary antibody, dilution 1:500 (ref: GFP-1020, Aves Labs). Slices were washed 3 times and incubated for 2h with a donkey anti-chicken-Alexa 488 secondary antibody at dilution 1:500 (ref: 703-545-155, Lot 100057, Jackson).

The slices were then mounted onto slides with Fluoromount mounting medium (Thermofisher) and protected with glass coverslips. Images were acquired with an epifluorescence microscope (Zeiss), and analyzed with Fiji software (see <https://imagej.net/software/fiji/>).

Part III

Results

Recordings in head-fixed rats during the instrumental task

6.1 General performance of head-fixed rats

To evaluate the motor parameters of each trial we measured the reaction time and movement time of the animal. The reaction time is the duration of the interval between the cue onset and the movement onset when the paw of the animal is not in contact with the metal pad anymore. The movement time is the duration of the interval between the movement onset and the lever press. It is important to mention that the first instant of the lever press does not mark the end of the movement, as the animal keeps grasping and pulling the lever for some time and then has to bring its paw back to its original position. Interestingly, analysis on individual trials reveal that the reaction time (RT) and the movement time (MT) do not show any sign of covariance (fig. 6.1 A), indicating that the two underlying neuronal process are likely independent. That being said, RT vs MT median comparison shows that when one parameter is shorter than usual, the other is often longer than usual. The RT is generally longer and more variable than the MT (median RT across all rats = 0.240s, RT std = 1.06, median MT across all rats = 0.185s, MT std = 0.52). Interestingly, the values of RT and MT were sensibly similar for contralateral and ipsilateral trials. In addition, each rat had different distributions of RT and MT (fig. 6.1 B and C). Amongst the three rats

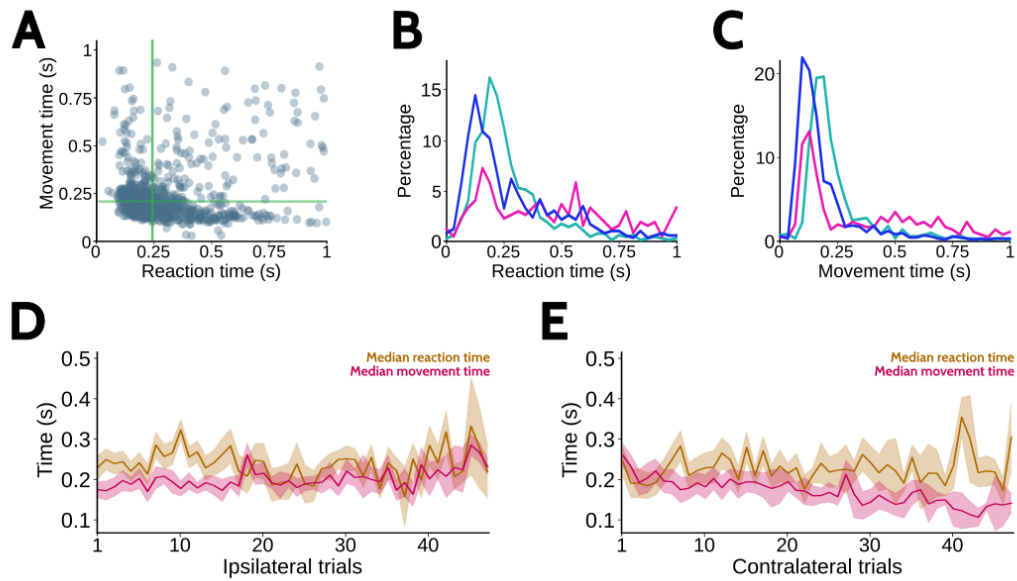


Figure 6.1: **Characteristics of the reaction time and movement time for the 3 rats included in fiber photometry analysis.** **A:** Scatter plots analysis of the corresponding reaction time and movement time. The green lines represent median values. **B-C:** Distribution of the reaction time (**B**) and movement times (**C**). **D-E:** Variation of the reaction time (orange) and movement time (pink) across trial during recording sessions, for ipsilateral (**D**) and contralateral trials (**E**). Median and SEM of all sessions.

represented here (the ones that were included in the fiber photometry analysis), two were females and had shorter RT and MT duration. There was no noticeable variation of those two parameters during recording sessions (fig. 6.1 D and E). If anything, the contralateral movement time tended to decrease toward the end of the session.

6.2 Dopaminergic signal in the SNc and DLS

In order to record the population activity of dopaminergic neurons during our reaching task, we used fiber photometry to record the activity of the axonal projections of dopaminergic neurons onto the dorso-lateral striatum (DLS). We targeted the DLS specifically because previous work done in the lab showed that DLS dopaminergic lesions, but not DMS lesions, impaired rats performance at an expert stage. Since DLS is considered to be the "motor striatum" and is usually implicated in habitual (or overtrained) behaviors, it seemed to be the best target to investigate the movement related dopaminergic activity. On a few instances we also implanted an optic fiber in the DMS of the animal but failed to record a quality signal.

In the DLS, we evaluated the quality of the signal across time from the injection with the response to unexpected rewards. Some visible response usually appeared around 10-15 days post injections and its amplitude increased until it reached a plateau around 20-30 days post injection (fig. 6.2 A) . We started recording data between 15 and 20 days post injection when the signal-to-noise ratio was optimal.

In one rat, we managed to record both the GCaMP activity in the SNc and the DLS. Although these recording were in different hemispheres, the signals were extremely correlated. Fig. 6.2 B shows a snippet of the GCaMP signal at both recording sites. The DLS signal was relatively stronger than the SNc (possibly due to a better placement of the optic fiber) but it was evident that transients in the SNc were followed by transients in the DLS. This slight delay between the two can be measured by computing a cross-correlogram (fig. 6.2 C), which shows that the DLS activity is delayed by approximately 20ms compared to the SNc activity. The higher correlation between -1 and 0s would suggest that the SNc transients tends to last longer than the DLS ones. This can be directly observed

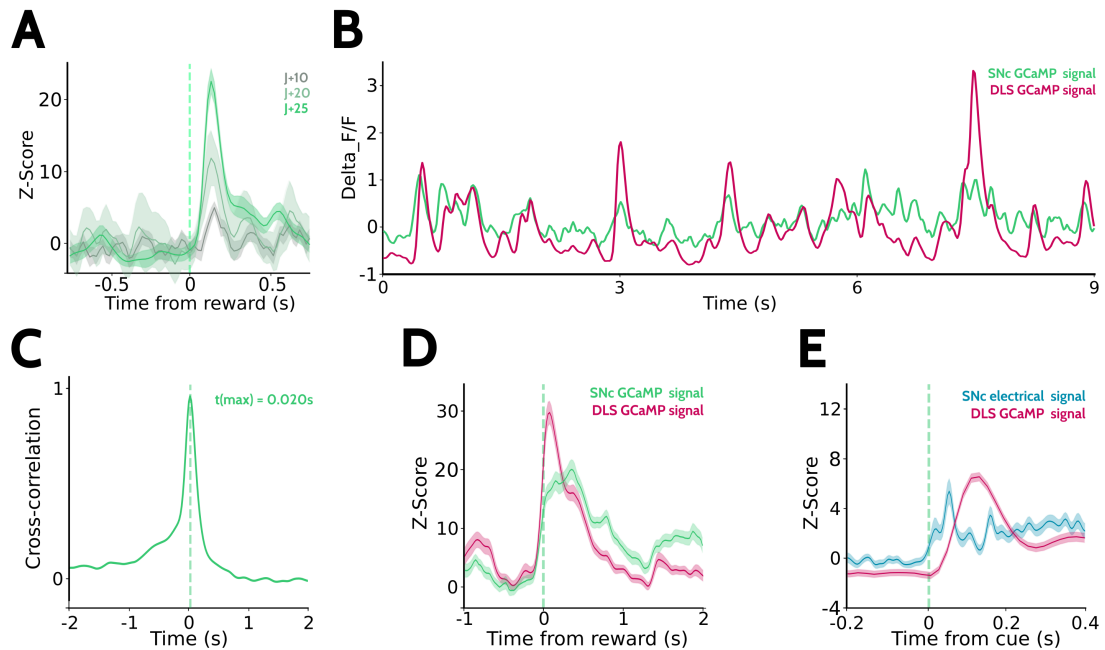


Figure 6.2: **GCaMP8f signal and kinetics.** **A:** Representative example showing the evolution of the DLS fiber photometry signal in response to an unexpected reward across time (10, 15 and 20 days post injection). **B:** Example traces of GCaMP signal recorded in the SNc of one hemisphere and the DLS of the other hemisphere. **C:** Cross-correlation of the SNc GCaMP signal with the DLS GCaMP signal of the same rat as in A. Maximum correlation at 20ms. **D:** Comparison of the SNc (green) and DLS (pink) GCaMP response to unexpected reward, in the same rat as in A and B, aligned on reward consumption. **E:** Comparison of the mean electrophysiological activity of putative dopaminergic neurons recorded in the SNc (blue) with the GCaMP signal recorded in the DLS (pink), during contralateral trials, aligned on the cue. $N = 3$ rats for each condition.

when comparing the two signals in response to unexpected rewards with the DLS response being narrower than the SNc one (fig. 6.2 D). Finally, we could also compare the electrophysiological activity of SNc putative dopaminergic neurons with the DLS GCaMP signal ($N = 3$ rats for each condition) by analysing the dopaminergic response to the cue in contralateral trials to determine the absolute delay of the DLS GCaMP signal (fig. 6.2 E). The electrophysiological response peaks around 50ms after the cue while the GCaMP response peaks 120ms after the cue, indicating that there is a delay of 70ms between the spiking activity of the

cell bodies and the GCaMP signal at the axonal terminals.

6.3 Identification of dopaminergic neurons

For electrophysiological recordings of dopaminergic neurons, we recorded mostly in the rostral part of the SNc as it is the region where the population is the most dense. The localisation of the recording electrode in the SNc was confirmed by histology for the last recording session by coating the electrode with a red dye (fig. 6.3 A).

Our first consideration when recording the electrophysiological activity of dopaminergic neurons was to be able to confidently identify them. Contrarily to extracellular glass electrode recordings, the action potential waveform shape did not allow us to easily identify dopaminergic neurons. If with glass electrodes their action potential lasts for 2-3ms, with silicone probes no action potential had a duration above 0.5ms. This difference could be due to a longer distance between the cell body and the recording site (high impedance glass electrodes recordings are usually very close to cell bodies while low impedance silicon probes can record neuronal activity from further away). Since we could not use the action potential duration as a clear indicator that a unit was dopaminergic, we used optogenetic to identify them through an optotagging strategy. A virus containing a CRE dependant ChR2 was injected in the SNc of TH-Cre rats to selectively express the ChR2 in dopaminergic neurons. During the recordings, we stimulated the neurons with short pulses of laser in order to elicit single action potentials. For the stimulation duration we tested 1, 5 and 10ms ; all three were able to stimulate the neurons but 1ms needed a higher stimulation intensity. We chose to only use 5ms as it was enough to elicit a consistent response and the artifacts due to the laser stimulation would not overlap with the triggered action potential. Indeed, light stimulation created artifacts on the raw electrophysiological signal, that were most likely due to population response or dendritic activation, as this artifact varied depending on the depth of the electrode. We chose to use repetition of single pulses of optogenetic stimulation in order to avoid triggering massive dopamine release that would have been reinforcing for the animal and could have influenced the motor behavior. The optogenetic stimulations were done during inter trial intervals

at random times across the session.

Fig. 6.3 B shows an optotagged example of dopaminergic neuron. We can see that the laser stimulation elicited a single action potential and that the neuron has a silent period afterwards, lasting around 40-60ms. When the laser stimulation occurred during the refractory period of the neuron it did not triggered an action potential. Clearly isolated clusters of dopamine neurons could not be identified and we thus classified the putative dopaminergic neurons in our recordings based on the firing properties measured in the optotagged neurons (firing frequency ranging from 1 to 20 Hz). Units discharging below 1 Hz were discarded from further analysis. Optotagged units above 20Hz were also discarded as they were likely the result of multiunits recordings. We also filtered the units based on their depth, but this measure could be influenced by variation in the depth of the cortical surface across days (caused by swelling and/or growth of scar tissue on to of the craniotomy), and in the depth of SNc localisation present between rats. Indeed, we observed a big variability in the location of dopaminergic neurons between rats and especially between males and females. The depth of dopaminergic neurons ranged from -6.5mm to -7.0mm in females and from 7.0mm to 7.8mm in males. As the dopaminergic neurons can sometimes be found in the SNr we did not put a maximum depth filter, but we did discard units above -6.4mm as some non-dopaminergic slow-firing units can be found above the SNc. The final filter discarded units that had a median inter-spike interval below 50ms at those units would be too bursty to be dopaminergic or they would be multi-units. Fig. 6.3 C shows the optotagged units (all of them before filtering) in blue, and in teal the putative dopaminergic units selected with the above described filters. This method of selection of putative dopaminergic units allowed us to obtain twice as many units as the one that were optotagged. This proportion is in accordance with juxtacellular experiments performed in rats injected with ChR2 showing that a significant fraction of dopaminergic neurons (up to 30%) do not respond to the optogenetic laser stimulations. As was previously stated, the dopaminergic neurons that we recorded did not exhibit the characteristic wide action potential waveform. Fig. 6.3 D shows the average action potential waveform for optotagged, putative dopaminergic and putative SNr units, and as we can see there is little difference between fast spiking SNr units and dopaminergic ones. SNr units tended to have

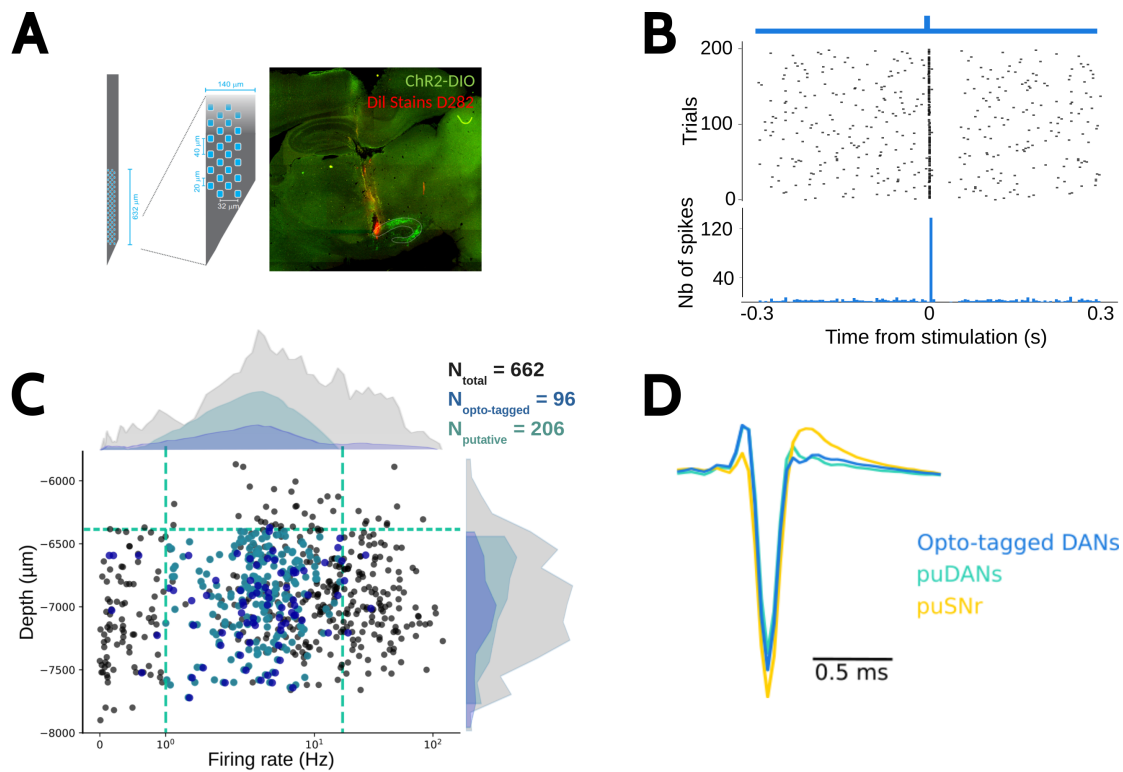


Figure 6.3: **Identification of dopaminergic neurons.** **A**: Recording arrangements of the 64 channels silicone probe (Cambridge Neurotech) and histological confirmation of the ChR2 expression in SNc neurons. On the last day of recordings, the probe was coated with a fluorescent red dye to verify the precise localisation of the recording electrode. **B**: Representative optotagging example of one dopaminergic neuron. The neuron was stimulated by a 5ms laser pulse at 10mW. Top: Raster plot of all the stimulations. Bottom: PSTH of the action potentials. **C**: Depth and firing rate of all recorded units. In blue are the opto-tagged units, in teal are the putative dopaminergic neurons and in grey are the non-dopaminergic neurons. The teal dashed lines represent the thresholds for the selection of putative dopaminergic neurons. The histograms on the top and the right represent the distribution of the units along the axis. **D**: Average waveforms of the opto-tagged dopaminergic neurons (blue), putative dopaminergic neurons (teal) and putative GABAergic SNr neurons (yellow).

spikes with higher amplitude as well as a bigger after-hyperpolarisation.

Dopaminergic activity during an instrumental skilled-reaching task

We will now characterize the dopaminergic activity pattern during a skilled reach-and-grasp motor task. We will compare the global calcium activity of dopaminergic axon terminals in the DLS and the single units activity in the SNc.

7.1 Global dopaminergic activity in the DLS during contralateral trials

To get a general idea of dopaminergic activity in the DLS during our motor task we looked at the activity during successful contra-lateral trials. The GCaMP signal was aligned on the cue or movement onset and averaged across all trials.

As we can see, dopamine responses are present at three moments during the task: after the cue, during the movement and during reward delivery and consumption (fig. 7.1 A). The response to the cue is the biggest in amplitude and tends to be shorter in duration. The activity during the movement starts at movement onset and reaches a maximum during the movement (fig. 7.1 B). When the reaction time increases, as can be seen at the bottom of the heatmap, the dopamine activation follows the movement onset (fig. 7.1 C).

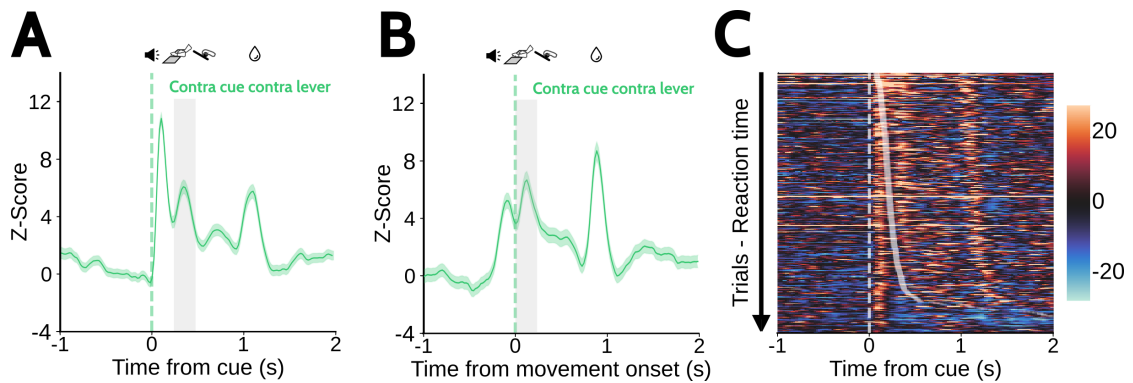


Figure 7.1: **Dopaminergic activity during contralateral trials.** **A-B** : Mean Z-Score of the GCaMP signal in the DLS during contralateral trials, aligned on the cue (**A**), movement onset (**B**). **C** : Heatmap of Z-Scored GCaMP signal during contralateral trials, aligned on cue and ordered by duration of the reaction time. The dashed line indicates the cue onset and the solid line indicates the movement onset

7.2 Single unit activity in the SNc during contralateral trials

The GCaMP dopamine axon activity in the DLS provide a good estimate of the dopamine dynamic during our motor task but does not provide information on how these responses are encoded at the single cell level. To address this issue, we then looked at the in vivo unit activity of our dopamine neurons during successful contralateral trials. To get an estimate of their activity pattern, we computed their instantaneous firing rate, averaged it over all trials and normalized it to a z-score. The units were then ordered depending on when during the trial their activity was at its maximum. Dopamine neurons were classified as 'cue unit' if they were most active during the reaction period, 'movement unit' if they were active during the movement and 'lever unit' if they were active after the lever press (fig. 7.2 A). In fig. 7.2 B, we can see an example of 'cue unit', with a phasic response around 60ms after the cue, followed by a pause in firing. Interestingly, some units did not exhibit a phasic activity in response to the cue, but fired a single action potential in a very reliable manner after the cue. This seems to represent a genuine response indicating that single spike synchrony more than a bursting activity amongst the

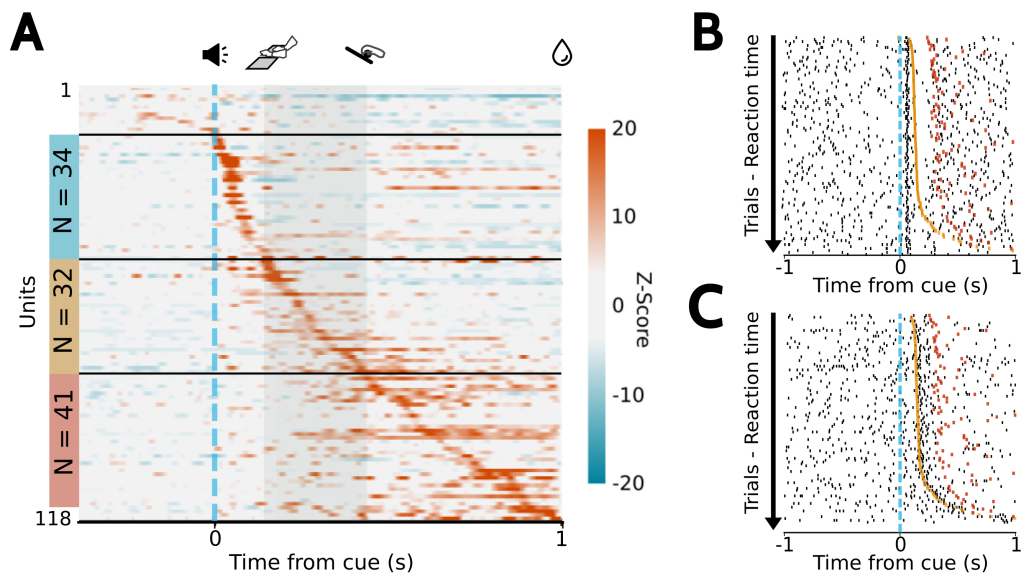


Figure 7.2: **Activity of putative SNc dopaminergic neurons during contralateral trials.** **A** : Heatmap of the mean z-scored activity of all putative dopaminergic units, aligned on the cue and ordered by maximum response latency. The blue dashed line represents the cue onset. The shaded area represents the median movement window. The horizontal black lines separate the 'cue units', 'movement units' and 'lever units'. **B-C** : Example cue unit (**B**) and movement unit(**C**), trials aligned on the cue and ordered by increasing reaction time. The blue dashed line represents the cue onset. The movement onset for each trial is represented by an orange dot and the lever press by a red dot.

population can elicit strong dopamine release at the axon terminal. That being said, technical limitations such as missing the smaller spikes occurring in a bursts when recording with silicon probes could also contribute to this phenomenon.

7.3 Lateralisation of the dopaminergic activity

As we saw in the introduction, a few studies have shown that dopaminergic activity could be lateralized during movement. For most of them however, there isn't a clear control and isolation of which limb is used for the movement. Our head-fixed task allow us to precisely isolate movements from one moving limb, and thus to clearly define the lateralisation of dopaminergic responses.

We first looked at the difference in dopaminergic activity in the DLS between successful contralateral and ipsilateral trials (fig. 7.3). The cue-evoked response is slightly lower for ipsilateral trials but this difference is not statistically significant. By contrast, we observed a large difference in the amplitude of the movement-related response, with contralateral movement being associated with a stronger dopaminergic activity. This difference is even more pronounced when the signal is aligned to movement onset (fig. 7.3 C). Interestingly, the response to reward delivery is identical for the two conditions. We then wondered what would happen of this lateralisation during trials when the animal presses the wrong lever. When the animal pressed the ipsilateral lever in response to a contralateral cue, the cue-evoked response is identical to the normal one, but we did not observe any dopaminergic activity at all during the movement window (fig. 7.3 D). This complete absence of movement-evoked dopaminergic is very puzzling, as we would have expected it to be similar to successful ipsilateral movement. One explanation could be sampling issue, although the data comes from three different rats and hundreds of trials. Another explanation could be that the rat somehow knows that its current movement is a mistake, which would trigger a negative RPE signal. A third possibility would be a signal akin to APE that would encode how much the current cue-movement association is different from the usual cue-movement association, but in the paper describing the APE signal (Greenstreet et al., 2022) a movement that is not usually associated with this particular cue triggers more dopamine response and not less as we do. By comparison, during trials with an

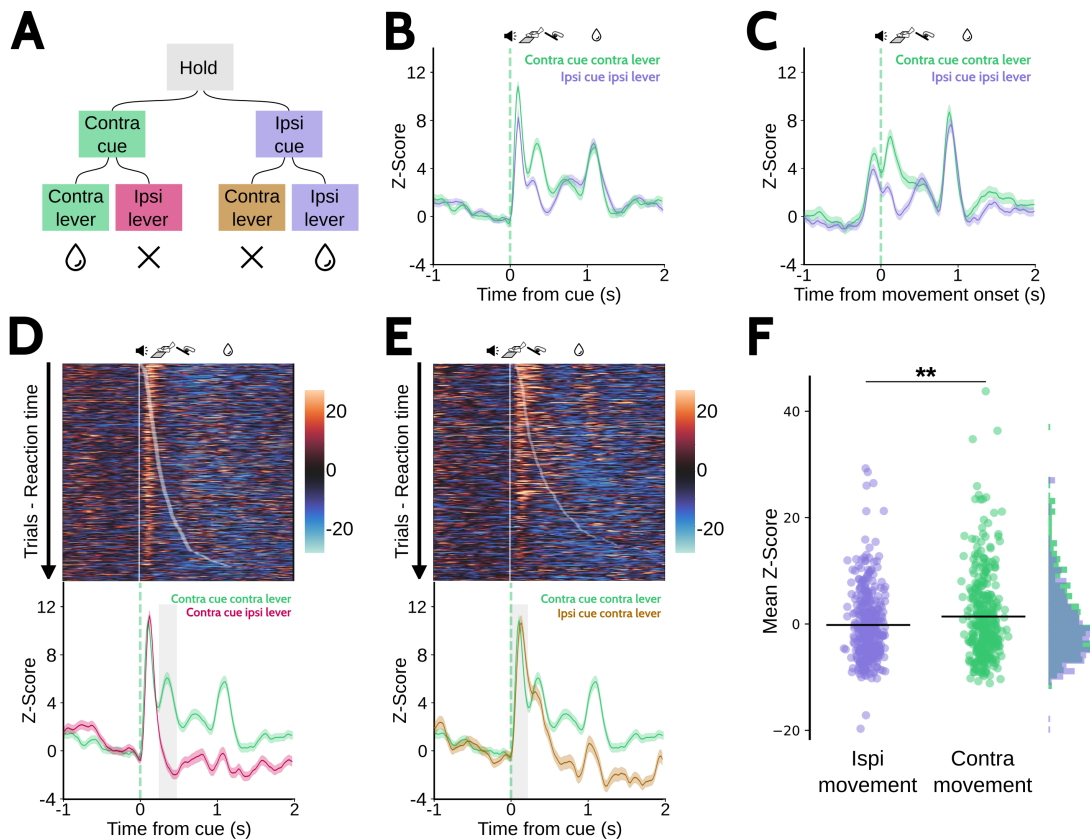


Figure 7.3: Lateralisation of dopaminergic signal. **A:** Possible outcomes of the task, excluding anticipatory presses and failure to press any lever. **B-C:** Mean Z-scored GCaMP activity during contralateral (green) and ipsilateral (purple) trials, aligned on the cue (**A**) and movement onset (**B**). **D:** Top: Heatmap of the Z-scored GCaMP activity during trials where the animal pressed the ipsilateral lever in response to a contralateral cue, aligned on the cue and ranked by increasing reaction time. The white lines represent the cue and movement onset. Bottom: Mean Z-scored GCaMP activity during trials where the animal pressed the contralateral lever in response to a contralateral cue (green) and trials where the animal pressed the ipsilateral lever in response to a contralateral cue (pink). The ‘Contra cue ipsi lever’ trials were not rewarded. The grey box represents the average movement window. **E:** Top: Heatmap of the Z-scored GCaMP activity during trials where the animal pressed the contralateral lever in response to an ipsilateral cue, aligned on the cue and ranked by increasing reaction time. The white lines represent the cue and movement onset. Bottom: Mean Z-scored GCaMP activity during trials where the animal pressed the contralateral lever in response to a contralateral cue (green) and trials where the animal pressed the contralateral lever in response to an ipsilateral cue (orange). The ‘Contra cue ipsi lever’ trials were not rewarded. The grey box represents the average movement window. **F:** Comparison of the mean Z-scored GCaMP activity during ipsilateral and contralateral movement. This analysis was restricted to trials with a reaction time superior to 300ms to avoid signal pollution coming from the cue response. One-way ANOVA, $p=0.004$.

ipsilateral cue and a contralateral movement, the movement-evoked activity seems to be present, albeit with a lower amplitude and reproducibility across trials (fig. 7.3 E), but see the next section for detailed quantification. The main difference in this condition is clearly seen after the lever press, when the rat receives feedback indicating the error, with a signal going back to baseline or even below baseline levels.

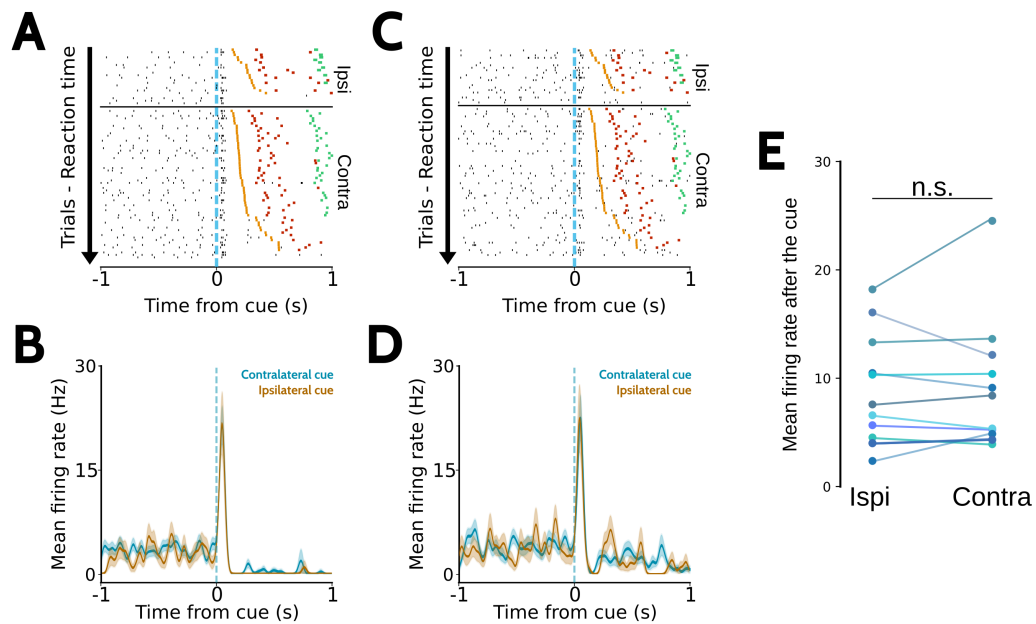


Figure 7.4: **Lateralisation of cue units.** **A:** Raster showing a representative example of a cue unit, for ipsilateral and contralateral trials, aligned on the cue and ordered by increasing reaction time. Orange markers represent movement onset, red markers lever press and green markers reward delivery. **B:** Mean instantaneous firing rate of the unit showed in A, during ipsilateral (orange) and contralateral (blue) trials, aligned on the cue. **C-D:** Same as in A-B but for another cue unit example. **E:** Comparison of the mean firing rate of each cue unit after the cue during ipsilateral and contralateral trials. Paired T-test, $p > 0.05$.

We then compared the *in vivo* electrical activity of the putative dopaminergic neurons during ipsilateral and contralateral trials. For this analysis we used the classification presented in fig.7.2. Cue units did not exhibit any lateralisation of their cue-evoked response (fig.7.4), neither at the individual level or when comparing all units. On the other hand, the activity of movement units was lateralised during

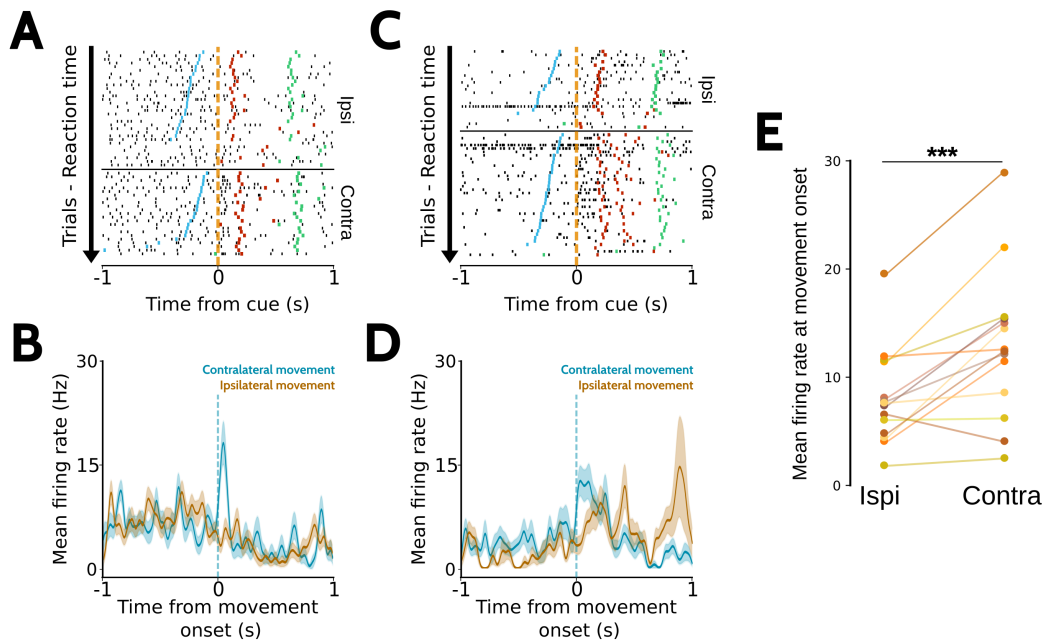


Figure 7.5: **Lateralisation of movement units.** **A:** Raster showing a representative example of a movement unit, for ipsilateral and contralateral trials, aligned on movement onset and ordered by increasing reaction time. Blue markers represent the cue, red markers lever press and green markers reward delivery. **B:** Mean instantaneous firing rate of the unit showed in A, during ipsilateral (orange) and contralateral (blue) trials, aligned on movement onset. **C-D:** Same as in A-B but for another movement unit example. **E:** Comparison of the mean firing rate of each movement unit during the first 100ms of ipsilateral and contralateral movement. Paired T-test, $p < 0.001$.

the movement (fig.7.5). This could be observed for units that had a phasic activity at movement onset (fig.7.5 A, B) and those who increased their activity during the entire movement (fig.7.5 C, D). The difference for all recorded movement units was statistically significant (fig.7.5 E).

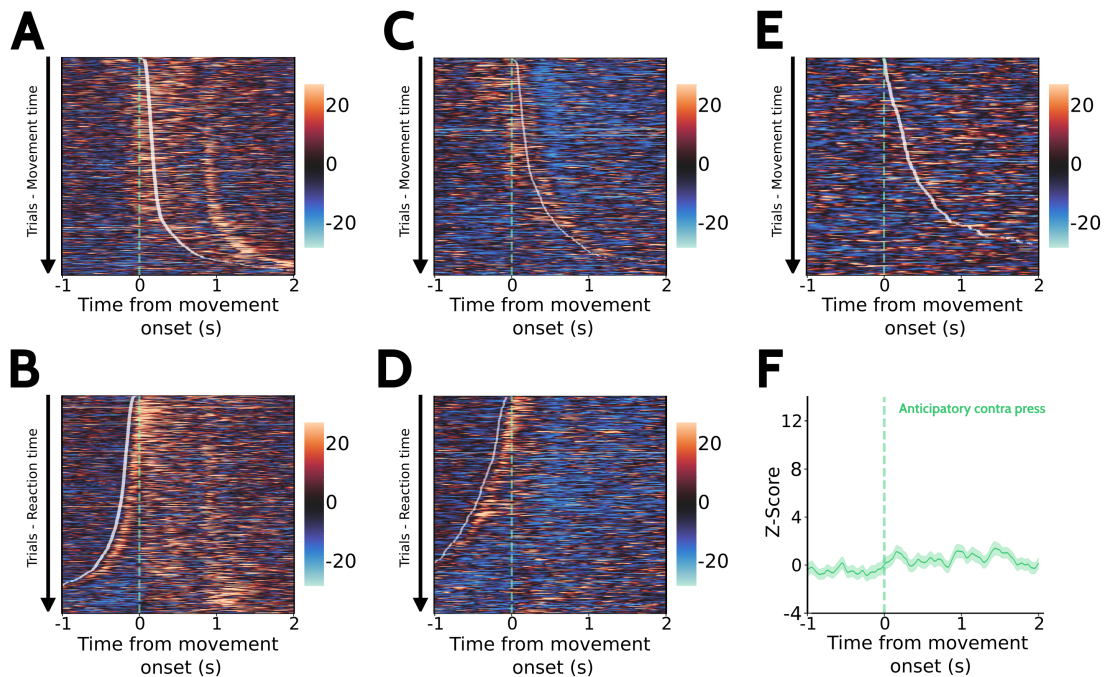


Figure 7.6: **Movement-evoked dopaminergic activity is dependant on context.** **A:** Heatmap of Z-Scored GCaMP signal during successful contralateral movements, aligned on movement onset and ordered by duration of the movement. The dashed green line indicates the movement onset and the solid line indicates the lever press. **B:** Same as A, ordered by reaction time (solid white line). **C:** Heatmap of Z-Scored GCaMP signal during contralateral movements after an ipsilateral cue, aligned on movement onset and ordered by duration of the movement. The dashed green line indicates the movement onset and the solid line indicates the lever press. **D:** Same as C, ordered by reaction time (solid white line). **E:** Heatmap of Z-Scored GCaMP signal during anticipatory contralateral movements, aligned on movement onset and ordered by duration of the movement. The dashed green line indicates the movement onset and the solid line indicates the lever press. **F:** Mean Z-Score of the GCaMP signal in the DLS during anticipatory contralateral movements, aligned on movement onset.

7.4 Similar movements do not trigger identical dopaminergic response

If during successful contralateral trials we clearly observed an increase of dopaminergic activity during movement, both during the reach and during the lever press (fig. 7.6 A and B), surprisingly, when the same movement was performed out of context, for example following the ipsilateral cue, the movement-related dopaminergic activity was reduced (fig. 7.6 C and D). In addition, during anticipatory lever presses (when animals executed the movement during the inter-trial interval or during the variable delay before the cue tone), the movement-related dopamine response was not present (fig. 7.6 E and F).

Since anticipatory lever presses are not triggered by a cue, we could have expected a similar type of activity has the one previously described for spontaneous movements (see 3.3.1), either an inhibition at movement onset (Coddington and Dudman, 2018; Dodson et al., 2016b) or an increase of activity around movement onset (Da Silva et al., 2018; Howe MW, 2016). On the opposite, we observed no activity around movement onset and only a slight activation after the lever press (fig. 7.6 E, even though no reward was delivered).

Taken altogether, these findings showing that similar movements performed in different contexts do not elicit identical dopamine responses are intriguing and have not been reported elsewhere. A possible explanation would be that this movement-related dopaminergic activity is not purely sensory-motor in nature, but rather could encode other information, such as reward expectation or Action Prediction Error (APE, see 3.3.2). In our case, it is unlikely that the movement-related dopaminergic activity encodes APE, since the APE should be higher during the 'ipsi cue contra lever' movements compared to normal contralateral movements (we will discuss this further in 8.3).

7.5 Dopamine and vigor

7.5.1 The cue-evoked and movement-evoked dopaminergic activity correlates with vigor

Since there has been a lot of controversy on whether or not dopamine is related to vigor, we wanted to investigate the potential correlations between dopaminergic activity in the DLS and vigor. As a proxy for vigor we used the reaction time and movement time of the animal, that could be precisely measured thanks to the metal pads on which the rat starts its trial on. There are many ways to correlate a signal to a parameter: should we use the mean activity level, the maximum amplitude, the area under the curve? For this analysis we chose the mean z-scored activity, which is close to an AUC since the baseline is centered on 0. We computed the mean z-scored activity during the reaction time and the movement time with no assumptions on the shape of the signal. This means that when the period (reaction or movement time) is unusually long, we still consider its whole duration, even though the main dopaminergic activity could be limited to the first hundred (or so) milliseconds. We excluded outsider values and only kept the trials when the reaction time and movement time were under 1s. Finally, because the vast majority of the points are very close to the median value of the reaction and movement time, we binned the trial in a logarithmic way in order to decrease the number of points and give less weight to very high values. Indeed, when computing the correlation on all trials (more than a thousand), the number of points makes the graph unreadable and the linear regression useless. Also, in this condition, the data is largely influenced by longer values of reaction time and movement time affecting the statistical significance of the analysis.

Fig. 7.7 shows the correlation between dopamine and vigor in four conditions: A compares the cue-evoked dopamine with the reaction time, B the cue-evoked dopamine with the movement time, C the movement-evoked dopamine with the reaction time and D the movement-evoked dopamine with the movement time. All of them show a negative correlation, meaning that the higher the dopaminergic activity in the DLS, the more vigorous the movement was. In C, the slope of the points for short reaction time is very steep, which is actually probably a

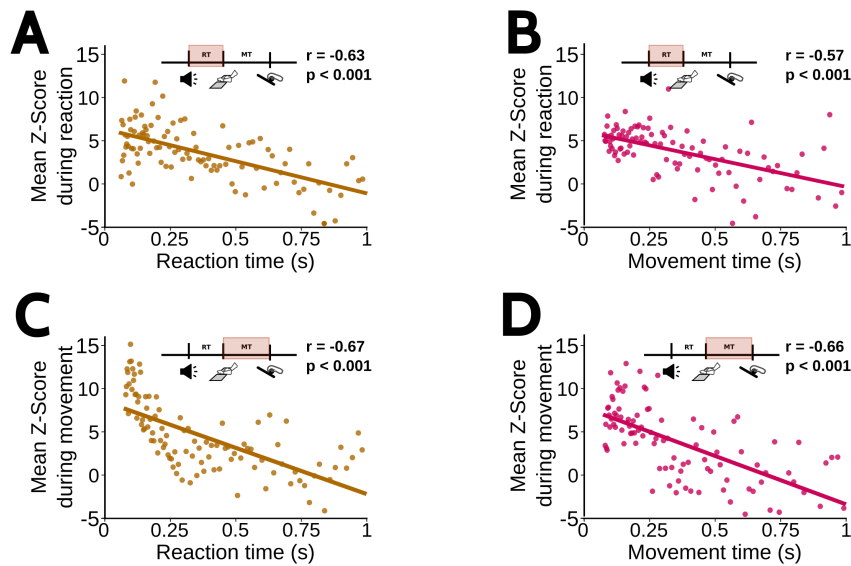


Figure 7.7: Correlation between dopaminergic activity in the DLS and the vigor of the movement. **A-B:** Correlation of the mean Z-scored GCaMP activity during the reaction time with the reaction (**A**) and movement time (**B**). The inset represents the part of the signal that was used. The averaged signal was then binned according to its corresponding reaction and movement time to decrease the number of points. The solid line represents the linear regression of the data. **C-D:** Correlation of the mean Z-scored GCaMP activity during the movement time with the reaction (**C**) and movement time (**D**). The inset represents the part of the signal that was used. The averaged signal was then binned according to its corresponding reaction and movement time to decrease the number of points. The solid line represents the linear regression of the data.

contamination of the movement-evoked activity by the cue-evoked activity; the reaction time being very short, the movement starts during the cue-evoked dopamine response, thus likely biasing the measure.

To get an idea of how this dopamine-vigor correlation looks on the signal, we separated all the contralateral trials in two depending on their reaction time (fig. 7.8 A). One group thus contained half the trials with the shortest reaction times and the other group the other half with the longest reaction time. The first thing to keep in mind with this method is that the group with the longest reaction times will be much more spread out, as the distribution of the reaction times is

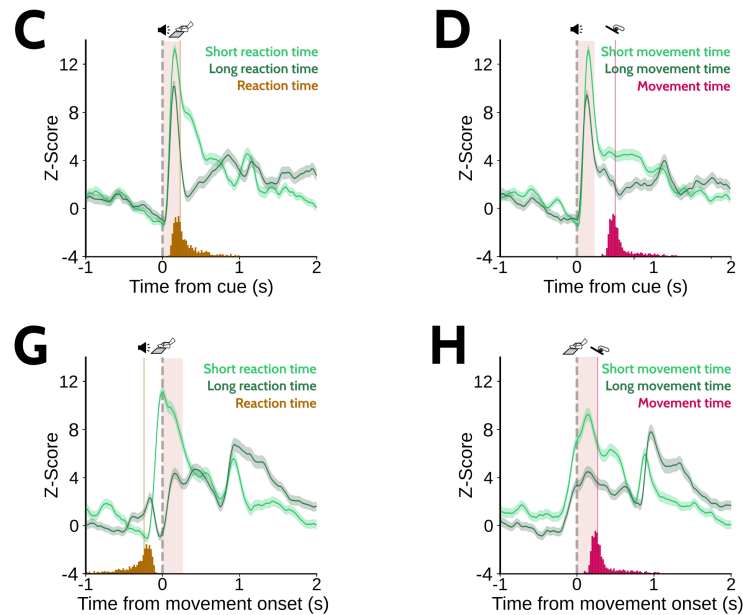


Figure 7.8: **Correlation between dopaminergic activity in the DLS and the vigor of the movement.** **A:** Mean Z-scored GCaMP activity during contralateral trials with a reaction time shorter than the median (‘Short reaction time’, light green), and trials with a reaction time longer than the median (‘Long reaction time’, dark green), aligned on the cue. The reaction time distribution is represented at the bottom (orange). The orange line represents the median reaction time, the red box represents the part of the signal that was averaged in A. **B:** Mean Z-scored GCaMP activity during contralateral trials with a movement time shorter than the median (‘Short movement time’, light green), and trials with a reaction time longer than the median (‘Long movement time’, dark green). The reaction time distribution is represented at the bottom (pink). The pink line represents the median movement time, the red box represents the part of the signal that was averaged in B. **C-D:** Same as A-B but aligned on movement onset.

poisson-like. When we plot the average signal for these two groups, we observe that indeed, when the reaction time is shorter, the cue-evoked dopaminergic activity has a higher amplitude. We can also see that with short reaction times, the activity during the movement is somewhat ‘merged’ with the cue-evoked activity, but that when the reaction time is longer, the dopamine activity comes back to near-baseline levels before increasing again. Then in fig. 7.8 B, we separated the trials based on their movement time. We can observe that, again, the cue-evoked response

seems to be higher for more vigorous movement. Fig. 7.8 C and D show the same analysis but the signal is now aligned on movement onset to get a better view of the impact of reaction time and movement time on the motor related activity. In C, we can clearly see that when the reaction time is very short, the movement starts during the cue-evoked dopamine response, so the dopamine activity is already high and decays slowly during the movement. However, when the reaction time is long, we can observe that the dopaminergic activity is back to baseline before the movement onset and that when the movement starts the dopamine activity increases. This plot is thus very important in that it allows us to really separate cue-evoked and movement evoked-activity (in the same way that Syed et al., 2016 did it). Finally, fig. 7.8 D separates short and long movement time. This plot is the hardest to interpret because there are many elements at play : Short movement time are associated with more cue-evoked activity and it can be seen here, as the activity level is higher at the time of movement onset. This difference in activity at the start of the movement is likely to have introduced a bias in fig. 7.7 D because the mean z-scored activity is going to be high for short movement time. However, if we look past the difference in starting point, the amplitude of the increase during the movement seems to be higher for short movement rather than long movement. Overall, analysing specifically the movement-related activity is never simple because of the proximity to the cue-evoked activity. It is thus very difficult to conclude as to whether the dopaminergic activity during the movement is really correlated to vigor. Addressing this point further would require optogenetic manipulation of dopamine transmission to artificially increase (opto-excitation) or decrease (opto-inhibition) the level of dopamine at specific timing of the motor task and see its causal impact on movement vigor.

7.5.2 The reward-evoked dopaminergic activity correlates with vigor

We next assessed if the dopaminergic activity after the lever press was linked to the vigor of movements. After successful lever presses, the reward was delivered with a delay of 500ms. Interestingly, the reward-evoked dopaminergic response scaled with movement time (fig. 7.9 A-B, correlation coefficient $r=0.59$, $p<0.001$)

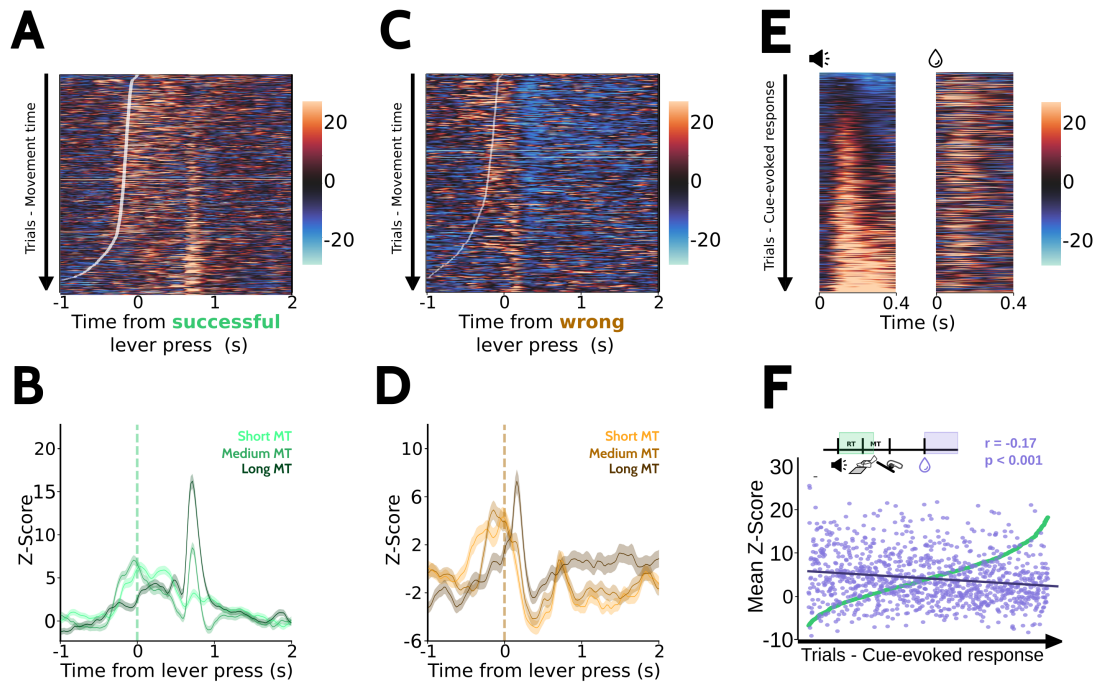


Figure 7.9: Correlation between movement speed and reward-evoked dopaminergic activity. **A:** Heatmap of Z-Scored GCaMP signal, aligned on successful contralateral lever press and ordered by duration of the movement. The reward is delivered 500ms after the lever press. The solid line indicates the movement onset. **B:** Mean Z-Scored GCaMP signal, aligned on successful contralateral lever press, depending on the movement time (Light green: 30% shortest MT, medium green: 30% medium MT, dark green: 30% longest MT). **C:** Heatmap of Z-Scored GCaMP signal, aligned on contralateral lever press after an ipsilateral cue, and ordered by duration of the movement. No reward was delivered and LED lights turned on after the wrong lever press. The solid line indicates the movement onset. **D:** Same as B with 'ipsi cue contra lever' trials (Light orange: 30% shortest MT, medium orange: 30% medium MT, dark orange: 30% longest MT) **E:** Z-Scored GCaMP signal after the cue (left) and the reward delivery (right) of successful contralateral trials. Trials are ordered by increasing mean cue-evoked responses. For each trial, the corresponding cue- and reward-evoked responses are shown. **F:** Correlation of the mean reward-evoked response with the cue-evoked response. Green: Cue-evoked response ordered from low to high. Purple: Reward-evoked response in function of the rank of the cue-evoked-response. The dark purple line represents the linear regression of the data, $r = -0.17$, $p < 0.001$. The inset represents the part of the signal that was used.

but not with reaction time. When the movement speed was high (short movement time), the reward-evoked dopaminergic response was almost nonexistent, while long movement time was associated with a strong reward-evoked dopaminergic response. On the other hand, after wrong lever presses ('ipsi cue contra lever' trials), no reward was delivered and LED lights turned on to warn the rat that it has made an error. Dopaminergic activity in the DLS was inhibited in response to the LEDs turning on, and the amplitude of this inhibition also correlated with the movement time (fig. 7.9 C-D, correlation coefficient $r=0.36$, $p<0.001$). Shorter movement were associated with a stronger inhibition after the lever press.

These two types of activity after the lever press could potentially signal Reward Prediction Error (RPE), quantifying how different the received reward was from the expected reward. During successful trials, the reward value is 1; if the expected reward value was also 1, there won't be an increase of dopaminergic activity. The observation that long movement time is associated with a strong reward-evoked dopaminergic response suggests that slow movements are associated with a lower expected reward value (leading to a positive RPE). Inversely, if the reward value is 0 (wrong lever pressed), the RPE is negative (actual reward value - expected reward value). In fig. 7.9 C-D, we can see this negative RPE signal around 200ms after the lever press. The weaker inhibition observed when the movement time is long might mean that the expected reward value was lower than trials with fast movement - between 0 and 1.

Why does this presumed RPE signal vary with the movement speed? These two measures might both depend on reward expectation. From the reward-evoked dopamine response we can deduce that reward expectation is approximately equal to 1 when the movement time is on the lower half of the median (top half of the heatmap). A first hypothesis would be that reward expectation is computed when the cue is played (as happens during pavlovian tasks). A 'high' reward expectation would invigorate movement, while 'lower' expectations would not. The trial by trial difference in reward expectation could vary based on past performance, thirst and energy levels of the animal, quality of the cue perception and/or confidence level of the animal. A second hypothesis would be that the reward expectation initially computed after the cue is updated during the trial depending on the motor performance of the animal. Hesitation, sub-optimal movement trajectory, low speed

etc. could update the reward expectation and thus change the ultimate reward prediction error. When comparing the cue-evoked response with its corresponding reward-evoked response during successful contralateral trial (fig. 7.9 E), we can observe that the reward-evoked response does not appear to be an inverted mirror of the cue-evoked response as can be seen during pavlovian tasks. There is however a slight correlation (fig. 7.9 F, $r=-0.17$, $p<0.001$). This would support the second hypothesis more than the first: the RPE is likely to be updated as the trial progresses.

7.6 Movement repression

7.6.1 Missed trials

During a recording session, it happened that the animal did not press any lever in response to the Go cue. These 'missed trials' were mostly present toward the end of the session when the animal was less thirsty or motivated, but could also happen at any point during the session. Some potential explanations for this behavior are 1) that the animal was frustrated because the previous trial(s) were unrewarded due to an error, 2) that the animal was not ready when the cue was played (this can happen if the animal licked the previous reward during the inter-trial interval and was surprised by the cue). Fig. 7.10 A displays the calcium activity in the DLS during such trials ('Contra cue no movement'). We can observe that the cue-evoked response is much smaller compared to the one on successful contralateral trials. This cue-evoked response is immediately followed by an inhibition of the dopaminergic activity, before a return to baseline. This pattern was very consistent across animals, days and time during the session.

To confirm that this activity was not restricted to end-session trials when the animal did not behave at all and all trials were missed trials, we compared the activity during missed trials that followed a rewarded trial and compared it to missed trials that followed another missed trial. The result of such analysis reveal that the activity is sensibly the same in both cases (fig. 7.10 B). This confirm that this activity is present since the first missed trial and is not a progressive degradation of the signal over multiple missed trials. Also, there seem to be no

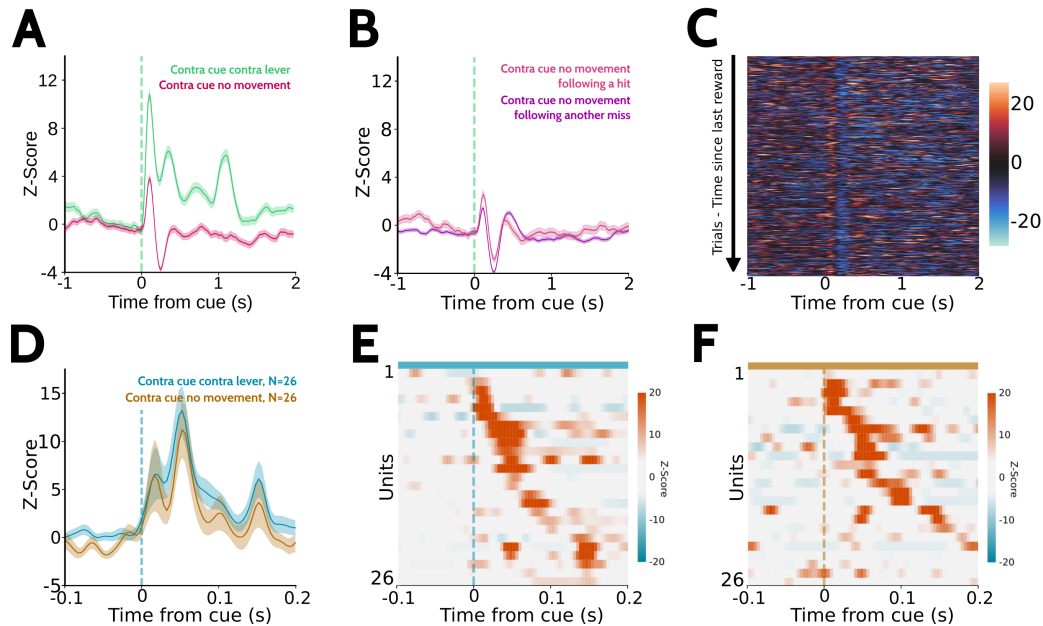


Figure 7.10: **Dopaminergic activity during trials when the animal did not press any lever.** **A:** Mean Z-scored GCaMP activity during good contralateral trials (green) and contralateral trials where the animal did not press any lever (pink), aligned on the cue. **B:** Mean Z-scored GCaMP activity during contralateral trials with no movement, if the previous trial was a successful (rewarded) one (light pink), or if the previous trial was also a trial where the animal did not attempt to press any lever (dark pink). **C:** Heatmap of the Z-scored GCaMP activity during missed trials, ordered by time elapsed since the last reward. **D:** Z-scored electrophysiological activity of 26 putative SNc dopaminergic neurons during good contralateral trials (blue) and contralateral trials with no movement (orange), aligned on the cue. **E-F:** Z-scored activity of the same 26 units during good contralateral trials (**E**) and contralateral trials with no movement (**F**), aligned on the cue and ordered by the maximum response latency.

variation of this response depending on the delay elapsed since the last reward (fig. 7.10 C).

In our electrophysiological recordings of SNc dopaminergic neurons, we did not observe the same type of activity (fig. 7.10 D-F). Units were responding to cue at the same level as during successful trials. Interestingly, amongst the units responding to the cue, not all had the same response latency. A subset of them responded around 15ms after the cue onset and another subset responded around

50ms after the cue onset. These units might receive different inputs that would explain the difference in response delay. Alternatively, since the cue lasts for 50ms, some units may respond to the cue offset. For this analysis we restricted ourselves to only compare units that have been recorded sufficiently in both conditions. Many units have been recorded in only one condition, but if we expand the analysis to include them the result is identical. We could not observe any inhibition of the recorded units during missed trials.

This difference between the electrophysiological and fiber photometry recordings could be due to a local modulation of the calcium activity of dopaminergic terminals in the striatum (by cholinergic interneurons for example). Alternatively, it could simply be explained by the important difference in scale between the two types of recordings : In one case we record around a few dozens of neurons for a few dozens of trials each, and in the other case we record the global activity of hundreds of neurons for hundreds of trials, so with electrophysiology we might miss some emerging properties that can only be observed at large-scales on the network dynamic.

This difference in dopaminergic activity between successful and missed trials is striking. Indeed, in both cases, the same cue was played and it is tempting to speculate that the cue-evoked activity in dopaminergic neurons would encode the value of the expected reward regardless of future action. Instead, we found that the amplitude of the cue-evoked activity was dependent on whether the animal would execute the action or not. Some studies have shown that dopamine could encode the value of the upcoming action (Morris et al., 2006), so the smaller cue-related dopaminergic response for missed trials could indicate that the value of 'not moving' is smaller than the value of pressing the lever.

7.6.2 NoGo trials

To further investigate the dopaminergic activity during movement inhibition, we added NoGo trials to our behavioral task. During NoGo trials, the cue played was a white noise and the animal has to refrain from pressing any lever during 800ms in order to receive the reward. This 800ms delay is much longer so the mean reaction time if the animal had wanted to respond and press a lever. Surprisingly,

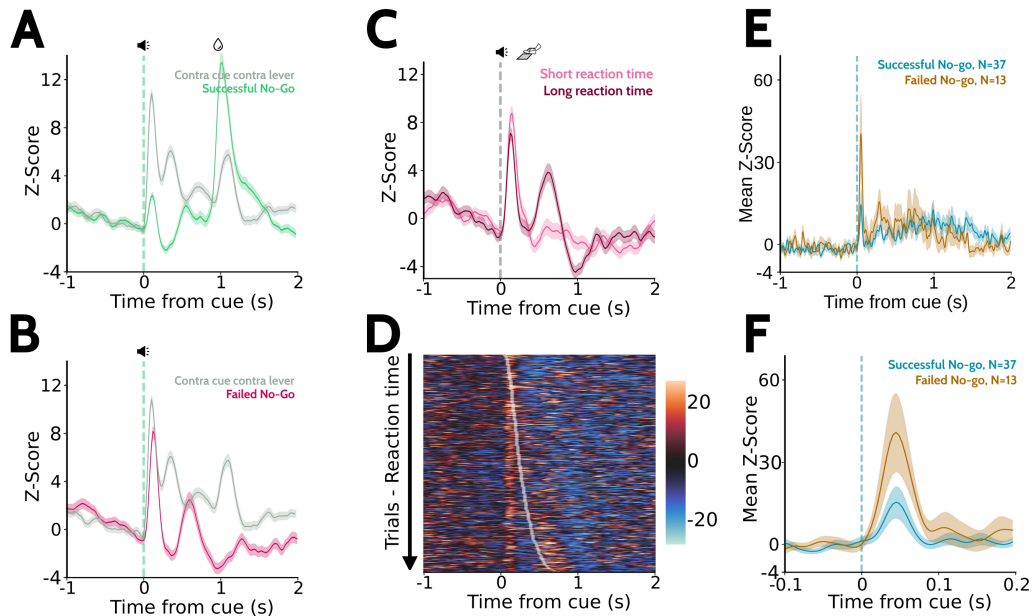


Figure 7.11: **Dopaminergic activity during NoGo trials.** **A-B:** Mean Z-scored GCaMP activity during successful (**A**, green) and failed (**B**, pink) NoGo trials. NoGo trials required the animal to not press any lever for 800ms to obtain the reward. A NoGo trial is failed if the animal pressed a lever in this interval. **C:** Mean Z-scored GCaMP activity during failed NoGo trials with a reaction time shorter than the median ('Short reaction time', light pink), and trials with a reaction time longer than the median ('Long reaction time', dark pink), aligned on the cue. **D:** Heatmap of all failed NoGo trials, ordered by increasing reaction time. The solid white line represents the movement onset. **E:** Z-scored electrophysiological activity of putative SNc dopaminergic neurons during successful NoGo trials (blue, N=37) and failed NoGo trials (orange, N=13). **F:** Zoom around the cue response of E.

the rats have relatively poor performances with this type of trial, and NoGo trials during the first third of the session are mostly failed. However, the animal still understood the NoGo as the probability of not pressing a lever during NoGo trials is still higher than compared to normal Go trials.

Fig. 7.11 A shows the dopaminergic activity during a successful NoGo trial. Similarly to what we observe during missed Go trials, there is a very small cue-evoked response that is followed by an inhibition. The inhibition here in NoGo trial lasts longer than the one in missed Go trials. During failed NoGo trials (7.11 B), when the animal pressed a lever before the end of the 800ms delay, the

dopaminergic cue-evoked response is stronger. It is however different from successful Go trials: 7.11 C and D show that the second increase in dopaminergic activity during failed NoGo trials is due to the movement onset, mostly during trials with a long reaction time. It is unclear why only trials with long reaction time exhibit this movement-related activity. An inhibition can be observed around the time of the lever press as the error is signaled to the rat (LED lights turning on).

Few units have been recorded in the SNc during enough NoGo trials. 7.11 E and F show their response around the NoGo cue. We can see that dopaminergic units tend to have stronger responses to the NoGo cue when the animals fails the trial and presses a lever. In this case the electrophysiological data is thus consistent with the fiber photometry data.

Part IV

Discussion

Discussion

During my PhD, we have developed a method to record simultaneously the electrophysiological activity of single dopaminergic neurons and the DLS global calcium activity of dopaminergic terminals, in head-fixed rats performing a complex reach-and-grasp instrumental task. We have observed that dopaminergic neurons were active during movement and that different neurons were responsible for the cue-evoked and movement-evoked dopamine transients. We have demonstrated that this movement-evoked dopaminergic activity was lateralized and stronger when the contralateral limb was used. Then, we investigated the correlation between this dopaminergic signal and the vigor of the movement and found that dopaminergic activity generally correlated with the reaction time and speed of the animal. Finally, we observed that when movement was repressed, dopaminergic activity in the DLS was inhibited, whether this absence of movement was associated with a reward or not.

8.1 Recording and identifying dopaminergic neurons in behaving head-fixed rats.

The choice of acute electrophysiological recordings instead of chronic implantation was driven by the scarcity of dopaminergic neurons. An implanted probe would at most record a few dopaminergic neurons per animal, which is too little considering the time needed to fully train a rat to perform the task. Here, the silicon probe was

lowered into a new area in the SNc for every recording session increasing the yield of dopamine neurons recordings per animal. The main downside of such acute head-fixed recordings in general, is the added training difficulty and the increased risk of death. Out of all implanted animals, less than 30% were ultimately successfully recorded. Another limitation of such acute preparation is the imprecision regarding the localization of the probe: with our long-term recordings, the bone of the skull tends to degrade and the stereotaxic landmarks could disappear. In addition, the surface of the brain at the recording site could swell over time, making depth estimation more difficult. A final downside of acute recording is the instability of the recordings, especially when the animal performs a behavioral task; units can abruptly switch from one recording channel to the next if the animal makes a sudden move.

8.2 Dopaminergic neurons activity is specialized for task events

By recording individual dopaminergic neurons during the instrumental task, we have shown that their activity is highly heterogeneous in the SNc, and that each neuron is only phasically active at one specific moment during the task. Almost all recorded neurons were active at a different moment, but sub-groups can be identified. In particular, we found that one subgroup of neurons was consistently more active between the cue and the movement onset (although even within this group they did not all respond with the same latency), another subgroup was the most active during the movement with some neurons firing phasically at movement onset and others increasing their activity for the entire duration of the movement. This heterogeneity in SNc neuronal response could be due to different inputs carrying sensory and/or proprioceptive information. This finding is consistent with a previous study showing that VTA neurons were encoding specific behavioral variables such as position, reward, cue, and kinematics (Engelhard et al., 2019). This specialized encoding of dopamine neurons has also been suggested by other studies (Pasquereau and Turner, 2015, Howe MW, 2016 and Mendonça et al., 2021) and was likely missed when recordings were performed during simple behavioral

tasks such as in Pavlovian conditioning.

It is worth noting that our recorded SNc neurons did not exhibit any response to reward delivery, whether expected or unexpected. It is plausible that this negative result is a consequence of technical recording limitations. During reward consumption, the electrophysiological activity became noisier and exhibited artifacts when the rat licked the reward port, which could have prevented us from isolating reward-related units. That being said, another explanation might be related to the fact that reward-related stimuli are preferentially processed by VTA neurons rather than SNc neurons (Morales and Margolis, 2017).

8.3 Dopaminergic neurons are active during movement

Both with electrophysiological recordings and fiber photometry recordings, we observed that dopaminergic neurons were active during the reaching movement of the rat. This activity usually followed and did not precede movement onset. The DLS GCaMP activity during the movement peaked around 140ms after movement onset, which is approximately the same delay as the cue-evoked response (120ms for DLS GCaMP). Depending on the animal, the GCaMP cue-evoked and movement-evoked response can be more or less isolated, due to variations in the signal as well as reaction time. By restricting our analysis to trials with a long reaction time, we can clearly identify a dopaminergic activity during the movement (fig. 7.8). These results are consistent with the literature, in particular, Syed et al., 2016 and Roitman et al., 2004, which showed the separation of the cue-evoked and movement-evoked responses. Many papers, however, argued that dopaminergic neurons were active before or around movement onset, but not during the movement (Coddington and Dudman, 2018; Da Silva et al., 2018; Dodson et al., 2016a). That said, these papers looked at spontaneous unskilled movements, and it can be more difficult to determine the precise timing of movement onset for behaviors such as locomotion. On the opposite, our behavioral task enabled us to precisely detect the moment at which the front paw lifted off the pad. This event is not the exact moment of movement onset though, especially since the animal has first to stabilize

its posture and starts contracting its limb muscles in an isometric fashion before it can actually start moving it. Looking at the video recordings of the forelimb movement, we could see that the limb started moving while staying in contact with the pad for up to 40ms before the movement onset was truly detected. With all these elements taken into consideration, a nice implementation in our experiments would be to measure the electromyography activity of the forelimb muscle which would provide the exact timing of movement onset. At present, it is plausible that we have a small 50ms lag in movement onset detection.

What is the physiological function of this movement-related activity? A recent preprint argued that movement-evoked dopaminergic activity in the tail of striatum encoded Action Prediction Error (Greenstreet et al., 2022, see 3.3.2). APE would signal how much the currently executed stimulus-action pair differs from the learned stimulus-action association. If the animal learned to do action A in response to stimulus A, when it does action B in response to stimulus A, there would be an action prediction error signal. Dopaminergic neurons would somehow predict that a certain action will follow a given stimulus and if this is not the case they will signal a prediction error, much like the RPE. In our data, we tend to observe an opposite effect to what the APE model would predict. Indeed, when the animal presses the wrong lever during a trial, the dopaminergic activity during movement should be higher than during correct movement. However, our recordings show that dopamine activity in the DLS is usually lower during wrong movements. The APE encoding of dopaminergic neurons might thus be restricted to the tail of the striatum.

8.4 Dopaminergic activity is selectively lateralized

A few studies have shown lateralization of dopaminergic activity, notably two recent preprints available on BioRxiv: Greenstreet et al., 2022 and Hart et al., 2022. Both papers recorded the dopaminergic activity in the striatum with the dopamine sensor dLight1.1. They observed that the dopamine release was stronger

in the hemisphere contralateral to the movement. However, it is unknown how this lateralization affects the activity of single dopaminergic neurons. Does contralateral movement recruit more neurons? Is the dopamine release modulated directly in the striatum? Mendonça et al., 2021 observed that different neurons were active during contralateral and ipsilateral movement. In our recording we found that the activity of dopaminergic terminals in the DLS was higher during contralateral movements, decreasing the likelihood that the difference in dopamine release is due to a control at the synapse. In addition, we recorded SNc dopaminergic neurons during both contralateral and ipsilateral movements. We found neurons that were active only during contralateral movement and neurons that were active both during ipsi- and contralateral movement (with a slightly higher mean firing rate during contralateral movement). We, however, did not find neurons only active for ipsilateral but not contralateral movement, thus contradicting the results from Mendonça et al., 2021. Interestingly, Hart et al., 2022 showed that the lateralization of movement-related activity appears during training since at the beginning of learning, both ipsi- and contralateral movements elicited dopamine release. In our data set, it is not possible to disentangle this lateralization issue or even dissect when movement-related activity starts to appear as we don't have access to recordings early in training.

Interestingly, in Greenstreet et al., 2022, the behavioral task involves the mice moving from a central port to a lateral port. It is surprising that a locomotion movement, that involves the whole body, triggers such a lateralization of dopaminergic activity. Overall, most instrumental tasks are little suited to the study of lateralization, since they rarely require an isolated movement of only one side of the body. Of course, no movement is ever truly unilateral: even moving one limb is going to require the activation of contralateral muscles in order to stabilize the body. However, our task enables the best isolation of movement among tasks described in the literature so far. The head fixation even decreases the need for contralateral stabilization as the body is already stabilized by the head.

8.5 Dopaminergic activity is correlated with movement vigor

The main goal of this PhD project was to investigate the role of dopamine in the control of movement. The idea being, that if dopaminergic neurons were active at movement onset they might influence the excitability of cortico-striatal synapses and facilitate movement or increase its speed. We were also interested in the role of the cue-evoked dopaminergic activity in the context of an instrumental task. Indeed, if in simple pavlovian tasks the cue-evoked activity only encodes information about the reward (value, probability, delay, see 2.2), in tasks where the reward is contingent upon an action from the animal, the cue-evoked activity might be involved in modulating movement kinematics.

To investigate this hypothesis we analyzed the correlation between the dopaminergic activity and the reaction time or movement time. We found a negative correlation between dopamine levels and reaction time as well as movement time. Surprisingly, both the cue-evoked and the movement-evoked dopaminergic activity correlated to the current trial's reaction and movement time. These unspecific results might indicate a bias in the analysis of the signal itself; for example, the proximity of the cue-evoked and the movement-evoked responses could bias the correlation between the movement-related activity and reaction time for the shorter reaction times (since short reaction time means that the movement starts during the cue-evoked dopaminergic activity). So, at this stage, it is hard to conclude which dopaminergic activity is the most likely to modulate movement vigor. In addition, we also found a correlation between the movement time and the reward-evoked dopaminergic response, as if the speed of the movement updated the reward prediction during the task. This could indicate an effect of movement vigor on dopaminergic activity.

A frequent critique addressed in the investigation of vigor control is that vigor lacks a clear definition. In addition, vigor and motivation are often used interchangeably and the distinction between the two is vague. Is vigor just speed? If so, why not just use the term 'speed'? Does vigor include reaction time, strength,

and amplitude of movement? Even in everyday usage, we can struggle to give an actual definition of vigor. The broad definition of vigor used for this study is the following: The energy used for a movement. We consider that this includes reaction time, speed, and strength. Due to the difficulty of measuring strength in our task, we only used reaction time and movement time as a proxy for vigor.

As for the distinction between vigor and motivation, we consider that motivation is something that influences whether you do an action or not in a given situation, while vigor defines how you do the action if you do initiate it.

A surprising result in our dataset is the fact that short reaction times are not usually associated with short movement time and that the opposite is more likely to happen (short reaction time with long movement time *et vice et versa*). If vigor as we define it is a combination of reaction time and speed, then why are the two not varying together? This decorrelation could suggest that in fact, reaction time and speed are separate processes that are controlled independently.

8.6 Future directions

As we have observed a correlation between dopaminergic activity and movement vigor, we now want to test if there is a causal link between the two. A first hypothesis would be that dopamine could influence vigor in an online manner (modulating the vigor of the movement immediately following the dopamine transient), as has been argued by Da Silva et al., 2018 and Mendonça et al., 2021. This online motor control would rely on the modification of MSNs excitability which is theoretically possible but we questioned its effectiveness since dopamine acts on G protein-coupled receptors whose signaling cascade is rather slow (Lohse et al., 2008). With this in mind, dopamine transmission might not be the ideal system to modulate online movement execution. A second hypothesis would be that dopamine can influence vigor but only through a learning/plasticity mechanism, over multiple trials. This type of effect is well described for the reward-evoked dopamine release: stimulating the dopaminergic neurons after an action reinforces this action, and is the fundamental principle of reinforcement learning (Lee et al., 2020). Here, we are more interested in the function of the cue-evoked and

movement-evoked dopamine release. If they do not impact the current movement, maybe they have an effect over repeated trials? To test these two hypotheses, we will perform optogenetic experiment to inhibit or stimulate SNc dopaminergic neurons after the cue or during movement, during random trials, or during subsequent trials. We will combine the optogenetic stimulations with GRAB-DA or dLight recordings in the DLS to 1) verify that the optogenetic stimulation is having an effect on dopamine release and 2) compare the dynamics of dopamine release during the task with the GCaMP recordings of dopaminergic terminals. Indeed, it is possible that the effective dopamine release do not match the activity of the cell bodies or DLS axon terminals (C. Liu et al., 2022; Mohebi et al., 2022).

Overall, our work has confirmed that skilled movements in an instrumental task will evoke an increase in dopaminergic activity and that this activity was lateralized depending on the limb used by the animal. We also report an unspecific correlation between the dopaminergic activity and the vigor of the movement, as well as an unexplained inhibition when movement is repressed. These results highlight the complex spatiotemporal dynamics of striatal dopamine transmission during skilled movements which meaning needs to be addressed through causal optogenetic perturbations. In conclusion, our work provides a complete overview of the activity of dopaminergic neurons in the SNc and the DLS during a complex instrumental motor task.

References

- Adams, C. D., & Dickinson, A. (1981). Instrumental responding following reinforcer devaluation. *The Quarterly Journal of Experimental Psychology Section B*, *33*(2), 109–121. <https://doi.org/10.1080/14640748108400816>
- Amo, R., Matias, S., Yamanaka, A., Tanaka, K. F., Uchida, N., & Watabe-uchida, M. (2022). A gradual temporal shift of dopamine responses mirrors the progression of temporal difference error in machine learning. *Nature Neuroscience*, *25*(8), 1082–1092. <https://doi.org/10.1038/s41593-022-01109-2.A>
- André, V. M., Cepeda, C., & Levine, M. S. (2010). Dopamine and Glutamate in Huntington’s Disease: A Balancing Act. *CNS Neuroscience & Therapeutics*, *16*(3), 163–178. <https://doi.org/https://doi.org/10.1111/j.1755-5949.2010.00134.x>
- Arber, S., & Costa, R. M. (2018). Connecting neuronal circuits for movement. *Science*, *360*(6396), 1403–1404. <https://doi.org/10.1126/science.aat5994>
- Balleine, B. W. (2005). Neural bases of food-seeking: Affect, arousal and reward in corticostriatolimbic circuits. *Physiology & Behavior*, *86*(5), 717–730. <https://doi.org/https://doi.org/10.1016/j.physbeh.2005.08.061>
Purdue University Ingestive Behavior Research Center Symposium. Dietary Influences on Obesity: Environment, Behavior and Biology
- Balleine, B. W., & Dickinson, A. (1998). The role of incentive learning in instrumental outcome revaluation by sensory-specific satiety. *Animal Learning and Behavior*, *26*(1), 46–59. <https://doi.org/10.3758/BF03199161>
- Bamford, N. S., Robinson, S., Palmiter, R. D., Joyce, J. A., Moore, C., & Meshul, C. K. (2004). Dopamine Modulates Release from Corticostriatal Terminals. *Journal of Neuroscience*, *24*(43), 9541–9552. <https://doi.org/10.1523/JNEUROSCI.2891-04.2004>
- Bamford, N. S., Wightman, R. M., & Sulzer, D. (2018). Dopamine’s Effects on Corticostriatal Synapses during Reward-Based Behaviors. *Neuron*, *97*(3), 494–510. <https://doi.org/10.1016/j.neuron.2018.01.006>
- Baraduc, P., Thobois, S., Gan, J., Broussolle, E., & Desmurget, M. (2013). A Common Optimization Principle for Motor Execution in Healthy Subjects

- and Parkinsonian Patients. *Journal of Neuroscience*, *33*(2), 665–677. <https://doi.org/10.1523/JNEUROSCI.1482-12.2013>
- Barnes, T. D., Kubota, Y., Hu, D., Jin, D. Z., & Graybiel, A. M. (2005). Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. *Nature*, *437*(7062), 1158–1161. <https://doi.org/10.1038/nature04053>
- Barter, J. W., Li, S., Lu, D., Bartholomew, R. A., Rossi, M. A., Shoemaker, C. T., Salas-Meza, D., Gaidis, E., & Yin, H. H. (2015). Beyond reward prediction errors: The role of dopamine in movement kinematics. *Frontiers in Integrative Neuroscience*, *9*(MAY), 1–22. <https://doi.org/10.3389/fnint.2015.00039>
- Bayer, H. M., & Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron*, *47*(1), 129–141. <https://doi.org/10.1016/j.neuron.2005.05.020>
- Beier, K. T., Steinberg, E. E., Deloach, K. E., Xie, S., Miyamichi, K., Schwarz, L., Gao, X. J., Kremer, E. J., Malenka, R. C., & Luo, L. (2015). Circuit Architecture of VTA Dopamine Neurons Revealed by Systematic Input-Output Mapping. *Cell*, *162*(3), 622–634. <https://doi.org/10.1016/j.cell.2015.07.015>
- Belin, D., Jonkman, S., Dickinson, A., Robbins, T. W., & Everitt, B. J. (2009). Parallel and interactive learning processes within the basal ganglia: Relevance for the understanding of addiction. *Behavioural Brain Research*, *199*(1), 89–102. <https://doi.org/https://doi.org/10.1016/j.bbr.2008.09.027>
Special issue on the role of the basal ganglia in learning and memory
- Bogacz, R. (2020). Dopamine role in learning and action inference. *eLife*, *9*, 1–33. <https://doi.org/10.7554/eLife.53262>
- Bova, A., Gaidica, M., Hurst, A., Iwai, Y., Hunter, J., & Leventha, D. K. (2020). Precisely-timed dopamine signals establish distinct kinematic representations of skilled movements. *eLife*, *9*, 1–141. <https://doi.org/10.7554/eLife.61591>
- Bromberg-Martin, E. S., & Hikosaka, O. (2009). Midbrain dopamine neurons signal preference for advance information about upcoming rewards. *Neuron*, *63*(1), 119–126. <https://doi.org/10.1016/j.neuron.2009.06.009>.Midbrain
- Buccino, A. P., Hurwitz, C. L., Garcia, S., Magland, J., Siegle, J. H., Hurwitz, R., & Hennig, M. H. (2020). SpikeInterface, a unified framework for spike sorting. *eLife*, *9*, e61834. <https://doi.org/10.7554/eLife.61834>
- Buhusi, C. V., & Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. <https://doi.org/10.1038/nrn1764>
- Bush, B. Y. R. R., & Mosteller, F. (1951). A MATHEMATICAL MODEL FOR SIMPLE LEARNING. *58*(5).
- Calabresi, P., Picconi, B., Tozzi, A., & Filippo, M. D. (2007). Dopamine-mediated regulation of corticostriatal synaptic plasticity. *Trends in Neurosciences*, *30*(5). <https://doi.org/10.1016/j.tins.2007.03.001>

- Coddington, L. T., & Dudman, J. T. (2018). The timing of action determines reward prediction signals in identified midbrain dopamine neurons. *Nature Neuroscience*, *21*(11), 1563–1573. <https://doi.org/10.1038/s41593-018-0245-7>
- Coddington, L. T., & Dudman, J. T. (2019). Learning from Action: Reconsidering Movement Signaling in Midbrain Dopamine Neuron Activity. *Neuron*, *104*(1), 63–77. <https://doi.org/10.1016/j.neuron.2019.08.036>
- 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), 1359–1359. https://doi.org/10.1007/978-3-540-68706-1_760
- Colwill, R. M., & Rescorla, R. A. (1985). Instrumental responding remains sensitive to reinforcer devaluation after extensive training. *Journal of Experimental Psychology: Animal Behavior Processes*, *11*(4), 520–536. <https://doi.org/10.1037//0097-7403.11.4.520>
- Corbit, L. H., & Balleine, B. W. (2005). Double Dissociation of Basolateral and Central Amygdala Lesions on the General and Outcome-Specific Forms of Pavlovian-Instrumental Transfer. *Journal of Neuroscience*, *25*(4), 962–970. <https://doi.org/10.1523/JNEUROSCI.4507-04.2005>
- Coyne, D., Marrelec, G., Perlberg, V., Péligrini-Issac, M., Van de Moortele, P. F., Ugurbil, K., Doyon, J., Benali, H., & Lehericy, S. (2010). Dynamics of motor-related functional integration during motor sequence learning. *NeuroImage*, *49*(1), 759–766. <https://doi.org/10.1016/j.neuroimage.2009.08.048>
- Cragg, S. J., Nicholson, C., Kume-Kick, J., Tao, L., & Rice, M. E. (2001). Dopamine-mediated volume transmission in midbrain is regulated by distinct extracellular geometry and uptake. *Journal of Neurophysiology*, *85*(4), 1761–1771. <https://doi.org/10.1152/jn.2001.85.4.1761>
- Cui, G., Jun, S. B., Jin, X., Pham, M. D., Vogel, S. S., Lovinger, D. M., & Costa, R. M. (2013). Concurrent activation of striatal direct and indirect pathways during action initiation. *Nature*, *494*(7436), 238–242. <https://doi.org/10.1038/nature11846>
- Da Silva, J. A., Tecuapetla, F., Paixão, V., & Costa, R. M. (2018). Dopamine neuron activity before action initiation gates and invigorates future movements. *Nature*, *554*(7691), 244–248. <https://doi.org/10.1038/nature25457>
- Dahlström, A., & Fuxe, K. (1964). Localization of monoamines in the lower brain stem. *Experientia*, *20*(7), 398–399. <https://doi.org/10.1007/BF02147990>
- Dayan, P. (2009). Goal-directed control and its antipodes. *Neural Networks*, *22*(3), 213–219. <https://doi.org/https://doi.org/10.1016/j.neunet.2009.03.004>
- Goal-Directed Neural Systems

- De Lafuente, V., & Romo, R. (2011). Dopamine neurons code subjective sensory experience and uncertainty of perceptual decisions. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(49), 19767–19771. <https://doi.org/10.1073/pnas.1117636108>
- Deniau, J. M., Menetrey, A., & Charpier, S. (1996). The lamellar organization of the rat substantia nigra pars reticulata: Segregated patterns of striatal afferents and relationship to the topography of corticostriatal projections. *Neuroscience*, *73*(3), 761–781. [https://doi.org/https://doi.org/10.1016/0306-4522\(96\)00088-7](https://doi.org/10.1016/0306-4522(96)00088-7)
- Dickinson, A., & Weiskrantz, L. (1985). Actions and habits: the development of behavioural autonomy. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *308*(1135), 67–78. <https://doi.org/10.1098/rstb.1985.0010>
- Dodson, P. D., Dreyer, J. K., Jennings, K. A., Syed, E. C., Wade-Martins, R., Cragg, S. J., Bolam, J. P., Magill, P. J., & Malenka, R. C. (2016a). Representation of spontaneous movement by dopaminergic neurons is cell-type selective and disrupted in parkinsonism. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(15), E2180–E2188. <https://doi.org/10.1073/pnas.1515941113>
spontaneous locomotion
- Dodson, P. D., Dreyer, J. K., Jennings, K. A., Syed, E. C., Wade-Martins, R., Cragg, S. J., Bolam, J. P., Magill, P. J., & Malenka, R. C. (2016b). Supporting Information : Representation of spontaneous movement by dopaminergic neurons is cell-type selective and disrupted in parkinsonism. *Proceedings of the National Academy of Sciences of the United States of America*, *2*(1), 1–5.
- Dudman, J. T., & Krakauer, J. W. (2016). The basal ganglia: from motor commands to the control of vigor. *Current Opinion in Neurobiology*, *37*, 158–166. [https://doi.org/https://doi.org/10.1016/j.conb.2016.02.005](https://doi.org/10.1016/j.conb.2016.02.005)
Neurobiology of cognitive behavior
- Dumartin, B., Doudnikoff, E., Gonon, F., & Bloch, B. (2007). Differences in ultrastructural localization of dopaminergic D1 receptors between dorsal striatum and nucleus accumbens in the rat. *Neuroscience Letters*, *419*(3), 273–277. [https://doi.org/https://doi.org/10.1016/j.neulet.2007.04.034](https://doi.org/10.1016/j.neulet.2007.04.034)
- Engelhard, B., Finkelstein, J., Cox, J., Fleming, W., Jang, H. J., Ornelas, S., Koay, S. A., Thiberge, S. Y., Daw, N. D., Tank, D. W., & Witten, I. B. (2019). Specialized coding of sensory, motor and cognitive variables in VTA dopamine neurons. *Nature*, *570*(7762), 509–513. <https://doi.org/10.1038/s41586-019-1261-9>

- Eshel, N., Tian, J., Bukwich, M., & Uchida, N. (2016). Dopamine neurons share common response function for reward prediction error. *Nature Neuroscience*, *118*(24), 6072–6078. <https://doi.org/10.1038/nn.4239>. Dopamine
- Eshel, N., Toupponse, G. C., Wang, A. R., Osterman, A. K., Shank, A. N., Groome, A. M., Taniguchi, L., Pinto, D. F. C., Tucciarone, J., Bentzley, B. S., & Malenka, R. C. (2022). Striatal dopamine integrates cost, benefit and motivation. *bioRxiv*, *3*(April), 49–58.
- Faure, A., Haberland, U., Condé, F., & Massioui, N. E. (2005). Lesion to the Nigrostriatal Dopamine System Disrupts Stimulus-Response Habit Formation. *Journal of Neuroscience*, *25*(11), 2771–2780. <https://doi.org/10.1523/JNEUROSCI.3894-04.2005>
- Fiorillo, C. D., Newsome, W. T., & Schultz, W. (2008). The temporal precision of reward prediction in dopamine neurons. *Nature Neuroscience*, *11*(8), 966–973. <https://doi.org/10.1038/nn.2159>
- Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*, *299*(5614), 1898–1902. <https://doi.org/10.1126/science.1077349>
- Flagel, S. B., Clark, J. J., Robinson, T. E., Mayo, L., Czuj, A., Willuhn, I., Akers, C. A., Clinton, S. M., Phillips, P. E. M., & Akil, H. (2011). A selective role for dopamine in reward learning. *Nature*, *176*(5), 139–148. <https://doi.org/10.1038/nature09588>.A
- Freeman, A. S., Meltzer, L. T., & Bunney, B. S. (1985). Firing properties of substantia nigra dopaminergic neurons in freely moving rats. *Life Sciences*, *36*(20), 1983–1994. [https://doi.org/https://doi.org/10.1016/0024-3205\(85\)90448-5](https://doi.org/https://doi.org/10.1016/0024-3205(85)90448-5)
- Gangarossa, G., Espallergues, J., Mailly, P., De Bundel, D., De Kerchove D’Exaerde, A., Hervé, D., Girault, J.-A., Valjent, E., & Krieger, P. (2013). Spatial distribution of D1R- and D2R-expressing medium-sized spiny neurons differs along the rostro-caudal axis of the mouse dorsal striatum. *Frontiers in Neural Circuits*, *7*. <https://doi.org/10.3389/fncir.2013.00124>
- Gao, C., Wedell, D. H., Green, J. J., Jia, X., Mao, X., Guo, C., & Shinkareva, S. V. (2018). Temporal dynamics of audiovisual affective processing. *Biological Psychology*, *139*, 59–72. <https://doi.org/https://doi.org/10.1016/j.biopsycho.2018.10.001>
- Gerdeman, G. L., Ronesi, J., & Lovinger, D. M. (2002). Postsynaptic endocannabinoid release is critical to long-term depression in the striatum. *Nature Neuroscience*, *5*(5), 1–6. <https://doi.org/10.1038/nn832>
- Gerfen, C. R., Engber, T. M., Mahan, L. C., Susel, Z., Chase, T. N., Monsma, F. J., & Sibley, D. R. (1990). D₁ and D₂ Dopamine Receptor-regulated Gene Expression of Striatonigral and Striatopallidal

- Neurons. *Science*, 250(4986), 1429–1432. <https://doi.org/10.1126/science.2147780>
- Gerfen, C. R., & Surmeier, D. J. (2011). Modulation of Striatal Projection Systems by Dopamine. *Annual Review of Neuroscience*, 34(1), 441–466. <https://doi.org/10.1146/annurev-neuro-061010-113641>
PMID: 21469956
- German, D. C., & Manaye, K. F. (1993). Midbrain dopaminergic neurons (nuclei A8, A9, and A10): Three-dimensional reconstruction in the rat. *Journal of Comparative Neurology*, 331(3), 297–309. <https://doi.org/https://doi.org/10.1002/cne.903310302>
- Gonon, F. (1997). Prolonged and extrasynaptic excitatory action of dopamine mediated by D1 receptors in the rat striatum in vivo. *Journal of Neuroscience*, 17(15), 5972–5978. <https://doi.org/10.1523/jneurosci.17-15-05972.1997>
- Goto, Y., & Grace, A. A. (2007). Dopamine Modulation of Hippocampal–Prefrontal Cortical Interaction Drives Memory-Guided Behavior. *Cerebral Cortex*, 18(6), 1407–1414. <https://doi.org/10.1093/cercor/bhm172>
- Grace, A. A., & Bunney, B. S. (1984). The control of firing pattern in nigral dopamine neurons: Burst firing. *Journal of Neuroscience*, 4(11), 2877–2890. <https://doi.org/10.1523/jneurosci.04-11-02877.1984>
- Grace, A. A., & Bunney, B. S. (1980). Nigral Dopamine Neurons: Intracellular Recording and Identification with L-Dopa Injection and Histofluorescence. *Science*, 210(4470), 654–656. <https://doi.org/10.1126/science.7433992>
- Graybiel, A. M. (2008). Habits, Rituals, and the Evaluative Brain. *Annual Review of Neuroscience*, 31(1), 359–387. <https://doi.org/10.1146/annurev.neuro.29.051605.112851>
PMID: 18558860
- Greenstreet, F., Vergara, H. M., Pati, S., Schwarz, L., Wisdom, M., Marbach, F., Johansson, Y., Rollik, L., & Moskovitz, T. (2022). Action prediction error : a value-free dopaminergic teaching signal that drives stable learning. *bioRxiv*.
- Guo, L., Walker, W. I., Ponvert, N. D., Penix, P. L., & Jaramillo, S. (2018). Stable representation of sounds in the posterior striatum during flexible auditory decisions. *Nature Communications*, 9(1). <https://doi.org/10.1038/s41467-018-03994-3>
- Haber, S. N., Fudge, J. L., & McFarland, N. R. (2000). Striatonigrostriatal Pathways in Primates Form an Ascending Spiral from the Shell to the Dorsolateral Striatum. *Journal of Neuroscience*, 20(6), 2369–2382. <https://doi.org/10.1523/JNEUROSCI.20-06-02369.2000>
- Håkansson, K., Galdi, S., Hendrick, J., Snyder, G., Greengard, P., & Fisone, G. (2006). Regulation of phosphorylation of the GluR1 AMPA receptor

- by dopamine D2 receptors. *Journal of Neurochemistry*, *96*(2), 482–488. <https://doi.org/https://doi.org/10.1111/j.1471-4159.2005.03558.x>
- Hamid, A. A., Pettibone, J. R., Mabrouk, O. S., Hetrick, V. L., Schmidt, R., Vander Weele, C. M., Kennedy, R. T., Aragona, B. J., & Berke, J. D. (2015). Mesolimbic dopamine signals the value of work. *Nature Neuroscience*, *19*(1), 117–126. <https://doi.org/10.1038/nn.4173>
- Hart, G., Burton, T. J., & Balleine, B. W. (2022). Striatal dopamine encodes the relationship between actions and reward. *bioRxiv*.
- Higley, M. J., & Sabatini, B. L. (2010). Competitive regulation of synaptic Ca²⁺ influx by D2 dopamine and A2A adenosine receptors. *Nature Publishing Group*, *13*(8), 958–966. <https://doi.org/10.1038/nn.2592>
- Hikosaka, O., Takikawa, Y., & Kawagoe, R. (2000). Role of the Basal Ganglia in the Control of Purposive Saccadic Eye Movements. *Physiological Reviews*, *80*(3), 953–978. <https://doi.org/10.1152/physrev.2000.80.3.953>
PMID: 10893428
- Hintiryan, H., Foster, N. N., Bowman, I., Bay, M., Song, M. Y., Gou, L., Yamashita, S., Bienkowski, M. S., Zingg, B., Zhu, M., Yang, X. W., Shih, J. C., Toga, A. W., & Dong, H. W. (2016). The mouse cortico-striatal projectome. *Nature Neuroscience*, *19*(8), 1100–1114. <https://doi.org/10.1038/nn.4332>
- Holland, P. C. (1980). CS-US interval as a determinant of the form of Pavlovian appetitive conditioned responses. *Journal of experimental psychology. Animal behavior processes*, *6* 2, 155–174.
- Hosp, J. A., Pekanovic, A., Rioult-Pedotti, M. S., & Luft, A. R. (2011). Dopaminergic Projections from Midbrain to Primary Motor Cortex Mediate Motor Skill Learning. *Journal of Neuroscience*, *31*(7), 2481–2487. <https://doi.org/10.1523/JNEUROSCI.5411-10.2011>
- Howard, C. D., Li, H., Geddes, C. E., & Jin, X. (2017). Dynamic Nigrostriatal Dopamine Biases Action Selection. *Neuron*, *93*(6), 1436–1450.e8. <https://doi.org/10.1016/j.neuron.2017.02.029>
- Howe, M. W., Patrick, L. T., Sandberg, S. G., Phillips, P. E., & Graybiel, A. M. (2014). Prolonged Dopamine Signalling in Striatum Signals Proximity and Value of Distant Rewards. *Nature*, *500*(7464), 575–579. <https://doi.org/10.1038/nature12475>.Prolonged
- Howe MW, D. D. (2016). Rapid signaling in distinct dopaminergic axons during locomotion and reward MW. *Nature*, *176*(10). <https://doi.org/10.1038/nature18942>.Rapid
- Hunger, L., Kumar, A., & Schmidt, R. (2020). Abundance compensates kinetics: Similar effect of dopamine signals on D1 and D2 receptor populations. *Journal of Neuroscience*, *40*(14), 2868–2881. <https://doi.org/10.1523/JNEUROSCI.1951-19.2019>

- Hunnicutt, B. J., Jongbloets, B. C., Birdsong, W. T., Gertz, K. J., Zhong, H., & Mao, T. (2016). A comprehensive excitatory input map of the striatum reveals novel functional organization. *eLife*, *5*, e19103. <https://doi.org/10.7554/eLife.19103>
- Hunter, J., Bova, A., Stevens, A., & Leventhal, D. K. (2022). Dopamine neuron stimulation induces context-dependent abnormal involuntary movements in healthy rats. *iScience*, *25*(3), 103974. <https://doi.org/10.1016/j.isci.2022.103974>
- Iino, Y., Sawada, T., Yamaguchi, K., Tajiri, M., Ishii, S., Kasai, H., & Yagishita, S. (2020). Dopamine D2 receptors in discrimination learning and spine enlargement. *Nature*, *579*(7800), 555–560. <https://doi.org/10.1038/s41586-020-2115-1>
- Jeong, H., Taylor, A., Floeder, J. R., Lohmann, M., Mihalas, S., Wu, B., Zhou, M., Burke, D. A., & Nambodiri, V. M. K. (2022). Mesolimbic dopamine release conveys causal associations. *Science*, *34*(3), 642–685. <https://doi.org/10.1126/science.abq6740>
- Jiang, H., & Kim, H. F. (2018). Anatomical Inputs From the Sensory and Value Structures to the Tail of the Rat Striatum. *Frontiers in Neuroanatomy*, *12*. <https://doi.org/10.3389/fnana.2018.00030>
- Jin, X., & Costa, R. M. (2010). Start/Stop Signals Emerge in Nigrostriatal Circuits during Sequence Learning. *Nature*, *23*(1), 1–7. <https://doi.org/10.1038/nature09263>
- Jin, X., Tecuapetla, F., & Costa, R. M. (2014). Basal ganglia subcircuits distinctively encode the parsing and concatenation of action sequences. *Nature Neuroscience*, *17*(3), 423–430. <https://doi.org/10.1038/nn.3632>
- Jocoy, E., André, V., Cummings, D., Rao, S., Wu, N., Ramsey, A., Caron, M., Cepeda, C., & Levine, M. (2011). Dissecting the Contribution of Individual Receptor Subunits to the Enhancement of N-methyl-d-Aspartate Currents by Dopamine D1 Receptor Activation in Striatum. *Frontiers in Systems Neuroscience*, *5*. <https://doi.org/10.3389/fnsys.2011.00028>
- Kebabian, J. W., & Calne, D. B. (1979). Multiple receptors for dopamine. *Nature*, *277*, 93–96.
- Kirkpatrick, D. C., McKinney, C. J., Manis, P. B., & Wightman, R. M. (2016). Expanding neurochemical investigations with multi-modal recording: simultaneous fast-scan cyclic voltammetry, iontophoresis, and patch clamp measurements. *Analyst*, *141*(16), 4902–4911. <https://doi.org/10.1039/C6AN00933F>
- Kobayashi, S., & Schultz, W. (2008). Influence of reward delays on responses of dopamine neurons. *Journal of Neuroscience*, *28*(31), 7837–7846. <https://doi.org/10.1523/JNEUROSCI.1600-08.2008>

- Krack, P., Hariz, M. I., Baunez, C., Guridi, J., & Obeso, J. A. (2010). Deep brain stimulation: From neurology to psychiatry? *Trends in Neurosciences*, *33*(10), 474–484. <https://doi.org/10.1016/j.tins.2010.07.002>
- Krausz, T. A., Comrie, A. E., Frank, L. M., Daw, N. D., & Berke, J. D. (2023). Dual credit assignment processes underlie dopamine signals in a complex spatial environment. *Neuron*, *(1)*, 1–23.
- Kravitz, A. V., Freeze, B. S., Parker, P. R. L., Kay, K., Myo, T., Deisseroth, K., & Kreitzer, A. C. (2010). Regulation of parkinsonian motor behaviors by optogenetic control of basal ganglia circuitry. *Nature*, *466*(7306), 622–626. <https://doi.org/10.1038/nature09159.Regulation>
- Kreitzer, A. C., & Malenka, R. C. (2007). Endocannabinoid-mediated rescue of striatal LTD and motor deficits in Parkinson ' s disease models. *Nature*, *445*(1), 643–647. <https://doi.org/10.1038/nature05506>
- Kupferschmidt, D. A., Juczewski, K., Cui, G., Johnson, K. A., & Lovinger, D. M. (2017). Parallel, but Dissociable, Processing in Discrete Corticostriatal Inputs Encodes Skill Learning. *Neuron*, *96*(2), 476–489.e5. <https://doi.org/10.1016/j.neuron.2017.09.040>
- Lahiri, A. K., & Bevan, M. D. (2020). Dopaminergic transmission rapidly and persistently enhances excitability of D1 receptor-expressing striatal projection neurons. *Neuron*, *14*(11), 22274–22330. <https://doi.org/10.1016/j.neuron.2020.01.028.Dopaminergic>
- Le Merrer, J., Detraux, B., Gandía, J., De Groote, A., Fonteneau, M., de Kerchove d'Exaerde, A., & Becker, J. A. J. (2023). Balance Between Projecting Neuronal Populations of the Nucleus Accumbens Controls Social Behavior in Mice. *Biological Psychiatry*. <https://doi.org/https://doi.org/10.1016/j.biopsych.2023.05.008>
- Lebowitz, J. J., & Khoshbouei, H. (2020). Heterogeneity of dopamine release sites in health and degeneration. *Neurobiology of Disease*, *134*, 1–26. <https://doi.org/10.1016/j.nbd.2019.104633>
- Lee, K., Claar, L. D., Hachisuka, A., Bakhurin, K. I., Trott, J. M., Gill, J. L., & Masmanidis, S. C. (2020). Temporally restricted dopaminergic control of reward- conditioned movements. *Nature Neuroscience*, *23*(2), 209–216. <https://doi.org/10.1038/s41593-019-0567-0.Temporally>
- Li, S. J., Sivam, S. P., McGinty, J. F., Huang, Y. S., & Hong, J. S. (1987). Dopaminergic regulation of tachykinin metabolism in the striatonigral pathway. *Journal of Pharmacology and Experimental Therapeutics*, *243*(2), 792–798. <https://jpet.aspetjournals.org/content/243/2/792>
- Liu, C., Cai, X., Ritzau-Jost, A., Kramer, P. F., Li, Y., Khaliq, Z. M., Hallermann, S., & Kaeser, P. S. (2022). An action potential initiation mechanism in

- distal axons for the control of dopamine release. *Science (New York, N.Y.)*, *375*(6587), 1378–1385.
- Liu, H., Melani, R., Sankaramanchi, A., Zeng, R., & Maltese, M. (2022). A permissive role for dopamine in the production of vigorous movements. *bioRxiv*, 1–19.
- Liu, P. G., & Kaeser, P. S. (2021). Spatial and temporal scales of dopamine transmission Changliang. *Physiology & behavior*, *176*(10), 139–148. <https://doi.org/10.1038/s41583-021-00455-7.Spatial>
- Lohse, M. J., Hein, P., Hoffmann, C., Nikolaev, V. O., Vilardaga, J.-P., & Bünemann, M. (2008). Kinetics of g-protein-coupled receptor signals in intact cells. *British Journal of Pharmacology*, *153*(S1), S125–S132. <https://doi.org/https://doi.org/10.1038/sj.bjp.0707656>
- Mackintosh, N. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, *82*, 276–298. <https://doi.org/10.1037/h0076778>
- Manohar, S. G., Chong, T. T., Apps, M. A., Batla, A., Stamelou, M., Jarman, P. R., Bhatia, K. P., & Husain, M. (2015). Reward Pays the Cost of Noise Reduction in Motor and Cognitive Control. *Current Biology*, *25*(13), 1707–1716. <https://doi.org/10.1016/j.cub.2015.05.038>
- Marcott, P. F., Mamaligas, A. A., & Ford, C. P. (2014). Phasic dopamine release drives rapid activation of striatal D2-receptors. *Neuron*, *84*(1), 164–176. <https://doi.org/10.1016/j.neuron.2014.08.058>
- Martinez, V. J., Asico, L. D., Jose, P. A., & Tiu, A. C. (2020). Lipid Rafts and Dopamine Receptor Signaling. *International Journal of Molecular Sciences*, *21*(23). <https://doi.org/10.3390/ijms21238909>
- Masayuki Matsumoto and Okihide Hikosaka. (2014). How do dopamine neurons represent positive and negative motivational events? *Nature*, *23*(1), 1–7. <https://doi.org/10.1038/nature08028.How>
- Matsumoto, M., & Hikosaka, O. (2009). Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature*, *459*(7248), 837–841. <https://doi.org/10.1038/nature08028>
- Mazzoni, P., Hristova, A., & Krakauer, J. W. (2007). Why Don't We Move Faster? Parkinson's Disease, Movement Vigor, and Implicit Motivation. *Journal of Neuroscience*, *27*(27), 7105–7116. <https://doi.org/10.1523/JNEUROSCI.0264-07.2007>
- Mendonça, M. D., Alves Da Silva, J., Hernandez, L. F., Castela, I., Obeso, J., & Costa, R. M. (2021). Transient dopamine neuron activity precedes and encodes the vigor of contralateral movements. *bioRxiv*, 2021.04.20.440527. <https://doi.org/10.1101/2021.04.20.440527>

- 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*, 1–26. <https://doi.org/10.7554/eLife.21886>
- Menegas, W., Bergan, J. F., Ogawa, S. K., Isogai, Y., Umadevi Venkataraju, K., Osten, P., Uchida, N., & Watabe-Uchida, M. (2015). Dopamine neurons projecting to the posterior striatum form an anatomically distinct subclass (S. B. Nelson, Ed.). *eLife*, *4*, e10032. <https://doi.org/10.7554/eLife.10032>
- Miller. (2019). Habits without Values. *psychol rev*, *176*(1), 139–148. <https://doi.org/10.1037/rev0000120.Habits>
- Miyamoto, Y., Katayama, S., Shigematsu, N., Nishi, A., & Fukuda, T. (2018). Striosome-based map of the mouse striatum that is conformable to both cortical afferent topography and uneven distributions of dopamine D1 and D2 receptor-expressing cells. *Brain Structure and Function*, *223*(9), 4275–4291. <https://doi.org/10.1007/s00429-018-1749-3>
- Mohebi, A., Collins, V. L., & Berke, J. D. (2022). Cholinergic interneurons drive motivation by promoting dopamine release in the nucleus accumbens. *bioRxiv*. <https://doi.org/10.1101/2022.11.06.515335>
- Mohebi, A., Pettibone, J. R., Hamid, A. A., Wong, J.-M. T., Vinson, L. T., Patriarchi, T., Tian, L., Kennedy, R. T., & Berke, J. D. (2019). Publisher Correction: Dissociable dopamine dynamics for learning and motivation. *Nature*, *571*(7763), E3–E3. <https://doi.org/10.1038/s41586-019-1353-6>
- Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *Journal of Neuroscience*, *16*(5), 1936–1947. <https://doi.org/10.1523/JNEUROSCI.16-05-01936.1996>
- Morales, M., & Margolis, E. B. (2017). Ventral tegmental area: Cellular heterogeneity, connectivity and behaviour. *Nature Reviews Neuroscience*, *18*(2), 73–85. <https://doi.org/10.1038/nrn.2016.165>
- Morrens, J., Aydin, Ç., Janse van Rensburg, A., Esquívelzeta Rabell, J., & Haesler, S. (2020). Cue-Evoked Dopamine Promotes Conditioned Responding during Learning. *Neuron*, *106*(1), 142–153.e7. <https://doi.org/10.1016/j.neuron.2020.01.012>
- Morris, G., Nevet, A., Arkadir, D., Vaadia, E., & Bergman, H. (2006). Midbrain dopamine neurons encode decisions for future action. *Nature Neuroscience*, *9*(8), 1057–1063. <https://doi.org/10.1038/nn1743>
- Nelson, A., & Killcross, S. (2006). Amphetamine Exposure Enhances Habit Formation. *Journal of Neuroscience*, *26*(14), 3805–3812. <https://doi.org/10.1523/JNEUROSCI.4305-05.2006>

- Nicola, S. M., Surmeier, D. J., & Malenka, R. C. (2000). Dopaminergic modulation of neuronal excitability in the striatum and nucleus accumbens. *Annual Review Neuroscience*, 185–215.
- Nicola, S. M., Yun, I. A., Wakabayashi, K. T., & Fields, H. L. (2004). Cue-Evoked Firing of Nucleus Accumbens Neurons Encodes Motivational Significance During a Discriminative Stimulus Task. *Journal of Neurophysiology*, 91(4), 1840–1865. <https://doi.org/10.1152/jn.00657.2003>
PMID: 14645377
- 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. <https://doi.org/10.1007/s00213-006-0502-4>
- Nomoto, K., Schultz, W., Watanabe, T., & Sakagami, M. (2010). Temporally extended dopamine responses to perceptually demanding reward-predictive stimuli. *Journal of Neuroscience*, 30(32), 10692–10702. <https://doi.org/10.1523/JNEUROSCI.4828-09.2010>
- Owesson-White, C., Belle, A. M., Herr, N. R., Peele, J. L., Gowrishankar, P., Carelli, R. M., & Mark Wightman, R. (2016). Cue-evoked dopamine release rapidly modulates D2 neurons in the nucleus accumbens during motivated behavior. *Journal of Neuroscience*, 36(22), 6011–6021. <https://doi.org/10.1523/JNEUROSCI.0393-16.2016>
- Palmiter, R. D. (2008). Dopamine Signaling in the Dorsal Striatum Is Essential for Motivated Behaviors. *Annals of the New York Academy of Sciences*, 1129(1), 35–46. <https://doi.org/https://doi.org/10.1196/annals.1417.003>
- Panigrahi, B., Martin, K. A., Li, Y., Graves, A. R., Vollmer, A., Olson, L., Mensh, B. D., Karpova, A. Y., & Dudman, J. T. (2015). Dopamine Is Required for the Neural Representation and Control of Movement Vigor. *Cell*, 162(6), 1418–1430. <https://doi.org/10.1016/j.cell.2015.08.014>
- Parker, N. F., Cameron, C. M., Taliaferro, J. P., Lee, J., Choi, J. Y., Davidson, T. J., Daw, N. D., & Ilana B. Witten, 3. (2016). Reward and choice encoding in terminals of midbrain dopamine neurons depends on striatal target Nathan. *Nature Neuroscience*, 176(1), 139–148. <https://doi.org/10.1038/nn.4287>.
Reward
- Pasquereau, B., & Turner, R. S. (2015). Dopamine neurons encode errors in predicting movement trigger occurrence. *Journal of Neurophysiology*, 113(4), 1110–1123. <https://doi.org/10.1152/jn.00401.2014>
- Phillips, P. E., & Stamford, J. A. (2000). Differential recruitment of N-, P- and Q-type voltage-operated calcium channels in striatal dopamine release evoked by 'regular' and 'burst' firing. *Brain Research*, 884(1-2), 139–146. [https://doi.org/10.1016/S0006-8993\(00\)02958-9](https://doi.org/10.1016/S0006-8993(00)02958-9)

- Planert, H., Berger, T. K., & Silberberg, G. (2013). Membrane properties of striatal direct and indirect pathway neurons in mouse and rat slices and their modulation by dopamine. *PLOS ONE*, *8*(3), 1–14. <https://doi.org/10.1371/journal.pone.0057054>
- Poulin, J. F., Caronia, G., Hofer, C., Cui, Q., Helm, B., Ramakrishnan, C., Chan, C. S., Dombeck, D. A., Deisseroth, K., & Awatramani, R. (2018). Mapping projections of molecularly defined dopamine neuron subtypes using intersectional genetic approaches. *Nature Neuroscience*, *21*(9), 1260–1271. <https://doi.org/10.1038/s41593-018-0203-4>
- Puig, M. V., Antzoulatos, E. G., & Miller, E. K. (2014). Prefrontal dopamine in associative learning and memory. *Neuroscience*, *282*, 217–229. <https://doi.org/https://doi.org/10.1016/j.neuroscience.2014.09.026>
- The Ventral Tegmentum and Dopamine: A New Wave of Diversity
- Rescorla, R. (1972). A theory of Pavlovian conditioning : Variations in the effectiveness of reinforcement and nonreinforcement.
- Reynolds, J. N. J., Hyland, B. I., & Wickens, J. R. (2001). A cellular mechanism of reward-related learning. *Nature*, 67–70.
- Richards, J., Mitchell, S., Wit, H., & Seiden, L. (1997). Richards JB, Mitchell SH, de Wit H, Seiden LS. Determination of discount functions in rats with an adjusting-amount procedure. *J Exp Anal Behav* 67: 353-366. *Journal of the experimental analysis of behavior*, *67*, 353–366. <https://doi.org/10.1901/jeab.1997.67-353>
- Richfield, E. K., Penney, J. B., & Young, A. B. (1989). Anatomical and affinity state comparisons between dopamine D1 and D2 receptors in the rat central nervous system. *Neuroscience*, *30*(3), 767–777. [https://doi.org/10.1016/0306-4522\(89\)90168-1](https://doi.org/10.1016/0306-4522(89)90168-1)
- Roesch, M. R., Calu, D. J., & Schoenbaum, G. (2007). Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nature Neuroscience*, *10*(12), 1615–1624. <https://doi.org/10.1038/nn2013>
- Dopamine
- Roitman, M. F., Stuber, G. D., Phillips, P. E., Wightman, R. M., & Carelli, R. M. (2004). Dopamine Operates as a Subsecond Modulator of Food Seeking. *Journal of Neuroscience*, *24*(6), 1265–1271. <https://doi.org/10.1523/JNEUROSCI.3823-03.2004>
- Romo, R., & Schultz, W. (1990). Dopamine neurons of the monkey midbrain: Contingencies of responses to active touch during self-initiated arm movements. *Journal of Neurophysiology*, *63*(3), 592–606. <https://doi.org/10.1152/jn.1990.63.3.592>

- Salamone, J. D., & Correa, M. (2012). The Mysterious Motivational Functions of Mesolimbic Dopamine. *Neuron*, *76*(3), 470–485. <https://doi.org/10.1016/j.neuron.2012.10.021>
- Salamone, J. D., Correa, M., Yang, J.-H., Rotolo, R., & Presby, R. (2018). Dopamine, Effort-Based Choice, and Behavioral Economics: Basic and Translational Research. *Frontiers in Behavioral Neuroscience*, *12*. <https://doi.org/10.3389/fnbeh.2018.00052>
- Satoh, T., Nakai, S., Sato, T., & Kimura, M. (2003). Correlated Coding of Motivation and Outcome of Decision by Dopamine Neurons. *Journal of Neuroscience*, *23*(30), 9913–9923. <https://doi.org/10.1523/jneurosci.23-30-09913.2003>
- Saunders, B. T., Richard, J. M., Margolis, E. B., & Janak, P. H. (2018). Dopamine neurons create Pavlovian conditioned stimuli with circuit-defined motivational properties. *Nat Neurosci*, *176*(1), 139–148. <https://doi.org/10.1038/s41593-018-0191-4.Dopamine>
- Schultz, W. (1986). Activity of pars reticulata neurons of monkey substantia nigra in relation to motor, sensory, and complex events. *Journal of Neurophysiology*, *55*(4), 660–677. <https://doi.org/10.1152/jn.1986.55.4.660>
- Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593–1599. <https://doi.org/10.1126/science.275.5306.1593>
- Schultz, W., & Romo, R. (1990). Dopamine neurons of the monkey midbrain: Contingencies of responses to stimuli eliciting immediate behavioral reactions. *Journal of Neurophysiology*, *63*(3), 607–624. <https://doi.org/10.1152/jn.1990.63.3.607>
- Schultz, W. (2007). Multiple dopamine functions at different time courses. *Annual Review of Neuroscience*, *30*, 259–288. <https://doi.org/10.1146/annurev.neuro.28.061604.135722>
- Schultz, W. (2016). Dopamine reward prediction-error signalling: a two-component response. *Nat Rev Neurosci.*, *176*(12), 139–148. <https://doi.org/10.1038/nrn.2015.26.Dopamine>
- Schultz, W. and A. Ruffieux, & Aebischer, P. (1983). The Activity of Pars Compacta Neurons of the Monkey Substantia Nigra in Relation to Motor Activation. *Experimental Brain Research*, 377–387.
- Shen, W., Flajolet, M., Greengard, P., & Surmeier, D. J. (2008). Dichotomous Dopaminergic Control of Striatal Synaptic Plasticity. *Science*, *23*(1), 1–7. <https://doi.org/10.1126/science.1160575.Dichotomous>
- Skinbjerg, M., Sibley, D. R., Javitch, J. A., & Abi-Dargham, A. (2012). Imaging the high-affinity state of the dopamine D2 receptor in vivo: Fact or fiction?

- Biochemical Pharmacology*, 83(2), 193–198. <https://doi.org/10.1016/j.bcp.2011.09.008>
- Smith, K. S., & Graybiel, A. M. (2013). A dual operator view of habitual behavior reflecting cortical and striatal dynamics. *Neuron*, 79(2), 361–374. <https://doi.org/10.1016/j.neuron.2013.05.038>
- Soares, S., Atallah, B. V., & Paton, J. J. (2016). Midbrain dopamine neurons control judgment of time. *Science*, 354(6317), 1273–1277. <https://doi.org/10.1126/science.aah5234>
- Summerside, E. M., Shadmehr, R., & Ahmed, A. A. (2018). Vigor of reaching movements: reward discounts the cost of effort. *Journal of Neurophysiology*, 119(6), 2347–2357. <https://doi.org/10.1152/jn.00872.2017>
PMID: 29537911
- Sutton, R. S., & Barto, A. G. (1981). Toward a Modern Theory of Adaptive Networks : Expectation and Prediction. *Psychology Reviews*, 88(2), 135–170.
- Sutton, R. S., & Barto, A. G. (1998). Reinforcement Learning: An Introduction. <http://incompleteideas.net/book/>
- 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. <https://doi.org/10.1038/nn.4187>
- Tecuapetla, F., Matias, S., Dugue, G. P., Mainen, Z. F., & Costa, R. M. (2014). Balanced activity in basal ganglia projection pathways is critical for contraversive movements. *Nature Communications*, 5, 1–10. <https://doi.org/10.1038/ncomms5315>
- Thorn, C. A., Atallah, H., Howe, M., & Graybiel, A. M. (2010). Differential Dynamics of Activity Changes in Dorsolateral and Dorsomedial Striatal Loops during Learning. *Neuron*, 66(5), 781–795. <https://doi.org/10.1016/j.neuron.2010.04.036>
- Thorndike, E. L. (1898). An experimental study of the associative processes in animals. *American Psychologist*. <https://doi.org/10.1038/nj7092-544a>
- Thura, D., Cos, I., Trung, J., & Cisek, P. (2014). Context-Dependent Urgency Influences Speed–Accuracy Trade-Offs in Decision-Making and Movement Execution. *Journal of Neuroscience*, 34(49), 16442–16454. <https://doi.org/10.1523/JNEUROSCI.0162-14.2014>
- Thurley, K., Senn, W., & Lüscher, H.-R. (2008). Dopamine Increases the Gain of the Input-Output Response of Rat Prefrontal Pyramidal Neurons. *Journal of Neurophysiology*, 99(6), 2985–2997. <https://doi.org/10.1152/jn.01098.2007>
PMID: 18400958

- Tobler, P. N., Fiorillo, C. D., & Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science*, *307*(5715), 1642–1645. <https://doi.org/10.1126/science.1105370>
- Tricomi, E., Balleine, B. W., & O’Doherty, J. P. (2009). A specific role for posterior dorsolateral striatum in human habit learning. *European Journal of Neuroscience*, *29*(11), 2225–2232. <https://doi.org/https://doi.org/10.1111/j.1460-9568.2009.06796.x>
- Tritsch, N. X., & Sabatini, B. L. (2012). Dopaminergic Modulation of Synaptic Transmission in Cortex and Striatum. *Neuron*, *76*(1), 33–50. <https://doi.org/10.1016/j.neuron.2012.09.023>
- Ungless, M. A., Magill, P. J., & Bolam, J. P. (2004). Uniform Inhibition of Dopamine Neurons in the Ventral Tegmental Area by Aversive Stimuli. *Science*, *303*(5666), 2040–2042. <https://doi.org/10.1126/science.1093360>
- Valjent, E., & Gangarossa, G. (2021). The Tail of the Striatum: From Anatomy to Connectivity and Function. *Trends in Neurosciences*, *44*(3), 203–214. <https://doi.org/10.1016/j.tins.2020.10.016>
- van Elzelingen, W., Warnaar, P., Matos, J., Bastet, W., Jonkman, R., Smulders, D., Goedhoop, J., Denys, D., Arbab, T., & Willuhn, I. (2022). Striatal dopamine signals are region specific and temporally stable across action-sequence habit formation. *Current Biology*, *32*(5), 1163–1174.e6. <https://doi.org/10.1016/j.cub.2021.12.027>
- Wang, W., Dever, D., Lowe, J., Storey, G. P., Bhansali, A., Eck, E. K., Nitulescu, I., Weimer, J., & Bamford, N. S. (2012). Regulation of prefrontal excitatory neurotransmission by dopamine in the nucleus accumbens core. *The Journal of Physiology*, *590*(16), 3743–3769. <https://doi.org/https://doi.org/10.1113/jphysiol.2012.235200>
- Wang, Y., Toyoshima, O., Kunimatsu, J., Yamada, H., & Matsumoto, M. (2021). Tonic firing mode of midbrain dopamine neurons continuously tracks reward values changing moment-by-moment. *eLife*, *10*, 1–17. <https://doi.org/10.7554/eLife.63166>
- 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. <https://doi.org/10.1016/j.neuron.2012.03.017>
- Wei, W., Mohebi, A., & Berke, J. D. (2021). Striatal dopamine pulses follow a temporal discounting spectrum. *BioRxiv*, 1–14.
- Westbrook, A., Frank, M. J., & Cools, R. (2021). A mosaic of cost–benefit control over cortico-striatal circuitry. *Trends in Cognitive Sciences*, *25*(8), 710–721. <https://doi.org/10.1016/j.tics.2021.04.007>
- Wickens, J. R., & Arbuthnott, G. W. (2005). Chapter IV Structural and functional interactions in the striatum at the receptor level. In S. B. Dunnett, M.

- Bentivoglio, A., Björklund, & T. Hökfelt (Eds.), *Dopamine* (pp. 199–236). Elsevier. [https://doi.org/https://doi.org/10.1016/S0924-8196\(05\)80008-9](https://doi.org/https://doi.org/10.1016/S0924-8196(05)80008-9)
- Wise, R. A. (2004). Dopamine, learning and motivation. *Nature Reviews Neuroscience*, *5*(6), 483–494. <https://doi.org/10.1038/nrn1406>
- Witten, I. B., Steinberg, E. E., Lee, S. Y., Davidson, T. J., Zalocusky, K. A., Brodsky, M., Yizhar, O., Cho, S. L., Gong, S., Ramakrishnan, C., Stuber, G. D., Tye, K. M., Janak, P. H., & Deisseroth, K. (2011). Recombinase-driver rat lines: tools, techniques, and optogenetic application to dopamine-mediated reinforcement. *Neuron*, *72*(5), 721–733. <https://doi.org/10.1016/j.neuron.2011.10.028>
- Wolff, S. B., Ko, R., & Ölveczky, B. P. (2022). Distinct roles for motor cortical and thalamic inputs to striatum during motor skill learning and execution. *Science Advances*, *8*(8), 1–15. <https://doi.org/10.1126/sciadv.abk0231>
- Wood, W., & Rünger, D. (2016). Psychology of Habit. *Annual Review of Psychology*, *67*(1), 289–314. <https://doi.org/10.1146/annurev-psych-122414-033417>
PMID: 26361052
- Xiong, Q., Znamenskiy, P., & Zador, A. M. (2015). Selective corticostriatal plasticity during acquisition of an auditory discrimination task. *Nature*, *521*(7552), 348–351. <https://doi.org/10.1038/nature14225>
- Yagishita, S., Hayashi-Takagi, A., Ellis-Davies, G., Urakubo, H., Ishii, S., & Kasai, H. (2014). A critical time window for dopamine actions on the structural plasticity of dendritic spines. *Science*, *61*(6), 515–525. <https://doi.org/10.1126/science.1255514.A>
- Yagishita, S. (2023). Cellular bases for reward-related dopamine actions. *Neuroscience Research*, *188*(November 2022), 1–9. <https://doi.org/10.1016/j.neures.2022.12.003>
- Yamaguchi, K., Maeda, Y., Sawada, T., Iino, Y., Tajiri, M., Nakazato, R., Ishii, S., Kasai, H., & Yagishita, S. (2022). A behavioural correlate of the synaptic eligibility trace in the nucleus accumbens. *Scientific Reports*, *12*(1). <https://doi.org/10.1038/s41598-022-05637-6>
- Yapo, C., Nair, A. G., Clement, L., Castro, L. R., Hellgren Kotaleski, J., & Vincent, P. (2017). Detection of phasic dopamine by D1 and D2 striatal medium spiny neurons. *Journal of Physiology*, *595*(24), 7451–7475. <https://doi.org/10.1113/JP274475>
- Yin, H. H., & Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nature Reviews Neuroscience*, *7*(6), 464–476. <https://doi.org/10.1038/nrn1919>
- Yin, H. H., Knowlton, B. J., & Balleine, B. W. (2004). Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in in-

- strumental learning. *European Journal of Neuroscience*, 19(1), 181–189. <https://doi.org/https://doi.org/10.1111/j.1460-9568.2004.03095.x>
- Yin, H. H., Mulcare, S. P., Hilário, M. R., Clouse, E., Holloway, T., Davis, M. I., Hansson, A. C., Lovinger, D. M., & Costa, R. M. (2009). Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. *Nature Neuroscience*, 12(3), 333–341. <https://doi.org/10.1038/nn.2261>
- Yin, H. H., Ostlund, S. B., Knowlton, B. J., & Balleine, B. W. (2005). The role of the dorsomedial striatum in instrumental conditioning. *European Journal of Neuroscience*, 22(2), 513–523. <https://doi.org/https://doi.org/10.1111/j.1460-9568.2005.04218.x>
- Young, W. S., Bonner, T. I., & Brann, M. R. (1986). Mesencephalic dopamine neurons regulate the expression of neuropeptide mRNAs in the rat forebrain. *Proceedings of the National Academy of Sciences of the United States of America*, 83(24), 9827–9831. <https://doi.org/10.1073/pnas.83.24.9827>
- Zhang, Y., Rózsa, M., Liang, Y., Bushey, D., Wei, Z., Zheng, J., Reep, D., Broussard, G. J., Tsang, A., Tsegaye, G., Narayan, S., Obara, C. J., Lim, J.-X., Patel, R., Zhang, R., Ahrens, M. B., Turner, G. C., Wang, S. S.-H., Korff, W. L., . . . Looger, L. L. (2021). Fast and sensitive gcamp calcium indicators for imaging neural populations. *bioRxiv*. <https://doi.org/10.1101/2021.11.08.467793>

Contents

Abstract	vii
Remerciements	ix
Acronyms	xi
Avant-propos	xiii
Sommaire	xvii
List of Tables	xix
List of Figures	xxi
I Introduction	1
1 Dopamine; from molecule to network	3
1.1 Dopamine as a neurotransmitter	3
1.2 Dopaminergic neurons	4
1.3 DANs electrophysiological properties	4
1.4 Dopamine release	6
1.5 Dopaminergic nuclei	7
1.5.1 VTA	7
1.5.2 SNc	7
1.6 Dopaminergic targets	8
1.6.1 Striatum	8
1.6.2 Direct vs indirect pathway	9
1.6.3 Striatal sub-regions	10
1.6.4 Contribution of the dorsal striatum to goal-directed and habitual behaviors	12
1.7 Dopamine receptors	14

1.7.1	The different effects of D1R and D2R activation on MSN activity	14
1.7.2	Different sensitivity of the D1R and D2R to dopamine levels	17
1.7.3	Functional effect of dopamine on MSNs	18
1.8	From cell plasticity to behavioral learning	19
1.9	Basal Ganglia and motor control	20
2	Dopamine and associative learning	23
2.1	Associative learning	23
2.1.1	Reinforcement learning	24
2.1.2	Rescorla–Wagner model	24
2.1.3	Temporal difference learning model	25
2.2	Dopamine and Reward Prediction Error	27
2.2.1	Value	28
2.2.2	Uncertainty	31
2.2.3	Timing	31
2.3	Alternative hypothesis to RPE	34
2.3.1	'Adjusted net contingency for causal relations' model	34
3	Dopamine and instrumental learning	41
3.1	What is instrumental learning?	41
3.2	Goal-directed vs habitual behavior	42
3.3	Dopaminergic contribution to instrumental behaviors	43
3.3.1	Dopaminergic activity during movement	44
3.3.2	Potential functions of the dopamine activity during movement	52
3.3.3	Cue-evoked dopaminergic response in instrumental tasks	57
3.3.4	Short vs long-term causal effects of dopamine on motor behavior	59
II	Methods	61
4	Data acquisition	63
4.1	Animals	63
4.2	Behavioral training prior to implantation	63
4.3	Surgery	65
4.3.1	Head-fixation implantation surgery	65
4.3.2	Viral injections and optic fiber implantation	65
4.4	Head-fixed behavioral training	67
4.4.1	Early training: spontaneous lever press	67
4.4.2	Increasing task difficulty	67
4.4.3	From training to recordings	68
4.5	Recordings	69

4.5.1 Electrophysiological recordings	69
4.5.2 Fiber photometry recordings	69
5 Data analysis	71
5.1 Behavior	71
5.2 Fiber photometry	71
5.3 Electrophysiology	72
5.4 Histology	72
III Results	75
6 Recordings in head-fixed rats during the instrumental task	77
6.1 General performance of head-fixed rats	77
6.2 Dopaminergic signal in the SNc and DLS	79
6.3 Identification of dopaminergic neurons	81
7 Dopaminergic activity during an instrumental skilled-reaching task	85
7.1 Global dopaminergic activity in the DLS during contralateral trials	85
7.2 Single unit activity in the SNc during contralateral trials	86
7.3 Lateralisation of the dopaminergic activity	88
7.4 Similar movements do not trigger identical dopaminergic response	93
7.5 Dopamine and vigor	94
7.5.1 The cue-evoked and movement-evoked dopaminergic activity correlates with vigor	94
7.5.2 The reward-evoked dopaminergic activity correlates with vigor	97
7.6 Movement repression	100
7.6.1 Missed trials	100
7.6.2 NoGo trials	102
IV Discussion	105
8 Discussion	107
8.1 Recording and identifying dopaminergic neurons in behaving head-fixed rats.	107
8.2 Dopaminergic neurons activity is specialized for task events	108
8.3 Dopaminergic neurons are active during movement	109
8.4 Dopaminergic activity is selectively lateralized	110
8.5 Dopaminergic activity is correlated with movement vigor	112
8.6 Future directions	113

References**115****Contents****133**

Activity of dopaminergic neurons during skilled reaching in head-fixed rats

Abstract

Midbrain dopaminergic neurons (DANs) are essential for associative learning and goal-directed movements. Several studies have recorded their activity during motor tasks and most have observed an increase of dopaminergic activity during movement or around its initiation. However, the functional role played by DANs during instrumental tasks is still unclear. It has been hypothesized that dopamine could modulate movement vigor in an online manner instead of only encoding reward prediction error, but causal experiments have yielded conflicting results. Other theories argue for a role of movement-related dopamine in habit formation or retrospective causal learning, but the extent of dopaminergic contribution to sensory-motor behaviors in physiological conditions is unknown. In this study, we sought to investigate the functional heterogeneity of DANs in the SNc and characterize their activity during a reach-and-grab motor task in TH-Cre rats. We recorded DANs electrical activity and GCaMP calcium signal in head-fixed rats performing the operant task using high-density silicon probes and fiber photometry, respectively. We observed that DANs increased their activity, as measured through global calcium signal, both after the auditory cue and during movement execution. Interestingly, the movement-evoked responses were highly lateralized depending on the limb used. We further dissected the neuronal mechanism causing these increases in calcium transients using *in vivo* electrophysiological recordings. We found that the activity of DANs neurons recorded in the SNc was functionally heterogeneous during the task with individual neurons responding to task events in the specialized-manner. Finally, we observed a correlation between dopaminergic activity and movement vigor, as well as an inhibition of dopaminergic activity during movement repression. Our results reveal the functional diversity present in DANs during a fine operant motor task, and highlight the close relationship between dopamine and motor execution.

Keywords: dopamine, snc, dorsolateral striatum, intrumental skilled-reaching behavior, *in vivo* electrophysiology & calcium imaging, vigor

Activité des neurones dopaminergiques pendant l'exécution d'une tâche comportementale de préhension chez des rats tête-restreinte

Résumé

Les neurones dopaminergiques sont essentiels à l'apprentissage associatif et aux mouvements volontaires. De multiples études ont étudié leur activité lors de tâches motrices et la plupart ont observé une augmentation de l'activité dopaminergique pendant le mouvement ou à son initiation. Cependant, le rôle fonctionnel joué par les neurones dopaminergiques pendant les tâches instrumentales est toujours inconnu. Si on leur attribue classiquement un rôle dans l'apprentissage ainsi que l'encodage de la valeur des récompenses attendues, une hypothèse dominante de cette dernière décennie suggère que les neurones dopaminergiques pourraient également moduler la vigueur des mouvements à venir, mais les preuves causales sont encore limitées. La contribution de la dopamine aux comportements instrumentaux complexes reste à élucider. Dans cette étude, nous avons enregistré des neurones dopaminergiques de la substance noire compacte afin de caractériser leur hétérogénéité fonctionnelle et leur activité lors d'une tâche motrice à choix. Pour cela nous avons mis en place une méthode de tête restreinte chez le rat qui nous permet d'effectuer des enregistrements électrophysiologiques aigus simultanément à des enregistrements de l'activité calcique des terminaisons dopaminergiques dans le striatum via fibre photométrie. À l'échelle de la population, nous avons observé une activité dopaminergique en réponse au signal sonore ainsi que durant le mouvement et à la consommation de la récompense. À l'échelle des neurones individuels, nous avons observé une grande spécialisation de leur activité. Nous avons montré que l'activité dopaminergique était latéralisée spécifiquement pendant le mouvement, et que cette activité ainsi que celle qui précède le mouvement était corrélée à sa vigueur. Nos résultats révèlent la diversité fonctionnelle des neurones dopaminergiques pendant une tâche comportementale complexe, et mettent en valeur les liens entre la dopamine et l'exécution motrice.

Mots clés : dopamine, snc, striatum dorsolatéral, comportement instrumental de préhension, électrophysiologie & imagerie calcique in vivo, vigueur
